English Historical Linguistics

Change in structure and meaning

EDITED BY

Bettelou Los
Claire Cowie
Patrick Honeybone
Graeme Trousdale

ENGLISH HISTORICAL LINGUISTICS

CURRENT ISSUES IN LINGUISTIC THEORY

General Editor

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 *benjamins.com/catalog/cilt*

Volume 358

Bettelou Los, Claire Cowie, Patrick Honeybone and Graeme Trousdale (eds.)

English Historical Linguistics. Change in structure and meaning
Papers from the XXth ICEHL

ENGLISH HISTORICAL LINGUISTICS

CHANGE IN STRUCTURE AND MEANING
PAPERS FROM THE XXTH ICEHL

Edited by

BETTELOU LOS
CLAIRE COWIE
PATRICK HONEYBONE
GRAEME TROUSDALE
University of Edinburgh

JOHN BENJAMINS PUBLISHING COMPANY
AMSTERDAM & PHILADELPHIA

∞™ The paper used in this publication meets the minimum requirements of
the American National Standard for Information Sciences – Permanence
of Paper for Printed Library Materials, ANSI z39.48-1984.

Table of contents

Acknowledgements

Sciences of the university for agreeing to opening the conference. Last but not least, the conference organization is deeply indebted to Steven McGauley of the School of Philosophy, Psychology and Language Sciences for financial advice, and Julie Anderson from the School's office for help with innumerable details, from setting up the ePay system to booking the buses for the social programme and ordering the contents of the conference packs.

Introduction

English Historical Linguistics at 20 ICEHLs

Bettelou Los and Patrick Honeybone
University of Edinburgh

Research on the history of English continues apace. Some of this work breaks new empirical ground, collecting novel evidence for change in the language from all stages of its existence, and other work reinterprets classic data, showing a new way to understand issues that have long intrigued English historical linguists. Some of this work is fundamentally philological, with its prime aim being to set out new discoveries about English, while other work aims wholeheartedly to interact with debates in general linguistics on how language can change in principle (both learning from and contributing to them). We are delighted that this volume contains research in all these areas (at the phonological, morphological, syntactic, semantic and pragmatic levels), nicely representing the diversity that exists in the current landscape of English historical linguistics. The articles gathered here are all based on presentations delivered at the 20th International Conference on English Historical Linguistics (ICEHL), which took place at the University of Edinburgh on 27–31 August 2018.

With the twentieth instalment of the conference, 2018 was a crown year for the ICEHL, and we were excited to be able to hold it in Edinburgh, which has a long association with research in English historical linguistics. It was Charles Jones who founded the conference series in 1979, when he had just left Edinburgh to take up an appointment as the Chair of English Language at Durham University. ICEHL complemented another conference series that had been (co-)founded by Charles Jones (with John Anderson), in 1973, when they were both at Edinburgh: the International Conference on Historical Linguistics (ICHL). Both conferences have thrived ever since, as has a third conference series also founded by Charles Jones (once he had returned to Edinburgh to take up the Forbes Chair of English Language), the International Conference on Late Modern English, the first instalment of which took place in Edinburgh in 2001.

That first ICEHL, in 1979, attracted around a dozen people, but clearly a good time was had by all, because the conference series took off, to become the large event that it is today – the central fixture on the English Historical Linguistics calendar.

Around 300 people attended the conference in 2018 in Edinburgh, from all of the world, showing both the popularity and the global reach of the ICEHL today. In time, the conference settled on a pattern of being held in even-numbered years (alternating with the ICHL in odd-numbered years), travelling round European universities, as set out in Table 1. At the time of writing, ICEHL-21 has had to break this biennial pattern because it needed to be postponed by a year due to the COVID-19 pandemic – it will now take place online, organised by colleagues at Leiden University. Writing this introduction in the midst of the pandemic, we are relieved that the Edinburgh conference occurred well before we had heard of coronaviruses (other than the common cold), and we very much look forward to a time when the ICEHL can return to meeting in person.

Table 1. Place and years of ICEHL conferences, with references to their proceedings

	Year	Place	Publications
ICEHL-1	1979	Durham	Blake & Jones (1984)
ICEHL-2	1981	Odense	Davenport et al. (1983)
ICEHL-3	1983	Sheffield	Crépin (1984)*
ICEHL-4	1985	Amsterdam	Eaton et al. (1985)
ICEHL-5	1987	Cambridge	Adamson et al. (1990)
ICEHL-6	1990	Helsinki	Rissanen et al. (1992)
ICEHL-7	1992	Valencia	Fernández et al. (1994)
ICEHL-8	1994	Edinburgh	Britton (1996)
ICEHL-9	1996	Poznań	Fisiak & Krygier (1998)
ICEHL-10	1998	Manchester	Bermúdez-Otero et al. (2000)
ICEHL-11	2000	Santiago de Compostela	Fanego et al. (2002a) Fanego et al. (2002b)
ICEHL-12	2002	Glasgow	Kay et al. (2004a) Kay et al. (2004b)
ICEHL-13	2004	Vienna	Ritt et al. (2006) Dalton-Puffer et al. (2006)
ICEHL-14	2006	Bergamo	Dossena et al. (2008) Dury et al. (2008) Gotti et al. (2008)
ICEHL-15	2008	Munich	Lenker et al. (2010) Sauer & Waxenberger (2012)
ICEHL-16	2010	Pécs	Hegedűs & Fodor (2012)
ICEHL-17	2012	Zurich	Jucker et al. (2013) Pfenninger et al. (2014)
ICEHL-18	2014	Leuven	Cuyckens et al. (2018)
ICEHL-19	2016	Essen	
ICEHL-20	2018	Edinburgh	
ICEHL-21	2021	Leiden	

* Only some of the papers gathered in this volume were delivered at the Sheffield ICEHL.

The conference has from its very beginnings been marked by collegiality and collaboration, reflecting the latest research into the history of English, from a wealth of different theoretical frameworks and approaches. It has spawned many publications, not only as conference proceedings from its general sessions (generally, but not invariably, published by John Benjamins, which has been a constant supporter and sponsor of ICEHL), but also as special journal issues and published volumes, based on papers given in its many workshops. Although the conference venues have so far been restricted to Europe, this is not true of its participants, who come from all over the world, and who testify to the worldwide interest in the English language, and in its history.

1.   Themes

The traditional heartlands of English historical linguistics were well represented at the 20th ICEHL conference, and papers from many of them have made it into this volume. There was representation of work on historical phonology (see, e.g., Gjertrud F. Stenbrenden's paper in this volume), on morphological change (see, e.g., Don Ringe & Charles Yang's paper included here), on morphosyntactic change including grammaticalisation, particularly of auxiliaries (see, e.g., Lilo Moessner's paper in this volume for a contribution to charting their development), on manuscript evidence, and on contact linguistics, as well as on the still relatively new fields of standardisation, prescriptivism, historical pragmatics and historical speech acts. This introduction does not have the space to mention all the 224 papers or the nine workshops that made up the conference at Edinburgh, but we will highlight some new trends in English historical linguistics, representing avenues of research in the field that are opening up (to accompany more established subfields). The clearest signs of new ground are evident in the themes of the workshops that formed part of the conference: interfaces, new methodologies and new tools, and a new field in English historical linguistics: paratextual studies.

Two workshops investigated interfaces – that between segmental phonology and prosody ("The foot in the phonological history of English", see, e.g., Elan Dresher & Aditi Lahiri's paper), and that between syntax and semantics ("Degree phenomena in the history of English"). A paper from a third workshop ("English as a syntactic outlier") contributes to an exploration of the interface between syntax, information structure and prosody (De Bastiani's paper in this volume).

Another trend, outside the workshops, was represented by a significant clutch of papers that reported on phenomena investigated over very large time frames, including the history of English in its entirety; these papers were able to take the long view because of the increasing quality and accessibility of digital tools. Two of the nine workshops focused on those tools: "Visualisations in Historical Linguistics",

which showcased what can be done with data-driven approaches like n-gramming and correspondence analysis, and with specific tools like Stylo, HistoBankVis, TVE2 (Text Variation Explorer) and Medusa (a spiderweb-visualisation of spellings and sounds, as a tool for exploring the phonology of under-researched languages), as well as that other powerful new resource, the *Historical Thesaurus of English*; and the workshop "Computational approaches to investigating meaning in the history of English", which similarly included work on data-driven methods (like Gerold Schneider's paper). A further workshop at the conference positioned itself as a re-action to the pervasive quantitative methodologies and broke a lance for developing a methodology for qualitative evidence ("Qualitative evidence and methodologies in historical linguistics"). There was also a trend among some papers in the general sessions in which new methods are applied to old problems, e.g., Aaron Ecay's use of insights from psycholinguistic priming studies to determine whether affirmative declarative *do* and the NICE contexts share an underlying syntax; and Don Ringe's plenary, where the replacement of regular past tenses and past participles (such as *stringed*) by irregular ones (such as *strung*) in Early Modern English is investigated using the mathematical model of Yang's Tolerance Principle (this formed the basis of the paper by Don Ringe & Charles Yang in this volume).

Finally, there was a workshop at the conference on 'paratextual communication' – the way in which layout conventions arose in printing, like the footnote, or text boxes in a pamphlet, which have developed communicative functions of their own. Some related papers in the general sessions were on the development of specific genres – for example, the patent specification genre, and minute-writing as a text-type.

2.   Other publications proceeding from ICEHL-20

Apart from a companion volume *English Historical Linguistics: Historical English in Contact*, edited by Bettelou Los, Chris Cummins, Lisa Gotthard, Alpo Honkapohja and Benjamin Molineaux, also published in the CILT series, some of the conference's workshops led to edited volumes and special journal issues:

Some of the papers presented at Workshop 3, "Computational approaches to investigating meaning in the history of the English language", will be published as a special issue of the *Transactions of the Philological Society*, edited by Susan Fitzmaurice and Seth Mehl.

Papers from Workshop 5, "Paratextual Communication in a Historical Linguistic Perspective", were published in 2020 as a volume edited by Matti Peikola and Birte Bös, under the title *The Dynamics of text and framing phenomena: Historical*

approaches to paratext and metadiscourse in English (Pragmatics & Beyond New Series, 317), by John Benjamins (https://benjamins.com/catalog/pbns.317).

Workshop 6, "Degree phenomena in the history of English" has been published (2021) as a double special issue of the *Journal of English Linguistics* (vol. 49, issues 1–2), with the title "Degree and related phenomena in the history of English", edited by Claudia Claridge and Merja Kytö.

Workshop 7, "Visualisations in Historical Linguistics", has been published as a special issue of the *Journal of Data Mining and Digital Humanities* in 2020, edited by Benjamin Molineaux, Bettelou Los and Martti Mäkinen (https://jdmdh.episciences. org/page/special-issue-on-visualisations-in-historical-linguistics).

3.   The papers in this volume

On the 224 papers that we presented at Edinburgh during the 20th ICEHL, 43 were submitted for publication in these general proceedings. A selection of these was made on the basis of extended abstracts, in order to create thematically coherent volumes, and these submissions were then subjected to peer review. Those papers which made it through this process make up this volume, and also its companion volume, *English Historical Linguistics: Historical English in Contact. Papers from the XXth ICEHL*, Volume 2.

The thirteen papers in this volume address various aspects of grammatical structure and linguistic meaning, and we have grouped them loosely into three parts: those which deal with aspects of phonology and morphology, those which deal with aspects of syntax, and those which deal with aspects of meaning (such as semantics and pragmatics). The first part on phonology and morphology contains four papers:

Gjertrud F. Stenbrenden's "Grimm's Law and Verner's Law: Towards a unified phonetic account" reconsiders the patterning and causation of Grimm's Law and Verner's Law, to investigate the extent to which a unified description of the changes involved can be given: are all the changes fundamentally the same kind of thing (and are they all lenitions)? And what exactly were the changes in the first place (and the pre-change states)? The author reviews a wide range of classic and current literature on the changes, from both phonological and phonetic perspectives, and fundamentally adopts the position (often known as 'Laryngeal Realism') that Germanic, like English and most other contemporary Germanic languages, marked the contrast between its two series of obstruents using the feature [spread glottis], not [voice]. Stenbrenden relates this, and the associated aspiration in fortis stops, to the change from Proto-Indo-European's pitch accent to Germanic stress accent,

arguing that this change was fundamental in setting off the chain of changes which she interprets Grimm's and Verner's Laws to be.

B. Elan Dresher & Aditi Lahiri's "The foot in the history of English: Challenges to metrical coherence" addresses the classic issue of change in the stress patterns of English, based on a consideration of the role of the foot: what exactly was the basic English foot during this period? How was it assigned? Which foot received primary stress? The authors have a broad focus, taking in changes that occurred over the course of many centuries, from Old English, through Middle English, to the Early Modern period when the effect of Romance loanwords was fully felt. The authors consider how the fundamental principles of stress assignment changed in English over this period, using evidence from a range of sources, including patterns in phonological change and transcriptions in pronouncing dictionaries, and engaging with issues in stress assignment from theoretical phonology. A key argument is that Yang's Tolerance Principle (also discussed by Ringe & Yang in Chapter 4 of this volume) explains why Romance loanwords did not have a fundamental effect on the phonology of the foot in English (and hence on stress assignment) until much later than is often thought.

Mieko Ogura & William S-Y. Wang's "Ambiguity resolution and the evolution of homophones in English" also considers change in the stress patterns of English, but from a very different perspective from that of the previous chapter. Ogura & Wang focus on the issue of the development of diatonic pairs (that is, stress doublets of the kind that exist in Present-Day English noun and verb forms of words like *permit*). The authors consider evidence that derives from pronouncing dictionaries from earlier periods of English, and also present novel contemporary neurolinguistic evidence (from near-infrared spectroscopy) which shows differences in the processing of nouns and verbs. They argue from a functionalist perspective that the noun-type non-final stress developed in diatonic pairs due to a pressure to avoid homophony, which they relate to functional pressures to comply with speakers' and hearers' needs, arguing that pressures from production led in this process in the 16th century, and that pressures from perception led in the process after the 17th century.

In "The threshold of productivity and the 'irregularization' of verbs in Early Modern English" by Don Ringe & Charles Yang, the interplay of language change and language acquisition in the domain of inflectional morphology is investigated, drawing on Yang (2016)'s Tolerance Principle (TP) as a statistical measure of a productivity threshold. In particular, it examines data from the Penn-Helsinki Parsed Corpus of Early Modern English (PPCEME; Kroch et al. 2004) relating to variation in the past tense form of verbs whose infinitive form ends in /-ɪŋ/, e.g., Modern English *bring-brought* vs. *sing-sang* vs. *ding-dinged* vs. *sting-stung*. Ringe & Yang explore

whether the TP can help to explain patterns in PPCEME, and conclude that the TP is useful for predicting likely productive morphological rules, while acknowledging complexities that arise from competition between closely related forms. They consider the role of adult speakers in the development of particular innovations.

Part 2, Syntax, contains the following five papers:

Chiara De Bastiani's "The reanalysis of VO in the history of English: Evidence for a language-internal account" investigates the change from OV to VO by considering the information-structural status (given or new) and weight of pre- or postverbal constituents in subclauses in selected Early Middle English texts. Pre-verbal elements turn out to be a more homogeneous set than post-verbal elements, as they are overwhelmingly given, while post-verbal elements are neither homogeneous in terms of given/new, or in terms of weight. Some Middle English dialects are clearly more advanced in terms of their progression to VO than others; and the investigation shows that charting this progress in the dialects can be obscured if the investigation fails to exclude texts that are Middle English copies of earlier material rather than authentic Middle English. The findings have consequences for earlier claims about the interaction of information status and weight as a trigger for particular word orders.

In "The role of (the avoidance of) centre embedding in the change from OV to VO in English" by Rodrigo Pérez Lorido, the object of the investigation is centre-embedded structures like (*The man* [*the boy* [*the woman saw*] *heard*] *left*). These structures are difficult to process, as they make a heavy demand on short-time memory. We find various escape hatches cross-linguistically; for instance, in OV languages with postnominal relatives, like Old English, object modifiers but also the object in its entirety can be extraposed. There have been various proposals in the literature that Relative Clause Extraposition as a means to avoid centre-embedded structures is relevant to the change of OV to VO in English. This investigation is the first to probe the limits of clausal embedding in Old English, based on an extensive corpus analysis. Centre embedding turns out to have been a viable grammatical option in Old English, and its selection appears to operate in much the same force field of factors that govern embedded relative objects in other modern SOV languages, like Modern German. The decrease in the frequency of preverbal relative objects from early to late Old English supports the claim that the change of OV to VO was already on the way in late Old English.

Centre-stage in Gerold Schneider's "Syntactic changes in verbal clauses and noun phrases from 1500 onwards" are data-driven methodologies on parsed diachronic corpora. Findings from the ARCHER corpus are validated against the Penn

Corpora, investigating frequency and creativity (i.e., the extent to which the language is 'chunked' and/or formulaic). The results align with what we know about syntactic trends in the history of English style; since the end of the Middle English period, the 'verbal' (sometimes called 'Doric') style is decreasing, while the 'nominal' ('Attic') style is increasing; there is a shift from a preponderance of finite clauses to non-finite clauses, and of parataxis to hypotaxis. Word order is increasingly becoming fixed, strengthening the principle of dependency length minimisation. The difficulty of data-driven approaches is how to interpret the results in terms of linguistic or stylistic change; in the case of English, the results could be matched, and validated, by what we know of EModE and LModE change from the literature. The advantage these approaches offer, Schneider argues, is a holistic perspective which is able to detect long-term diachronic trends in the face of short-term synchronic linguistic variation, and hence opens up new avenues of research.

Eva Zehentner & Marianne Hundt's "Prepositions in Early Modern English: Argument structure and beyond" examines the development of prepositional arguments, focussing mainly on English in modern times. Zehentner & Hundt consider the various roles that prepositional phrases (PPs) play in early Modern English, and in contemporary world Englishes. Using a range of corpora, they provide a quantitative analysis of the frequency of prepositions at various stages in the language, and across varieties. The findings suggest that PPs are increasingly used as verbal arguments (with integration resulting in the development of 'prepositional verbs' such as *rely on*), but that there is also significant lexical variation. In addition, while certain patterns in World Englishes reflect uses attested in early modern corpora, there does not seem to be evidence that speakers of English as a second or other language are more likely to use PP variants (over NP alternatives) compared to native speakers.

Lilo Moessner's "*Should* with non-past reference: A corpus-based diachronic study" explores the rise of *should* in mandative constructions, as an alternative to the subjunctive (compare *It's vital that John be there* vs. *It's vital that John should be there*) in the history of English, focussing particularly on uses of *should* in two main corpora (the Helsinki Corpus of English Texts, and the Middle English Dictionary). The rise of the mandative use of *should* represents a semantic split between the present and the past tense forms of this modal, so that this paper contributes to the literature on the grammaticalization of the modals in English. Moessner argues that this development is strongly connected to other developments in the history of *should*, specifically a set of semantic changes relating to the marking of obligation vs. non-factivity, which calls for a consideration of the range of syntactic contexts in which the subjunctive appeared historically. Her findings suggest that the competition between the subjunctive and *should* in mandative constructions is more recent

than has been assumed in the literature, as it cannot be traced back any earlier than the late Middle English period.

The final part of the volume, Semantics and Pragmatics, contains four papers:

Gabriella Mazzon's "Shifting responsibility in passing information: Stance-taking in Sir Thomas Bodley's diplomatic correspondence" explores diplomatic correspondence at the turn of the 16th century by investigating the letters of Thomas Bodley, master spy to Queen Elizabeth I, concerning his missions to the Netherlands, Germany, Denmark and France, in which he reports on Protestant revolts and their implications for trade. It was important for Bodley to make his reports credible and useful, but at the same time, he needed to not sound overconfident, so as not to lose face if he was disproved. In some cases he names a source, and in others he alludes to rumour or the "bruit". For scholars this is an excellent opportunity to explore the Early Modern expression of stance towards information, and stance towards sources. Mazzon surveys a wide range of epistemic and evidential markers which Bodley deploys to this end: epistemic modals, main verbs of cognition, hedges, direct and indirect speech. Extensive quotation shows the density of these markers, which is a distinction of the register.

James Hyett & Carol Percy's "Theatrical practices and grammatical standardization in eighteenth-century Britain: YOU WAS and YOU WERE" contributes to our knowledge of YOU WAS, a short-lived construction of the first half of the eighteenth century, which plays a role in the shift of the YOU pronoun from plural to singular, and of YOU WERE from plural to singular. YOU WAS shares some of the profile of THOU as the low status variant. The peak and rapid dropping off of YOU WAS around the 1750s has been observed in letters and fiction, before its proscription by figures such as Bishop Lowth in the 1760s. Hyett & Percy focus on its appearance in drama between 1740 and 1760. The authors innovatively adapt large corpora such as LION and ECCO (not specifically designed for (historical) linguistics) to build a subcorpus of mostly comedies for their purposes. Given their preoccupations with the accumulation of wealth, and the acquisition of social status, the comedies are full of social stereotypes, and YOU WAS is used extensively to index low status characters, but also figures in attempts by high status characters to get close to these characters. The detailed qualitative analysis of the comedies reveals that YOU WAS and YOU WERE vary within plays, within scenes, and within characters, highlighting the usefulness of this genre for historical sociolinguistic research.

Anne-Christine Gardner's "Towards a companionate marriage in Late Modern England? Two critical episodes in Mary Hamilton's courtship letters to John Dickenson" analyses letters from the Hamilton-Dickenson courtship, and shows how the couple's exchanges respond to eighteenth century notions of equality and

companionship in courtship and marriage. This genre has been explored before, but the article highlights interesting differences between this couple and their contemporaries. The power balance between the two is negotiated through epistolary 'crises', shown here in two episodes at different stages of the courtship. Gardner is concerned not just with the text but the material written object, particularly Anne's self-censorship at the time of writing, and many years later in the construction of her personal archive. We also see how the mechanics of letter writing and the timing of exchanges shape the relationship. The Critical Discourse Analysis approach that Gardner adopts allows the micro- and macro-contexts to be brought together through concepts like intertextuality and materiality.

Ekkehard König & Letizia Vezzosi's "On the development of OE *swā* to ModE *so* and related changes in an atypical group of demonstratives" builds on previous comparative studies and comparative reconstructions, as well as corpus investigations, to trace the syntactic and semantic development of OE *swā, swylc* and *þus* to Modern English *so, such* and *thus*. The chapter argues that these lexemes form a distinct set, best described as 'demonstratives of manner, quality and degree'. Starting out from a basic exophoric (gestural) use and its typical extensions to anaphoric and cataphoric uses, these expressions develop into a wide variety of grammatical markers in ModE.

4.   Conclusion

The papers in this volume show very clearly that research on the history of English continues apace, both in terms of investigations using classic, well-tried frameworks and methodologies and in terms of research using novel, often interdisciplinary methods. All periods remain popular objects of research; this volume contains work on, in effect, the complete history of the language, from Germanic through to Old English, Middle English, Early Modern English and Late Modern English. The chapters showcase a range of frameworks, with linguistic change investigated through the lenses of generative theory, functional approaches, computational approaches, pragmatics, and Critical Discourse Analysis. We see this diversity as a great strength, and we look forward to the next twenty ICEHLs (and to all the ICEHLs after that, too) with excitement.

References

Adamson, Sylvia M., Vivien A. Law, Nigel Vincent & Susan Wright (eds.). 1990. *Papers from the 5th International Conference on English Historical Linguistics* (Current Issues in Linguistic Theory 65). Amsterdam: John Benjamins. https://doi.org/10.1075/cilt.65

Bermúdez-Otero, Ricardo, David Denison, Richard M. Hogg & C. B. McCully (eds.). 2000. *Generative theory and corpus studies: A dialogue from 10ICEHL*. Berlin: Mouton de Gruyter. https://doi.org/10.1515/9783110814699

Blake, Norman F. & Charles Jones (eds.). 1984. *English historical linguistics: Studies in development* (CECTAL Conference Papers Series 3). Sheffield: University of *Sheffield*.

Britton, Derek (ed.). 1996. English Historical Linguistics 1994: Papers from the 8th *International Conference on English Historical Linguistics*, Edinburgh, 19–23 September 1994. Amsterdam: John Benjamins. https://doi.org/10.1075/cilt.135

Crépin, André (ed.). 1984. *Linguistic and stylistic studies in medieval English*. Paris: Association des Médiévistes anglicistes de l'enseignement supérieur (publication 10).

Cuyckens, Hubert, Hendrik De Smet, Liesbet Heyvaert & Charlotte Maekelberghe (eds.). 2018. *Explorations in English historical syntax* (Studies in Language Companion Series 198). Amsterdam: John Benjamins. https://doi.org/10.1075/slcs.198

Dalton-Puffer, Christiane, Dieter Kastovsky, Herbert Schendl & Nikolaus Ritt (eds.). 2006. *Syntax, style and grammatical norms* (Linguistic insights 13). Frankfurt: Lang. https://doi.org/10.3726/978-3-0351-0272-7

Davenport, Michael, Erik Hansen & Hans Frede Nielsen (eds.). 1983. *Current topics in English historical linguistics: Proceedings of the Second International Conference on English Historical Linguistics held at Odense University, 13–15 April 1981* (Odense University studies in English 4). Odense: Odense University.

Dossena, Marina, Richard Dury & Maurizio Gotti (eds.). 2008. *English historical linguistics 2006: Selected papers from the fourteenth International Conference on English Historical Linguistics (ICEHL 14), Bergamo, 21–25 August 2006*. Volume 3: Geo-Historical Variation in English (Current Issues in Linguistic Theory 297). Amsterdam: John Benjamins. https://doi.org/10.1075/cilt.297

Dury, Richard, Maurizio Gotti & Marina Dossena (eds.) 2008. *English historical linguistics 2006: Selected papers from the fourteenth International Conference on English Historical Linguistics (ICEHL 14), Bergamo, 21–25 August 2006*. Volume 2: Lexical and Semantic Change (Current Issues in Linguistic Theory 296). Amsterdam: John Benjamins. https://doi.org/10.1075/cilt.296

Eaton, Roger, Olga Fischer, Willem F. Koopman & Frederike van der Leek (eds.). 1985. *Papers from the 4th International Conference on English Historical Linguistics, Amsterdam, April 10–13, 1985* (Current Issues in Linguistic Theory 41). Amsterdam: John Benjamins. https://doi.org/10.1075/cilt.41

Fanego, Teresa, Javier Pérez-Guerra & María José López-Couso (eds.). 2002a. *English historical syntax and morphology: Selected papers from 11 ICEHL, Santiago de Compostela, 7–11 September 2000*, Volume 1. (Current Issues in Linguistic Theory 223). Amsterdam: John Benjamins. https://doi.org/10.1075/cilt.223

Fanego, Teresa, Belén Méndez-Naya & Elena Seoane (eds.). 2002b. *Sounds, words, texts and change: Selected papers from 11 ICEHL, Santiago de Compostela, 7–11 September 2000*, Volume 2. (Current Issues in Linguistic Theory 224). Amsterdam: John Benjamins. https://doi.org/10.1075/cilt.224

Fernández, Francisco, Miguel Fuster Márquez & Juan José Calvo (eds.). 1994. Papers from the 7th International Conference on English Historical Linguistics, Valencia, 22–26 September 1992 (Current Issues in Linguistic Theory 113). Amsterdam: John Benjamins. https://doi.org/10.1075/cilt.113

Fisiak, Jacek & Marcin Krygier. 1998. Advances in English Historical Linguistics (Trends in Linguistics, Studies and Monographs 112). Berlin: Mouton de Gruyter.

Gotti, Maurizio, Marina Dossena & Richard Dury (eds.) 2008. *English historical linguistics 2006: Selected papers from the fourteenth International Conference on English Historical Linguistics (ICEHL 14), Bergamo, 21–25 August 2006.* Volume 1: Syntax and Morphology (Current Issues in Linguistic Theory 295). Amsterdam: John Benjamins. https://doi.org/10.1075/cilt.295

Hegedűs, Irén & Alexandra Fodor (eds.). 2012. *English historical linguistics 2010: Selected Papers from the Sixteenth International Conference on English Historical Linguistics (ICEHL 16), Pécs, 23–27 August 2010* (Current Issues in Linguistic Theory 325). Amsterdam: John Benjamins. https://doi.org/10.1075/cilt.325

Jucker, Andreas H., Daniela Landert, Annina Seiler & Nicole Studer-Joho (eds.). 2013. *Meaning in the history of English: Words and texts in context* (Studies in Language Companion Series 148). Amsterdam: John Benjamins. https://doi.org/10.1075/slcs.148

Kay, Christian, Simon Horobin & Jeremy J. Smith (eds.). 2004a. *New perspectives on English historical linguistics: Selected papers from 12 ICEHL, Glasgow, 21–26 August 2002.* Volume 1: Syntax and Morphology (Current Issues in Linguistic Theory 251). Amsterdam: John Benjamins.

Kay, Christian, Carole Hough & Irené Wotherspoon (eds.). 2004b. *New perspectives on English historical linguistics: Selected papers from 12 ICEHL, Glasgow, 21–26 August 2002.* Volume 2: Lexis and Transmission (Current Issues in Linguistic Theory 252). Amsterdam: John Benjamins.

Kroch, Anthony, Beatrice Santorini & Lauren Delfs. 2004. Penn-Helsinki Parsed Corpus of Early Modern English. http://www.ling.upenn.edu/hist-corpora/

Lenker, Ursula, Judith Huber & Robert Mailhammer (eds.). 2010. *English Historical Linguistics 2008: Selected papers from the fifteenth International Conference on English Historical Linguistics (ICEHL 15), Munich, 24–30 August 2008.* Volume 1: The history of English verbal and nominal constructions (Current Issues in Linguistic Theory 314). Amsterdam: John Benjamins. https://doi.org/10.1075/cilt.314

Pfenninger, Simone E., Olga Timofeeva, Anne-Christine Gardner, Alpo Honkapohja, Marianne Hundt & Daniel Schreier (eds.) 2014. *Contact, variation, and change in the history of English* (Studies in Language Companion Series 159). Amsterdam: John Benjamins. https://doi.org/10.1075/slcs.159

Rissanen, Matti, Ossi Ihalainen & Terttu Nevalainen (eds.). 1992. *History of Englishes. New methods and interpretations in historical linguistics* (Topics in English Linguistics 10). Berlin: Mouton de Gruyter. https://doi.org/10.1515/9783110877007

Ritt, Nikolaus, Herbert Schendl, Christiane Dalton-Puffer & Dieter Kastovsky (eds.). 2006. *Medieval English and its heritage: Structure, meaning and mechanisms of change* (Studies in English Medieval Language and Literature 16). Frankfurt: Lang.

Sauer, Hans & Gaby Waxenberger (eds.). 2012. *English Historical Linguistics 2008: Selected papers from the fifteenth International Conference on English Historical Linguistics (ICEHL 15), Munich, 24–30 August 2008.* Volume 2: Words, texts and genres (Current Issues in Linguistic Theory 324). Amsterdam: John Benjamins. https://doi.org/10.1075/cilt.324

Yang, Charles. 2016. *The price of linguistic productivity: How children learn to break the rules of language.* Cambridge (MA): MIT Press. https://doi.org/10.7551/mitpress/9780262035323.001.0001

Phonology and morphology

Grimm's Law and Verner's Law

Towards a unified phonetic account

Gjertrud F. Stenbrenden
University of Oslo / Inland Norway University of Applied Sciences

This paper gives a unified account of Grimm's and Verner's Laws in light of findings from experimental phonetics. The Germanic stress shift and stress placement shift are separate phenomena, and I argue that Iverson & Salmons' (2003) shift in 'articulatory setting' corresponds to the former, and that the shift in how prosodic emphasis was expressed, from high pitch to dynamic stress, set Grimm's Law in motion, because a phonetic correlate of dynamic stress is higher subglottal pressure. Increased subglottal pressure induced aspiration, affrication and spirantisation in Proto-Indo-European (PIE) voiceless stops and devoicing in voiced stops. The voiced aspirates became fricatives, which were allophones of voiced stops; these fricative allophones later fell together with the main plosive allophones to produce Germanic voiced stops.

Keywords: Grimm's Law, Verner's Law, Germanic enhancement, aspiration, experimental phonetics

1. Introduction

The systematic changes to stop consonants in Germanic (Gmc) were among the first to be formulated as sound-laws, i.e., 'Grimm's Law' and 'Verner's Law', and their nature has long since been recognised (Rask 1818; Grimm 1822–1837; Verner 1875).

Essentially, Grimm's Law states that (a) Indo-European (IE) voiceless stops became Gmc voiceless fricatives, (b) IE voiced stops became Gmc voiceless stops, and (c) IE voiced aspirated stops became Gmc voiced (unaspirated) stops; Grimm's Law did not apply when the consonants were preceded by *s*. It is usually assumed that the IE voiceless stops were aspirated. Thus, Grimm's Law clearly affected the stops' manner and force of articulation, whereas their place of articulation remained unchanged. Grimm's Law is laid out in (1), and in a simplified manner in (2) below.

(1)   Grimm's Law:
      IE voiceless stops /pʰ tʰ kʰ kʰʷ/ > Gmc voiceless fricatives /f θ x xʷ/
      IE voiced stops /b d g gʷ/ > Gmc voiceless stops /p t k kʷ/
      IE voiced aspirated stops /bʰ dʰ gʰ gʰʷ/ > Gmc voiced stops /b d g gʷ/

(2)   voiced aspirates > voiced stops > voiceless stops > voiceless fricatives
      /bʰ dʰ gʰ gʷʰ/   /b d g gʷ/   /pʰ tʰ kʰ kʰʷ/   /f θ x xʷ/

Verner's Law states that something more happened to the Gmc reflexes of the IE voiceless stops when the consonant was non-initial and the preceding vowel was not stressed. When these two conditions were met, the IE voiceless stops appear as voiced stops in the Gmc historical record; additionally, IE intervocalic /s/ became /r/ via /z/ in the same environment (rhotacism). Thus, using the coronals as examples, the process is IE *tʰ* > Gmc *θ* (Grimm's Law) > *ð* > *d* (Verner's Law). Verner's Law changes are thus not exceptions to Grimm's Law, but additional and later changes.

   Another essential change in (Pre-)Gmc, which obscured the mechanism behind Verner's Law, was the change in how prosodic emphasis was realised. IE had a pitch accent, with high tone on whatever syllables were stressed, whereas Gmc developed dynamic stress on the root syllable of all words. Dynamic stress has a number of phonetic correlates besides higher pitch: more air from the lungs, greater loudness, and longer duration (Ladefoged 2005: 20–25). As Verner's Law was triggered by stress differences, the law must have operated before the fixing of the stress on the root syllable.

   The mechanism at work in Verner's Law is phonetic lenition in unstressed environments. The term 'lenition' is going to play a major role in this paper, so a definition is provided here: it is traditionally defined as change which involves weakening of segmental strength, such as opening or sonorisation, whereas 'fortition' has the opposite effect (Honeybone 2002: 39–43, 2008, 2012). Lass (1984: 177) refers to lenition and fortition as "involving change both in stricture and glottal state", and states that lenition "increases the permeability of the vocal tract to airflow". Voicing is usually treated as an instance of lenition,[1] and stress placement may in fact change the voicing of segments: consider Present-Day English (PDE) <x>, which is pronounced as /ks/ when the stress falls on the preceding vowel (*exit, exercise, execute*), but as /gz/ when the stress falls on the following vowel (*exam, exert,*

---

1.   Lass (1984: 177) comments that it is not entirely clear why changes in glottal state should be grouped with changes in stricture as instances of lenition/fortition, but that "the frequency with which the change voiceless → voiced is a precursor to opening of stricture argues for an essential similarity". He concludes (1984: 178) that "Perhaps the best way to look at lenition/fortition overall is in terms of two strength scales, one of **openness** and one of **sonority**: movement down the first involves decreased resistance to airflow, movement down the second an increase in the output of periodic acoustic energy" (original emphasis).

executive).[2] Verner's Law changes are hence combinative, whereas the changes in Grimm's Law appear to be isolative.

There is no dearth of research on Grimm's Law and Verner's Law, but the present chapter does not attempt to review all or even most of the vast literature on the topic. The best and most comprehensive account to date is that of Iverson & Salmons (2003), and the aim of this paper is to use findings from modern experimental phonetics to throw light on some unresolved issues in their account, especially the processes which defy characterisation as lenition, and the problem of the prime mover. Specifically, it will be argued that the shift from IE pitch to Gmc stress was the first change, which set the changes subsumed under Grimm's Law in motion, because it introduced aspiration as an active laryngeal feature, which caused affrication of the inherited voiceless stops. In using insights from experimental phonetics, this paper continues the work carried out by Ohala (1974, 2003), searching for phonetic answers to historical phonological problems, connecting phonetics and phonology through elements of physiology and aerodynamics (Ohala 1973, 1983, 1997). Section 2 deals with problems associated with the traditional formulation of Grimm's and Verner's Laws; Section 3 examines terms that are relevant to providing a unified account of Grimm's and Verner's Laws; Section 4 gives a summary of Iverson & Salmons's exposition; Section 5 addresses some aspects of Grimm's and Verner's Laws which remain problematic and proposes a chronology of the changes involved; Section 6 summarises my conclusions.

2.   Problems with the traditional account

2.1   The original IE/Gmc obstruent inventory

One essential problem pertains to the assumed Proto-Indo-European obstruent inventory: as conventionally outlined, it is (thought to be) typologically marked, since it contains two series of voiced stops and one series of voiceless stops, which goes against the typical state of affairs in modern languages (Henton et al. 1992).[3]

Competing views have therefore been proposed. Emonds (1972) suggests that it is in fact Gmc which had the most conservative stop system, and that it was the other, 'classical', IE languages which had changed; Emonds's hypothesis appears not to have many adherents at present. The Glottalic Theory (Gamkrelidze & Ivanov

---

2.   There is variation among speakers with regard to some of these words, especially *exit*, but the point still stands.

3.   Besides, the /b/ of the plain voiced series is extremely rare; this problem will not be addressed here, but see Lehmann (1993: 97) for an account of the issues involved.

1973; Hopper 1973), as reported by Salmons (1993), proposes alternative IE and Gmc obstruent inventories, most notably that IE or Pre-Gmc had ejectives in lieu of the traditional plain voiced stops (cf. Vennemann 1984: 19; Noske 2012). Ejectives tend to be voiceless, so the 'glottalic scenario' implies that the ejectives may have been phonemically unspecified for voice, and that Grimm's Law entailed their changing from ejectives to plain voiceless plosives. If so, Grimm's Law may be as shown in (3), leaving out the labiovelar consonants: aspiration was non-contrastive, and the key oppositions were hence voiceless vs. ejective vs. voiced.

(3)   Grimm's Law:
      IE voiceless stops /p t k/ > Gmc voiceless fricatives /f θ x/
      IE ejectives /b' d' g'/ > Gmc voiceless stops /p t k/
      (IE voiced aspirated stops /b d g/ > Gmc voiced stops /b d g/)

This is not enough, however, to make the proposed inventory unmarked: in their phonological research on modern languages, Henton et al. (1992: 78–79, 95) examine stop inventories in a very high number of languages, and consequently suggest the following implicational series of stops: voiceless (unaspirated) > voiced > voiceless aspirated > voiceless ejectives > voiced implosives.[4] None of the proposed IE or Pre-Gmc stop inventories agrees completely with these findings: whether or not ejectives are included, the inventories are somehow marked.[5]

     Kümmel (2015: 12), after assessing data from a large number of languages, concludes that "synchronic typology favours 'glottalic' models, but diachronic typology rather contradicts this", and that "the diachronic typology of systemic developments clearly favours the traditional reconstruction of the plosives as against all 'glottalic' models", i.e. voiceless unaspirated stops, voiced stops, and voiced aspirated stops.[6] Moreover, most accounts written within the glottalic framework now assume intermediate changes to the stop inventory between IE and Pre-Gmc, or between Pre-Gmc and PGmc, such that the direct inputs to Grimm's Law were in fact the

---

4.   In other words, if a language has only one set of phonemically distinct stops, that set almost always consists of voiceless (unaspirated) stops; if a language has two sets, the voiceless stops will be supplemented by a series of voiced stops, etc.

5.   Noske (2012: 71) believes that with the alternative "model, the typological problems concerning the obstruent inventory have been resolved" as "there are no longer voiced aspirates", but this seems to be merely a result of the model: it is only if we accept that the traditional voiced stops were ejectives that we can re-*label* the traditional voiced aspirates as simple voiced stops. However, ejectives are not much more common than voiced aspirates, so the original problem has not disappeared.

6.   As will become clear, I assume that aspiration was acquired between Pre-Gmc and the onset of Grimm's Law.

conventional series of stops, i.e. voiceless, voiced and voiced aspirated stops (e.g., Salmons 1993: 59–61; Vennemann 2006: 131). Iverson & Salmons (1995, 2003: 55 fn. 15) also take the traditional series as their point of departure.

Basing his scenario on observed changes in obstruent inventories, Kümmel (2015) proposes an alternative inventory, i.e., that PIE originally had implosives. These are "non-explosive" stops: they have ingressive pulmonic airstream and for that reason, there is no release.[7] These implosives changed into simple voiced stops, and the original voiced stops acquired breathy voice, to produce the traditional inventory of voiced stops and voiced aspirated stops (Kümmel 2015: 14), besides the voiceless stops. For these reasons, I shall assume the conventional stop inventory as the immediate input to Grimm's Law.

2.2   Chronology and intermediate stages

Another issue is the internal chronology of Grimm's Law and Verner's Law. Early accounts assumed that Grimm's Law predated Verner's Law, whereas more recent work often states that identical results obtain with any chronology. However, Verner's Law clearly affected voiceless fricatives as a *group*, since PGmc /s/ is found among the affected consonants. Since Gmc voiceless fricatives, with the exception of /s/, arose as a result of the operation of Grimm's Law, it follows that the part of Grimm's Law which affected the IE voiceless stops must have preceded Verner's Law, which then affected all the Gmc voiceless fricatives, regardless of origin.

Vennemann (1984: 20–22) argues that Verner's Law took place before the affrication and spirantisation of the etymological voiceless stops, partly because such a chronology gets rid of the problem of occlusivisation. That is, in Vennemann's chronology, as well as that of, e.g., Kortlandt (1985, 1988) and Noske (2012), Verner's Law changed PIE /t/ directly to /d/. However, such a chronology does not explain why Verner's Law seems to have operated on voiceless *fricatives* as a group – in other words, why /s/ should be likewise affected. Besides, the essence of Verner's Law is not stopping, despite Noske (2012),[8] but voicing/lenition in unstressed environments. The traditional scenario is simpler in that it assumes that Grimm's Law operated on all voiceless stops in all environments (except after *s*), and that Verner's

---

7.   Kümmel (2015: 12) calls them non-obstruent stops, but they do have a complete obstruction of the air, so they are more correctly classified as *unreleased* obstruents.

8.   Noske (2012: 68) additionally finds it problematic that *s* did not undergo stopping. As stated in the preceding, however, the essence of Verner's Law is not stopping, but voicing/lenition, which of course also affected *s*. It could be argued that rhotacism of *z* (< *s*) is in fact a kind of fortition, at least if the product was a tongue-tip trill or tap (the resulting *z* could not stop to a coronal plosive, because it would then have merged with pre-existing /d/).

Law caused voicing of a subset of the output in certain environments. The alternative chronology, however, has to posit a bifurcating process (Noske 2012: 74–75; cf. Vennemann 1985) in which the voiceless stops were affricated and spirantised only in contexts in which Verner's Law did not operate, making Grimm's Law context-sensitive as well.

Concerning the other two Grimm's Law changes, things are less clear, as their chronology may be connected to the issue of possible intermediate stages, especially with respect to the process that changed the IE voiced aspirates into Gmc plain voiced stops; I will return to this later.

Voiceless aspirated stops have a long history of changing into voiceless affricates and fricatives: Grimm's Law is but one example; Greek *ph* > *f* is another (Allen 1987: 22–26). 'Liverpool lenition' refers to an ongoing process, in which the voiceless aspirated stops /p t k/ in Scouse are undergoing affrication to [pɸ tθ kx], and then spirantisation to [ɸ θ x]; it is reported to depend on, e.g., prosody and word position (Honeybone 2007, and references cited there).[9] Affrication and spirantisation of voiceless plosives is also seen in the Second Consonant Shift (SCS), which affected High German, and which changed the then voiceless plosives (< IE voiced stops) into affricates and fricatives, after which the then voiced stops devoiced (Moulton 1954: 33–42).[10] It therefore seems highly probable that the IE voiceless stops were first affricated in Gmc in the first stage of Grimm's Law, and then spirantised. Now, if changes were symmetrical in the voiceless and voiced series, the IE voiced aspirated stops *bʰ dʰ gʰ* would also have undergone affrication (> *bᵝ dᵟ gᵞ*) and spirantisation (> *β ð ɣ*) before finally stopping to *b d g*; I will return to this in § 5. If so, the last change involving stopping of voiced fricatives is identical to the last stage in Verner's Law (e.g., *ð* > *d*). A possible chronology is given in (4) below; the internal chronology of stages (ii) and (iii) is of little consequence if (iii) entailed affrication and/or spirantisation; if it did not, (ii) must have preceded (iii).

(4)   The stages of Grimm's Law (i–iii) and Verner's Law (iv):
   i.   IE voiceless stops > Gmc voiceless affricates > Gmc voiceless fricatives
   ii.   IE voiced stops > Gmc voiceless stops
   iii. IE voiced aspirated stops > Gmc voiced affricates > Gmc voiced fricatives
   iv.   IE voiceless fricatives (from (i)) > voiced fricatives (which merge with voiced fricatives from (iii) above) > voiced stops

---

9. For instance, affricated allophones are found initially, whereas spirants occur word-finally (Honeybone 2007).

10. Goblirsch (2001) views this process as a consonant shift in itself, and posits different stages of similar shifts for High German, Danish and Icelandic.

The devoicing in stage (ii) and the stopping in the last stage in (iv) are the changes that stand out, defying characterisation as processes of lenition, which brings us to the next point. The chronology of Grimm's and Verner's Laws, and the question of intermediate stages, will be further discussed in §§ 4 and 5.

2.3   A unified phonetic account

Another problem is that of arriving at a unified phonetic account of the changes covered by Grimm's and Verner's Laws, as they seem to go in opposite directions, whether described articulatorily or in terms of lenition. Articulatorily, Grimm's Law involves affrication, spirantisation, devoicing, de-aspiration (if there were no intermediate stages in the change IE voiced aspirates > Gmc voiced stops) or stopping (if there were intermediate stages), whereas Verner's Law involves voicing and stopping. Under the definition of 'lenition' and 'fortition' given in § 1, affrication, spirantisation, de-aspiration and voicing are processes of lenition, whereas devoicing and stopping are instances of fortition. However, according to Kager et al.'s definitions (2007), so-called 'devoicing' can really be aspiration (and 'de-aspiration' is a change from long-lag VOT to short-lag VOT). Besides, Lass & Laing (2013: 99 fn. 6) define lenition rather as an "increase of trans-oral airflow", which means that aspiration is in fact lenition, not fortition, because the duration of the complete closure for the plosive is reduced (so also Lass 1984: 178). Additionally, affrication and spirantisation are likely due to the fact that heavily aspirated stops have shorter and less complete closures (Iverson & Salmons 2003: 58) – they can be seen as target undershoot. If this is accepted, all the changes subsumed under Grimm's and Verner's Laws are instances of lenition, except the stopping process.[11]

It could be, however, that Grimm's and Verner's Law are not best explained as processes of lenition/fortition, and this raises the following question: is it possible to find one common 'phonetic' mechanism which triggered Grimm's and Verner's Laws? I believe the answer is in the affirmative, but in order to identify this mechanism, we need to arrive at a better understanding of the acoustic-articulatory phenomena involved, i.e., voicing, aspiration, and voice onset time (VOT), as well as of other relevant articulatory and prosodic parameters, such as sub-glottal and trans-oral pressure, duration of closure, gestural timing, stress, etc. These are the topic of the next section.

---

11. In this view, de-aspiration can consequently be defined as either lenition or fortition, which is a problem in so far as it threatens to render 'lenition' vacuous as an explanatory term.

3.   Air-flow, pressure, voicing, aspiration and laryngeal specification

The state of the glottis varies considerably and forms a continuum, from fully closed for a glottal stop to fully open for breathing; the "continuum reflects differing amounts of airflow at the glottis" (Henton et al. 1992: 72–73; 66 Figure 1). Transglottal airflow rates thus vary with the type of laryngeal constriction.[12]

Henton et al. (1992) conclude that so-called 'voiced aspirated' stops are in fact voiced stops with breathy voice, and are intermediate between aspirated stops and voiced stops.[13] As for the timing of the glottal opening, it is "characterized by an abduction of the vocal cords at the time of the release of the stop closure" (Henton et al. 1992: 82), which is a feature they share with voiceless aspirated stops. The findings of Lisker & Abramson (1964: 403, 418–419) and Schwarz et al. (2019) also suggest that voiced aspirated stops are in fact both voiced and aspirated.

"Pressure built up in the mouth depends on air pressure from the lungs and organic pressure at the constriction. The more intense the air pressure, the stronger the organic pressure must be to stop it" (Delattre 1965: 118; cf. Morton manuscript: 23–24). This fact has at least two consequences: it is "[h]ard to maintain voicing during a stop due to air pressure equalizing in the oral and subglottal cavities" (Henton et al. 1992: 72, referring to Ohala 1983; cf. Morton manuscript: 24–25); and voiced stops have a narrower tongue constriction because "the pharyngeal cavity must expand during closure to keep voicing going", simply because there is a complete closure (Shadle 2010, referring to Westbury 1983).

According to Browman & Goldstein (1989: 69), an articulatory gesture is

> identified with the formation (and release) of a characteristic constriction within one of the relatively independent articulatory subsystems of the vocal tract (i.e., oral, laryngeal, velic). Within the oral subsystem, constrictions can be formed by the action of one of three relatively independent sets of articulators: the lips, the tongue tip/blade and the tongue body. As actions, gestures have some intrinsic time associated with them – they are characterizations of movements through space and over time.

---

12. Henton et al. (1992: 73) find that aspirated stops have an airflow rate of up to 1,000 ml/s, whereas the rate for stops pronounced with breathy voice is close to 500 ml/s, and that for voiced stops is 120 ml/s.

13. Breathy voiced stops in Hindi have (1) less activity of the cricothyroid muscle than voiced stops because the vocal folds are slack, (2) a "[m]oderately open glottis (about 50% of that used during aspiration)", (3) a "high rate of oral airflow following release", (4) a quick decrease in subglottal air pressure, and (5) voice present (Henton et al. 1992: 80–81).

It follows that any consonant involves one or more gestures, and that the relative timing of their formation and release may vary; this is called gestural timing. A consonant has three stages: the form stage, the hold stage, and the release stage (cf. Henton et al. 1992: 66). The duration of closure refers to the hold stage and varies systematically for so-called voiced and voiceless obstruents: the closure is longer for the voiceless stops than for their voiced counterparts.

Phonetically, voicing or phonation refers to the vibration of the vocal cords. However, there is a difference between phonetic voice and phonemic voice. Traditionally, the phonemic difference between obstruent pairs such as /t/ and /d/ has been referred to as a difference in 'voice', but in Gmc, phonemic 'voice' does not typically correlate with phonetic voice. That is, the difference between obstruent pairs in Gmc rarely involves phonation, but rather the closure duration and aspiration (Delattre 1965): the 'voiceless' obstruents have a longer complete closure and are aspirated, whereas the so-called 'voiced' obstruents have a shorter closure and no aspiration. Aspiration in turn refers to spread glottis, i.e., abducted vocal cords, which entails a delay in the onset of the voicing of the following segment: "Aspiration occurs when the open [glottal] state is prolonged past the moment of oral release" (Henton et al. 1992: 72).

This difference in how phonemic 'voice' is realised has given rise to a dichotomy between 'voicing languages' (e.g., Romance) and 'aspiration languages' (e.g., Germanic languages except Dutch),[14] which refers to a difference in laryngeal specification, following Lisker & Abramson's seminal article (1964). The phonetic realisations of laryngeal contrasts vary cross-linguistically. The main acoustic and perceptual cue is voice onset time (VOT), or the time between the release of a stop consonant and the beginning of phonation (vocal cord vibration). Gmc languages seem to have a two-way contrast, voice and aspiration. In Dutch, the voice onset

---

14. I assume, in agreement with most of the research literature, that (standard varieties of) the Gmc languages are aspirating languages, except Dutch (but see Hunnicutt & Morris 2016: 216). This is not to say that variation is non-existent: a great deal of variation is certainly found within accents of English regarding (pre-)voicing and (degree of) aspiration, for instance in Welsh English (Wells 1982: 388) or Scots (Wells 1982: 409, 413); this synchronic variation may reflect a diatopic and diachronic situation in which the status of aspiration varied, although in some cases it is probably a Celtic interference/substrate effect, and is so explained by Wells. That Gmc languages are aspirating languages is maintained in, e.g., Goblirsch (2001) for Icelandic and Danish; Moosmüller & Ringen (2004) for German and Austrian Standard German; Beckman et al. (2013) for German, English, Danish and Icelandic; van de Weijer & van der Torre (2007) for German and English; Schluter et al. (2017) for English. Helgason & Ringen (2008) find that Central Standard Swedish seems to employ a two-way contrast, pre-voicing and post-aspiration, i.e., both [voice] and [spread glottis] are active laryngeal features; the same appears to hold for Southern American English (Hunnicutt & Morris 2016).

starts simultaneously, or even before, the release, which makes it a (pre)voicing language with "voicing lead" (Lisker & Abramson 1964: 389),[15] whereas in the other Gmc languages, there is a delay in VOT for both traditional 'voiced' (short-lag VOT) and voiceless obstruents (long-lag VOT), which suggests that they are aspiration languages. Lisker & Abramson (1964: 407) find that "the distribution of [VOT] values is essentially tri-modal, corresponding generally to the ranges centering at −100, +10 and +75 msec" for the traditional 'voiced', 'voiceless unaspirated' and 'voiceless aspirated' stops, respectively.

Linguists have disagreed as to the specification of laryngeal features. Some researchers espouse the Single Feature Hypothesis (SFH), claiming that there is one feature for laryngeal representation, [voice], but they differ in their view of this as a binary feature [±voice] or a monovalent or privative feature [voice] (Kager et al. 2007: 42–44). Others find that there are multiple monovalent features (e.g., Iverson & Salmons 1995), and that one of these features is the active one in the different languages, i.e., [voice] for (pre)voicing languages and [spread glottis] for aspiration languages; this is the Multiple Feature Hypothesis (MFH). The MFH is now referred to as "laryngeal realism" (Honeybone 2005: 345).

In Dutch, [voice] is the active feature, and it thus has, e.g., the phonemes /p/ vs. /b/, whereas [spread glottis] or aspiration is the active laryngeal feature in aspirating languages, and voicing is hardly contrastive, which gives /pʰ/ vs. /b̥/ (or /pº/, cf. Honeybone 2005: 332). The findings of Kager et al. (2007) strongly support 'laryngeal realism'; the laryngeal specifications for the three languages examined (Dutch, German and English) are as shown in Table 1 (reproduced from Kager et al. 2007: 45).

Table 1. Laryngeal features in three Gmc languages (Kager et al. 2007: 45)

	Voicing lead	Short-lag VOT	Long-lag VOT
Dutch	[voice]	[ ]	
German		[ ]	[spread glottis]
English		[ ]	[spread glottis]

Thus, in Dutch and other true voice languages, voiceless consonants lack specification, whereas voiced stops are specified for [voice], which means that the unmarked (unspecified) category is voicelessness. In aspiration languages, the contrast is marked with a [spread glottis] specification for aspirated stops, whereas plain

---

15. The same is true of certain accents of English in Scotland (e.g., those of Aberdeen and Shetland, cf. Hughes et al. 2012: 133, 163) and elsewhere (e.g., the Scottish-English border, cf. Docherty et al. 2011).

unaspirated stops lack laryngeal specification, making them the unmarked category (Beckman et al. 2013).[16] The distinction between the SFH and the MFH (or between tradition (i) and tradition (ii) in Honeybone's (2005) terminology) is therefore essential, because it encompasses a crucial difference in what category is unmarked (and changes tend to go in the direction of the unmarked category), and because it carries with it predictions about possible and impossible phonological changes (Honeybone 2005).

Essential characteristics of the larynx have often been identified in terms of the features [voice], [spread glottis] and [larynx height], but another way to represent laryngeal specification has been put forth more recently – Avery & Idsardi's 'Laryngeal Dimensions' (2001). They propose three dimensions of the larynx: Glottal Tension (with the features [slack] and [stiff]), Glottal Width (with the features [spread] and [constricted]), and Larynx Height (with the features [raised] and [lowered]). Voicing languages thus have Glottal Tension as their essential laryngeal specification (and the default feature for obstruents tends to be [slack]), whereas aspiration languages have Glottal Width (and the default feature tends to be [spread]; Iverson & Salmons 2003: 47–48). In any one language, one laryngeal dimension is active, and only "one member of an antagonistic gestural pair is used contrastively in a given system" (Iverson & Salmons 2003: 47). Glottal Width has the antagonistic pair [spread] and [constricted], and in English and German, the phonologically active gesture is [spread]: they thus "contrast [spread] voiceless aspirated stops with laryngeally unmarked lenis stops", though it is possible that "the other member may be invoked as a phonetic embellishment, or 'enhancement', of a contrast" (2003: 47).

The difference between these two representations of 'phonemic voice' or laryngeal specification is very much in focus in Iverson & Salmons's (2003) account of Grimm's Law, which is the topic of the next section. A crucial question concerns when (Pre-)Gmc went from being a voicing language to being an aspiration language, i.e., when Glottal Tension was replaced by Glottal Width as the active laryngeal specification.

---

16. Schluter et al. (2017), however, find in favour of the single feature hypothesis, since they obtain the same Mismatch Negativity results in experimental research involving native speakers of English (aspirating language), Arabic (supposedly voicing language), and Russian (true voicing language). Their findings are incompatible with the laryngeal realism paradigm, but support a model that "posits that two-way laryngeal contrasts are represented in long-term memory in a format that abstracts away from their precise articulatory details" (2017: 23), and that the active laryngeal feature for *all* languages is in fact [-voice] or [spread glottis].

4.   'Germanic enhancement'

In essence, Iverson & Salmons (2003: 56) state that the introduction of aspiration in Gmc set the whole shift (Grimm's Law) in motion in a pull-chain. In other words, rather than simply assuming *a priori* that the voiceless stops in Gmc were aspirated, they see the acquisition of aspiration as an innovation and a "persistent articulatory constraint" (2003: 44, 53 ff.), which they call 'Germanic Enhancement'.[17] PIE was a voicing language, and the acquisition of aspiration was the result of some change in 'articulatory setting' in Gmc speakers (2003: 68).[18] Whatever the nature of this change in setting, it introduced (a) aspiration as a characteristic of the voiceless plosives, and (b) [spread glottis] as an active laryngeal feature. It was this newly acquired aspiration that triggered the affrication and spirantisation of the voiceless plosives, after which the voiced stops "devoice to form a new unmarked series", i.e., lose their marking for Glottal Tension and become aspirated, and the voiced aspirates "simplify to voiced", through the "loss of the now superfluous Glottal Width dimension" (2003: 56). In the end, [voice] was removed from the system, and Gmc no longer had Glottal Tension as its essential laryngeal specification, but Glottal Width (2003: 63).

The IE accent was a pitch accent with high pitch, i.e., stiff vocal cords, on stressed vowels. Slack vocal cords entail low pitch and in fact induce passive voicing of obstruents (Iverson & Salmons 2003: 44), especially of laryngeally unmarked fricatives (2003: 48–50). Among the sonorants, of course, Glottal Tension is an active laryngeal feature. Verner's Law is thus simply passive or spontaneous voicing, the voicing "bleeding" from a preceding *unaccented* vowel;[19] stressed vowels are marked as having [stiff] vocal cords, which inhibit voicing/vocal cord vibration because of low transglottal airflow (2003: 60); hence Verner's Law did not take effect when the preceding vowel was stressed (see also Johnsen 2011). Verner's Law, then, is "not an independent innovation", nor indeed a "law" (Iverson & Salmons 2003: 44, 59), but the contextual phenomenon of passive obstruent voicing.

---

17.   Keyser & Stevens (2006: 42) define enhancement as making a distinction more salient.

18.   Iverson & Salmons (2003: 68) see "articulatory setting" as the way in which all the articulators are aligned or "set" in native speakers of a given language; this setting differs between, e.g., Dutch, Norwegian and Urdu speakers, and is part of the reason why, e.g., Norwegian speakers may be identified as such when they attempt to speak another language.

19.   That is, "glottal vibration ensues automatically with sufficient airflow across vocal folds which are in a neutral state of abduction, as in the case with ordinary sonorants of all types, consonants as well as vowels" (Iverson & Salmons 2003: 50); further, "the phenomenon of 'passive voicing' in obstruents can be understood as the extension of spontaneous voicing into a neighbouring segment not already specified for a laryngeal quality. This penetration operates rightward from a segment provided with Glottal Tension to a following laryngeally empty one, which, in English … is any of the lenis obstruents" (2003: 51).

Thus, in Iverson & Salmons's account, there was no stopping of fricatives in the Grimm's Law changes to IE voiced aspirates, simply because the voiced aspirates did not affricate or spirantise in the first stage – they simply lost their aspiration (Iverson & Salmons 2003: 56). Fricative allophones developed later as a result of "passive spirantisation", which produced voiced fricatives "which contrast with the medial voiceless fricatives in Germanic, i.e., those that escaped voicing via Verner's Law due to preceding accent" (2003: 58). The voiced fricatives from Verner's Law do appear as stops in the written record,[20] so it is possible that they merged with the voiced fricatives that arose through passive spirantisation of voiced stops (< IE voiced aspirates), but Iverson & Salmons (2003) do not state this explicitly; this may be the weak point of their account, and I will return to it in § 5.

There is some evidence supporting the claim that Gmc went from being a voicing language in its very earliest stages, to becoming an aspiration language: i.e., the failure of Grimm's Law to take effect when the stops were preceded by *s*.[21] For

---

20. Honeybone (2005: 333) notes that the letters <b, d, g> have "been used to represent more than one kind of phonological segment", i.e., voiced /b d g/ in voicing languages, and /pº tº kº/ in aspiration languages. This is a natural consequence of the adoption of the Roman alphabet, which was devised for a voicing language. The argument could, however, be extended: the fact that the outputs of Verner's Law (and of the development of the voiced aspirates for that matter) appear as <b, d, g> need not mean that the sounds they correspond to were stops.

21. Besides, Kager et al. (2007) argue that the difference in laryngeal specifications for voicing and aspiration languages has consequences for predictions regarding acquisition errors to do with devoicing and voicing in L1 children: children "tend toward the unmarked value" (Kager et al. 2007: 46). The unmarked values are /p t k/ for voicing languages, /b̥ d̥ g̥/ for aspiration languages. Specifically, the MFH predicts that children acquiring voicing languages should show errors involving devoicing (Kager et al. 2007: 47), whereas children acquiring aspiration languages, on the other hand, should show de-aspiration errors (ibid.). The predictions of both hypotheses are reproduced below (from Kager et al. 2007: 47).

	Voicing language	Aspiration language
SFH:	/b/ → [p]	/b̥/ → [pʰ]
MFH:	/b/ → [p]	/pʰ/ → [p]

It is the laryngeal error predictions of the Single Feature Hypothesis, as well as those of the Multiple Feature Hypothesis for *voicing* languages (and not for aspiration languages), which seem to be borne out in the Grimm's Law change involving the devoicing of IE voiced stops (/b d g/) to Gmc voiceless stops (/p t k/), i.e., one of the changes which stand out as fortition processes. In both models, this change involves the simple deletion of the laryngeal specification [voice] in a voicing language, i.e., a change towards the unmarked consonant. This appears to support the hypothesis that the Gmc languages were still voicing languages in their very earliest stages, but that they later became aspiration languages. This does not refute Iverson & Salmons's hypothesis: even if aspiration was acquired phonetically very early in Gmc and set Grimm's Law in motion, it may not have become a distinctive (i.e., phonemic) feature until much later. To use a parallel: Gmc languages still have phonetic voicing, but its distinctiveness and application vary greatly.

PDE, Iverson & Salmons (1995) are able to give a unified account of both the lack of aspiration of voiceless stops after /s/ and devoicing of sonorants after voiceless obstruents, under the assumptions (a) that the laryngeal feature in English is [spread glottis], (b) that the feature [spread glottis] may be shared by obstruent-initial consonant clusters, and (c) that [spread glottis] has a constant temporal duration. Only the lack of aspiration after /s/ will be considered in this context. In the words of Beckman et al. (2013: 267), "since the peak of glottal opening for a voiceless stop occurs relatively late in a single stop, but in a fricative coincides with the beginning of oral constriction, … the vocal folds will be coming together for voicing earlier" in the clusters /sp/, /st/, /sk/ "than in a singleton stop, and hence there will be no aspiration of the stop". The fact that Grimm's Law failed to apply when the stops in question were preceded by *s* must be taken as relatively certain proof that aspiration, i.e., [spread glottis], had become an active laryngeal feature by the time Grimm's Law started. At the same time, aspiration of the voiceless stops should not be assumed *a priori* for PIE or Pre-Gmc, as PIE is believed to have been a voicing language, in which [spread glottis] would not have been an active laryngeal feature.

5.   Remaining problems

5.1   The voiced aspirates and fricative stopping

It was noted in § 2.2 that if changes in the so-called 'voiceless' and 'voiced' series are parallel, then the voiced aspirates must also have been affricated and spirantised before stopping. Such a scenario is envisaged by Moulton (1954: 38–39), because it seems in fact to entail a less complicated process towards the final attested results than does simple de-aspiration. The last two stages of this process are identical to Verner's Law, and the terminal products of both processes merged. In Iverson & Salmons's account (2003: 56), however, the IE voiced aspirates were not affricated or spirantised, and then stopped, but simply lost their aspiration. Thus, the stopping process which stubbornly resists characterisation as an instance of lenition is a pseudo-problem. Passive spirantisation was a secondary change and merely allophonic; the main allophone was always [d].

Moulton's account (1954) is an interesting alternative to Iverson & Salmons's view that the change to IE voiced aspirates consisted of simple de-aspiration. Moulton, basing his claims on a thorough assessment of all the earliest attestations of Gmc, finds that voice held primacy over occlusion type as a distinctive feature in the earliest Gmc, which entails that the voiced fricatives were allophones of the voiced stops: /t/ and /θ/ vs. /d/ [d, ð]. Later, occlusion type became the primary distinctive feature, which means the voiced fricatives were allophones of the voiceless

fricatives: /t/ and /d/ vs. /θ/ [θ, ð]. Even later, aspiration seems to have acquired primacy, certainly in OHG. If this is accepted, the end product of the changes to, e.g., IE /dʰ/ may in fact have been [ð] (not [d]), which later fell together with /d/ <d>, *only because voiced fricatives were allophones of voiced stops in early Gmc*; i.e., the allophone [ð] of /d/ merged with the main allophone [d], which in fact agrees with Iverson & Salmons's account. Such a scenario would also make the stopping process in Grimm's/Verner's Law a pseudo-problem: it would merely be the result of the system of allophones and phonemes.

However, Verner's Law did entail fricative stopping, and there have been later processes of stopping of the dental fricatives in most Gmc languages, so some discussion of the issue is in order. Fricative stopping in general is not common, but TH-Stopping has taken place historically in all the Gmc languages except (standard) English and Icelandic, and is happening in varieties of PDE, e.g., in Liverpool (Honeybone 2007), New York City and Jamaica (Wells 1982: 515–6, 565–6, 575); besides, many non-standard varieties of English have lost the dental fricatives (accents of Scots, accents of Irish English), possibly as a substrate effect (Wells 1982). Cross-linguistically, dental fricatives are rare, 'unnatural' or marked (Wells 1982: 96), simply because they are difficult to produce. Stopping of other fricatives is very rare, but does occur in the speech of small children.[22] This suggests that stops are easier to pronounce for children because a complete closure does not require the fine-motor skills necessary for fricatives, and that stopping may in fact be target overshoot: a complete closure, instead of a considerable narrowing, is made between the upper and lower articulators.

Iverson & Salmons (2003) and Moulton (1954) argue cogently for their hypotheses, and it is difficult to determine which provides the more likely account. What the two accounts have in common is the consequence, namely that the development of the IE voiced aspirates in Gmc did not involve stopping *per se*. The later processes of TH-Stopping are not directly relevant to the topic discussed here, but stopping in Verner's Law is relevant: they all seem to be adequately explained by markedness and language acquisition research as articulatory simplification. It is also possible, as outlined in the preceding discussion, that the voiced fricatives resulting from Verner's Law merged with those arising from passive spirantisation of early Gmc voiced stops, for which the main allophones were stops. Now, if the

---

22. In L1 acquisition, children acquire stops before continuants, and L1 acquisition errors include fricative stopping: "There is a lot of variation in the chronology of stopping. Stopping of /f/ is prevalent in children between 2;00–2;06 years and is typically eradicated by 3;00 years. … stopping of its voiced counterpart /v/ persists a little longer, until around 3;06 years. … Finally, stopping of the dentals /θ/ and /ð/ may persist until 5;00 years of age" (Williamson 2016).

IE voiced aspirates did not undergo affrication and spirantisation, the chronology and nature of Grimm's Law must be revised, as in (5).

(5) The stages of Grimm's Law (i–iii) and Verner's Law (iv)
    i.   IE voiceless aspirated stops > Gmc voiceless affricates > Gmc voiceless fricatives
    ii. IE voiced stops > Gmc voiceless stops
    iii. IE voiced aspirated stops > Gmc voiced stops (> allophonic voiced fricatives)
    iv. IE voiceless fricatives (from (i)) > voiced fricatives (also from (iii)) > voiced stops

5.2   Devoicing

A number of issues connected to stops and voicing emerge from aerodynamic and physiological research: a crucial finding is that it is in fact difficult to maintain "voicing during a stop … because the air flowing through the glottis accumulates in the oral cavity, causing oral pressure to approach subglottal pressure", and then "the air flowing through the glottis gradually diminishes and voicing is extinguished" (Ohala 1983: 194).

Moreover, a number of factors affect voicing and VOT: duration, stress, word/syllable position, and place of articulation. With respect to duration, Ohala (1983: 195) finds that long stops tend to be devoiced, whereas short stops become voiced (Ohala 1983: 195). As for stress, "[m]ain stress affects the duration of voicing in stops. The physiological concomitants of stress include increased respiratory muscle activity, resulting in greater subglottal pressure", according to Henton et al. (1992: 79), who also find that stops have different acoustic qualities in different word or syllable positions, and that final stops tend to be voiceless (1992: 80). Yao (2009) concludes that VOT may be affected by *linguistic* factors such as speaking rate (shorter VOT with increased speech rate), the following vowel (longer VOT before high vowels, and longer VOT when the consonant is followed by a vowel), utterance position (longer VOT in final position), and place of articulation (longer VOT the further back the consonant is pronounced), as well as by *extra-linguistic* factors such as age (older speakers have shorter VOT) and sex (females have longer VOT).[23] Evidently, voicing and VOT are unstable features that depend on a number of variables (cf. Iverson & Salmons 1995; Beckman et al. 2013).

---

23. Similar findings are reported by Ladefoged (2005: 52–53 Figure 6.3), Henton et al. (1992: 79–80), and Wells (1905: 526). If these phonetic facts are applied on the Gmc obstruent inventory, one might venture to suggest that Grimm's Law started in final consonants: aspirated voiceless stops affricated and spirantised, and voiced stops devoiced; the process may have started with the velars, and young women led the way.

As for typological facts on stops, Henton et al. state very clearly that all languages have stops (1992: 65) and that all languages have voiceless stops (1992: 79).[24] If so, the process IE voiced stops > Gmc voiceless stops may also have been systemically conditioned to supply voiceless stops at a stage when none existed,[25] and also to avoid merger between the inherited IE voiced stops and the new voiced stops from IE voiced aspirates.[26] Iverson & Salmons see "demurmuring" and devoicing as chain-shift effects of spirantisation (2003: 58).

5.3   A Gmc consonant shift cycle?

Iverson & Salmons explicitly see Grimm's Law as a chain-shift (2003: 69). Certainly, Grimm's Law, Verner's Law and the SCS instantiate certain repeat changes: the SCS affrication and/or spirantisation of voiceless plosives and the subsequent devoicing of voiced plosives are identical to two of the processes in Grimm's Law, i.e., stages (i) and (ii) in (5) above (cf. Goblirsch 2001). What emerges is a consonant shift cycle, laid out in (6), with a low number of recurring processes: de-stopping and stopping, de-aspiration and aspiration; it appears to involve lenition mostly, but with entry points for stops in the final stages.[27]

(6) a. Voiceless aspirated plosive > voiceless affricate > voiceless fricative > voiced fricative > voiced plosive > voiceless plosive
     b. Voiced aspirated plosive > voiced plosive > voiceless aspirated plosive > voiceless affricate > voiceless fricative > voiced fricative > voiceless/voiced plosive

---

24. In Ruhlen's study (1975) of 706 languages, only 4 had voiced stops but not voiceless stops, and Henton et al. show that most languages have "voiceless" and "voiced" stops (1992: 67 Table 1; 73). It should be noted, however, that such databases must be used with caution, as the information they contain may be inaccurate.

25. A related process seems to be affecting Scouse: the voiceless aspirated stops are affricated and spirantised, whereas the dental fricatives are undergoing stopping; Honeybone (2007) ascribes the latter to interference from Irish English – which is likely – but it is interesting nonetheless that the two processes should be going on at the same time.

26. See also Noske (2012) for an account of the Grimm's Law chain-shift within Contrast Preservation Theory.

27. If we allow, for the sake of argument, that the aspect of Grimm's Law affecting IE voiced aspirates entailed affrication, spirantisation and stopping, the last stage is identical to Verner's Law (and may have been the same process). This cycle would be: voiced aspirated plosive > voiced affricate > voiced fricative > voiced plosive > voiceless plosive.

If the majority of the sound-changes observed in Grimm's Law and Verner's Law are processes of lenition, the question is why stops are produced. The answer is probably the same as that provided for devoicing in the preceding section: all languages have stops (Henton et al. 1992: 65), and so stops simply tend to be supplied in cases where changes have left the system without plosives, cf. the discussion of markedness and 'impossible' phonological changes in Honeybone (2005: 318–320). Stops may also be the result of target overshoot and articulatory simplification (of fricatives).

5.4   The prime mover

Some questions still remain unanswered: Which was first, the accent shift or the introduction of aspiration? What was the change in articulatory setting?

Logically, the shift in stress *placement* must have postdated Verner's Law, since the latter depends on the IE variable stress; the same is not true of the shift in prosodic emphasis itself, i.e., the change from IE pitch-dominant prosody to Gmc dynamic stress prosody. It is crucial to distinguish between the stress shift and the stress *placement* shift; they are not the same (cf. Iverson & Salmons 2003: 65). The stress/prosodic shift was a prosodic revolution: Gmc went from marking prosodic importance by means of high pitch (stiff vocal folds) to marking such importance by means of dynamic stress, whose phonetic correlates are manifold, e.g., greater respiratory effort, longer duration, greater loudness and intensity, and higher pitch. The stress *placement* shift entailed a change from lexical, 'free' stress to fixed stress on the root syllable of all words (barring prefixes).[28] It would not be surprising if two such crucial changes had phonetic consequences; moreover, there may have been an overlap between the two systems for some time.

The Gmc stress shift may have been more directly "involved" in Grimm's Law than has been assumed, as it caused a "considerable disturbance in the coordination of the factors of speech" (Wells 1905: 524). Gestural coordination may in fact be disturbed with increased speech rate, which often follows for unstressed syllables: in languages with isochronous stress, stressed syllables come at roughly equal intervals, even in natural speech, and so unstressed syllables are spoken at a greater speed. This suggests that Grimm's Law may have been a co-articulation phenomenon. There is some research on gestural timing in relation to aspiration and VOT. Ladefoged & Cho (2001: 7–8) suggest

---

28. Quoting Halle (1997), Iverson & Salmons state that "default locus of accentuation in Indo-European is word-initial, so that removal of lexically marked accent would automatically result in the observed Germanic pattern" of stress on the root syllable of all words (2003: 64).

that there is a phonetic parameter, which we will call Articulatory VOT, definable in terms of the difference in time between the initiation of the articulatory gesture responsible for the release of a closure and the initiation of the laryngeal gesture responsible for vocal fold vibration.

Lending support to Lisker & Abramson's conclusions (1964: 422), Ridouane (manuscript) finds that "it is the timing of laryngeal and supralaryngeal articulations that control aspiration rather than the laryngeal gesture per se" (manuscript: 5); further, that "aspiration is not a function of glottal opening per se, but rather a function of the degree of glottal opening at stop release" (manuscript: 6); and finally, that there is "temporal alignment between laryngeal and supralaryngeal gestures" (manuscript: 7). In other words, there has to be coordination between an articulatory (supra-laryngeal) gesture aiming at opening the closure and a laryngeal gesture aiming at initiating phonation, and they do not always coincide perfectly, causing a delay in VOT, the magnitude of which depends on the degree of spread in the glottis.[29] It is possible that the Gmc stress shift seriously affected this gestural timing, especially after the introduction of aspiration as an active laryngeal feature, and that the shift in how prosodic emphasis was realised in fact *was* the prime mover and is identical to the change in 'articulatory setting' proposed by Iverson & Salmons (2003: 68). The acquisition of aspiration on voiceless plosives may in fact have been accompanied by a parallel development of aspiration on the original PIE voiced stops, as in Kümmel's view (2015: 14), differentiating the latter from the 'new' PIE voiced stops (< implosives); this would produce the three series of voiceless aspirated plosives, voiced aspirated plosives and voiced plosives as the direct input to Grimm's Law.

In an attempt at establishing a chronology of changes, I therefore submit that the shift from IE pitch-dominant prosody to Gmc dynamic stress prosody was the very first change to affect Gmc, but that the stress placement remained, as required by Verner's Law, and possibly the high tone also. In other words, the earliest Gmc may have developed dynamic stress (with all its phonetic correlates), but kept the IE pitch differences for some time, realised as high tone on stressed syllables. As

---

29. Ridouane (manuscript: 7) elaborates: "The definition proposed identifies an articulatory continuum (spreading of the glottis and the arytenoids cartilages) that corresponds to a relatively stable region in the acoustic output (aspiration noise generated following the release). During the 50-odd ms following the release, the glottis moves toward a configuration appropriate for the voicing of the following vowel. In this time interval, aspiration occurs without glottal vibration since the glottal spreading inhibits glottal [vibration]. Spread glottal opening is the articulatory definition of the feature and aspiration noise duration the acoustic definition. It is not clear, however, how should [sic] the temporal alignment between laryngeal and supralaryngeal gestures (necessary for the production of aspiration) be represented within this definition. As shown above, glottal opening per se is not enough to account for aspiration." See also Ridouane et al. (2011).

stated above, the "physiological concomitants of stress include increased respiratory muscle activity, resulting in greater subglottal pressure" (Henton et al. 1992: 79). This increased respiratory effort may have been the direct reason for the acquisition of aspiration, which crucially triggered affrication of the voiceless stops and hence set Grimm's Law in motion.[30] The chronology would thus be as in (7).

(7)   IE pitch > Gmc stress (= increased respiration) > aspiration > affrication, etc.

Verner's Law is often explained almost as an exception, i.e., as voicing only in a certain position and only with a certain stress placement. Is it possible that Grimm's Law depended on these variables too, i.e., the position of the obstruent and the stress placement? In such a scenario, the shift in prosodic emphasis (tone > dynamic stress) caused the Grimm's Law changes, in so far as the changed realisation of prosodic emphasis produced new obstruent allophones depending on a number of factors: the consonants' position,[31] the presence/absence of stress on the syllable to which the obstruent belonged and hence the speech rate, the presence of stress on the preceding or following vowel, (the nature of) any preceding/following consonants, etc. With regard to laterals and preceding or following stress, Scobbie & Wrench (2003: 1872) found that "Word-medial intervocalic /l/ was also consonantal preceding stress, but following lexical stress some vocalised … forms were found". Thus, it may be that the consonantal gesture is reduced in certain environments, i.e., after lexical stress.

There is also a connection between pitch accent, high tone and obstruents. Ohala (1973: 6) explains that "a high position of the tongue creates a slight pull on the larynx which is translated into increased vocal cord tension. This results in the widely-noted slightly higher average pitch for high vowels [i, u] and slightly lower pitch for low vowels [æ, a, ɔ]", but also that "the vocal cords are stiff during the production of voiceless obstruents and are slack during voiced obstruents" which results in "pitch variations accompanying voiced and voiceless obstruents" (1973: 7). Ohala also finds that "Comparing the fundamental frequency after the release of the [b], [p], [pʰ], it is obvious that the pitch is a little higher for the voiceless stops [p] and [pʰ], than it is for the voiced stop [b]" (Ohala 1978: 365; cf. Iverson & Salmons 2003: 44).

---

30. Vennemann (1984: 20) also sees the spirantisation of the voiceless aspirated plosives as the most characteristic of the Gmc sound-changes: "Sie nenne ich die Urgermanische Lautverschiebung, da sie für das Germanische als vor allem anderen charakteristisch angesehen wird". It is implied that this was the first stage in Grimm's Law (although chronologically, Vennemann puts Verner's Law before Grimm's Law).

31. For instance, Honeybone (2007) shows that in Liverpool lenition, affricated allophones are found word-initially, whereas spirants occur word-finally.

I believe that Iverson & Salmons's proposed change in "articulatory setting" corresponds to the shift from pitch stress to dynamic stress, and that if the Grimm's Law changes were part of a unified process, that process may have been increased respiratory effort and heightened sub-glottal pressure, which would cause an increase in transglottal and trans-oral airflow. Such a change would turn (a) voiceless stops into aspirates, (b) voiceless and voiced aspirates into affricates (> fricatives), and (c) voiced plosives into devoiced plosives, because the heightened pressure would extinguish voicing sooner (Ohala 1983: 194).[32] Alternatively for (c), voiced plosives might become breathy voiced, which might be re-interpreted as voiceless aspirates, as they have certain phonetic characteristics in common, including a higher rate of oral airflow, more abducted glottis and a rapid drop in subglottal pressure (Henton et al. 1992: 80–82). This, in essence, is Grimm's Law. The change that is not explained is that affecting IE voiced aspirates if they did not affricate; Iverson & Salmons (2003: 56) propose that they simply lost their aspiration because it was redundant in a system specified for Glottal Width.

The phonetic arguments given above are supported by systemic arguments. It is likely that the voiceless aspirated plosives changed first, due to the acquisition of aspiration in Gmc. The moot point is stage (ii), i.e., whether the IE plain voiced stops changed next, or the IE voiced aspirates. If the voiced aspirates changed first, losing their aspiration, this would leave the inherited plain voiced plosives in danger of merging with the new voiced plosives, and this might have contributed to the devoicing of the IE voiced stops, which may have been helped by the physiological fact that voicing is hard to maintain in stops. Alternatively, the voiced stops devoiced first in order to supply voiceless plosives to the system after stage (i). Such chain-shift arguments find support in the changes of the SCS, in which the voiceless plosives affricated and/or spirantised first, upon which the voiced stops devoiced, and secondly in the seemingly universal requirement that languages must have voiceless stops. The changes to IE voiced stops and IE voiced aspirates may equally have been more or less simultaneous. In fact, stages (i)–(iii) may all have been simultaneous,[33] but stage (iv) (Verner's Law) must have taken place after the voiceless plosives had spirantised.

---

32. In addition to the natural pitch difference between high and non-high vowels noted in the preceding, there was phonemic high pitch on stressed vowels in IE. Stiff vocal cords and high pitch on accented vowels in IE had an effect on stops, inhibiting voicing (Iverson & Salmons 2003: 44). It is possible that stiff vocal cords for stressed vowels also caused devoicing in following obstruents (Iverson & Salmons 2003: 66).

33. In my examination of the other famous chain-shift, the Great Vowel Shift in English, I conclude that the changes to the high and high-mid long vowels were likely simultaneous (Stenbrenden 2016).

6.   Conclusion

This paper has attempted to provide a unified account of Grimm's Law and Verner's Law. Although most of the changes in question are leniting processes, 'lenition' fails to explain the changes which involve stopping and devoicing. I have therefore tried to identify the one common phonetic mechanism that would trigger the changes observed in Grimm's Law. I agree with Iverson & Salmons (2003) that the acquisition of aspiration in Gmc, and hence of Glottal Width as the primary laryngeal specification, was the change that set the Grimm's Law changes in motion. Aspiration of the voiceless plosives induced affrication and spirantisation, because of the shorter and weaker closure for heavily aspirated stops. This was followed by devoicing of the IE voiced stops, either for systemic reasons – to supply voiceless stops to the system – or for phonetic reasons – because voicing in plosives is difficult to maintain. Finally, the voiced aspirates either affricated and spirantised, or simply lost their aspiration, possibly to supply the system with voiced stops, or because aspiration became redundant when Glottal Tension gave way to Glottal Width.

Iverson & Salmons (2003: 68) propose that the acquisition of aspiration was the result of some change in articulatory setting, which was thus the prime mover. I have suggested here that the prime mover may have been the widely-noted shift in prosodic emphasis, from IE high pitch on stressed vowels, to Gmc dynamic stress, with its phonetic correlates including longer duration, as well as heightened respiration, intensity and pitch. This prosodic shift may have resulted in increased respiratory effort and heightened sub-glottal pressure, which would cause an increase in transglottal and trans-oral airflow. In turn, increased pressure and airflow would induce aspiration. In other words, the key to Grimm's Law lies in the expression of prosodic emphasis.

Acknowledgements

References

Allen, W. Sidney. 1987 [1968]. *Vox Graeca: A guide to the pronunciation of Classical Greek*. Cambridge: Cambridge University Press.

Avery, Peter & William J. Idsardi. 2001. Laryngeal dimensions, completion and enhancement. In T. Alan Hall (ed.). *Distinctive feature theory*, 41–70. Berlin: de Gruyter. https://doi.org/10.1515/9783110886672.41

Beckman, Jill, Michael Jessen & Catherine Ringen. 2013. Empirical evidence for laryngeal features: Aspirating vs. true voice languages. *Journal of Linguistics* 49. 259–284. https://doi.org/10.1017/S0022226712000424

Browman, Catherine P. & Louis Goldstein. 1989. Articulatory gestures as phonological units. *Haskins Laboratories Status Report on Speech Research*, SR-99/100, 69–101. https://doi.org/10.1017/S0952675700001019

Delattre, Pierre. 1965. *Comparing the phonetic features of English, French, German and Spanish: An interim report*. London: Harrap.

Docherty, Gerard, Dominic Watt, Carmen Llamas, Damien Hall & Jennifer Nycz. 2011. Variation in voice onset time along the Scottish-English border. *ICPhS* XVII, 591–594.

Emonds, Joseph. 1972. A reformulation of Grimm's Law. In Michael Brame (ed.), *Contributions to generative phonology*, 108–122. Austin: University of Texas Press.

Gamkrelidze, Thomas & Vjaceslav Ivanov. 1973. Sprachtypologie und die Rekonstruktion der gemeinindogermanischen Verschlüsse: Vorläufiger Bericht. *Phonetica* 27. 150–156. https://doi.org/10.1159/000259433

Goblirsch, Kurt G. 2001. The Icelandic Consonant Shift in its Germanic context. *Arkiv för nordisk filologi* 116. 117–133.

Grimm, Jacob. 1822–1837. *Deutsche Grammatik*. Göttingen: Dieterichschen Buchhandlung.

Halle, Morris. 1997. On stress and accent in Indo-European. *Language* 73. 275–313.

Helgason, Pétur & Catherine Ringen. 2008. Voicing and aspiration in Swedish stops. *Journal of Phonetics* 36. 607–628. https://doi.org/10.1016/j.wocn.2008.02.003

Henton, Caroline, Peter Ladefoged & Ian Maddieson. 1992. Stops in the World's languages. *Phonetica* 49. 65–101. https://doi.org/10.1159/000261905

Honeybone, Patrick. 2002. Germanic obstruent lenition: Some mutual implications of theoretical and historical phonology. Newcastle upon Tyne: University of Newcastle upon Tyne PhD thesis.

Honeybone, Patrick. 2005. Diachronic evidence in segmental phonology: The case of obstruent laryngeal specification. In Marc van Oostendorp & Jeroen van de Weijer (eds.), *The internal organization of phonological segments*, 319–354. Berlin: Mouton de Gruyter. https://doi.org/10.1515/9783110890402.317

Honeybone, Patrick. 2007. New-dialect formation in nineteenth century Liverpool: A brief history of Scouse. In Anthony Grant & Clive Grey (eds.), *The Mersey Sound: Liverpool's language, people and places*, 106–140. Liverpool: Open House Press.

Honeybone, Patrick. 2008. Lenition, weakening and consonantal strength: Tracing concepts through the history of phonology. In Joaquim Brandão de Carvalho, Tobias Scheer & Philippe Ségéral (eds.), *Lenition and fortition*, 9–92. Berlin: Mouton de Gruyter. https://doi.org/10.1515/9783110211443.1.9

Honeybone, Patrick. 2012. Lenition in English. In Terttu Nevalainen & Elizabeth C. Traugott (eds.), *The Oxford handbook of the history of English*, 1–26. Oxford: Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199922765.013.0064 (1 August, 2018)

Hopper, Paul J. 1973. Glottalized and murmured occlusives in Indo-European. *Glossa* 7. 141–166.

Hughes, Arthur, Peter Trudgill & Dominic Watt. 2012. *English accents and dialects* (5th edn.). London: Routledge.

Hunnicutt, Leigh & Paul A. Morris. 2016. Prevoicing and aspiration in southern American English. *University of Pennsylvania Working Papers in Linguistics* 22(1). 215–224.

Iverson, Gregory K. & Joseph C. Salmons. 1995. Aspiration and laryngeal representation in Germanic. *Phonology* 12(3). 369–396. https://doi.org/10.1017/S0952675700002566

Iverson, Gregory K. & Joseph C. Salmons. 2003. Laryngeal enhancement in early Germanic. *Phonology* 20. 43–74. https://doi.org/10.1017/S0952675703004469

Johnsen, Sverre Stausland. 2011. The phonetics and phonologization of Verner's law. In Thomas Krisch & Thomas Lindner (eds.), *Indogermanistik und Linguistik im Dialog*, 232–241. Wiesbaden: Reichert.

Kager, René, Suzanne van der Feest, Paula Fikkert, Annemarie Kerkhoff & Tania S. Zamuner. 2007. Representations of [voice]: Evidence from acquisition. In Jeroen van de Weijer & Erik Jan van der Torre (eds.), *Voicing in Dutch: (de)voicing – phonology, phonetics, and psycholinguistics*, 41–80. Amsterdam: John Benjamins. https://doi.org/10.1075/cilt.286.03kag

Keyser, Samuel J. & Kenneth N. Stevens. 2006. Enhancement and overlap in the speech chain. *Language* 82(1). 33–63.

Kortlandt, Frederik. 1985. Proto-Indo-European glottalic stops: The comparative evidence. *Folia Linguistica Historica* 6. 183–201. https://doi.org/10.1515/flih.1985.6.2.183

Kortlandt, Frederik. 1988. Proto-Germanic obstruents. *Amsterdamer Beiträge zur älteren Germanistik* 27. 3–10. https://doi.org/10.1163/18756719-027-01-90000003

Kümmel, Martin. 2015. The role of typology in historical phonology. In Patrick Honeybone & Joseph Salmons (eds.), *The Oxford handbook of historical phonology*, 121–132. Oxford: Oxford University Press.

Ladefoged, Peter. 2005. *Vowels and consonants*. Malden, MA: Blackwell.

Ladefoged, Peter & Taehong Cho. 2001. Linking linguistic contrasts to reality: The case of VOT. In Nina Gronnum & Jørgen Rischel (eds.), *Travaux du Cercle Linguistique de Copenhague* XXXI, 212–225. Copenhagen: C.A. Reitzel.

Lass, Roger. 1984. *Phonology: An introduction to basic concepts*. Cambridge: Cambridge University Press.

Lass, Roger & Margaret Laing. 2013. The early Middle English reflexes of Germanic *ik 'I': Unpacking the changes. *Folia Linguistica Historica* 34(1). 93–114.

Lehmann, Winfred P. 1993. *Theoretical bases of Indo-European linguistics*. London: Routledge.

Lisker, Leigh & Arthur S. Abramson. 1964. A cross-language study of voicing in initial stops: Acoustical measurements. *Word* 20(3). 384–422. https://doi.org/10.1080/00437956.1964.11659830

Moosmüller, Sylvia & Catherine Ringen. 2004. Voice and aspiration in Austrian German plosives. *Folia Linguistica* 38(1–2). 43–62. https://doi.org/10.1515/flin.2004.38.1-2.43

Morton, Katherine. Manuscript. *Experimental Phonology and Phonetics*. 1984. Available here: http://www.morton-tatham.co.uk/publications/to_1994/Morton_1984a.pdf (1 August, 2018)

Moulton, William G. 1954. The stops and spirants of early Germanic. *Language* 30(1). 1–42. https://doi.org/10.2307/410217

Noske, Roland. 2012. The Grimm-Verner push chain and Contrast Preservation Theory. In Bert Botma & Roland Noske (eds.), *Phonological explorations: Empirical, theoretical and diachronic issues*, 63–86. Berlin: De Gruyter. https://doi.org/10.1515/9783110295177.63

Ohala, John J. 1973. The physiology of tone. *Consonant types and tone. Southern California Occasional Papers I Linguistics* 1. 2–14.

Ohala, John J. 1974. Experimental historical phonology. In John Anderson & Charles Jones (eds.), *Historical linguistics II: Theory and description in phonology*, 353–389. Amsterdam: North Holland.

Ohala, John J. 1978. Experimental historical phonology. In Jeri J. Jaeger et al. (eds.), *Proceedings of the Fourth Annual Meeting of the Berkeley Linguistics Society*, 353–387. Berkeley: Berkeley Linguistics Society.

Ohala, John J. 1983. The origin of sound patterns in vocal tract constraints. In Peter F. MacNeilage (ed.), *The production of speech*, 189–216. New York: Springer.
https://doi.org/10.1007/978-1-4613-8202-7_9

Ohala, John J. 1997. Aerodynamics of phonology. *Proceedings of the 4th Seoul International Conference on Linguistics*, 92–97.

Ohala, John J. 2003. Phonetics and historical phonology. In Brian D. Joseph & Richard D. Janda (eds.), *The handbook of historical linguistics*, 667–686. Malden, MA: Blackwell.
https://doi.org/10.1002/9780470756393.ch22

Rask, Rasmus. 1818. *Undersøgelse om det gamle nordiske eller islandske Sprogs Oprindelse*. Copenhagen: Gyldendalske.

Ridouane, Rachid. Manuscript. On the feature [spread glottis]. 2006. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.504.3088&rep=rep1&type=pdf (1 August, 2018)

Ridouane, Rachid, G. N. Clements & Rajesh Khatiwada. 2011. Language-independent bases of distinctive features. In John Goldsmith, Elizabeth Hume & Leo Wetzels (eds.), *Tones and Features*, 260–287. Berlin: Mouton de Gruyter.

Ruhlen, Merritt. 1975. *A guide to the languages of the world*. Stanford: Languages Universals Project.

Salmons, Joseph. 1993. *The Glottalic Theory: Survey and synthesis*. (Journal of Indo-European Studies, Monograph Series 10.) Washington, DC: Institute for the Study of Man.

Schluter, Kevin T., Stephen Politzer-Ahles, Meera Al Kaabi & Diogo Almeida. 2017. Laryngeal features are phonetically abstract: Mismatch negativity evidence from Arabic, English, and Russian. *Frontiers in Psychology* 8. 746. https://doi.org/10.3389/fpsyg.2017.00746

Schwarz, Martha, Morgan Sonderegger & Heather Goad. 2019. Realization and representation of Nepali laryngeal contrasts: Voiced aspirates and laryngeal realism. *Journal of Phonetics* 73. 1–33. https://doi.org/10.1016/j.wocn.2018.12.007

Scobbie, James M. & Alan A. Wrench. 2003. An articulatory investigation of word final /l/ and /l/-sandhi in three dialects of English. *Proceedings of the International Congress of Phonetic Sciences* 15. 1871–1874.

Shadle, Christine H. 2010. Aerodynamics of speech, and the puzzle of voiced fricatives. Paper presented at the Conference on Phonetic Universals, Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany, October 29–30. https://www.eva.mpg.de/lingua/conference/10-PhoneticUniversals/pdf/Shadle_Abstract_fnl.pdf (1 August, 2018)

Stenbrenden, Gjertrud. F. 2016. *Long-vowel shifts in English, c. 1050–1700: Evidence from spelling*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9781107295469

van de Weijer, Jeroen & Erik Jan van der Torre (eds.). 2007. *Voicing in Dutch: (De)voicing – phonology, phonetics, and psycholinguistics*. Amsterdam: John Benjamins. https://doi.org/10.1075/cilt.286

Vennemann, Theo. 1984. Hochgermanisch und Niedergermanisch. *Beiträge zur Geschichte der deutschen Sprache* 106. 1–45.

Vennemann, Theo. 1985. The bifurcation theory of the Germanic and German consonant shifts synopsis and some further thoughts. In Jacek Fisiak (ed.), *Papers from the 6th International Conference on Historical Linguistics*, 527–547. Amsterdam: John Benjamins. https://doi.org/10.1075/cilt.34.36ven

Vennemann, Theo. 2006. Grimm's Law and loanwords. *Transactions of the Philological Society* 104(2). 129–166. https://doi.org/10.1111/j.1467-968X.2006.00170.x

Verner, Karl. 1875. Eine Ausnahme der ersten Lautverschiebung. *Zeitschrift für vergleichende Sprachforschung auf dem Gebiete der indogermanischen Sprachen* 23. 97–130.

Wells, Frederic L. 1905. Experimental phonetics and Verner's Law. *The Journal of English and Germanic Philology* 5/4. 522–527.

Wells, John C. 1982. *Accents of English*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511611759

Westbury, John. 1983. Enlargement of the supraglottal cavity and its relation to stop consonant voicing. *Journal of the Acoustical Society of America* 73. 1322–1336. https://doi.org/10.1121/1.389236

Williamson, Graham. 2016. Stopping. *SLTinfo*. https://www.sltinfo.com/phon101-stopping/ (1 August, 2018)

Yao, Yao. 2009. Understanding VOT Variation in Spontaneous Speech. *UC Berkeley Phonology Lab Annual Report 2009*, 29–43.

CHAPTER 2

The foot in the history of English

Challenges to metrical coherence

B. Elan Dresher and Aditi Lahiri
University of Toronto / University of Oxford

Dresher & Lahiri (1991) propose that Old English displays 'metrical coherence': different phonological processes are sensitive to the same metrical structure. We consider how English has dealt with challenges to metrical coherence. We show that the resolved moraic trochee, assumed to characterize the early Old English foot (Bermúdez-Otero manuscript; Goering 2016a, b), became untenable after the shortening of unstressed vowels, arguing that this stage of Old English, at least, requires the Germanic Foot, an extended and resolved trochee. After 1570 (Lahiri 2015) the direction of parsing changed from left-to-right to right-to-left when the number of Latin loanwords with stress-affecting suffixes had passed a threshold derived from Yang's Tolerance Principle (Yang 2016). This change re-established the metrical coherence that had been disrupted by these words.

Keywords: metrical coherence, Germanic Foot, moraic trochee, stress, Tolerance Principle

1. Introduction

Dresher & Lahiri (1991) argued that Old English displays a property called 'metrical coherence', whereby different phonological and poetic processes are all sensitive to the same metrical structure. A grammar is a complex system, and over time it is affected by phonological or morphological changes that result in patterns that may not be consistent with the previous metrical structures. Such changes pose a threat to the metrical coherence of the grammar. In this paper we will look at how the grammar of English has dealt with some such challenges to metrical coherence. In some cases the challenges have come from within, as when a sound change makes a previous analysis untenable; other challenges have come from without, in the forms of large-scale borrowing of lexical items that are not compatible with the prevailing metrical system. We will show how the grammar has met these challenges, either

by modifying the metrical system itself, or by modifying the segmental phonology to bring it in line with the metrical system, or, for words that cannot be easily integrated into the new system, by treating them as exceptions subject to special treatment of some kind.

2.   Change in the early Old English metrical system

Many have proposed that the Old English foot was a moraic trochee (Keyser & O'Neil 1985; Halle et al. 1993; Idsardi 1994; Hutton 1998; Bermúdez-Otero & Hogg 2003; Bermúdez-Otero manuscript; Goering 2016a, b). We have argued rather for what we have called the Germanic Foot (Dresher & Lahiri 1991), a resolved and extended trochee. As the name suggests, we proposed that this foot was inherited by Old English from its Germanic ancestor. Here we will review the moraic trochee analysis and show why we think it is untenable for Old English as it is attested in the written record.

The moraic trochee analysis is most plausible when applied to a stage of the language in which some inflectional vowels still retained length inherited from Proto-Germanic, as illustrated by (1), which gives the Pre-Old English paradigm of the neuter *a*-stem nouns, according to Hogg & Fulk (2011).

(1)   Pre-Old English paradigms of neuter *a*-stem nouns (Hogg & Fulk 2011: 15)

NOM/ACC.SG	*STEM-Ø	NOM/ACC.PL	*STEM-u
GEN.SG	*STEM-as	GEN.PL	*STEM -ō̃
DAT.SG	*STEM-ǣ	DAT.PL	*STEM-um

The rule of High Vowel Deletion (HVD) deleted short *u* and *i* under certain conditions when in an open syllable. Bermúdez-Otero (manuscript) and Goering (2016a, b) show that the contexts for HVD emerge clearly from positing moraic trochees assigned from left to right, as in (2).

(2)   Pre-Old English moraic trochees and HVD (Bermúdez-Otero manuscript; Goering 2016a, b)
    a.   Stem-initial syllable is heavy: "head"

	Pre-Old English	Old English
NOM/ACC.SG	(xau) (ƀud)	hēafud
GEN.SG	(xau) ƀu (das)	hēafdes
DAT.SG	(xau) ƀu (dǣ)	hēafde
NOM/ACC.PL	(xau) (ƀu du)	hēafudu
GEN.PL	(xau) ƀu (dō̃)	hēafda
DAT.PL	(xau) ƀu (dum)	hēafdum

    b.   Stem-initial syllable is light: "army"

	Pre-Old English	Old English
NOM/ACC.SG	(we rud)	weorud
GEN.SG	(we ru) (das)	weorudes
DAT.SG	(we ru) (dǽ)	weorude
NOM/ACC.PL	(we ru) du	weorud
GEN.PL	(we ru) (dõ̃)	weoruda
DAT.PL	(we ru) (dum)	weorudum

The vowels affected by HVD in heavy stems (2a) are those that cannot be parsed into a moraic trochee, as in all the oblique cases. In the GEN.SG, for example, the first syllable *xau* is heavy (as it is throughout the paradigm) and forms a moraic trochee by itself; the final syllable, *das*, is also heavy and forms a foot on its own, leaving the medial light syllable, *ƀu*, on its own and unparsed into a foot, and subject to deletion. By contrast, the second syllable in the NOM/ACC.SG is heavy and not subject to HVD. The NOM/ACC.PL is the only form in this paradigm in which the final syllable is light; hence, the medial and final syllable can make up a moraic trochee, and both *u* are retained in OE *hēafudu*.

    This analysis has the merit of being able to explain why we find *hēafudu* rather than **hēafdu* as the earliest attested form of the NOM/ACC.PL (Fulk 2010), in contrast to the other oblique forms which show HVD. Notice that in this analysis deletion depends not only on the weight of the syllable *preceding* the unstressed high vowel, but also on the weight of the syllable that *follows* it.

    The Pre-Old English foot was more complicated than a simple moraic trochee, however. Old English allowed stems with initial light syllables, as in (2b), and these had to be parsed into licit metrical feet. In (2b), the second vowel joins the initial one to form a moraic trochee (*wé ru*) in all the inflected forms. In the NOM/ACC. PL, the result is that the final suffixal *-u* is left unfooted and deletes. Note that the NOM/ACC.SG *weorud* shows that it is not the case that all feet have two moras: if a stem-initial syllable is light,[1] it must form a foot with whatever syllable follows, creating in this case a tri-moraic trochee (*wéo rud*). In verse this phenomenon is called 'resolution' (Sievers 1893; Kuryłowicz 1949, 1970; Russom 1987), whereby the two syllables together make up a single metrical position. Therefore, if the pre-Old English foot was a moraic trochee, it must have been a 'resolved' moraic trochee that could have up to three moras.

---

1.   Though the phonetics of the 'short diphthongs' <ea, eo, io> has been disputed, there is broad agreement that they functioned as short vowels (Hogg 1992: 16–20); hence the first syllable of *weo.rud* is light. The 'long diphthongs', as in the first syllable of *hēa.fud*, have always been assumed to pattern with the long vowels in making their syllable heavy.

The shortening of unstressed vowels in early Old English made the metrical analysis just presented problematic, particularly in the heavy stem paradigm, as can be seen in (3).

(3)   Expected moraic trochees after shortening of unstressed vowels

	Pre-Old English	Old English
NOM/ACC.SG	(hēa) (fud)	hēafud
GEN.SG	(hēa) fu (des)	hēafdes
DAT.SG	(hēa) (fu de)	*hēafude
NOM/ACC.PL	(hēa) (fu du)	hēafudu
GEN.PL	(hēa) (fu da)	*hēafuda
DAT.PL	(hēa) fu (dum)	hēafdum

The shortening of the inflectional vowels in the DAT.SG and GEN.PL lead us to expect that the medial high vowel should be the head of a binary foot and hence not subject to HVD, but this result is not observed in Old English. Rather, the medial vowel continues to delete in all the inflected cases, except for the NOM/ACC.PL. These latter cases now appear to be an exception to a new, and simpler, generalization governing HVD: an unstressed high vowel in an open syllable deletes when it directly follows a heavy syllable or a sequence of two light syllables. That is, the weight of the syllable that follows the potential HVD target is no longer relevant to the operation of HVD.

Therefore, after the shortening of unstressed vowels in early Old English, the moraic trochee became untenable. Rather, the synchronic facts of Old English are what motivated Dresher & Lahiri (1991) to propose what we called the Germanic Foot, given in (4) (see also Lahiri et al. 1999; Fikkert et al. 2006). Some sample parsings are shown in (5).

(4)   Old English metrical analysis (Dresher & Lahiri 1991; Lahiri et al. 1999; Fikkert et al. 2006)
   a.   Germanic Foot: From left to right, construct a resolved and expanded moraic trochee of the form (|head| dependent), where the head must consist of at least two moras and the dependent may have at most one mora.
   b.   Main stress is on the leftmost foot.
   c.   Defoot a foot (|ẍ|) that does not carry the main stress, is final in the word, and has no dependent.

(5)   Old English stress: sample parsings (the head of the foot is indicated by |x|)

   a.   "word GEN.PL"   b.   "army GEN.PL"   c.   "king DAT.SG"
          x   x   x
        (|x| .)   (|x | .)   (|x | .)
        [H   L]ω   [L   L   L]ω   [L   H   L]ω
        wór   da   wéo ru da   cý nin ge

   d.   "dwelling NOM.PL"   e.   "other NOM.SG"   f.   "other ACC.SG"
          x   x   x
        (|x |)   (|x|)   (|x̶|)   (|x|)   (|x| .)
        [L L]ω   [H   H]ω   H   H   L
        ló fu   ó   þer   ó   þèr ne

Comparing the Germanic Foot with the moraic trochee, we can see that the mo-
raic trochee corresponds to the head of the Germanic Foot, and that an unfooted
light syllable that follows a moraic trochee is incorporated as a weak branch into
the Germanic Foot. HVD now deletes a high vowel in the weak branch of a foot,
as shown in (6).

(6)   High Vowel Deletion in the weak branch of a foot

   a.   "head NOM.SG"   b.   "head DAT.SG"   c.   "head DAT.PL"
          x   x   x
        (|x|) (|x̶|)   (|x| .)   .   (|x| .) (|x̶|)
        [H   H]ω   [H   L   L]ω   [H   L]   H]ω
        héa   fud   héa fu̶ de   héa fu̶ dum

   d.   "army NOM.SG"   e.   "army NOM.PL"   f.   "word NOM.PL"
          x   x   x
        (|x |)   (|x | .)   (|x| .)
        [L   H]ω   [L   L   L]ω   [H   L]ω
        wéo   rud   wéo ru du̶   wór du̶

   g.   "journey NOM.SG"   h.   "journey GEN.PL"   i.   "journey DAT.PL"
          x   x   x
        (|x | .)   (|x | .)   (|x |) (|x̶|)
        [L H   L]ω   [L   H   L]ω   [L H   H]ω
        fǽ rel du̶   fǽ rel da   fǽ rel dum

HVD applies in a straightforward fashion to all forms in the *a*-stem paradigms except for *hēafudu*. Putting that form aside for the moment, we observe in (6b, c) that HVD applies consistently to the stem vowel *u* when it is in an open syllable following a heavy syllable, regardless of the weight of the following syllable. In (6a), *u* is in a closed syllable and does not delete. In (6d–e), the stem-internal *u* follows a light syllable and must be part of the head of the foot, where it does not delete; an inflectional final *-u*, however, is in the weak branch of the foot and deletes (6e). In (6f) the head of the foot is a heavy syllable and the inflectional *-u* deletes in the weak branch of the foot. In (6g–i), the stem-initial light syllable of the masculine *u*-stem (or neuter *a*-stem) noun *fǽreld* (*fǽrelt*) must join with the following heavy syllable to make up the head of the foot, leaving room for an additional light syllable in the weak branch; a final *-u* deletes in this position (6g), as it does after a single heavy syllable (6f) and after two light syllables (6e).

The form *hēafudu* stands out, in that it does not fit the pattern in (6). We have seen diachronically how this form came to be the odd man out in its paradigm, and indeed it appears to have been problematic even for Old English speakers, as shown by the fact that such forms show considerable variation across dialects (Fulk 2010); it is clear that the same analysis will not be adequate for all Old English dialects. We also find intra-dialect variation, sometimes in the same document. Because of all this variation it is difficult to disentangle the effects of phonology from analogy. Nevertheless, Fulk (2010) argues that the phonologically expected outcome in early Old English is indeed *hēafudu*, a form that appears (though not the only form that appears) in the Mercian *Vespasian Psalter* (*Ps(A)*). We will here focus on that dialect.

This form and other nominal forms with inflectional *-u* require special treatment in any analysis. Bermúdez-Otero & Hogg (2003: 22) and Bermúdez-Otero (manuscript: 7), for example, propose that the nominal *a*-stem inflectional affixes have become 'phonologically stratified': "the neut.nom/acc.pl. ending is added at the stem level, whilst other *a*-stem noun affixes are word-level." Similar in spirit is the analysis of Dresher (1985), who posits a special boundary before nominal inflectional *-u*. However, in that analysis the levels are the reverse of Bermúdez-Otero's: Dresher (1985) posits that *-u* is a word-level affix, while the other affixes are stem level. This analysis is consistent with that of Dresher (1993), shown in (7), which proposes that in Mercian, verbal inflectional agreement affixes are word level, and everything else – roots, stems, and noun affixes – are stem level, what is called there the extended stem, or E-stem level, because it includes stem extensions. Dresher's (1993) analysis of levels does not consider the problem of Vowel Deletion, but is based on the behaviour of several other rules that affect stressed vowels.

(7)   Mercian lexical phonology (Dresher 1993: 333)

	MORPHOLOGY	PHONOLOGY
E-stem level:	Roots, stems, nominal affixes	Breaking, Retraction, Back Mutation, Smoothing, i-Monophthongization, i-Mutation
Word level:	Verb AGR	Breaking, Retraction, Back Mutation

In terms of that stratification, the nominal -*u* inflection falls in with the word-level affixes. The rationale is that HVD applies to the second *u* of *hēafudu* as if the final -*u* were not present, treating *hēafudu* as if it were *hēafud*, as shown in (8a); compare *hēafde* in (8b).

(8)   High Vowel Deletion at the extended stem level (based on Dresher 1985)

    a.   "head NOM/ACC.PL"   b.   "head DAT.SG"

```
       x                            x
     (|x|)  (|x|)      .          (|x|   .)    .
     [H      H  ]E-stem   L          [L    L    L ]E-stem
     hḗa    fud        u          hḗa   fu̶   de
```

There is other evidence for the special status of -*u* in *Ps(A)*. In (9a), we find a stem-internal vowel *e* when the stem precedes a consonant, a word boundary, and -*u*; the vowel does not appear before other suffixes that begin with a vowel. Dresher (1985) argues that this stem has been reanalyzed in the *Ps(A)* dialect as deriving from an underlying monosyllable, with *e* the result of epenthesis when a vowel does not follow the stem; inflectional -*u* acts as if it is not present when epenthesis applies. A similar distribution occurs in (9b) "many".

(9) Adjectives in *Ps(A)*

a. miċel- ~ micl- "great"

	FEMININE		NEUTER
NOM.SG	miċelu (3x) ~ miċel (1x)	NOM.PL	miċelu (5x) ~ miċel (3x)
DAT.SG	miċelre	GEN.PL	miċelra
NOM.PL	micle	DAT.PL	miclum

b. moniġ- ~ monġ- "many"

	MASCULINE	FEMININE	NEUTER
NOM.PL	monġe	monġe (1x) ~ monġ (1x)	moniġ (1x)
ACC.PL	monġe	monġe	monigu (1x)
GEN.PL	moniġra	moniġra	
DAT.PL	mongum		

The application of HVD and Epenthesis to stem-medial vowels is very consistent in *Ps(A)*. The same is not the case for the deletion of the final *-u* when it follows a stem of more than one syllable. We can see this already in (9), where we find *miċel* varying with *miċelu*, and *moniġ* varying with *monigu*. The presence of *-u* is unexpected no matter what the underlying representation of these stems is (Fulk 2010 suggests that these adjectives may be influenced by *lȳtelu* "little", where retention of *-u* after a heavy-light sequence is what we expect).

We also find variation in nouns. A sample of such variation in *Ps(A)* is given in (10).

(10) Variation between *-u* ~ *-Ø* in NEUTER NOM/ACC.PL nouns in *Ps(A)*

a. hēafudu (2x) ~ hēafud (5x)   "head"
b. wolcenu (2x) ~ wolcen (6x)   "cloud"
c. calferu (1x) ~ calfur (2x)   "calf"
d. lomberu (1x) ~ lombur (1x)   "lamb"

The variation after surface disyllables is in contrast with the regular behaviour of *-u* after surface monosyllables in (11):

(11) Behaviour of *-u* after surface monosyllables in *Ps(A)*

a. *-u* after light syllable in neuter *a*-nouns: *lofu* "glory", *geatu* "gate", etc.
b. *-Ø* after heavy syllable in neuter *a*-nouns: *word* "word", *gōd* "good thing", etc.
c. *-u* after heavy syllable in neuter *ja*-nouns: *rīċu* "dominion", *styċċu* "piece", *wītu* "punishment", etc.

Evidently, there was some uncertainty as to how -*u* was incorporated into the metrical system at the word level in the former cases, but not in the latter ones; see further Dresher & Lahiri (1991: 279–281) for why HVD became opaque when following two surface syllables in *Ps(A)*. This situation was unstable, and in other dialects of Old English it was regularized in various ways (Fulk 2010; Bermúdez-Otero manuscript).[2]

The above discussion has assumed that the advocates of a moraic trochee were correct for pre-Old English, and that the shortening of unstressed inflectional vowels made the original moraic trochees an unsuitable environment for HVD, necessitating their replacement by the Germanic Foot. This scenario raises the question of the status of the latter: did it begin as a pan-Germanic foot, give way to the moraic trochee in pre-Old English, and then return in Old English, or was the moraic trochee in fact the Germanic foot, with the 'Germanic Foot' arriving on the scene only in Old English? This question merits a separate study, and here we will only make the following points.

First, Dresher & Lahiri (1991: 264–269) argue that Sievers's Law in Gothic and High Vowel Deletion in Old High German provide evidence for the Germanic Foot (see also Lahiri 1982 and Fikkert et al. 2006), which suggests an early origin for this metrical constituent. Second, we observe that in many forms, the moraic trochee and the Germanic Foot yield similar results, making it hard to discern which one is correct. Thus, if we apply the principles in (4) and the parsings in (5) and (6) to the pre-Old English paradigms in (2), we find that we can account for all the cases of HVD just as for Old English. Again, the NOM/ACC.SG of "head" requires some special treatment. This raises the question of how early this special treatment began.[3] An answer to this question requires a wider study of other relevant forms, particularly the *ja*-nouns mentioned in (11c).

To sum up this section, whether the moraic trochee or the Germanic Foot was present in pre-Old English, the metrical system had to make an accommodation for data that did not fit. In the first scenario, the shortening of unstressed inflectional vowels posed a challenge to the metrical coherence of the early Old English

---

2. The nouns in (11c) have NOM/ACC.SG *rīċe, styċċe, wīte*, with a final *e* derived from *j* or *i* that is not deleted by HVD, whereas the NOM/ACC.PL *rīċu, styċċu, wītu* consistently have final *u* that is also not deleted by HVD, despite its surface occurrence after a heavy syllable. The history and synchronic analysis of these nouns have been controversial: see Lahiri (1982) and Fikkert et al. (2006) for synchronic analyses, and Fulk (2010) for discussion of their diachrony.

3. This analysis is consistent with the observations by Hogg (2000) that HVD was to some extent morphologically conditioned, and may have been from an early point, and was in other respects problematic and subject to analogical pressures and much variation. Contrary to what his title might suggest, he does not in fact argue that HVD did not exist.

metrical system by making the original moraic trochees an unsuitable environment for HVD. We propose that metrical coherence was restored by adding one mora to the trochee, and treating the formerly regular *-u* suffix of forms like *hĕafudu* as exceptional. On the second scenario, the Germanic Foot was already in place in pre-Old English, as was the exceptional treatment of forms like *xauƀudu > *hĕafudu*.

3.   The change in directionality

Foot form is only one aspect of English foot-related metrical structure that has changed over time. Present Day English metrical structure resembles that of Latin, shown in (12)–(13): a moraic trochee is computed from the right edge, and main stress is assigned to the rightmost foot (with various exceptions).[4]

(12)   Latin main stress (Roca 1999)
   a.   Stress the penultimate syllable if it is heavy: *amī́cus* "friend", *reféctus* "restored".
   b.   Otherwise, stress the antepenultimate syllable, if there is one: *dóminus* "master", *fémina* "woman", *refíciunt* "they…restore".
   c.   Otherwise, stress the first syllable: *vénīs* "you-SG. come", *cónsul* "consul".

(13)   Latin stress: metrical analysis
   a.   A final syllable is extrametrical.
   b.   Build quantity-sensitive trochees from the right edge of the word.
   c.   Main stress falls on the rightmost foot in the word.

As in other Germanic languages, this shift in the English metrical system occurred under the influence of Romance loanwords, but was not abrupt. We summarize our proposed chronology in (14).

(14)   Approximate dates of changes in direction and position of English stress
   a.   Gmc.–Middle Eng.:   Foot direction *left*, main stress *left*;
   b.   c1570:   Foot direction is changing to *right*.
   c.   c1660–:   Main stress changes to *right* in stages.

---

4.   We present the Latin stress system because this system characterized the Latin words that were imported into English (French borrowings followed a different rule; see Halle & Keyser 1971 and Dresher & Lahiri 2005). The Present Day English stress system is more complicated, being the result of how the Latin system was adapted to English, including the complications discussed below.

It can be shown that the influx of Anglo-Norman and Old French words following the Norman conquest had little impact on Middle English prosody (Minkova 1997; Redford 2003; Dresher & Lahiri 2005; Lahiri 2015), contrary to what has sometimes been claimed.

Rather, the extended trochee survived long and the direction of parsing changed from left-to-right to right-to-left only in early Modern English (after 1570: Dresher & Lahiri 2015; Lahiri 2015), when the number of Latin loanwords with stress-affecting suffixes, shown in (15), had passed a threshold which we can derive from Yang's Tolerance Principle (Yang 2016) (16). In (15), we compare the number of stress-affecting Latinate suffixes in 1400 and in 1570.[5]

(15)   Latinate words with stress-affecting suffixes in English (*L*) in 1400 and 1570

	L	*1400*	*1570*	*% δ*
a.	*-able*	204	906	344%
b.	*-al* (adj)	163	745	357%
c.	*-an* (adj)	64	313	389%
d.	*-ar* (adj)	41	104	154%
e.	*-ation*	242	957	295%
f.	*-efy*	3	10	233%
g.	*-etude*	2	3	50%
h.	*-ety*	19	40	111%
i.	*-ible*	40	146	265%
j.	*-ic*	87	279	221%
k.	*-ify*	26	80	208%
l.	*-ile*	35	69	97%
m.	*-ion*	507	1,717	239%
n.	*-ison*	34	52	53%
o.	*-itude*	9	41	356%
p.	*-ity*	144	563	291%
q.	*-ous*	168	657	291%

(16)   Tolerance Principle (Yang 2005, 2016)
Let *R* be a rule that is applicable to *N* items, of which *e* are exceptions. *R* is productive if and only iff $e \leq \theta_N$ where

$$\theta_N = \frac{N}{\ln N}$$

Rough numbers are shown in (17).

---

5.   See Dresher & Lahiri (2015) for further discussion of Yang's principle and of the rationale behind our calculations. The figures in (15) are based on searches of the *Oxford English Dictionary* (*OED*) done by December 2015.

(17)   Latinate suffixes and the Yang Threshold, *Y*

Category	1400	1570
a. All words (*N*)	30,568	69,364
b. ln *N*	10.33	11.15
c. *N*/ln *N* = *Y*	2,960	6,223
d. Latin suffixes (*L*)	1,788	6,682
e. *L*/*Y*	60.4%	107.4%

In 1400, the number of words with Latinate stress-affecting suffixes (*L*) is 60% of the Yang Threshold, *Y*: they can be treated as exceptions to the stress rule. In 1570, *L* is 7% over *Y*: these words can affect the directionality of the English stress rule, changing it from *left* to *right*.

The shift in directionality reestablished a degree of metrical coherence that had been disrupted by the increasing number of Latin loanwords that were inconsistent with the old system. This is because many native words were also compatible with parsing from the right; again, special adjustments were required to incorporate some words into the new system.

An interesting snapshot of the stress patterns in the period close to where we believe the change in directionality was taking hold is provided by Peter Levins' *Manipulus Vocabulorum*, published in 1570. It is a reverse (rhyming) dictionary and indicates main stress in many words. Some words are shown in (18).

(18)   Levins' *Manipulus Vocabulorum* (1570): some stress patterns

Non-initial stress			
FINAL		**PENULT**	**ANTEPENULT**
quarrél	(*v*)	oriéntall	antíquitie
rebéll	(*v*)	aduénture	seuéritie
lamént	(*v*)	recógnise	memóriall
flagón	(*n*)	conféssour	opínion

Initial stress			
2 σ		**4, 5 σ**	**3 σ**
quárel	(*n*)	díuisible	túrpentine
rébel	(*n*)	húmidity	défectiue
députe	(*n*)	bárbaritie	cánonise
récorde	(*n*)	príncipalitie	mármalad

Levins' list leaves little doubt that the grammar of stress was in flux, though how to interpret this is not obvious (see Halle & Keyser 1971: 109–123 for a detailed study in a different framework). We propose that main stress is still mainly on the

left, and that much of the variability in Levins can be accounted for by two ongoing changes summarized in (19).

(19) Changes in Levins' grammar of stress (Lahiri 2015)
   a. An increase in the number of morphologically governed stress-alternating doublets, consistent with parsing from the *right*, with the final syllable of nouns becoming extrametrical.
   b. Words with Latin stress-affecting suffixes show variation: the direction of parsing is changing from *left* to *right*.

The first change is an increase in the number of morphologically governed stress-alternating doublets: we find verbs *quarrél, rebéll, depúte, recórde* contrasting with nouns *quárel, rébel, députe, récorde* (a few nouns like *flagón* are exceptions). It becomes increasingly difficult to regard the first syllables of these verbs as unstressed prefixes, suggesting a right-edge oriented reanalysis as in (20a). Nouns can continue to be parsed in the old way, or can accommodate to the shift in the verbs by making their final syllable extrametrical, as in (20b).

(20) Noun~verb pairs in Levins

a. "quarrel VERB"
```
            x
     .   (|x|)
   [ L     H ]_VERB
   qua    rél
```

b. "quarrel NOUN"
```
            x
       (|x|)
   [  L   <H> ]_NOUN
   quá     rel
```

Here we will focus on the second change in (19b). This change concerns the treatment of complex Latin words with stress-affecting suffixes. In the earlier period such words had come in as simplex forms, and were assimilated to the native pattern of stress assignment from the left. Such words remained in the grammar. Hence, we find *dívisible* and *húmiditie*, which must be parsed from the left edge, as in (21); there is no plausible way to get main stress on the first syllable parsing from the right.

(21) Words in Levins with Main stress *left*, Direction *left* (older grammar)

a. "dívisible"
```
     x
   (|x   |  .)   .
   [ L  L  L   L]_ADJECTIVE
    di  ui  si  ble
```

b. "húmidity"
```
       x
   (|x   |   .)   .
   [ L   L   L   L]_NOUN
    hu  mi  di  tie
```

Nevertheless, unlike the earlier period where *all* words were parsed from the left, we now find words like *sevéritie* and *opínion* which must be parsed from the right (22).

(22) Words in Levins with Main stress *left*, Direction *right* (newer grammar)

    a. "sevérity"   b. "opínion"

```
        x                              x
     . (|x   |  .)                   . (|x   |  .)
     [L  L  L  L]_NOUN              [L  L  L  L]_NOUN
     se  ve  ri tie                 o   pi  ni on
```

Many words with initial stress have ambiguous directionality, because the main stress parameter remains set to *left*, and Levins does not indicate secondary stresses; therefore, words like *bárbaritie, príncipalitie* can be parsed from either direction (23).

(23) Words in Levins with Main stress *left*, Direction ambiguous

    a. "bárbarity" from the *left* b. "bárbarity" from the *right*

```
        x                              x
     (|x|  .)  |x   |)               (|x|)  (|x  |  .)
     [ H  L  L  L]_NOUN             [ H     L  L  L]_NOUN
      bar  ba ri tie                 bar    ba ri tie
```

These facts suggest that the direction of parsing is changing from *left* to *right*, while main stress remains set to *left*. To say that a grammar is changing is shorthand for what is really going on, which in our case could be a number of things. It is important to bear in mind that a written record may not be a perfect reflection of any individual's grammar, but could include forms generated by the grammars of a previous generation, as well as from slightly different contemporaneous grammars. An individual might acquire such forms as exceptions to the prevailing rule, or as a special class subject to their own rule. Or, as has been proposed by Kroch (1989) and Kroch & Taylor (1997), individual speakers might internalize two or more different grammars.

Any of these scenarios may apply in Levins' case: thus, *húmidity* might be a word still in common use that he learned as an exception to his working grammar, or perhaps as one of a class of exceptions that must be parsed from the left; or he might himself have internalized two different grammars, one with stress parsed from the left and one with stress parsed from the right. While the exact state of Levins' grammar may not be knowable, with the benefit of hindsight we can say that, over time, the older grammar with stress from the left eventually ceased to be acquired by learners of the language, and the newer grammar with stress computed from the right came to prevail.

4.   The change in the edge of main stress

While the change in directionality of foot parsing was well underway by 1600, the change in the position of main stress (14c), from *left* to *right*, started later. As pointed out by Dresher & Lahiri (2005), Latin/French borrowings remained subject to the Countertonic Principle (Danielsson 1948, who attributes the observation to Walker 1791), whereby the tonic and countertonic of the Latin/French original were switched when Englished to maintain main stress on the left even when the directionality was from the right; e.g., Latin *àcadémia* became English *ácadèmy*; French *acàdemíe* became English *acádemỳ*.

While the immediate trigger for the change in main stress is not as clear as for the shift in directionality, in Dresher & Lahiri (2005) we speculated that the change may have begun around 1660, the year which, according to Danielsson (1948: 29) was the "turning point" when French words kept their final accent in English, as with suffixes like those in (24).[6]

(24)   Suffixes retaining main stress
       -ade, -ee, -eer, -esque, -ette, -oon

(25)   Words with final stressed suffixes in Present Day English
       cannoneer (1562), *arabesque* (1611), *parade* (1656), *grenadier* (1676), *payee* (1758), *musette* (1811)

While the addition of these suffixes would have helped to push main stress to the right, Lahiri (2015) finds that the change to main stress *right* took a long time to complete, and may not be entirely completed yet.

John Walker's 1791 *A critical pronouncing dictionary and expositor of the English language* is a justly celebrated account of English stress in his time (see Halle & Keyser 1971 for discussion). We observe that the change in parsing direction that was ongoing in Levins appears to have been completed in Walker. Thus, we observed that the words with initial stress in the penultimate column of (18), parsed in (21) and (23), were either parsed from the left, or had ambiguous directionality in Levins. These words are stressed in Walker on the antepenultimate; all these words are now consistent with a direction of parsing from the right edge.

(26)   Words in Walker (1791) with Main stress *right*, Direction *right*
       divísible, humídity, barbárity, principálity

---

6.   1660 marked the start of the Restoration with the return of King Charles II from exile. According to Blake (1996: 238): "The antipathy towards anything foreign, particularly if it had a papist tinge, shown by the Puritans was replaced by the wish to emulate all that was sophisticated and modern in France in particular. Latin loanwords became less frequent as French loans proliferated." See further Dresher (2013).

That is, *divísible* and *humídity* can now be parsed like *sevérity* and *opínion* in (22). Moreover, *barbárity* and *principálity* also show that main stress is on the right (27).

(27)   Walker: Main stress *right*, Direction *right*

a.   "barbárity"   b.   "principálity"

```
            x                              x
(|x|) (|x   |   .)        (|x|   .) (|x   |   .)
  H    L   L   L            H   L   L   L   L
 bar   ba  ri  ty         prin ci  pa  li  ty
```

Nevertheless, the old rule of putting main stress on the left continued to have influence. Walker writes (1791: 67):

> nor has even the interposition of two consonants been always able to keep the accent from mounting up to the antepenultimate syllable, as we may see in *minister, sinister, character, magistrate*, &c. and this may be said to be *the favourite accent of our language*.   [emphasis added]

Note that the antepenultimate syllable in these words is the initial syllable.

We still have words like in (28), such as *mátrimony, húrricane, láboratory*, etc., with main stress on the left, against the now general rule. These words require some sort of special treatment in the modern language: in terms of Halle & Keyser, a series of stress retraction rules, for example.

(28)   Words that continue to have main stress on the *left*
       mátrimony, húrricane, ánecdote, tábernacle, cávalcade, brígantine, túrpentine, láboratory (N. American)[7]

Like *hēafudu* long before, words that were once mainstream in the old grammar persist as exceptions in the new one.

Acknowledgements

---

7.   The *OED* comments that the British pronunciation with stress on the second syllable seems to have been first noted in 1895.

Funding

References

Bermúdez-Otero, Ricardo. Manuscript. *A*-stem nouns in West Saxon: Synchrony. *From: The life cycle of constraint rankings: Studies in early English morphophonology*. 2005. http://www.bermudez-otero.com/lifecycle_chapter4.pdf. (30 June, 2021).

Bermúdez-Otero, Ricardo & Richard M. Hogg. 2003. The actuation problem in Optimality Theory: Phonologization, rule inversion, and rule loss. In D. Eric Holt (ed.), *Optimality Theory and language change*, 91–120. Dordrecht: Kluwer Academic Publishers. https://doi.org/10.1007/978-94-010-0195-3_4

Blake, Norman F. 1996. *A history of the English language*. London: Macmillan Press. https://doi.org/10.1007/978-1-349-24954-1

Danielsson, Bror. 1948. *Studies on accentuation of polysyllabic Latin, Greek, and Romance loanwords in English, with special reference to those ending in -able, -ate, -ator, -ible, -ic, -ical, and -ize*. Stockholm: Almqvist & Wiksells.

Dresher, B. Elan. 1985. *Old English and the theory of phonology*. New York: Garland. Reissued 2019. Oxford: Routledge.

Dresher, B. Elan. 1993. The chronology and status of Anglian smoothing. In Sharon Hargus & Ellen Kaisse (eds.), *Studies in Lexical Phonology: Phonetics and Phonology*, Volume 4, 325–341. New York: Academic Press. https://doi.org/10.1016/B978-0-12-325071-1.50018-2

Dresher, B. Elan. 2013. The influence of loanwords on Norwegian and English stress. *Nordlyd* 39(2). 24–43. Special Issue: Sylvia Blaho, Martin Krämer & Bruce Morén-Duolljá (eds.), *A Festschrift on the occasion of X years of CASTL phonology and Curt Rice's Lth birthday*. https://septentrio.uit.no/index.php/nordlyd/. https://doi.org/10.7557/12.2475

Dresher, B. Elan & Aditi Lahiri. 1991. The Germanic Foot: Metrical coherence in Old English. *Linguistic Inquiry* 22(2). 251–286.

Dresher, B. Elan & Aditi Lahiri. 2005. Main stress left in early Middle English. In Michael Fortescue, Eva Skafte Jensen, Jens Erik Mogensen & Lene Schøsler (eds.), *Historical Linguistics 2003. Selected papers from the 16th International Conference on Historical Linguistics, Copenhagen, 10–15 August 2003*, 75–85. Amsterdam: John Benjamins. https://doi.org/10.1075/cilt.257.07dre

Dresher, B. Elan & Aditi Lahiri. 2015. Romance loanwords and stress shift in English: A quantitative approach. Paper presented at the Second Edinburgh Symposium on Historical Phonology, University of Edinburgh, December 3–4. https://dresher.artsci.utoronto.ca/talks/Dresher-Lahiri-ESohPh_talk_pub..pdf. (30 June, 2021).

Fikkert, Paula, B. Elan Dresher & Aditi Lahiri. 2006. Prosodic preferences: From Old English to Early Modern English. In Ans van Kemenade & Bettelou Los (eds.), *The handbook of the history of English*, 125–150. Oxford: Blackwell. https://doi.org/10.1002/9780470757048.ch6

Fulk, Robert D. 2010. The roles of phonology and analogy in Old English high vowel deletion. *Transactions of the Philological Society* 108(2). 126–144. https://doi.org/10.1111/j.1467-968X.2010.01233.x

Goering, Nelson. 2016a. Early Old English foot structure. *Transactions of the Philological Society* 114(2). 171–197. https://doi.org/10.1111/1467-968X.12079

Goering, Nelson. 2016b. The linguistic elements of Old Germanic metre: Phonology, metrical theory, and the development of alliterative verse. Oxford: University of Oxford PhD thesis.

Halle, Morris & Samuel Jay Keyser. 1971. *English stress: Its form, its growth, and its role in verse*. New York: Harper & Row.

Halle, Morris, Wayne O'Neil & Jean-Roger Vergnaud. 1993. Metrical coherence in Old English without the Germanic Foot. *Linguistic Inquiry* 24(3). 529–539.

Hogg, Richard M. 1992. *A grammar of Old English. Volume 1: Phonology*. Oxford: Blackwell.

Hogg, Richard M. 2000. On the (non-)existence of High Vowel Deletion. In Aditi Lahiri (ed.), *Analogy, levelling, markedness*, 353–377. Berlin: Mouton de Gruyter. https://doi.org/10.1515/9783110808933.353

Hogg, Richard M. & Robert D. Fulk. 2011. *A grammar of Old English. Volume 2: Morphology*. Oxford: Wiley-Blackwell.

Hutton, John. 1998. The development of secondary stress in Old English. In Linda van Bergen & Richard M. Hogg (eds.), *Historical Linguistics 1995, Volume 2: Germanic linguistics*, 115–130. Amsterdam: John Benjamins. https://doi.org/10.1075/cilt.162.10hut

Idsardi, William J. 1994. Open and closed feet in Old English. *Linguistic Inquiry* 25(3). 522–533.

Keyser, Samuel Jay & Wayne O'Neil. 1985. *Rule generalization and optionality in language change*. Dordrecht: Foris Publications. https://doi.org/10.1515/9783110250541

Kroch, Anthony. 1989. Reflexes of grammar in patterns of language change. *Language Variation and Change* 1. 199–244. https://doi.org/10.1017/S0954394500000168

Kroch, Anthony & Ann Taylor. 1997. The syntax of verb movement in Middle English: Dialect variation and language contact. In Ans van Kemenade & Nigel Vincent (eds.), *Parameters of morphosyntactic change*, 297–325. Cambridge: Cambridge University Press.

Kuryłowicz, Jerzy. 1949. Latin and Germanic metre. *English and Germanic Studies* 2. 34–38. Reprinted 1950 in *Biuletyn Polskiego Towarzystwa Językoznawczego X and in Jerzy Kuryłowicz. 1973. Esquisses linguistiques I*. Munich: Wilhelm Fink.

Kuryłowicz, Jerzy. 1970. Die sprachlichen Grundlagen der altgermanischen Metrik. *Innsbrucker Beiträge zur Sprachwissenschaft* Vorträge 1. Innsbruck. Reprinted in Jerzy Kuryłowicz. 1975. *Esquisses linguistiques II*. Munich: Wilhelm Fink.

Lahiri, Aditi. 1982. Theoretical implications of analogical change: Evidence from Germanic languages. Providence, RI: Brown University PhD thesis.

Lahiri, Aditi. 2015. Change in word prosody: Stress and quantity. In Patrick Honeybone & Joseph Salmons (eds.), *The handbook of historical phonology*, 219–244. Oxford: Oxford University Press.

Lahiri, Aditi, Tomas Riad & Haike Jacobs. 1999. Diachronic prosody. In Harry van der Hulst (ed.), *Word prosodic systems in the languages of Europe*, 335–422. Berlin: Mouton de Gruyter. https://doi.org/10.1515/9783110197082.1.335

Levins, Peter. 1570. *Manipulus vocabulorum: A dictionary of English and Latin words, arranged in the order of the last syllables, by Peter Levins (1570)*. Re-edited, with a preface and alphabetical index, by Henry B. Wheatley. 1867. Westminister: Nichols & Sons for the Camden Society.

Minkova, Donka. 1997. Constraint ranking in Middle English stress-shifting. *English Language and Linguistics* 1. 135–175. https://doi.org/10.1017/S1360674300000393

Redford, Michael. 2003. Middle English stress doubles: New evidence from Chaucer's meter. In Paula Fikkert & Haike Jacobs (eds.), *Development in Prosodic systems (Studies in Generative Grammar 58)*, 159–196. Berlin: Mouton de Gruyter. https://doi.org/10.1515/9783110894530.159

Roca, Iggy M. 1999. Stress in the Romance languages. In Harry van der Hulst (ed.), *Word prosodic systems in the languages of Europe*, 659–811. Berlin: Mouton de Gruyter. https://doi.org/10.1515/9783110197082.2.659

Russom, Geoffrey. 1987. *Old English meter and linguistic theory*. Cambridge: Cambridge University Press.

Sievers, Eduard. 1893. *Altgermanische Metrik*. Halle: Niemeyer.

Walker, John. 1791. *A critical pronouncing dictionary and expositor of the English language*. London: Robinson.

Yang, Charles. 2005. On productivity. *Linguistic Variation Yearbook* 5(1). 265–302. https://doi.org/10.1075/livy.5.09yan

Yang, Charles. 2016. *The price of linguistic productivity: How children learn to break the rules of language*. Cambridge, MA: MIT Press. https://doi.org/10.7551/mitpress/9780262035323.001.0001

Ambiguity resolution and the evolution of homophones in English

Mieko Ogura and William S-Y. Wang

Tsurumi University / Hong Kong Polytechnic University

Based on a quantitative study of the evolution of homophones in English, we present an argument about why homophones occur. Zipf's law, which states that word frequency decreases as a power law of its rank, can be seen as the outcome of form-meaning associations, adopted in order to comply with listener and speaker needs. This implies that one form can correspond to many meanings (i.e., polysemy and homophony). We argue that homophony is a desirable feature in communication systems, is stable, and increases through time. When a large number of homophones emerge, however, an impetus to avoid homophones comes into play. We suggest that the evolution of diatones is a case of the avoidance of homophony. Related to this, we examine the neural substrates of bisyllabic noun-verb homophones, using near-infrared spectroscopy. We show that noun and verb categories are represented in different neural substrates in the left hemisphere, and relate this to our historical data, explaining why the actuation of diatone-formation was connected with production in frequent homophones in the 16th century, but was connected with perception in infrequent words after the 17th century.

Keywords: linguistic evolution, homophones, ambiguity, Zipf's law, word frequency, diatones, neural substrates

1. Introductory remarks

The existence and abundance of ambiguity in languages has long intrigued linguists. If we view language as a system to encode meanings as signals, it would seem that language is not ideal, because in an ideal code one signal should correspond to exactly one meaning. Each linguistic form should map bijectively to one meaning, so that listeners do not need to expend effort inferring what the speaker intends to convey (Ogura & Wang 2006, Wang 2011).

When there are one-to-many correspondences between form and meaning, as in homophony and polysemy, lexical ambiguities arise. Following Piantadosi et al. (2012), we argue that all efficient communication systems are ambiguous, assuming that context is informative about meaning. In the neural substrates, all meanings become active initially and the context leads to the appropriate meaning, or frequency-based dominance and contextual congruity lead to the appropriate meaning (Zempleni et al. 2007).

Anttila (1989: § 9.5) argues that all languages have homophony to different degrees, and that we can never predict with complete confidence when a community or speaker will find it inconvenient enough to remove it. Even when the avoidance or correction of homophony does occur, there is no way of telling by what mechanism of change it will happen.

In this paper, we offer an argument for the existence of homophones, based on a quantitative study of the evolution of homophones in English. We argue that homophony is a desirable feature in communication system and that it likely results from the ubiquitous pressure for efficient communication. However, when a large number of homophones emerge, a force which aims to avoid homophones comes into play. We examine the evolution of diatones, i.e., noun-verb pairs where the stress falls on the first syllable for the noun but the second syllable for the verb, as a case of the avoidance of homophony. In addition to this, we investigate the neural substrates of homophones and diatones in English.

2.   The evolution of homophones

2.1   Why do homophones exist?

Normal communication is not much hampered by homophones, because we generally do not process sentences in isolation from contexts. Our quantitative claims could be criticised on the basis that instances of homophonic clash cause ambiguity and must be examined on a case-by-case basis. However, as shown in Table 1, the number of homophones decreases as the number of words in a homophonous set increases, and the percentage of homophones for each number of syllables decreases as the syllable number increases.

Zipf (1949: Chapter 2) argues that the balance of speakers' and listeners' interests is met in a balance between word frequency and the number of words (see below). He further extends his thinking to the balance between the average number of meanings per word and the word rank, showing that frequent words are more ambiguous. Zipf (1949: Chapter 3) also discusses the relation between word frequency and the length of words in terms of the number of syllables, showing that short words are more frequent. We can observe in the number of syllables and the

number of words in the homophonous set in Table 1 (together with the number of syllables and word frequency in Table 2 below), that the smaller the number of syllables, i.e., the shorter the words, the higher the word frequency and the larger the number of homophonous sets. Also, the larger the number of words in a homophonous set, the more ambiguous the homophonous set is, and therefore the more ambiguous the shorter words are.

Piantadosi et al. (2012) quantify Zipf's (1949) proposal for trading off between ease of comprehension and ease of production by means of the concepts of 'clarity' and 'ease'. A clear communication system is one in which the intended meaning can be recovered from the signal with high probability. An easy communication system is one in which signals are efficiently produced, communicated, and processed. There are many factors which likely determine ease for human language: for instance, words which are easy to process are likely short, frequent, and phonotactically well-formed. We can observe the trade-off between two communicative pressures (clarity and ease) in the number of syllables and the number of words in the homophonous set in Table 1. The smaller the number of syllables (i.e., the shorter the word) the easier the word is to produce. The larger the number of words in a homophonous set, the smaller the probability with which the intended meaning can be recovered from the signal.

Based on the CELEX lexical database of English (version 2.5, Baayen et al. 1995), we give, in Table 1, the number of homophones in English (classified according to the number of syllables) in the columns, and the number of words in a homophonous set in the rows. The total number of types classified according to number of syllables is also given in parentheses. 11,980 (22.8%) of the 52,447 types are homophones in Present-Day English. 4,743 (70.2%) of the 6,761 one-syllable words and 4,509 (24.3%) of the 18,564 two-syllable words are homophones. In total, they form 77.2% of all homophones.[1] That is, the majority of homophones are short and easy to produce, and they are frequent words. The 6,761 one syllable words and 18,564 two syllable words make up 48.3% of all 52,447 words. The distribution of homophones suggests a threshold of homophones that can be tolerated: around 20% of all words, which is the boundary of high frequency words (see below and § 2.2).[2]

---

1. We have also analyzed the LDC Japanese Lexicon (1997) and found similar results in Present-Day Japanese: 8,827 (17.2%) of the 51,274 types are homophones; 145 (58.3%) of the 252 one-syllable words and 1.232 (40.9%) of the 2,946 two-syllable words are homophones (Ogura & Wang 2006).

2. Strang (1980) collected some 1700 monosyllabic words from the *OED*, and found that the level of exploitation of homophones is higher than might be thought tolerable in several forms: /biː/, /bɛː/, /bei/, /bʌt/, /bʌk/. For example, /bʌt/ has had 21 homophones, including 16 nouns and 4 verbs, with 12 surviving till the 19th century.

Table 1. The number of homophones in Present-Day English

	1 syl.	2 syls.	3 syls.	4 syls.	5 syls.	6 syls.	7 syls.	8–12 syls.	Total
2 words	3068	3888	1792	588	132	8	6	0	9482
3 words	900	477	123	60	3	0	0	0	1563
4 words	460	104	16	0	0	0	0	0	580
5 words	175	40	0	0	0	0	0	0	215
6 words	96	0	0	0	0	0	0	0	96
7 words	28	0	0	0	0	0	0	0	28
8 words	16	0	0	0	0	0	0	0	16
total	4743	4509	1931	648	135	8	6	0	11980
	(6761)	(18564)	(15195)	(7970)	(3000)	(711)	(188)	(58)	(52447)

Why do homophones occur, despite the fact that, in an ideal code, one signal corresponds to one meaning? Zipf (1949: Chapter 2) suggests this is due to the simultaneous minimization of the two opposing forces from listener and speaker for form and meaning association. One form for all meanings is an ideal code for the speaker, while one-to-one correspondence between form and meaning is an ideal code for the listener. Zipf's law, which states that word frequencies decrease as a power law of its rank, may be seen as the outcome of form-meaning association adopted in order to comply with listener and speaker needs. Arranging signals according to Zipf's law is the optimal solution for maximizing the referential power while respecting the constraint regarding effort for a speaker. Zipf's law implies that there should be one form to many meanings, i.e., polysemy and homophony. Polysemy and homophony are therefore the necessary conditions for symbolic systems (Ferrer i Cancho & Solé 2003).

Zipf argues that the balance of speakers' and listeners' interests is observed in a balance between word frequency and the number of words. If word frequencies are tabulated and the words ranked from most frequent to least frequent, a simple formula describes the relation. This relation is known as Zipf's law. If p_r is the probability of the rth most frequent word, then:

$$Pr = \frac{0.1}{r}$$

When plotted on log-log coordinates, this function gives a straight line with a slope of -1 (Zipf 1949: Chapter 2; Miller 1981: 107). It shows a heavy-tailed distribution, and we may draw a boundary between the most frequent 2,000 words and the remaining 8,000 words in a set of 10,000 words (see § 2.2 below).

Ferrer i Cancho & Solé (2003) provide a formal backing for Zipf's intuitive explanation, showing that the power law distribution arises when information-theoretic difficulty for speakers and listeners is appropriately balanced. Piantadosi et al.

(2012) argue that ambiguity can be understood as the trade-off between two communicative pressures which are inherent to any communicative system: clarity and ease (see above). For further discussion on the trade-off between speaker and listener needs in the neural substrates, see § 4.1.

2.2   Word frequencies and homophones in Present-Day English

Table 2 shows the average word frequencies of homophones, and of all words, classified according to their number of syllables, in the CELEX database of English, where the frequency information is taken from the 17.9 million token COBUILD/ Birmingham Corpus. We also give the percentage of homophones (for the number of homophones and of all words, see Table 1). Table 3 gives the average word frequencies ranked from the most frequent to the least frequent in the CELEX database, and Figure 1 plots the power-law distribution of word frequencies as a function of word rank on log-log coordinates. This graph shows a similar slope

Table 2. Average word frequencies of homophones and of all words, and the percentage of homophones

	1 syl.	2 syls.	3 syls.	4 syls.	5 syls.	6 syls.	7 syls.	8–12 syls.
homophones	2082.5	320.6	101.26	69.91	60.17	1.75	3.5	
all words	2009.8	182.13	72.39	49.86	34.32	13.49	6.63	0.79
% of homophones	70.2%	24.3%	12.7%	8.1%	4.5%	1.1%	3.2%	0%

Table 3. Average word frequencies ranked from the most frequent to the least frequent in the CELEX database

Word rank	Average frequency
1–500	26768.54
501–1000	2627.33
1001–2000	1262.48
2001–3000	674.13
3001–4000	430.35
4001–5000	294.31
5001–10000	140.55
10001–15000	53.26
15001–20000	26.24
20001–25000	13.57
25001–30000	6.77
30001–35000	2.97
35001–38731	1.13
38732–52447	0.00
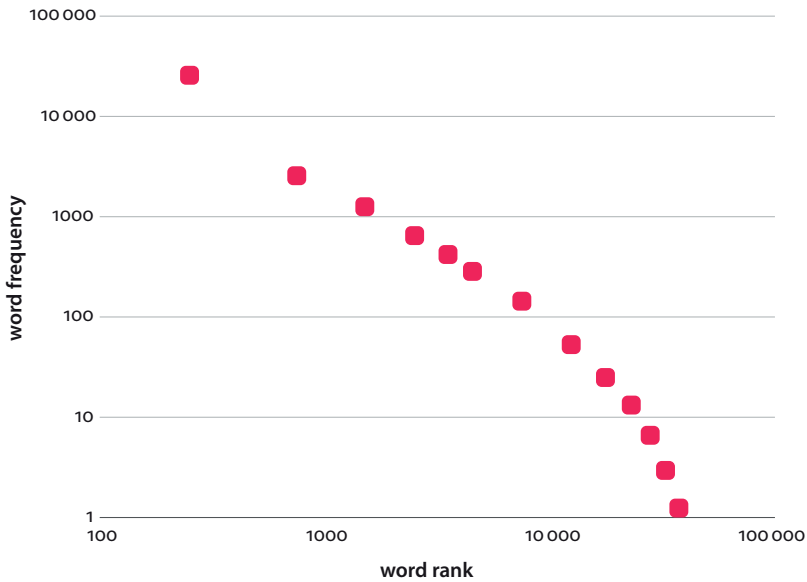
Figure 1. Power-law distribution of word frequency in the CELEX database

to Zipf's law, which gives a straight line with a slope of −1. We find that within each number-of-syllables group, average word frequency of homophones is higher than all words except the 6 and 7 syllable words. The average frequencies of homophones in 1, 2 and 3 syllable words are in the rank of 1001–2000, 4001–5000 and 5001–10000, respectively. Homophones with 1, 2 or 3 syllables form the majority of homophones (4743 + 4509 + 1931 out of 11980, i.e., 93%), and the average frequencies are in the top 10,000 most frequent words out of 52,447 words. If we draw a boundary of highly frequent words around the top 20% most frequent words (see § 2.1 above), the threshold of homophones could also be expected to be at around 20% of all words, because the homophones tend to be formed with highly frequent words.

There are subtypes of iconicity, isomorphism (that is, the one-form-one-meaning condition) and automorphism (which holds that linguistic elements which are alike semantically should also resemble one another formally; Haiman 1985: 4). The former type of iconicity underlies homophony, while the latter type underlies polysemy. Because of the automorphism of iconicity, it is plausible that there are many polysemous words. In WordNet we find 5,196 (46.0%) out of 11306 verbs, and 15,279 (13.3%) out of 114,513 nouns are polysemous. We also find that, the greater the number of meanings, the higher is word frequency within each lexical category (Ogura & Wang 2012a). As for homophony, people avoid homophones because of isomorphism of iconicity. Thus we assume that there is some threshold

of homophony. The distribution of homophones in Present-Day English in Table 1 shows this situation.

We would like to add that most of the homophones are composed of words from different lexical categories in English. Based on the first 4,919 samples out of 11,980 homophones from the CELEX database, we find that the number of occurrence of words whose lexical categories occur once in each homophonous set is 4,263 (86.7%) of 4,919 homophones. Noun-verb ambiguous pairs form a large portion of these inter-category homophones in English.[3]

We argue that homophony is a desirable feature in a communication system and that it likely results from a ubiquitous pressure for efficient communication. However, when a large number of homophones emerge, the balance of speakers' and listeners' interests observed in the balance between word frequency and the number of words is lost, and the force to avoid homophones comes into play. One such case is the evolution of diatones in English. Before we discuss the evolution of homophones and diatones after the 16th century (in § 3), we consider the evolution of homophones in Old English (OE) and Middle English (ME).

2.3   Homophones in Old English and Middle English

2.3.1   *The stability of OE homophones*
The Brooklyn-Geneva-Amsterdam-Helsinki Parsed Corpus of Old English (Pintzuk et al. 2000) contains 18,629 types, and 1205 types (6.5%) of them are homophones. The scoring criterion used here, and definition of homophony, is different from that used for the CELEX database of English and the LDC Japanese Lexicon, where a lexeme (a set of forms taken by a single root word) is a basic unit of scoring. However in OE, any differences in forms are counted as different types except scribal variations, because the inflectional system of OE was quite complex and distinct in form in nouns, personal pronouns, adjectives, and verbs (Moore 1967: § 26). Thus, for example, we see *sum* (which has become *some* in Present-Day English), as having the following 8 homophonous sets: *sum, suman, sume, sumere, sumne, sumon, sumre, sumum* as adjective and pronoun. These are treated as one homophonous set ("some") as an adjective and pronoun in the CELEX database. If a root is used as a representative form, OE has a very restricted number of root words that can stand alone as a word. On the other hand, Present-Day English has little inflection and a tendency to have words that are identical to their roots.

---

3.   In Japanese, however, most of the homophones are composed of words from the same lexical categories. The number of occurrences of words whose lexical categories occur once in each homophonous set is 365 (4.1%) of 8,902 homophones (Ogura & Wang 2006).

In OE double consonants were typically geminates. In late OE, because of the open syllable lengthening of disyllabic words, long/short consonants and long/short vowels occur in complementary distribution: long consonants after short vowels, and short consonants after long vowels. Thus long consonants are no longer phonemically long, but we use the phonetically long pronunciation, because our calculation for homophonous forms is based on phonetically different forms. Thus, for example, the homophones *al* as adjective and pronoun are treated as different from the homophones *all* as adjective and pronoun (both of which remain as *all* in Present-Day English).

Zipf (1949: Chapters 2, 3) argues that all words which differ phonetically in the fully inflected form in which they occur should be counted as separate words (thus the forms, *give, gives, gave, given, giving, giver, gift* represent seven different words and not one word in seven different forms). A word is defined as any sequence of successive letters bounded by spaces in the text. He employs the criterion of inflection when he analyzes the rank-frequency distribution of words based on James Joyce's *Ulysses*, and especially in the treatment of highly inflected languages, such as Nootka, German, Gothic and Ælfric's Old English which all use inflections virtually throughout their entire structure. However, he uses the *Thorndike-Century Dictionary*, where lexical units, i.e., words in non-inflected form are listed, when he analyzes the meaning-frequency distribution of words. He states that because of the low degree of inflection of words in Present-Day English, the difference between a rank-frequency distribution of lexical units and that of words in fully inflected form is not considerable.

780 (65%) of 1,205 OE homophones are still used in Present-Day English, and 425 have become obsolete. Table 4 gives the number of types that became obsolete in a set of periods. The dates of obsolescence were checked in the *Oxford English Dictionary, Version 2.0 on CD-ROM (OED2)*. Many of the 144 homophones that became obsolete in the 19th century are archaic, literary or dialectal in Present-Day English. We also find that in 487 (85%) of 572 homophonous sets, at least one type exists in Present-Day English. Thus we may state that OE homophones have been stable over time.

Table 4. Number of OE homophones that became obsolete

OE	37
12th c.	13
13th c.	68
14th c.	43
15th c.	49
16th c.	23
17th c.	35
18th c.	13
19th c.	144

2.3.2 *Stability of ME homophones*

We have attempted to confirm the above observation about OE homophones in terms of ME homophones, too, based on the *Penn-Helsinki Parsed Corpus of Middle English* (Kroch & Taylor 2000). This corpus contains 48,725 types (2,003 clitics are excluded), and we find that 4,653 (9.6%) of them are homophones.

2,691 types (= 1,966 types of OE origin + 725 types of ME origin), which is 57.8% of 4,653 ME homophones (= 2,981 types of OE origin + 1,672 types of ME origin) are used in Present-Day English. Table 5 gives the number of types that became obsolete from the 12th to the 19th century, classified according to OE origin and ME origin. The total of the types of OE origin that became obsolete in the 19th century and those that are still used in Present-Day English are 2,669 types (703 types + 1966 types), which is 89.5% of the 2,981 homophones of OE origin. The total of the types of ME origin that became obsolete in the 19th century and those that are still used in Present-Day English are 1,377 types (652 types + 725 types), which is 82.4% of 1,672 homophones of ME origin. It is not clear how many homophones of OE origin became obsolete in OE from Table 5, but we can confirm that homophones of both OE and ME origin are stable, and that homophones of ME origin add to those of OE origin.

Table 5. Number of ME homophones that became obsolete

	OE origin	**ME origin**
12th c.	8	7
13th c.	61	29
14th c.	48	33
15th c.	73	49
16th c.	57	48
17th c.	52	92
18th c.	13	37
19th c.	703	652

2.3.3   *Word frequency of OE and ME homophones*

Tables 6 and 7 give the average word frequencies (ranked from the most frequent to the least frequent) in the OE and ME databases, respectively. Figure 2 plots the power-law distribution of word frequencies in the OE and ME database on log-log coordinates. The slopes of OE and ME graphs are parallel, and they are approximately parallel to the graph for Present-Day English in Figure 1, except the heavy tailed part at the bottom. The average frequencies of OE and ME homophones are 30.7 and 87.9 respectively, and they are high frequency words in the word rank of 251–500 of the OE database in Table 6 and the word rank of 501–1000 of ME database in Table 7, respectively.

Table 6. Average word frequencies ranked from the most frequent
to the least frequent in the OE database

Word rank	Average frequency
1–250	216.53
251–500	29.09
501–1000	15.39
1001–1500	9.29
1501–2000	6.62
2001–3000	4.61
3001–4000	3.24
4001–5000	2.45
5001–10000	1.43
10001–18629	1.00

Table 7. Average word frequencies ranked from the most frequent
to the least frequent in the ME database

Word rank	Word frequency
1–250	1520.26
251–500	164.99
501–1000	78.76
1001–1500	43.21
1501–2000	29.52
2001–3000	19.80
3001–4000	13.30
4001–5000	9.88
5001–10000	5.62
10001–15000	2.89
15001–20000	2.00
20001–25000	1.17
25001–48725	1.00
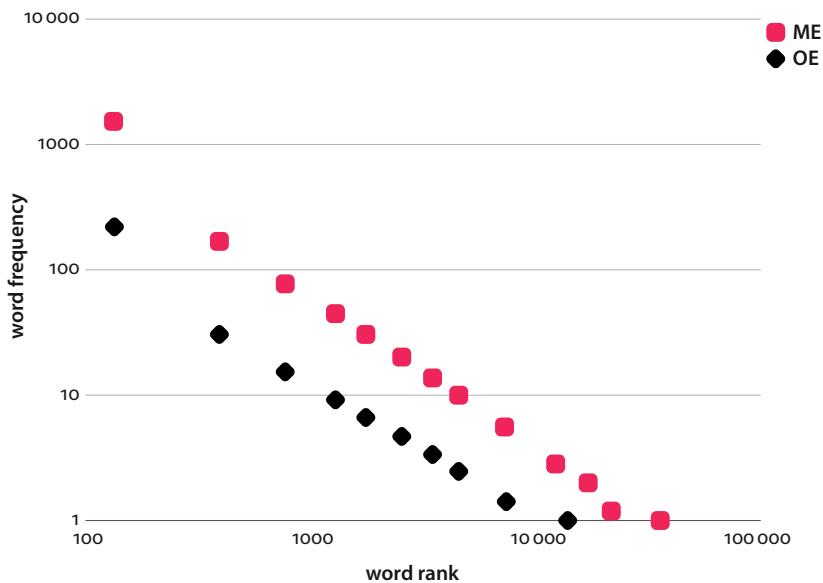
Figure 2. Power-law distribution of word frequency in the OE and ME database

Table 8 gives the average frequencies of the obsolete words from OE to the 19th century in the OE database. We also give the average frequency of the words that remain in the 20th century. Table 9 gives the average frequencies for the obsolete words from the 12th to the 19th century in the ME database, classified according to OE origin and ME origin. We also give the average frequency for the words that remain in the 20th century. We find that words that became obsolete in OE in the OE database in Table 8 are low frequency words. This is to be expected: rare things are forgotten most easily. From the ME database in Table 9, we find that the frequency of words of OE origin is higher than that of words of ME origin. This is because the low frequency words became obsolete in OE. We also find that there is a strong correlation between dates of obsolescence and word frequencies. Words that became obsolete in the 19th century and those that remain in the 20th century are by far the most frequent words both in terms of OE and ME origin. The average word frequencies from the 12th to the 18th century and those in the 19th and 20th century for words of OE origin and ME origin are correlated by 0.999 of the Pearson correlation coefficient at the 5 percent level. We note that the average frequency of words that became obsolete in the 13th century in the OE database is high as compared with that in the ME database. This may be due to the change in which the relative pronouns *þe* and *se*, which were frequently used in OE, were replaced by *þat* in the 13th century.

Table 8. Average frequencies of obsolete words from OE to the 19th century and those of words that remain in the 20th century from the OE database

OE	2.27
12th c.	29.08
13th c.	89.06
14th c.	45.58
15th c.	41.04
16th c.	10.65
17th c.	11.23
18th c.	39.77
19th c.	19.45
20th c.	28.98

Table 9. Average frequencies of obsolete words from the 12th to the 19th century and those of words that remain in the 20th century from the ME database

	OE origin	ME origin
12th c.	17.63	5.71
13th c.	9.10	3.14
14th c.	5.00	3.79
15th c.	47.84	19.27
16th c.	22.16	3.42
17th c.	21.58	3.40
18th c.	5.92	2.97
19th c.	66.15	33.55
20th c.	160.56	22.63

3.   Evolution of diatones

3.1   Why do diatones evolve?

As we have argued in § 2, homophony is a desirable feature in a communication system, it is stable and accumulates over time. But when a large number of homophones emerge, a force to avoid homophones comes into play. We suggest that when a large number of homophones emerged in the 16th century, diatones first emerged in order to avoid the existence of homophones. Based on Sherman's (1975) list of noun-verb oxytonic homophones, we give the number of homophones that emerged from the 16th to the 20th century and the total number of homophones in each century in Table 10. We can see that a large number of homophones emerged in the 16th century.

Table 10. The number of noun-verb oxytonic homophones that emerged from the 16th and 20th century and the total number of homophones in each century

	16th c.	17th c.	18th c.	19th c.	20th c.
number of homophones that emerged	129	33	19	29	−17
total number of homophones	**129**	**162**	**181**	**210**	**193**

Diatones are noun-verb pairs where the stress falls on the first syllable (paroxytonic) for the noun but on the second syllable (oxytonic) for the verb, e.g., *address, permit, subject, contract*, etc. Sherman (1975) gives a catalogue of a large number of historical sources of diatones. For each N-V pair, he gives accentual evidence from dictionaries and grammars, ranging in date from 1570 (Levins) to 1798 (Jones). There are two early dictionaries with relatively complete accent markings. Of these two, only Levins (1570) provides reliable evidence for any N-V diatones, and he gives three such words: *outlaw, rebel, record*. Twelve years later, Mulcaster (1582) provides evidence for five additional disyllabic diatones (*convert, desert, incense, present, refuse*). This scarcity of sources changes in the 18th century, which sees the publication of nearly 250 dictionaries. Sherman's consultation of 30 of these shows that diatones emerged in the latter half of the 16th century, although he does not have an explanation for their emergence at this particular time.

We find several prerequisite linguistic environments for the evolution of homophones and diatones in the 16th century. In OE there was a predominance of word-initial stress, but as Campbell (1959: 30) notes, there were noun-verb pairs (not homographs) which exhibited stress alternation, with word-initial stress being shown by the noun and non-initial stress by the verb as in *'æfþunca* "source of offence" vs. *of'þyncan* "to displease", *'bīgenga* "inhabitant" vs. *be'gān* "to occupy". It is plausible that these pairs supplied a model for the productive process of diatone-creation which emerged later in Early Modern English (EModE).

In ME and EModE we find the development of two stress rules that have survived to the present: the Germanic Stress Rule and the Romance Stress Rule. This is especially the case with disyllabic words, as is evident from alliterative and rhyme poetry in LME. Abundant double-stressed words occur whether they are native or of Romance origin, or whether they are unaffixed, suffixed, or complex words as shown in: *a'leye~'aleye, bo'som~'bosum, de'sire~'dezyre, fol'low~'folow, mes'sage~'message, ly'vynge~'lyvyng, out'lawe~'outlawe, ac'cesse~'access, bi'leve~'bilyue, con'clude~'conclude* (Nakao 1978; Lass 1992).

Another linguistic phenomenon is the gradual leveling and loss of inflections from Late OE through Late ME. As a result of this, nouns and verbs fell into one and the same form in many of their inflected forms, and the morphological process of conversion operated more frequently. Yet another linguistic factor is the great

influx of French and Latin vocabulary in ME and EModE. In addition to their effect on the stress system, the loanwords themselves provided a rich source of disyllabic words with morphological structure of predominantly prefix plus base.

The present study is based on the 237 diatone data from the late 16th to the early 20th century: Sherman's 147 diatones,[4] and the additional data for the 19th and the 20th centuries from Hotta's (2013) 90 diatones. Hotta collected the data from Sherman's oxytonic portion of the noun-verb isotone inventory, Fudge (1984: 189–92), Jerspersen (1954: 173–82), and a number of dictionaries in the 19th and 20th centuries.

The Appendix to this article shows the evolution of 72 diatones that arose from a large number of homophones from the late 16th to the late 18th century. The 72 diatonic pairs are arranged according to the earliest period at which explicit lexicographic evidence is found for stress alternation in the second column. 16c, 17b and 18c, etc. mean the late 16th, middle 17th, and late 18th century respectively. The third column shows the distinction between N(oun) and V(erb).

The fourth column of the Appendix shows the etymologies based on *OED2*. The abbreviations are as follows: OF. Old French; F. French; mod. F. Modern French; AF./AFr. Anglo-French; ONF. Old Northern French; L. Latin; med. L. medieval Latin; mod. L. Modern Latin; ON. Old Norse; OE. Old English; lOE. late Old English; Com. Teut. Common Teutonic; ppl. a. participial adjective; < means 'comes from'; conversions are also shown, such as *relapse* (v), which means *relapse* (n) comes from conversion of *relapse* (v). In the Appendix, we can see a great influx of French and Latin vocabulary in ME and Early Modern English (EModE), which are a rich source of disyllabic words. The fifth column shows the date of the first attestation of each noun and verb based on *OED2*. The sixth column shows the period of the first attestation of each isotonic or homophonous N-V pair from which a diatone emerged. We find that a large number of homophones appeared in the 16th and 17th centuries. This is due to a great influx of French and Latin vocabulary in ME and EModE which occurred with gradual inflectional leveling and loss from LOE through LME.

The seventh column shows the patterns of the accentuation, O(xytonic), P(aroxytonic), and D(iatonic) which Sherman lists based on some 30 dictionaries and grammars, ranging in date from 1570 (Levins) to 1789 (Jones). We may assume that originally some isotonic N-V pairs had the stress on the second syllable (O(xytone)), some on the first syllable (P(aroxytone)), and others both (O and P), and that diatones developed from these three patterns. The three patterns developed

---

4. Sherman lists 150 diatone pairs, but we left out three (*recast, repeat, upright*), because Sherman does not give historical description to them and their diatone status is not attested in contemporary references.

from the double-stressed disyllabic words where one and the same word manifests final stress at one time and initial stress at another time in ME and EModE. The eighth column shows the word frequency of each diatone pair, based on the *American Heritage Word Frequency Book* (Carroll et al.1971).

3.2 Word frequency and diatones

As we follow the course of the evolution of diatones from the 16th to the 20th century, we can see the gradual diffusion of stress-alternation, reaching more and more noun-verb homographic pairs, thereby creating an ever larger number of diatones. The inventory of diatones expanded to 32 by 1700, 72 by 1800, 153 by 1900, 237 by 2000 (here the homograph is counted as diatonic if any source gives a diatonic pattern, even secondarily). If we plot the steady growth of the number of diatonic N-V homographs as a function of time, the resulting graph is a chronological profile of lexical diffusion, as shown in Figure 3. The change is still in progress at the present. Lass (1980: 75–80) argues against homophonic clash or avoidance of homophony as an explanation for certain (usually sporadic) changes. Our work is based on the quantitative examinations of an adaptation through the amplification of and competition between diatones, paroxytone pairs and oxytone pairs, which leads to growth in complexity in language over time.

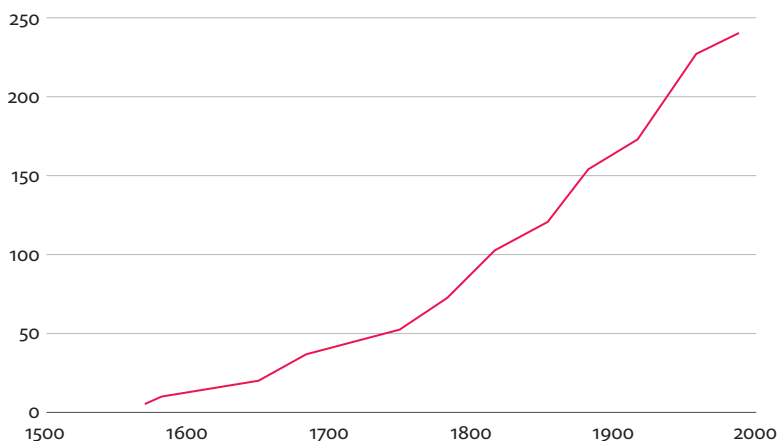**Figure 3.** Chronological profile of the lexical diffusion of N-V diatones

Table 11 gives the average word frequencies for the disyllabic diatones and noun-verb homophones from the 16th to the 20th century. The diatone data from the late 16th to the early 20th century are based on Sherman's 147 diatones and the additional data from Hotta's (2013) 90 diatones. The disyllabic noun-verb homophones are

Table 11. Average word frequencies of disyllabic diatones and noun-verb homophones from the 16th to the 20th century

	16th c.	17th c.	18th c.	19th c.	20th c.
diatone	44.28(8)	16.99(28)	7.45(36)	4.56(81)	5.27(84)
homophone	13.29(129)	13.14(162)	12.44(181)	11.28(210)	11.21(193)

based on Sherman's oxytonic noun-verb isotone inventory. Sherman lists 213 pairs (though he claims 215 pairs), from which 3 pairs and 17 pairs changed to diatones in the 19th and 20th centuries respectively, and they are in Hotta's list. The word frequency (frequency-per-million) of each diatone pair and oxytone pair is based on the *American Heritage Word Frequency Book* (Carroll et al. 1971). The number of occurrences of new diatones and oxytonic homophones are given in parentheses for each century.

We find that the formation of diatones started in the high-frequency homophones in the 16th century, but after the 18th century it gradually diffused from the low-frequency homophones. The 17th century is the transitional period. As shown in the Appendix, most of the diatones that appeared in the late 16th (16c) to middle 17th (17b) century have their sources of both noun and verb or either of them in ME and have higher frequency than those that appeared after the late 17th century (17c), whose sources are mostly French and Latin words that were borrowed after the 16th century.

Just after a large number of homophones emerged in the 16th century, the avoidance of homophony strongly operated in the highly frequent homophones, because, we argue, speakers aimed to avoid homophones intentionally to keep the distinction of the forms and meanings, which were kept distinct in ME before they became homophonous in the 16th century. But in the homophones which remained as they are in the 16th and 17th century, diatones evolved from the less frequent words. Diatone formation is perceptually or cognitively motivated change by the listener, and it is in the less frequent words that the change starts (Phillips 1984).[5]

---

5.   Phillips (1984) groups Sherman's disyllabic noun-verb diatones and oxytonic homophones on the basis of the prefix, and compares their average word frequencies. However, we cannot see any effect of the prefix on word frequency. Moreover, her calculation is inexact, because Sherman's data are sometimes missing in her data.

4. Neural bases of the evolution of homophones and diatones

4.1 Neural substrates of bisyllabic noun-verb homophones in English

Sereno & Jongman (1995) investigate systematic acoustic differences in bisyllabic noun-verb ambiguous pairs which do not exhibit stress differences. 5 speakers read 16 bisyllabic words in both noun and verb contexts. The 16 bisyllabic homophones consist of four categories: 4 pairs which occur more frequently as forestressed nouns (*favor, poison, practice, struggle*); 4 pairs which occur more frequently as backstressed nouns (*control, debate, dispute, report*); 4 pairs which occur more frequently as forestressed verbs (*handle, notice, rescue, welcome*); and 4 pairs which occur more frequently as backstressed verbs (*embrace, escape, neglect, reply*).

Sereno & Jongman (1995) find that stimuli that are more frequent as nouns in English (e.g., *poison, debate*) show significantly different amplitude ratios than word stimuli that are more frequent as verbs (e.g., *notice, escape*) (pp. 68–69). In other words, the amplitude of the first syllable of word forms that are usually used as nouns is higher than that of words usually used as verbs, regardless of whether the form in question is being pronounced as a noun or a verb (Phillips 2014: 89). Highly frequent nouns produced a large number of isotonic paroxytone pairs due to speakers' ease of producing those words. Sereno & Jongman (1995) conclude that the significant dominance effects suggest that speakers maximize the difference between noun- and verb-dominant words in conformity with the lexical distribution of English in which the majority of bisyllabic nouns are stressed on the first syllable and the majority of bisyllabic verbs on the second syllable (p. 69). The above noun-verb homophonous pairs except the pair for *struggle* emerged in the 16th century, and we may assume that 16th century English shows a similar situation in terms of homophony to Present-Day English. The homophonous pair *struggle* appeared in the 17th century.

We investigate the neural substrates of bisyllabic noun-verb homophones in English using near-infrared spectroscopy (NIRS), inquiring into their interaction with speech production based on the acoustic data by Sereno & Jongman (1995). We used 16 bisyllabic noun-verb ambiguous pairs from Sereno & Jongman (1995). The experimental sentences were constructed such that the noun-context and the verb-context conditions were minimally different prior to the critical ambiguous word, by including sentence initial phrases like *I had/wanted/needed/planned/continued a (no, the)/to* (Fedorenko et al. 2012). Note that the grammatical categories of the words are never ambiguous in the sentence contexts. They are similar in meaning.

The stimuli were presented randomly to six adult (both male and female) native English speakers. The subjects were instructed to listen to the auditorily presented sentences silently. All the sentences were within 5s in length. The interval between

the sentences was 19s to allow the hemodynamic response to return to the baseline before initiating the following trial.

The changes in hemoglobin (Hb) concentrations and their oxygeneration levels in the frontal and temporal lobes in the left hemisphere and the right homologous areas were recorded using NIRS systems (ETG-4000; Hitachi, Tokyo, Japan). NIRS measures the Hb concentration changes of the optical paths in the brain between the nearest pairs of incident and detection probes (Watanabe et al. 1996; Yamashita et al. 1996). The system emits two wave-lengths, approximately 780 and 830 nm, of continuous near-infrared lasers, modulated at different frequencies depending on the channels and the wavelengths, and detected with the sharp frequency filters of lock-in amplifiers (Watanabe et al. 1996). The probe geometry is represented in Figure 4. The probes are placed in their respective areas with an emitter-detector separation length of 25 mm (Watanabe et al. 1996; Yamashita et al. 1996). This separation enables us to measure hemodynamic changes in the brain 2.5–3 cm deep from the head surface, which corresponds to the gray matter on the outer surface of the brain (Fukui et al. 2003).

The probe pads were positioned onto the subjects' bilateral frontal and temporal areas as shown in Figure 4. Green squares (1–22 without circles round) indicate the channel positions. Red probes (probes 11–18, 21–28 in grey circles) are emitters, and blue ones (probes 11–18, 21–28 in black circles) are detectors. The numbers of channels 1–22 correspond to the numbers 1–22 of the time course of hemoglobin responses in the left and right hemispheres in Figures 5–8.
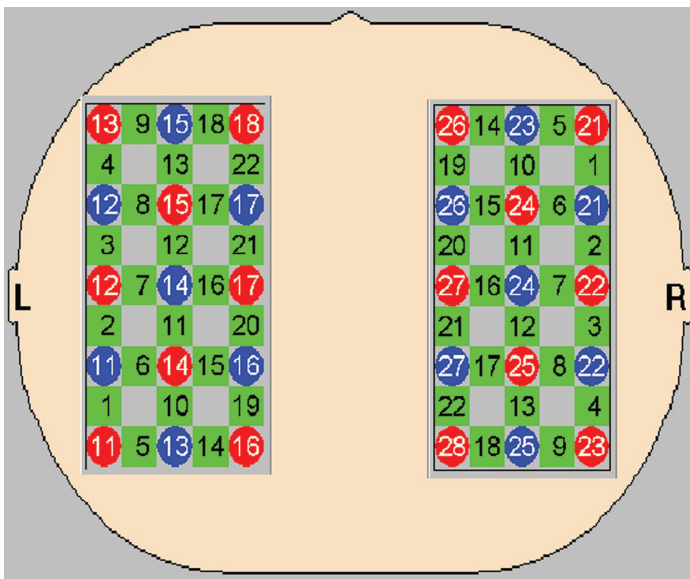
Figure 4. Probe arrangements

Figure 5 shows the average values of the oxy-Hb changes of the 6 English subjects in the frontal and temporal lobes in the left hemisphere and the homologous areas in the right hemisphere for the forestressed nouns in red (solid) lines and the fore-stressed verbs in blue (broken) lines in noun-dominant words in the range of ±0.10 mM*mm. We find anatomical-behavioral correlations, with a left frontal cortical network activated for verbs, which is marked with a solid circle, and greater activation in the temporal regions for nouns, which is marked with a broken circle. In the neuroimaging literature, there is a substantial corpus of studies asking whether nouns and verbs are represented in different neural substrates, or whether both categories are processed in the same anatomical area but with functional differences between them (Crepaldi et al. 2011; Tsigka, et al. 2014; Vonk et al. 2015).
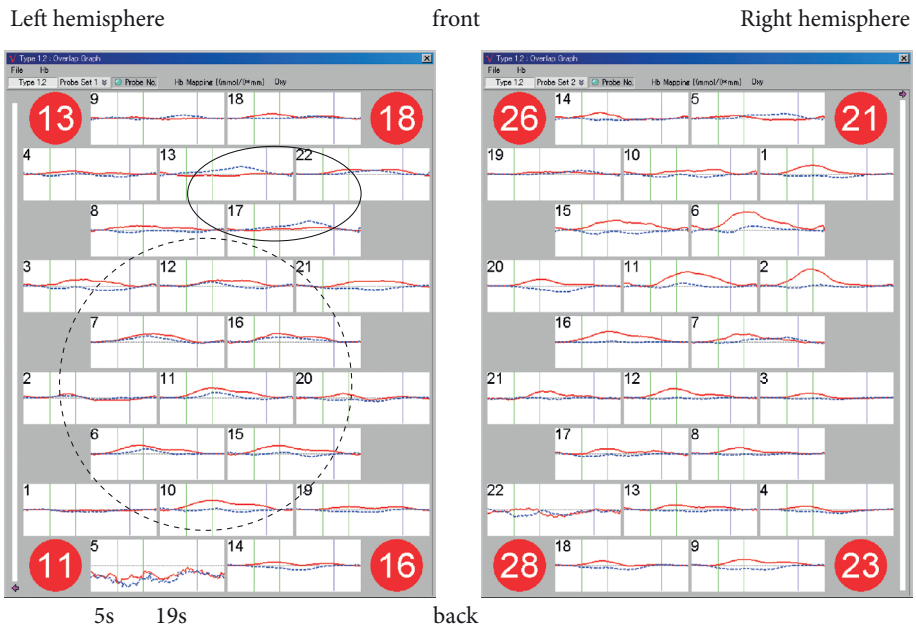
Figure 5. Noun-dominant forestressed nouns and verbs

Our results show that noun and verb categories are represented in different neural substrates in the left hemisphere. We argue, following Shapiro & Caramazza (2003), that information about a word's grammatical category is represented independently of its meaning. Semantic information about nouns and verbs is stored in the frontal and temporal lobes, in which categorical distinctions are explicitly represented when nouns and verbs are activated: there is greater activation of nouns than verbs in the temporal regions and greater activation of verbs than nouns in the frontal regions. Statistically significant differences are observed between nouns and verbs by a t-test at the $p \leq 0.05$ level.

In the right hemisphere, the activation of highly frequent nouns in noun-dominant pairs occurs in the frontal and temporal regions. We note that highly frequent nouns activate in the frontal lobe and temporal lobe in the right hemisphere, but less frequent verbs that show greater activation than the highly frequent nouns in the frontal lobe in the left hemisphere do not activate in the frontal lobe in the right hemisphere. We may assume that when highly frequent words activate in the left hemisphere, the activation occurs in the homologous area in the right hemisphere.

Figure 6 shows the verb dominant backstressed nouns in red (solid) lines and verbs in blue (broken) lines. Verb-dominant forestressed nouns and verbs show similar patterns of activation. We find the activation of both forestressed and backstressed verbs in the frontal lobe of the left hemisphere, which is marked with a solid circle. Furthermore, we find the decrease to the negative value in the oxy-Hb for both forestressed and backstressed verbs in the frontal lobe, which is marked with a broken circle. Task-induced deactivation occurs because certain types of neural processes active during passive states are interrupted when subjects are engaged in effortful tasks (Binder et al. 2009). We note that the deactivation occurs in the area where the activation of less frequent verbs in the forestressed noun-dominant pairs occurs (see Figure 5). In the right hemisphere the activation of highly frequent verbs in verb-dominant pairs occurs both in positive and negative values in the frontal regions. We find statistically greater activation of verbs than nouns by a t-test at the $p \leq 0.05$ level in stressed syllables in the frontal lobes in verb-dominant pairs.
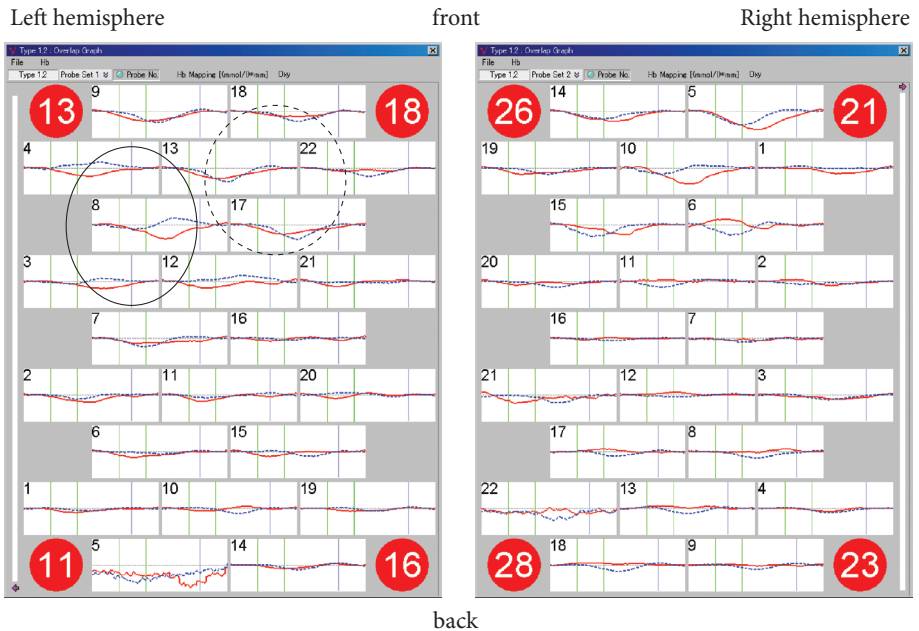

Figure 6. Verb-dominant backstressed nouns and verbs

We argue that listeners try to maximize the difference between noun- and verb-dominance for the ease of perception, though the amplitudes in stressed syllables in verb-dominant pairs are lower than the amplitude of the first syllable of the forestressed noun-dominant pairs in production. We suggest that the cortical representation of speech does not merely reflect the external acoustic environment, but instead gives rise to the perceptual aspects relevant for the listener's intended goal. In 16th century English, we assume that the noun-verb pairs which exhibit stress alternation from OE and ME (see § 3.1) supplied a model for maximizing the difference between noun- and verb-dominant pairs. Here we can see the trade-off between the speaker's ease of production and the listener's ease of perception of homophones.

Ogura & Wang (2018) have further suggested that nouns and verbs are originally processed in the same anatomical area of the brain when the basic word order of a language is SOV, and discourse organization is closely interwoven with syntactic organization as shown in OE and Japanese. When the SOV order is changed to SVO through embedding, a strictly syntactic organization of the clause appears, and the noun/verb distinction is represented in different neural substrates in the brain, as shown after ME (for details, see Ogura & Wang 2012b, 2014, 2018; Ogura forthcoming).

4.2   Neural substrates of diatones

In the experiment, the sentence pairs are organized in terms of noun-plus-verb frequency: four high-frequency pairs (*a subject, to subject; a project, to project; a present, to present; progress, to progress*), and four low-frequency pairs (*a compress, to compress; an imprint, to imprint; a digest, to digest; an insult, to insult*). Phillips (2014) demonstrates noun-plus-verb frequency is more influential in the shift to diatones than noun or verb frequency alone.

Figure 7 shows the average values of the oxy-Hb changes for the six English subjects in the frontal and temporal lobes at the 22 channels in the left hemisphere and the homologous areas in the right hemisphere for high-frequency diatones: nouns are shown with red (solid) lines and verbs with blue (broken) lines. Figure 7 shows the activation of verbs in the frontal regions in the left hemisphere, both with positive and negative values, which is circled in a solid and broken line respectively. In the right hemisphere we find the activation of verbs in the frontal regions.

Figure 8 shows the average values of the oxy-Hb changes for the six English subjects in the frontal and temporal lobes at the 22 channels in the left hemisphere and the homologous areas in the right hemisphere for low-frequency diatones: again, nouns are shown with red (solid) lines and verbs with blue (broken) lines. Figure 8 shows the activation of verbs in the frontal and temporal areas in the
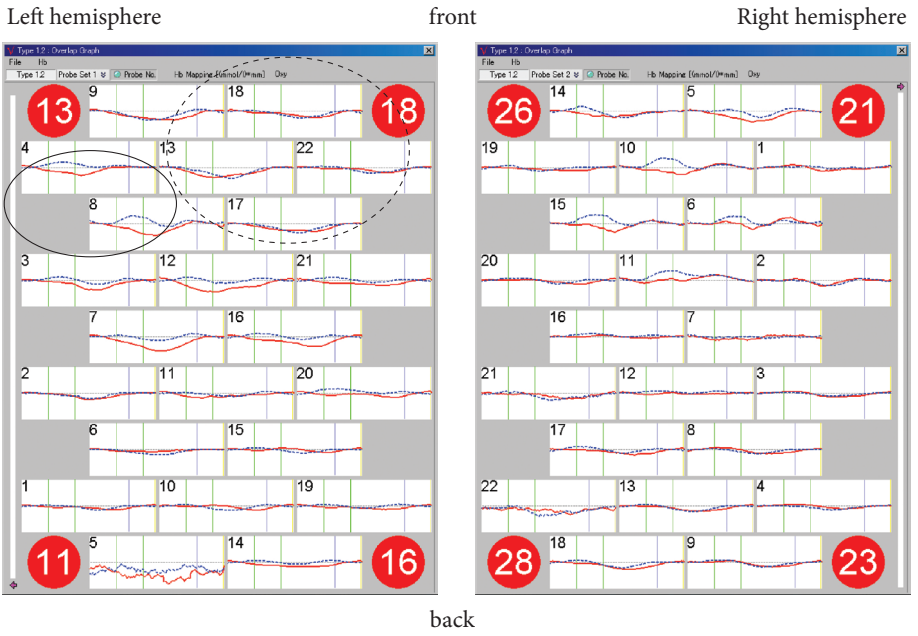
Left hemisphere   front   Right hemisphere

Figure 7. High-frequency diatones

Left hemisphere   front   Right hemisphere
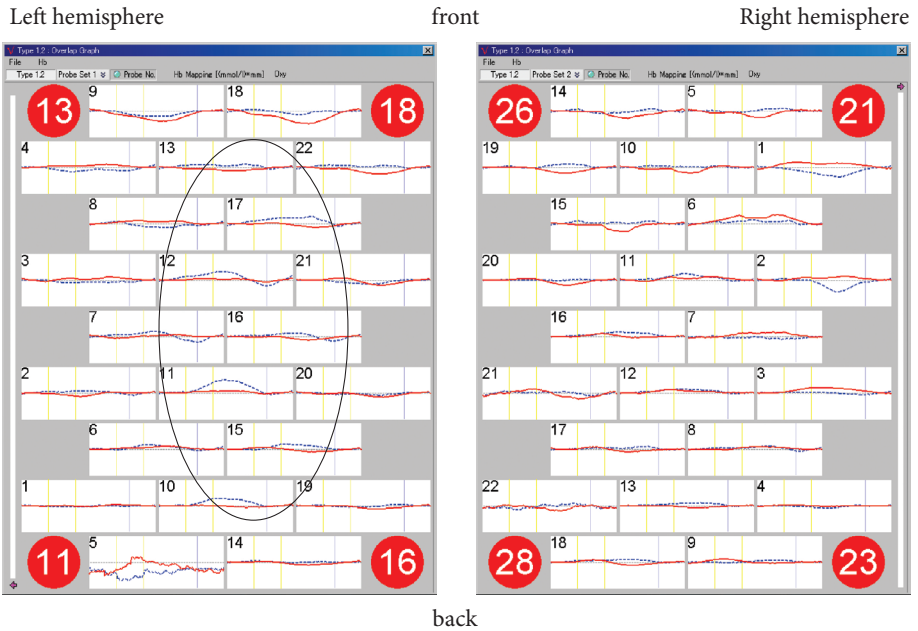

Figure 8. Low-frequency diatones

left hemisphere, where semantic information about nouns and verbs are stored. This is marked with a solid circle. In the right hemisphere there is little activation of verbs. It is interesting that the activation of verbs occurs categorically in the high-frequency diatones, while it occurs semantically in the low-frequency diatones in the left hemisphere.

Categorical activation of the verbs of diatones in the frontal lobe reflects speakers' intention to keep a distinction between forms and meanings in the production of high-frequency homophones, while the semantic activation of the verbs of diatones in the frontal and temporal lobes reflects perceptually or cognitively motivated change by listeners for low-frequency homophones. The neural substrates of the diatones explain why the actuation of diatone-formation was connected with production in highly frequent homophones in the 16th century, but was connected with perception in infrequent words after the 17th century.

5.   Concluding remarks

Based on a quantitative study of the evolution of homophones in English, we have shown why homophones occur. Zipf's law, which states that word frequency decreases as a power law of its rank, can be seen as the outcome of form-meaning associations adopted in order to comply with listeners' and speakers' needs. This implies one form to many meanings, i.e., polysemy and homophony. Homophony is a desirable feature in a communication system, and it is stable and increases through time. But when a large number of homophones emerge, a force to avoid homophones comes into play. We have suggested that the evolution of diatones is a case of avoidance of homophony.

Furthermore, we have examined the neural substrates of comprehension of bisyllabic noun-verb homophones in English. The evolution of homophones is a result of the interaction between a speaker's production and a listener's perception, and the cortical representation of speech does not merely reflect the external acoustic environment. We have suggested that noun and verb categories are represented in different neural substrates in Present-Day English. We have further shown that the activation of verbs occurs categorically in high-frequency diatones, while it occurs semantically in low-frequency diatones in the left hemisphere. The neural substrates of the diatones explain why the actuation of diatone-formation was connected with production in frequent homophones in the 16th century, but was connected with perception in infrequent words after the 17th century.

Acknowledgements

Funding

Sources

CELEX Lexical Database of English (version 2.5). 1995. Centre for Lexical Information, Max Planck Institute for Psycholinguistics, Nijmegen. (Baayen et al. 1995).

LDC Japanese Lexicon. 1997. *Linguistic Data Consortium*, University of Pennsylvania, Philadelphia.

The Brooklyn-Geneva-Amsterdam-Helsinki Parsed Corpus of Old English. 2000. (Pintzuk et al. 2000).

The Penn-Helsinki Parsed Corpus of Middle English. 2000. (Kroch & Taylor 2000).

References

Antilla, Raimo. 1989. *Historical and comparative linguistics* (2nd revised edn.). Amsterdam: John Benjamins. https://doi.org/10.1075/cilt.6

Baayen, R. Harald, Richard Piepenbrock & Léon Gulikers. 1995. *The CELEX lexical database [webcelex]*. Philadelphia, PA: University of Pennsylvania, Linguistic Data Consortium.

Binder, Jeffrey R., Rutvik H. Desai, William W. Graves & Lisa L. Conant. 2009. Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cerebral Cortex* 19. 2767–2796. https://doi.org/10.1093/cercor/bhp055

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

Carroll, John B., Peter Davies & Barry Richman. 1971. *American heritage word frequency book*. Boston: Houghton Mifflin.

Crepaldi, Davide, Manuela Berlingeri, Eraldo Paulesu & Claudio Luzzatti. 2011. A place for nouns and a place for verbs? A critical review of neurocognitive data on grammatical-class effects. *Brain & Language* 116. 33–49. https://doi.org/10.1016/j.bandl.2010.09.005

Fedorenko, Evelina G., Steven Thomas Piantadosi & Edward A. Gibson. 2012. The interaction of syntactic and lexical information sources in language processing: The case of the noun-verb ambiguity. *Journal of Cognitive Science* 13. 249–285. https://doi.org/10.17791/jcs.2012.13.3.249

Ferre i Cancho, Ramon & Richard V. Solé. 2003. Least effort and the origins of scaling in human language. *Proceedings of the National Academy of Sciences* 100. 788–791. https://doi.org/10.1073/pnas.0335980100

Fudge, Erik. 1984. *English word-stress*. London: Allen Unwin.

Fukui, Yuichi, Yusaku Ajichi & Eiji Okada. 2003. Monte Carlo prediction of nearinfrared light propagation in realistic adult and neonatal head models. *Applied Optics* 42. 2881–2887. https://doi.org/10.1364/AO.42.002881

Haiman, John (ed.). 1985. *Iconicity in syntax*. Amsterdam: John Benjamins. https://doi.org/10.1075/tsl.6

Hotta, Ryuichi. 2013. The diatonic stress shift in Modern English. *Studies in Modern English* 29. 1–20.

Jespersen, Otto. 1954. *A Modern English grammar on historical principles*, Part 1. London: Routledge.

Kroch, Anthony & Ann Taylor. 2000. Penn-Helsinki Parsed Corpus of Middle English (2nd edn.). http://www.ling.upenn.edu/hist-corpora/PPCME2-RELEASE-4/index.html

Lass, Roger. 1980. *On explaining language change*. Cambridge: Cambridge University Press.

Lass, Roger. 1992. Phonology and morphology. In Norman Blake (ed.), *The Cambridge history of the English language, Volume 2: 1066–1476*, 23–155. Cambridge: Cambridge University Press. https://doi.org/10.1017/CHOL9780521264754.003

Levins, Peter. 1570. *Manipulus Vocabulorum* (Early English Text Society, UK edition, 1973).

Miller, George A. 1981. *Language and speech*. San Francisco: W. H. Freeman.

Moore, Samuel. 1967. *Historical outlines of English sounds and inflections*. Revised by Albert H. Marckwardt. Ann Arbor: George Wahr.

Mulcaster, Richard. 1582. *The First Part of the Elementarie which entreateth chefelie of the right writing of our English tung* (Scolar Press; Facsimile of 1582 edition, 1970).

Nakao, Toshio. 1978. *The prosodic phonology of Late Middle English*. Tokyo: Shinozaki Shorin.

Ogura, Mieko. Forthcoming. *Language evolution as a complex adaptive system: A multidisciplinary approach to the history of English*. New York: Oxford University Press.

Ogura, Mieko & William S-Y. Wang. 2006. Ambiguity and language evolution: Evidence of homophones and syllable number of words. *Studia Anglica Posnaniensia* 42. 3–30.

Ogura, Mieko & William S-Y. Wang. 2012a. The global organization of the English lexicon and its evolution. In Hans Sauer & Gaby Waxenberger (eds.), *English historical linguistics 2008, Volume 2: Words, texts and genres*, 65–83. Amsterdam: John Benjamins. https://doi.org/10.1075/cilt.324.04ogu

Ogura, Mieko & William S-Y. Wang. 2012b. Ambiguity resolution and evolution of word order. In Thomas C. Scott-Phillips, Monica Tamariz, Erica A. Cartmill & James R. Hurford (eds.), *The evolution of language: Proceedings of 9th International Conference*, 274–281. Singapore: World Scientific. https://doi.org/10.1142/9789814401500_0036

Ogura, Mieko & William S-Y. Wang. 2014. Evolution of tense and aspect. In Erica A. Cartmill, Sean Roberts, Heidi Lyn & Hannah Cornish (eds.), *The evolution of language: Proceedings of the 10th international conference*, 213–220. Singapore: World Scientific. https://doi.org/10.1142/9789814603638_0027

Ogura, Mieko & William S-Y. Wang. 2018. Evolution of homophones and syntactic categories noun and verb. In Christine Cuskley, Molly Flaherty, Hannah Little, Luke McCrohon, Andrea Ravignani & Tessa Verhoef (eds.), *The evolution of language: Proceedings of the 12th international conference*, 355–363. Torun: Nicolaus Copernicus University.

Phillips, Betty S. 1984. Word frequency and the actuation of sound change. *Language* 60. 320–342. https://doi.org/10.2307/413643

Phillips, Betty S. 2014. Gradience in an abrupt change: Stress shift in English disyllabic noun-verb pairs. In Simone E. Pfenninger et al. (eds.), *Contact, variation, and change in the history of English*, 83–94. Amsterdam: John Benjamins. https://doi.org/10.1075/slcs.159.05phi

Piantadosi, Steven T. Harry Tily & Edward Gibson. 2012. The Communicative function of ambiguity in language. *Cognition* 122. 280–291. https://doi.org/10.1016/j.cognition.2011.10.004

Pintzuk, Susan, Eric Haeberli, Ans van Kemenade, Willem Koopman & Frank Beths. 2000. The Brooklyn-Geneva-Amsterdam-Helsinki Parsed Corpus of Old English. https://www-users.york.ac.uk/~sp20/corpus.html

Sereno, Joan A. & Allard Jongman. 1995. Acoustic correlates of grammatical class. *Language and Speech* 38(1). 57–76. https://doi.org/10.1177/002383099503800103

Shapiro, Kevin & Alfonso Caramazza. 2003. The representation of grammatical categories in the brain. *Trends in Cognitive Sciences* 7(5). 201–206. https://doi.org/10.1016/S1364-6613(03)00060-3

Sherman, Donald. 1975. Noun-verb stress alternation: An example of the lexical diffusion of sound change in English. *Linguistics* 159. 43–71. https://doi.org/10.1515/ling.1975.13.159.43

Strang, Barbara M. H. 1980. The ecology of the English monosyllable. In Sydney Greenbaum, Geoffrey Leech & Jan Svartvik (eds.), *Studies in English linguistics: For Randolph Quirk*, 277–293. London: Longman.

Tsigka, Styliani, Christos Papadelis, Christoph Braun & Gabriele Miceli. 2014. Distinguishable neural correlates of verbs and nouns: A MEG study on homonyms. *Neuropsychologia* 54. 87–97. https://doi.org/10.1016/j.neuropsychologia.2013.12.018

Vonk, Jet M. J., Roel Jonkers & Loraine K. Obler. 2015. Semantic subcategories of nouns and verbs: A neurolinguistic review on healthy adults and patients with Alzheimer's disease. In Corine Astésano & Mélanie Jucla (eds.), *Neuropsycholinguistic perspectives on language cognition: Essays in honour of Jean-Luc Nespoulous*, 61–74. London: Psychology Press.

Wang, William S-Y. 2011. Ambiguity in language. *Korean Journal of Chinese Language and Literature* 1. 3–20.

Watanabe, Eiju, Yuichi Yamashita, Atsushi Maki, Yoshitoshi Ito & Hideaki Koizumi. 1996. Non-invasive functional mapping with multi-channel near infra-red spectroscopic topography in humans. *Neuroscience Letters* 205. 41–44. https://doi.org/10.1016/0304-3940(96)12376-4

Yamashita, Yuichi, Atsushi Maki & Hideaki Koizumi. 1996. Near-infrared topographic measurement system: Imaging of absorbers localized in scattering medium. *Review of Scientific Instruments* 67. 730–732. https://doi.org/10.1063/1.1146849

Zempleni, Monika-Zita, Remco Renken, John C. J. Hoeks, Johannes M. Hoogduin & Laurie A. Stowe. 2007. Semantic ambiguity processing in sentence context: Evidence from event-related fMRI. *NeuroImage* 34. 1270–1279. https://doi.org/10.1016/j.neuroimage.2006.09.048

Zipf, Kingsley. 1949. *Human behavior and the principle of least effort: An introduction to human ecology*. Cambridge, MA: Addison-Wesley.

Appendix. The evolution of diatones in the 16th, 17th, and 18th century

Word	Diatone	N/V	OED source	OED first citation	Homo-phone	O/P/D	Word frequency
outlaw	16c	v	LOE	10..	16	D, P	2.7833
outlaw	16c	n	LOE	924	16	D, P	2.7833
rebel	16c	n	F	1340	16	D	3.0981
rebel	16c	v	F	1340	16	D	3.0981
record	16c	n	OF	a1300	16	D, O	115.97

Word	Diatone	N/V	OED source	OED first citation	Homo-phone	O/P/D	Word frequency
record	16c	v	OF	a1225	16	D, O	115.97
convert	16c	n	convert (v)	1561	16	D	5.8238
convert	16c	v	OF	a1300	16	D	5.8238
desert	16c	n	OF	a1225	16	D	78.870
desert	16c	v	mod. F.	1539	16	D	78.870
incense	16c	n	OF	c1290	16	O, D	0.8155
incense	16c	v	F	1303	16	O, D	0.8155
present	16c	n	OF	a1225	16	D	137.90
present	16c	v	OF	c1290	16	D	137.90
refuse	16c	n	OF	c1440	16	O, D	8.9941
refuse	16c	v	F	13..	16	O, D	8.9941
accent	17b	n	Fr	1538	16	P, D	19.386
accent	17b	v	Fr	1530	16	P, D	19.386
collect	17b	n	F, L	a1225	16	D	31.686
collect	17b	v	OF, L	1573	16	D	31.686
compound	17b	n	compound (adj)	1530	16	D, O	58.832
compound	17b	v	OF	c1374	16	D, O	58.832
conduct	17b	n	OF or L	c1290	16	D, P	11.899
conduct	17b	v	F or L	c1400	16	D, P	11.899
contract	17b	n	OF	c1315	16	O, D	14.346
contract	17b	v	L	1530	16	O, D	14.346
convoy	17b	n	F	1500–20	16	D, P	0.4747
convoy	17b	v	F	1375	16	D, P	0.4747
object	17b	n	partly ppl.a, partly L	c1380	16	D, P	117.35
object	17b	v	L	c1400	16	D, P	117.35
relapse	17b	n	relapse (v)	1533–4	16	O, D	0.0322
relapse	17b	v	L	1568	16	O, D	0.0322
torment	17b	n	OF, ONF	c1290	16	D	0.8642
torment	17b	v	OF	c1290	16	D	0.8642
abstract	17c	n	L	1528	16	D	2.7941
abstract	17c	v	abstract (ppl. a)	1542	16	D	2.7941
cement	17c	n	OF	c1300	16	D, O, P	13.289
cement	17c	v	cement (n)	1340	16	D, O, P	13.289
compact 1	17c	n	ppl a(1) < L	1601	17	D	7.2348
compact 1	17c	v	compact (a) < L	1530	17	D	7.2348
confine	17c	n	F	c1400	16	D	0.6545
confine	17c	v	F	1523	16	D	0.6545
conflict	17c	n	L	c1430	16	D, P	16.476
conflict	17c	v	L	1432–50	16	D, P	16.476
conserve	17c	n	F	1393	16	O, D, P	2.5129
conserve	17c	v	F	c1374	16	O, D, P	2.5129

Word	Diatone	N/V	OED source	OED first citation	Homo-phone	O/P/D	Word frequency
consort	17c	n	F	1419	16	D, P	0
consort	17c	v	consort (n1)	1588	16	D, P	0
contest	17c	n	contest (v)	1643	17	D, O	22.184
contest	17c	v	F	1579	17	D, O	22.184
converse	17c	n	converse (v)	1610	17	D, P, O	2.5234
converse	17c	v	F	1340	17	D, P, O	2.5234
convict	17c	n	convict (ppl. a.) < L	1530–1	16	O*, D	0.1133
convict	17c	v	L	c1366	16	D	0.1133
essay	17c	n	OF	1597	16	D, O, P	4.5664
essay	17c	v	ASSAY (< L)	1483	16	D, O, P	4.5664
exile	17c	n	OF	1300	16	P, P, D	2.0964
exile	17c	v	OF	1330	16	P, D	2.0964
extract	17c	n	L	1549	16	O, D	3.4555
extract	17c	v	L	c1489	16	O, D	3.4555
ferment	17c	n	Fr	c1420	16	D	0.4827
ferment	17c	v	F	1398	16	D	0.4827
insult	17c	n	F	1603	17	D, O	2.0882
insult	17c	v	L	1570–6	17	D, O	2.0882
outcast	17c	n	out- < OE + cast(v) < ON	13..	16	D, P*, O*	0.4082
outcast	17c	v	out- < OE + cast(v) < ON	a1300	16	D	0.4082
project	17c	n	L	a1400–50	16	D	21.216
project	17c	v	L	c1477	16	D	21.216
subject	17c	n	OF	13..	16	D, P	112.47
subject	17c	v	OF or L	1382	16	D, P	112.47
transport	17c	n	transport (v)	1456	16	D, O	6.2411
transport	17c	v	F or L	c1374	16	D, O	6.2411
bombard	18b	n	OF	1393	16	P, O, D	0
bombard	18b	v	F	1598	16	P, O, D	0
compress	18b	n	F	1599	16	O, D	0.6516
compress	18b	v	OF	1398	16	O, D	0.6516
concert	18b	n	F	1665	17	O, D, P	9.5770
concert	18b	v	F	1598	17	O, D, P	9.5770
concrete	18b	n	L	1528–1725	17	O*, D	14.067
concrete	18b	v	concrete (a) < L	1635	17	D	14.067
confect	18b	n	med. L	1587	16	P, O, D	0
confect	18b	v	L	1545	16	P, O, D	0
contrast	18b	n	F	1597	16	O, P, D	22.787
contrast	18b	v	OF	1489	16	O, P, D	22.787

Word	Diatone	N/V	OED source	OED first citation	Homo-phone	O/P/D	Word frequency
discord	18b	n	OF	1297	16	P*, P, D	0.4495
discord	18b	v	OF	a1300	16	P, D	0.4495
discount	18b	n	F	1622	17	O, P, D	1.9460
discount	18b	v	OF	1629	17	O, P, D	1.9460
export	18b	n	export (v)	1690	17	O, D	4.4660
export	18b	v	L	c1485	17	O, O*, D	4.4660
import	18b	n	import (v)	1588	16	O, D	4.2570
import	18b	v	L	c1430	16	O, D	4.2570
impress 1	18b	n	impress (v1)	1590	16	O, D	2.4433
impress 1	18b	v	L	c1374	16	O, D	2.4433
prelude	18b	n	F	1561	17	O, D	0.1668
prelude	18b	v	L	1655	17	O, D	0.1668
produce	18b	n	produce (v)	1699	17	D, O	100.37
produce	18b	v	L	1499	17	D, O	100.37
survey	18b	n	survey (v)	1535	16	O, D	9.6180
survey	18b	v	AF	1467–8	16	O, D	9.6180
undress	18b	n	un- < OE + dress (n) < dress(v)	1683	17	O, D	0.8570
undress	18b	v	un- < OE + dress(v) < OF	1596	17	O, D	0.8570
affix	18c	n	Fr	1612	17	O, D	0.2770
affix	18c	v	med. L	1533	17	O, D	0.2770
decrease	18c	n	OF	1383	16	O, D	3.3580
decrease	18c	v	OF	1382	16	O, D	3.3580
defile 1	18c	n	F	1685	18	O, D	0.4014
defile 1	18c	v	F	1705	18	O, D	0.4014
descant	18c	n	OF	c1380	16	O, P, D	0.2587
descant	18c	v	OF	c1510	16	O, P, D	0.2587
digest	18c	n	L	1387	16	O, P, D	2.0360
digest	18c	v	L	c1450	16	O, P, D	2.0360
increase	18c	n	increase (v)	c1374	16	O, D	43.379
increase	18c	v	AF	13..	16	O, D	43.379
inlay	18c	n	inlay (v)	1656	17	O, D	0
inlay	18c	v	in- < OE + lay (v) < OE	1596	17	O, D	0
outleap	18c	n	out- < OE + leap (n) < OE	c1250	16	O, D	0
outleap	18c	v	out- < OE + leap (v) < Com.Teut.	1600	16	O, D, O*	0
outwork	18c	n	out- < OE + work (n) < OE	c1615	17	P*, D, O*	0

Word	Diatone	N/V	OED source	OED first citation	Homo-phone	O/P/D	Word frequency
outwork	18c	v	out- < OE + work (v) < OE	c1250	17	D	0
perfume	18c	n	F	1533	16	O, D	8.9841
perfume	18c	v	F	1538	16	O, D	8.9841
permit	18c	n	permit (v)	1714	18	O, P, D	12.290
permit	18c	v	L	1489	18	O, P, D	12.290
prefix	18c	n	mod. L	[1614] 1646	17	O, D	10.702
prefix	18c	v	OF	c1420	17	O, D	10.702
presage	18c	n	F	1390	16	P, O, D	0
presage	18c	v	F	1562	16	P, O, D	0
protest	18c	n	F	c1400	16	O, D	6.4360
protest	18c	v	F	1440	16	O, D	6.4360
purport	18c	n	AF	[1278] 1455	16	P, D	0.1127
purport	18c	v	AF	[1300] 1528	16	P, D	0.1127
regress	18c	n	L	c1375	16	P*, O*, O, D	0
regress	18c	v	L	1552	16	P*, O, D	0
reprint	18c	n	reprint (v)	1611	17	O*, D	0
reprint	18c	v	re- < L + print (v) < print (n) < OF	1551	17	D	0
surcharge	18c	n	surcharge (v)	1569	16	O, D	0
surcharge	18c	v	OF	1429	16	O, D	0
transfer	18c	n	transfer (v)	1674	17	P, D	7.9900
transfer	18c	v	F or L	1382	17	P*, P, D	7.9900
transverse	18c	n	L	1596	16	D	0.2062
transverse	18c	v	OF	1377	16	D	0.2062
uprise	18c	n	up- < OE + rise (n) < rise(v) < Com. Teut.	a1300	16	O, D	0
uprise	18c	v	up- < OE + rise(v) < Com.Teut.	a1300	16	O, D	0

CHAPTER 4

The threshold of productivity
and the 'irregularization' of verbs
in Early Modern English

Don Ringe and Charles Yang
University of Pennsylvania

This paper tests whether Yang's Tolerance Principle (TP) is useful in explaining
one type of morphological change in the history of English. We suggest that
some, but not all, innovative non-default past tenses that appear in the written
record between 1500 and 1700 can be explained by the TP and thus could be the
results of innovation in native language acquisition (NLA).

Keywords: Tolerance Principle, native language acquisition, Early Modern English,
irregularization

1.   The Tolerance Principle, its range and limits

Yang (2016: 8–9, 60–66) proposes and argues for a threshold for the productivity
of linguistic rules of all kinds in native language acquisition (NLA), as follows. Let
N be the number of lexemes which meet the structural description of the rule and
e the number of exceptions (i.e., the number of lexemes which meet the structural
description but are not subject to the rule). There is then a threshold

$$\theta_N = \frac{N}{\ln N}$$

such that e must be less than or equal to θ_N for the rule to be productive; that is,
the threshold is the number of lexemes that could obey the rule divided by the
natural logarithm of the same. If that threshold is not exceeded, so that the rule
is productive, the native learner will extend the range of the rule's application, re-
moving exceptions one by one. If the threshold of productivity is exceeded by the
number of exceptions, no change is expected; if change is observed, it must proceed
by some other mechanism.

   This 'Tolerance Principle' (TP), specifying the maximum number of excep-
tions which a rule can tolerate in order to be productive, is compatible with a wide

range of phenomena observed in NLA, including the formation of the past tense in English (Yang 2016: 26–34, 81–91), various details of verb inflection in German and Spanish (Yang: 34–37, 80–81), and the scope of syntactic constructions. In this paper we focus on the productivity of lexical classes, i.e., arbitrary inflectional classes of major words which compete for membership, in NLA and otherwise.

Though the TP makes a prediction about any rule a native learner can come up with, even a 'crazy' rule, it is not a hypothesis about the origins of all linguistic innovations. It cannot assess an innovation which is not rule-governed, and it has nothing to say about the innovative behavior of adult native speakers. Interestingly, there is some evidence that innovative non-default past tenses in English might be constructed by adults. An especially clear example relevant to the present study emerged from the celebrated 'wug' study (Berko 1958). Berko presented both children and adults with nonce verbs that strongly resembled existing English 'irregulars', e.g., *gling*, potentially of the type *ring, rang* or *sling, slung*, and *spow*, potentially of the type *grow, grew*. The past tenses for these verbs produced by children were almost exclusively default pasts in *-ed*, but a significant number of adults suggested *glang* or *glung* (See Yang (2016: 31–40) for a crosslinguistic review of how children's acquisition of morphology reflects a categorical distinction between productive and unproductive processes; this is just one example of many). This dichotomy between changes in the scope of rules in NLA, governed by the TP, and other processes of change will turn out to be important in our investigation. We return briefly to adult native speaker innovations at the end of § 9.

We also emphasize that this line of work is fundamentally different from earlier approaches, in that it hypothesizes a specific PROCESS in the course of NLA that gives rise to some, but not all, linguistic changes. It contrasts sharply with the approach of Strik (2015), for example, who treats 'analogy' as a kind of black box and attempts to determine what is going on inside using purely statistical methods.[1] That and the (extensive) other work investigating frequency effects in linguistic change is not directly relevant to our inquiry because it does not deal with the TP, nor with NLA at all. In short, we here pursue a new line of inquiry based on radically shifted premises. For a more extensive demonstration of the explanatory power of the TP in dealing with the linguistic past see Kodner (2020).

---

1.   Some of Strik's closing observations, which apparently puzzled him, point to NLA as a source of change. Overall resistance to morphological change (Strik 2015: 183) is a consequence of the fact that children are remarkably good at learning complex inflectional systems; the fact that both very frequent lexemes and rare lexemes tend to resist change (Strik 2015: 184) results from the fact that the former are constantly reinforced, while many of the latter are not acquired early enough in NLA (see § 5 below).

2.   English verbs in *-ing*

An especially clear illustration of the TP is furnished by the past tense of (standard) English verbs ending in /-ɪŋ/. They fall into four groups:

i.   *bring*, past *brought* /bɹɔt/, unique in this subset of verbs;
ii.   *ring, sing, spring*, with pasts in /-æŋ/ (*rang, sang, sprang*);
iii.   *cling, fling, sling, sting, string, swing, wring*, with pasts in /-ʌŋ/ (*clung*, etc.); and
iv.   a handful of verbs, all uncommon or rare, that belong to the huge default class. The ones that occur in the senior author's speech are *ding, king* (in checkers), *ping, wing*, pasts *dinged* /dɪŋd/, etc.; at least some speakers also use *ring* "put a ring around", past *ringed*, a reviewer notes colloquial British *ming* "be disgusting", and recently *bling* "be flamboyant" and *bing* "use Bing" have begun to appear.

The largest group is those with pasts in /-ʌŋ/, and a 'common sense' approach to morphological change would suggest that that group should gradually acquire new members. In fact that was long the usual view: "A form which is statistically predominant is also likely to be productive for new combinations" (Nida 1949: 45). However, new past tenses in /-ʌŋ/ are conspicuous mostly by their absence in contemporary standard English, and Yang's hypothesis offers a reason why the situation is stable. In the senior author's speech N is 15, and the natural logarithm of 15 is 2.70805. The quotient of those numbers is 5.539, and the TP states that no rule of past tense formation for this group of verbs should become productive if the number of exceptions to it exceeds five. In fact, no matter which rule we choose, the number of exceptions is substantially greater than five: there are eight exceptions to the *ung*-rule, eleven to the default past tense rule, and of course even more to the *ang*-rule. Thus it is not surprising that there is no change in progress. That remains true even for speakers whose only 'regular' verb in this group is *wing* (the least rare of those with default past tenses). For them N is 12, and its natural logarithm is 2.48491; the quotient is 4.829, and there are still five exceptions (*bring, ring, sing, spring, wing*) to the rule yielding past tenses in /-ʌŋ/.

At this point it may be convenient to list the productivity thresholds for lexical classes of sizes 3 through 17:

N	θ_N	N	θ_N	N	θ_N
3	2	8	3	13	5
4	2	9	4	14	5
5	3	10	4	15	5
6	3	11	4	16	5
7	3	12	4	17	6

Note that when *N* is 12 the number of exceptions to the *ung*-rule is very close to the TP threshold. What if a native learner acquires no verbs in /-ɪŋ/ with default past tenses? For such a learner *N* is 11, but there are only four exceptions to the rule yielding past tenses in /-ʌŋ/ (*bring, ring, sing, spring*); thus the rule ought to be productive if there are no verbs in /-ɪŋ/ with default past tenses at all. In short, the 'ung-rule' is a borderline case which might provide an opportunity to test the TP.

Actually one innovative past in /-ʌŋ/ has appeared recently in colloquial English, exemplified in the phrase *they sprung him out of prison* (Margaret Laing p.c. 27 August 2018). The phrase seems to be U.S. slang, not much more than a century old (see the *Oxford English Dictionary* (*OED*) s.v. *spring*, v.1 II.26). We cannot demonstrate that that finite past arose in the speech of a learner who used no verbs in /-ɪŋ/ with default past tenses, but the distribution of facts makes it likely.

Other innovative past tenses that have not caught on can also be explained by the TP. Yang notes that "the productivity of a rule may change during the course of language acquisition" (2016: 70), depending on the vocabulary that the native learner has learned at any given point. To judge from the 6-million-word corpus of child-directed English in the *CHILDES* database (MacWhinney 2000), the commonest verbs in this class, in descending order, are *bring* (past *brought*), *sing* (*sang*), and *ring* (*rang*), and it is reasonable to suppose that children learn those verbs first more often than not. The seven verbs with pasts in /-ʌŋ/ and *spring* (*sprang*) are subsequently learned in an order that must vary from child to child; verbs with default pasts such as *wing* (*winged*) must be learned comparatively late, given their rarity. The TP predicts that the productivity of the lexical classes should change as follows. When the only verbs in *-ing* which the child knows are

> *bring, brought; sing, sang; ring, rang,*

the lone exception to the *ang*-rule, *brought*, does not exceed the threshold for *N* = 3, so we expect that rule to be productive, with at least a few children producing past *brang* early in NLA. In fact *brang* does occasionally appear in the acquisition of standard English. If the fourth verb learned is *spring, sprang*, the situation will not change (see the table of thresholds above); if instead the fourth verb is one with a past tense in *-ung* – say, *swing, swung* – the class in *-ang* will remain productive; it will remain productive if *swing* and *spring* are learned in that order as the fourth and fifth verbs in *-ing*, and even if another verb with a past in *-ung* is acquired sixth, so that the verbs in this class are (for instance)

> *bring, brought; sing, sang; ring, rang; spring, sprang; swing, swung;*
> *sting, stung,*

given that when $N = 6$ a rule can tolerate three exceptions. However, if the fourth, fifth, and sixth verbs learned all have pasts in *-ung* (and *spring* is acquired only later), it will be the *ung*-rule that is productive. In that case it will make a difference whether the seventh verb acquired is *spring* or another verb with a past in *-ung;* in the former case the *ang*-rule will again become productive, but in the latter case the *ung*-rule will continue to be productive, since when $N = 7$ a rule can tolerate only three exceptions. When the eighth verb is acquired only the *ung*-rule can be productive, unless and until at least one verb in *-ing* with a default past is acquired. Thus the TP offers an easy explanation for the fact that innovative pasts tenses like *brung* and *sung* are produced less rarely than *brang* in the acquisition of standard English.[2]

It may be helpful to put this discussion in a broader context. Children learning English virtually always produce default past tenses of 'irregular' verbs (see, e.g., Pinker 1999: 189–210) but seldom produce innovative non-default past tenses (Xu & Pinker 1995); examples of the latter are rare. The point here is that most of those rare exceptions obey the *ung*-rule.[3] For instance, in the entire collection of North American child English transcribed in the *CHILDES* database, the past tense of *bring* is correct *brought* 95 times, default *bringed* 6 times, and *brang* or *brung* 9 times. In line with the thought experiment of the preceding paragraph, the non-default but incorrect pasts would not be possible for a speaker with a vocabulary large enough to include most of the verbs in *-ing*, but they are predicted to be spontaneously available to children who have a much smaller vocabulary. In short, we expect *brang* or *brung* to be produced at an early stage of NLA when a learner has not learned enough vocabulary to destroy the transient productivity of the rule. If it is asked why *bringed* is produced comparatively often, given that verbs in *-ing* with default pasts are all rare, we remind the reader of two further facts. In the first place, different learners can be expected to hypothesize different lexical classes; a class of verbs in *-ing* is a universally available hypothesis for children learning English, but it need not be chosen by every learner. Secondly, it is well known that when native learners of English discover the default formation of the past tense they at first overgeneralize it, rejecting the non-default inflections that they have begun to learn, only to return to them a bit later (Marcus et al. 1992). That errors of any kind are rare is a straightforward result of the well-known fact that children are astonishingly good at NLA.

---

2. The fact that the past tenses of *sing, ring, spring* are *sung, rung, sprung* in some dialects of English comparatively remote from the standard can be explained in more than one way; see fn. 10 below.

3. See Yang (2016: Chapter 1) for a crosslinguistic review of these and similar findings.

Of course demonstrating in detail that this is the process leading to the production of innovative past tenses for these verbs will involve further and more intensive research on NLA. However, there is another dataset with which the TP can be confronted, namely observed historical changes in a language. Since the history of English over the past six or seven centuries is especially well documented and intensively studied, we believe that it is worthwhile to see whether the TP can help to explain the English historical data. That is the question this paper will address.

Before we proceed, it seems advisable to clarify a methodological point. We noted above that different native learners can be expected to hypothesize different lexical classes, though some will be hypothesized more often than others; only experimental work on NLA can shed further light on that point. Because we are dealing with the distant past, we are constrained to adopt a different strategy: we start from changes that did occur in the history of English and attempt to retrofit the analysis so as to explain those developments in accordance with the TP. This approach is complementary to simple computational models of inductive learning, such as those constructed by Yip & Sussman (1997) and Albright & Hayes (2003), which operate on the phonological representations of words to produce rules such as the *-ung* rule. In either case some further hypothesis, such as the TP, is needed to assess the productivity of rules.

If we could explain everything that we observe by our strategy, we could say nothing about the limits of the TP. However, it will be seen that there are observed developments which the TP probably cannot explain; most importantly, they are precisely the developments which do not seem to be rule-governed and might have been initiated by adults.

3.   The 16th and 17th centuries

Over the course of the 16th and 17th centuries written English becomes increasingly uniform, but there is still a good deal of variation. In the context of that variation several noteworthy innovations in past tense formation occurred. *Stringed* was replaced by *strung*, suggesting that the *-ung* rule was productive for some speakers for at least part of that period; more surprisingly, *digged* was replaced by *dug* and *sticked* was replaced by *stuck*, suggesting that the rule was actually broader, taking as input verbs with /ɪ/ in the root followed by any single velar consonant (or homorganic velar cluster – see further below). We need to determine whether such a rule could have been productive according to the TP, or whether some other type of change must have occurred.

The two principal resources for investigating the development of English during this period are the *OED* and the *Penn-Helsinki Parsed Corpus of Early Modern*

English (*PPCEME;* Kroch et al. 2004), which covers the period 1500–1710. The *OED* records every form attested and attempts to give the earliest instance of each attested form, usually with (at least approximate) success. The *PPCEME*, though it includes more than 1.7 million words of running text, is not large enough to offer examples of every lexeme, let alone every form; on the other hand, it gives a fair indication of which lexemes occur most frequently, and for common lexemes it provides good information about the relative frequency of competing forms.

We also consulted the *Parsed Corpus of Early English Correspondence* (*PCEEC;* Nevalainen et al. 2006), but it contained too few relevant forms to contribute significantly to the relevant statistics; for instance, though *sing* is one of the commonest relevant verbs in the *PPCEME*, its finite past is unattested in the *PCEEC*, which also contains only two examples of its past participle *sung*. Forms from the *PCEEC* will be cited in the footnotes where relevant. Other potential corpora will also be mentioned where relevant.

4. New past tense forms in Early Modern English

The currently ascertainable facts regarding the new past tense forms noted in the last section are the following.

The *OED* records a single instance of the verb *string* from late Middle English (ME), *with bowes gode wel y-strenged* "with good bows well strung", from the *Laud Troy Book*, ca. 1425; the verb *to string* (a bow or a musical instrument) next appears in the 16th c., the earliest *OED* quotations being from 1530 and 1545.[4] Since the verb is clearly derived from the noun *string* (which was inherited from Old English (OE), first attested in the 8th- or 9th-c. poem *Andreas*), one would expect it to inflect according to the default pattern, with a past and past participle *stringed*. In fact *stringed* and *strung* are both quoted from the 16th c. onward, though *stringed* is gradually eliminated, falling out of use in the 18th c. In the 16th c. *stringed* is first quoted from 1530, and again from 1548 (past participle in both instances); *strung* (likewise past participle) is first quoted from Spenser (*Virgil's Gnat*, 1591), and from Donne (1612), Chapman (1613), Shakespeare (before 1616), and obscure authors of the 1590's and 1600's. The past and past participle of this verb are not attested in the *PPCEME*. Though the attestation is too sparse for absolute certainty, it appears

---

4.   There are also a couple of quotations each in the more obscure meanings "to furnish (a garment) with ties" (first in 1548) and "to remove the notochord from (a lamprey)" (1508). It is not clear whether those meanings are relevant; in any case they are rare, and the *OED* quotes no past tense for either. The now common meanings "to put (things) serially on a string" and "to arrange in a line" both appear first in the 17th c.

that *stringed* was in place early in the 16th c. and began to be replaced by *strung* in literary English sometime in the second half of the century.

For *dig* the record is a bit fuller and the situation slightly clearer. The verb was borrowed from French around 1300, and a default past tense *diggede* and past participle *digged* are well attested in late ME. Such forms continue to occur throughout the 16th and 17th centuries; the *PPCEME* has examples from Tyndale's bible (1520's and '30's, adopted without change in the King James translation of 1611), from the works of Sir Francis Bacon, William Clowes, and Richard Hooker (all late 16th c.), from the letters of Lady Katherine Paston, John Taylor, and Henry Oxinden (all latest 16th and early 17th c.), and from statute books throughout the period. Early examples of *dug* happen not to occur in the corpus, but the *OED* (under *dig*, § II.4) gives an example of past participle *dug* from 1580, and another occurs in the *PPCEME* sample from Pepys' diary (1660's). The finite past *dug* does not seem to be attested before the beginning of the 18th c.; the *PPCEME* records *dug* in both functions in the memoirs of Celia Fiennes, completed ca. 1702. It appears that the past participle *dug* was created around the same time as *strung* but either spread more slowly or was resisted in written English; so far as our evidence goes, the finite past *dug* could have been created in the 17th c.[5] Both had become acceptable among the literate class by the end of that century.

For the past of *stick* we have much more information. It is the descendant of OE *stician* "to pierce; to adhere", whose past tense *sticode* would be expected to develop into a default past tense *sticked*. Sure enough, a past *sti(c)ked(e)* is well attested in ME and continues to be attested in the 16th c. The *PPCEME* includes two examples, both finite pasts: *stycked* from a letter of Sir Thomas More,[6] who was born in 1478, and *sticked* from the works of William Clowes, who was born in the early 1540's. But *stuck* is attested significantly earlier than *strung* or *dug*. The earliest attestation in the *PPCEME* is from *Ralph Roister Doister*, a play written by Nicholas Udall (born 1504) in about 1552; it is a finite past *stucke*, rhyming with *Gawyn Goodlucke* (a major figure in the play). Another seven examples, both finite past and past participle, from later in the century occur in the corpus.[7] It seems clear that by the early 17th c. the innovative strong forms were normal.

---

5.   This is a principal reason for using 17th-c. data from the *PPCEME*, even though *dug* makes its first appearance late in the 16th c.

6.   Also in the *PCEEC*.

7.   The *PCEEC* examples are all from the 17th c.

In the case of *stick* there is a further fact which complicates the picture. A synonymous strong verb *steke(n)* also appears in ME; it is not attested in OE,[8] yet it appears to be cognate with Old Frisian *steka*, Old Saxon *stekan*, and Old High German *stehhan*, all meaning "pierce". In late ME the past tense of this verb was *stake*, well attested in the works of Malory (also spelled *stak* and *stack;* the reconstructable pronunciation is /sta:k/), and the *OED* records a past participle *stoken ~ ystoke;* the inflection of the verb was precisely parallel to that of *break* (*brake, broken*). Just as the vowel of the latter verb's past participle was levelled into its finite past, yielding *broke*, one would expect to find a past *stoke*, and the *OED* records such a form from the 15th and 16th centuries. However, NONE OF THESE FORMS is likely to be the direct ancestor of *stuck*. If a form survives intact from one generation to the next, the only kind of change it undergoes is 'sound change', i.e., spontaneous changes in pronunciation. The historical record shows that sound change is overwhelmingly regular (and modern sociolinguists have identified a process of sound change that is completely regular; see, e.g., Labov (1994, 2020), Fruehwald (2013)). Apparently irregular changes in pronunciation which we cannot explain do occasionally appear in the record (though much more rarely than nonspecialists suppose), but to posit such a change without overwhelming evidence is always to flout the odds.[9] No professional gambler would do that, and no historical linguist should. Since none of the vowels in the various forms of *steke* regularly became /ʌ/, we must conclude that none of them can be the ancestor of *stuck*. Whether the existence of a synonymous inherited strong verb is relevant to the creation of *stuck* in some other way will be considered below.

5.   The productivity of /ʌ/ in Early Modern English: The case of *strung*

Making sense of the changes under consideration involves dealing with several relevant questions more or less simultaneously. We present the discussion in an order which we hope will make the situation maximally intelligible.

First of all, we need to address the complicating issue of past participles, because the situation among the past participles is not the same as the situation among

---

8.   That is probably because it had been lost in all the OE dialects for which we have adequate attestation but survived in one or more dialects that are unattested (or nearly so). For various historical reasons the dialects of almost the whole Midlands area are very sparsely attested or completely unattested in the OE period.

9.   This is why we discount the suggestion of Dobson (1968: 509, 676) that *stroke* could be the ancestor of *struck* (on which see further below); of course that is not completely impossible, but the odds are against it.

the finite pasts in standard Modern English. Since finite pasts in -*ang* (*sang, rang, sprang*) correspond to past participles in -*ung* (*rung, sung, sprung*), there are actually 10 past participles in -*ung* in standard Modern English – so even for a speaker who has 15 verbs in -*ing*, we might expect the *ung*-rule to be productive for past participles, since the TP predicts that the *ung*-rule should be able to tolerate five exceptions (in the senior author's case, *brought, winged, dinged, kinged*, and *pinged*). In nonstandard English the *ung*-rule probably is marginally productive for past participles, since *brung* does occur, but in standard English it has not replaced *brought*, and past participles like '*wung*' (for *winged*) are not attested anywhere (so far as we know). Since the TP otherwise accounts well for the productivity of morphological rules, there must be a specific reason for its inapplicability in this case. We might suggest that the paradigmatic relationship between the finite past and the past participle is responsible; in effect, since finite past *winged* is not under pressure from '*wung*', neither is past participle *winged*. The great frequency of *brought* might also be a factor; so might the default status of *winged*, etc. More research on this problem is needed, but we do not see that it jeopardizes our investigation into the finite pasts under discussion.

Secondly, while the past participles of verbs in -*ing* in the 16th c. were essentially identical to their modern descendants (with the exception of *stringed*), the finite past tenses were not. The finite pasts *sang, sprang, rang* were in competition with *sung, sprung, rung*, and that competition was vigorous.[10] The *PPCEME* includes twelve examples of *sang(e)* and twelve of finite past *sung* from the beginning of the 18th c. or earlier; most remarkably, their excerpts from the diary of Samuel Pepys – written by a single individual over the course of about a decade – include two of each. Though the finite past tense of *spring* is not as common, the situation is comparable: there are six examples of *sprang(e)* (ignoring a duplicate in the King James bible) and four of *sprung*. The finite past *rang* does not occur in that corpus, but *rung* occurs twice. It seems clear that for at least some speakers no fewer than ten verbs had both finite pasts and past participles in -*ung*; the only exceptions were *bring, string*, and a handful of very rare verbs like *wing* "to dismember (a partridge)". For those learners who encountered only one or two verbs in -*ing* with default pasts

---

10. This competition arose in the Middle English period. In Old English these verbs had two finite past stems; for instance, the indicative 1sg. and 3sg. of *singan* were *sang*, but the 2sg. was *sunge*, the plural was *sungon*, and the subjunctives were sg. *sunge*, pl. *sungen*. In the north *sang* became the only finite past in the 13th c., but further south both stems survived for another century or more, and different dialects levelled them in one direction or the other (see, e.g., Brunner 1948: 76–77). London English acquired both finite pasts because of the massive influx of migrants from other areas of England leading to substantial dialect mixture (Ekwall 1956; for further discussion see especially Lass 1994, Cheshire 1994, both with references.)

other than *string*, *N* was 12 or 13 and there were only three or four exceptions to the *-ung* rule (*bring*, *string*, and the other default-past verb(s)); in their acquisition of English the replacement of *stringed* by *strung* is exactly what the TP would lead us to expect, since a rule with twelve inputs will tolerate four exceptions and a rule with thirteen inputs will tolerate five. Even learners who heard both *sang* and *sung*, etc., as finite pasts in the speech of some adults (like Samuel Pepys) MIGHT have been moved to replace *stringed* by *strung* if the TP is not inhibited by competing forms (a question which we are not yet in a position to answer definitively, though see Sneller et al. 2019 for suggestions).

In short, the TP actually leads us to expect replacement of *stringed* by *strung* in the 16th c. (if not before). Presumably *brought* escaped replacement by *brung* because its extreme frequency made it easier to learn and more resistant to change; rarer verbs with pasts in *-inged* were perhaps not learned during NLA, but only later in life.

However, when *sang, sprang, rang* won their respective competitions in standard English – toward the end of the 17th c., to judge from the available evidence – the *ung*-rule could no longer be productive in that dialect, because the number of finite pasts in *-ung* dropped to only 7, while the number of input verbs remained the same (or perhaps increased); as the table of thresholds indicates, a rule with 12 inputs should not tolerate 5 exceptions (12–7), nor should a rule with 13 inputs tolerate 6 exceptions (13–7). The *ung*-rule might have remained productive in other dialects, and occasional speakers continue to produce forms in *-ung* even now, but there was no longer any chance that such forms would 'catch on' in the standard language.

Innovative *strung*, then, is fully consistent with the TP. *Stuck* and *dug* are more complicated cases.

6. *Stuck*

While there were many verbs with finite past tenses in *-ung* in London English in the years before 1500, there were apparently none with finite past tenses in *-uck*. Therefore if *stuck* arose by productive rule, the rule must have applied to more than verbs with roots in *-ick*. The most conservative alternative, requiring the fewest additional hypotheses, would seem to be verbs with roots in which /ɪ/ was followed by a single velar consonant, but there is a complication. ModE word-final /-ŋ/ reflects early ME /-ng/ [-ŋg], and we need to consider when word-final *-g* after a nasal was lost. Unfortunately the loss seems to have begun in eastern Midlands dialects in the 14th c., but not to have been complete in educated English until the late 16th

(Dobson 1968: 963–4).[11] Since we cannot pinpoint exactly when and where *stuck* was created, and since it could have spread into London English from a dialect in which the phonotactics were different, we must consider whether the TP can account for its creation under two different scenarios, one based on applying the *ung*-rule to verbs with roots in /ɪ/ plus any single velar (for dialects in which /-ŋ/ already existed) and one based on applying the rule to verbs with roots in /ɪ/ plus any single velar or homorganic velar cluster (i.e., /-ng/ or /-nk/). We present the options in that order.

Figuring out which verbs in /-ɪK/ were present in the speech heard by a typical native learner in, say, 1510 is probably beyond the capability of philology; we will need to employ reasonable proxies. One plausible proxy is the collection of verbs whose finite past tenses and past participles are actually attested in the 16th- and 17th-c. samples in the *PPCEME*.[12] Here is the list, with numbers of attestations:

> *brought* (scores of exx., sometimes more than 10 in a single sample);
> *digged* (11x), *picked* (7x), *pricked* (6x), *thicked* (4x, meaning "thickened"), *sticked* (2x), *nicked* (1x), *rigged* (1x);
> *sang/sung* (36x), *sprang/sprung* (12x), *flung* (10x), *stung* (5x) / *stinged* (1x),[13] *rung* (3x), *slung* (2x), *wrung* (2x)
> – plus *stuck* (8x) and *dug* (3x), not present at the beginning of the period.

In addition (as a reviewer reminds us), there is *strick*, a rare variant of *strike*, which for at least one author who used it had a past *strake* but for others had a past *struck* (see § 7 below).

Thus for this calculation there are 15 or 16 verbs altogether, and only 7 or 8 follow the rule that was extended to yield *stuck* and *dug*, so that there are 8 or 9

---

11. We are grateful to an anonymous reviewer for pointing this out – Dobson's summary of the facts is indispensable, but some details of his discussion need revision. To speak of a 'standard' pronunciation of English even now is to idealize a good deal, and for the 16th c. any such idealization seems misconceived (Milroy 1994). Nor is the fact that the early orthoepists do not recognize the velar nasal as a separate sound decisive; Dobson recognizes that their interpretations of spellings are not entirely trustworthy in his discussion of *-mb* (p. 960), and the same reserve seems all the more appropriate in a discussion of *-ng*, given that that was the only conceivable way of spelling /ŋ/.

12. A reviewer points out that one might also construct such a list from a concordance to Shakespeare's plays, which are largely in dialogue, or the *Corpus of English dialogues 1560–1760* (Kytö et al. 2006). Both should be attempted, but it is not clear that either would be a better fit to child-directed speech, which tends to include a language's most frequent lexemes overall – making the sheer size of a corpus the most important variable.

13. *Stinged* must be a nonce form, an outlying dialect form, or a personal idiosyncrasy; the irregular *stung* is clearly usual, and had been for some generations.

exceptions. Since a rule potentially applying to 15 or 16 lexemes should tolerate only 5 exceptions, those 8 or 9 should easily be enough to prevent it from becoming productive. However, it is plausible to suppose that not all these forms were acquired in NLA. In particular, *thicked* occurs only in discussions of the preparation of cloth (in statutes of the period), *rigged* describes the outfitting of a ship, and a reviewer observes that *nick* is also quite rare in 16th- and 17th-c. documents. If we remove those verbs, the total is 12 or 13; a class of 13 should tolerate only 5 exceptions, and a class of 12 only 4. In the latter case we have 5 exceptions, namely *brought, digged, picked, pricked*, and *sticked;* in the former (including *strick*) we have 6 exceptions (including *strake*) or 5 (if the past was *struck*). Only under the last scenario does the TP predict the creation of *stuck*. This is a borderline result: only if *struck* was already in existence does the TP predict *stuck*. We do not really know whether that was the case or not, but the chronology of attestations suggests that *stuck* was created first, in which case the TP fails in this case too.

Of course we might be using the wrong proxy. An alternative proxy would be a relevant modern list of lexemes in child-directed speech, adjusted so as to account for the differences in inflection between early 16th-c. English and present-day English. In the 6-million-word *CHILDES* corpus of child-directed American English a dozen relevant verbs appear in the past tense at least once per million tokens each; we give them here with their early-16th-c. past forms:[14]

1	bring, brought	7	fling, flung
2	sing, sang/sung	8	kick, kicked
3	ring, rang/rung	9	lick, licked
4	spring, sprang/sprung	10	pick, picked
5	swing, swung	11	stick, sticked
6	sting, stung	12	dig, digged

Half of these dozen verbs had past tenses in *-ung* in the early 16th c., but according to the TP the six verbs in this list that have past tenses of other kinds should have been enough to prevent the formation of *stuck* in NLA, or at least to prevent it from catching on, if these verbs are a representative sample of child-directed speech in the 16th c.

But what if *stuck* arose in a dialect in which *bring*, etc. were still [bɹɪŋɡ], etc.? In that case verbs in *-ink* need to be included in the potential lexical class membership. Here is the list of past tenses attested in the *PPCEME* with verbs in *-ink* included (but *stinged* omitted, see above):

---

14. These and other classes of English verbs in the *CHILDES* corpus were collected and extensively used to account for English NLA in Yang (2016); see Yang (2016 passim).

brought, thought (both very common);
digged (11x), *picked* (7x), *pricked* (6x), *linked* (5x), *thicked* (4x), *sticked* (2x), *nicked* (1x), *rigged* (1x), *winked* (1x), *blinkt* (1x);
drank/drunk (44x),[15] *sang/sung* (36x), *sprang/sprung* (12x), *flung* (10x), *stung* (5x), *rung* (3x), *shrank/shrunk* (3x), *slung* (2x), *wrung* (2x), *stank/stunk* (2x), *sank/*sunk* (2x)[16]
 – plus *stuck* (8x) and *dug* (3x), not present at the beginning of the period, and possibly *strake* (see above).

We now have a class of 23 or 24 verbs, so any rule should tolerate 7 exceptions; but there are at least 11 exceptions to any potential rule. If we exclude *thicked, rigged, nicked* (see above) and *blinkt* – the single token of the last means "turned sour", not likely to occur in child-directed speech – we still have 19 or 20 verbs, so any rule should tolerate 6 exceptions, but there are still at least 8. Once again the TP does not predict *stuck*.

Finally, let us repeat the experiment using relevant verbs from the CHILDES database that are not rare, this time including verbs in -*ink*, with their 16th-c. past tenses, namely:

1	bring, brought	9	sting, stung
2	think, thought	10	fling, flung
3	sing, sang/sung	11	kick, kicked
4	ring, rang/rung	12	lick, licked
5	spring, sprang/sprung	13	pick, picked
6	drink, drank/drunk	14	stick, sticked
7	sink, sank/sunk	15	dig, digged
8	swing, swung	16	blink, blinked

The last of these verbs might not have been nearly so common in the 16th c., but *wink, winked* was probably commoner, since it could mean "keep the eyes closed"; the numbers should be approximately correct in any case. We now have 16 verbs, among which any rule should tolerate 5 exceptions, but the *ung*-rule applies to only half of them – there are still 8 exceptions.

In short, we have not certainly been able to derive *stuck* via the application of the TP in NLA under any reasonable assumptions about what 16th-c. learners of English were hearing (though there is a bare possibility that we might be able to).

---

15. Though both these forms occur as finite pasts, *drank* appears 41 times and *drunk* only 3; thus the balance is very different than for other verbs with such forms in competition.

16. *Sunk* happens not to occur as a finite past in the *PPCEME*, though *OED* citations show that such a form was in use. (*Sunk* does occur as a past participle 14x in the *PPCEME*.)

It seems at least as likely that *stuck* was created by some other process, and we can propose one. As we noted at the end of § 4, there was beside the weak verb *stick* (past *sticked*, participle *sticked*) a synonymous strong verb *steke* (*stake, stoken;* later *stoke, stoken*). It seems reasonable to suggest that some speakers – not necessarily native learners – confused the two and were therefore prompted to create a strong past and past participle for *stick*. But given that the vowel in the root of the later was /ɪ/ (throughout the word's attested history), a strong past or past participle could only be either *stack* /stak/ (which would develop into /stæk/ by regular sound change) or *stuck* /stʊk/ (which would develop into /stʌk/). In the event *stuck* was created in both functions and caught on. This could be called lexical analogy, though it is lexical analogy of a special kind: not the influence of a lexeme on a quite different lexeme that happens to resemble it in various ways, but the perhaps mutual influence of very similar (and in fact etymologically related) lexemes competing for position as the default representation of a meaning in the basic vocabulary.[17] Whether such a process can be modelled as rule-governed is doubtful; we have not been able to come up with a rule which the TP could test.

There is another piece of evidence which points in the same direction. We have had to deal with it in the calculations in this section, but we have not discussed it in detail because it adds a further layer of complexity to the picture. We must do so before we turn to *dug*.

7.   *Strike / strick*, past *stroke / strake / struck*

The OE verb *strīcan* "to stroke, to wipe; to dash, to run" is the direct ancestor of Modern English *strike* (the meaning has evolved considerably over time). Though only the present is attested in OE, the past 3sg. can only have been *strāc and the past participle *stricen, parallel to *wrītan, wrāt, writen*, because the inflection of OE strong verbs with *ī* in the root is uniform. The result should have been *strike, stroke, stricken* in the 16th c., parallel to *write, wrote, written*, and in fact that paradigm is well attested. The present stem *strike* is common; perusal of about 40% of the *PPCEME* yielded 21 examples. Past *stroke* appears seven times in the same corpus, and the examples are distributed throughout the 16th c. and down to 1630. Past

---

17.   A more typical example of lexical analogy from the history of English verb inflection is the early ME creation of past *cauȝte* (beside *cacchede*) to the French loanword *cacchen* on the model of *lauȝte*, the past of the synonymous native verb *lacchen* (see the *OED* s.vv. *catch*, v. and *latch*, v.1). In principle such straightforward lexical analogies can be modelled as rules. Of course a rule covering only two lexemes is comparatively inefficient in an adult lexicon, but in the restricted lexicon of a native learner that need not be the case.

participle *stricken* occurs with some frequency (and survives in Modern English, though more often as an adjective than as a participle).[18]

But there are other forms in competition with all of the above in the *PPCEME*. There is a rare present *strick* beside *strike*. It too was a descendant of OE *strīcan*, but one in which the vowel had been shortened; that is known to have happened to a handful of class I strong verbs in some dialects of OE (Seebold 1966). One example of *strick* in the *PPCEME* occurs in the work of Thomas Harman, published in the 1560's:

> "A vaunt verlet," quoth this vpright man, and letes dryue with all his force at this hosteler, and after halfe a dosen blowes, he *strycks* his staffe out of his hande, …
> (*PPCEME*, HARMAN-E1-P1, 64.7–9)

Much more surprisingly, there are two examples from John Locke's well-known treatise on education, published in 1693. Since Locke was born in Somerset in 1632, it is possible that *strick* is a dialect form that he never abandoned, but apparently it was not stigmatized. So far as our evidence goes, *strick* persisted as a variant of *strike* in some areas, or in the speech of some individuals, into the 18th c.[19]

It seems clear that *struck* must have been created as a past tense to *strick* in much the same way that *stuck* was formed to *stick* – and since *strike / strick* was already a strong verb, *struck* presumably replaced *stroke* and *stricken*. But already by the 16th c. the pattern of variation was more complex. The passages from Locke's works in the *PPCEME* contain no example of a past tense or past participle of *strick*, but an online search of Locke's works on Project Gutenberg turns up examples of *struck* in both functions, as one would expect. However, for Thomas Harman the past of *strick* was *strake;* the passage quoted above continues:

> and as this hosteler stept backe to haue taken vp his staffe agayne, his glymmeringe Morte flinges a great stone at him, and *strake* him one the heade that downe hee fales, wyth the bloud about his eares.   (*PPCEME*, HARMAN-E1-P1, 64.200–1)

The past *strake* also occurs in two *PPCEME* passages from the chronicle of Robert Fabyan, who died in 1516, beside a past participle *stryken* (i.e., *stricken*); his present for this verb is not attested in the *PPCEME*, but an online search of the Open Library copy yields four examples of *stryke* (as well as many more of past participle *stryken*). The *PPCEME* attests both present *strike* and past *strake* from the biography of Sir John Perrott, written in the 1590's. But many authors whose present stem is

---

18. Past participle *stroke* also occurs, e.g., in one of the Cely letters from the late 15th c. and in a 17th-c. letter of John Chamberlain, both in the *PCEEC*.

19. The *PCEEC* seems to have only one example, from a letter of Nicholas Bacon in the second half of the 16th c.

strike already have a past *struck*, exactly as in Modern English. The diary of Henry Machyn, written in the 1550's and '60's, exhibits a paradigm *stryke, struck, stryken*. The examples of *struck* next in chronological order are from the chronicle of John Hayward, written late in the 16th c. or early in the 17th, and from a comedy by Thomas Middleton published in 1630; both are past participles. No finite past other than *struck* appears in the texts of the *PPCEME* from after 1630, though the past participles *struck* and *stricken* continue to compete in general use until at least the 1660's.[20] Samuel Pepys and Celia Fiennes both have the modern paradigm *strike, struck, struck*.

The reader who finds this picture confusing is no more clueless than we are; it seems clear enough that by the middle of the 16th c. the competition between stems of *strike / strick* could be described as a free-for-all. That is relevant because influence of *stick* on *strike / strick* and vice versa is a real possibility, given the similarity of form and meaning that the two verbs exhibit: in effect, the entirely new strong past *stuck* and the remodelled strong past *struck* might have supported one another. The fact that *strike* seems to have replaced *strick* while on the other hand *struck* replaced *stroke* makes the entire situation difficult to judge. In any case there seems to be no rule that the TP could be used to assess, and once again we cannot be certain that these innovations occurred in NLA at all; it is possible that adult native speakers were responsible for them.

8. *Dug*

The pattern of evidence suggests, though it does not quite establish, that *dug* was created later than either *struck* or *stuck;* it might or might not have been created later than *strung*. Let us therefore revisit the first calculation under § 6, with *thicked, rigged*, and *stinged* (see fn. 13) excluded but *struck* and *stuck* included. We have the following set of relevant forms for native learners who hear *strick* in use:

> *brought;*
> *digged, picked, pricked, nicked;*
> *sang/sung, sprang/sprung, flung, struck, stuck, stung, rung, slung, wrung.*

In this case *N* is 14, which should tolerate five exceptions; since we have only five exceptions to the rule "replace /ɪ/ with /ʊ/", we expect the rule to be productive. The only surprise is that *digged*, the commonest default-past verb of this set in our

---

20. Forman's diary, which covers the second half the 16th c., yields a variant *stroken*. Past *strake* and past participle *strucken* occur several times in a wide range of letters in the *PCEEC*, adding to the picture of extensive variation in several dimensions.

sample, has been targeted for replacement; that will have to be laid to contingent factors no longer recoverable. For native learners hearing *strike* in use, we will have to omit *struck* from the above list, and it is possible that *nicked* should be omitted even early in the 17th c. In either case *N* is then 13, which also tolerates five exceptions, or, if both are omitted, *N* is 12, which tolerates four, so the rule should have been productive even for those learners.

 In short, if *dug* was the last of these innovative forms to be constructed, the TP predicts its appearance. Moreover, just as the victory of *sang, sprang, rang* late in the 17th c. explains why *winged* was not replaced by '*wung*', it also explains why *rigged* did not yield to '*rug*' – nor *picked* to '*puck*', and so on, in accordance with the TP. Note that the verbs in *-ick* which entered English after the victory of *-ang* over *-ung* have all adopted and retained the default inflectional pattern; to judge from the *OED* entries, they include at least *click, flick, tick*, and *trick*. Finally, the fact that the calculations in this section worked out so neatly suggests that a list of forms most frequently attested in non-specialized texts is a reasonable proxy for the unobtainable list of input forms for NLA in centuries long past; for relevant discussion see Kodner (2019, 2020). Of course this conclusion regarding the origin of *dug* cannot strictly be proved, but it is a construct *où tout se tient*.

9.   Looking forward

The results just described are clear-cut enough and interesting enough to prompt testing of the TP against other instances of past tense 'irregularization' in English. While a full study of any of the relevant verbs is beyond the scope of this paper, two cases seem clear enough that suggestions might be useful.

 The replacement of *dived* by *dove* in 19th-c. North American English appears to be easily explainable by the TP. So far as we can discover, only four verbs in *-ive* /-ɑɪv/ were common enough to be learned regularly by small children at that time and place;[21] their finite pasts were:

> *drove, strove;*
> *throve ~ thrived;*
> *dived.*

In addition, there were two verbs of more specialized meaning, whose finite pasts were:

> *hived (bees), rived (cedar shakes with a froe).*

---

21. *Shrive* and *swive* were no longer in everyday use; *jive* had not yet been coined.

Even if small children learned the specialized verbs, pasts in *-ove* could have become productive in the NLA of those learners who heard *throve*, according to the table in § 3; for children who did not learn the specialized verbs early, the *-ove* rule could have become productive even if they heard *thrived*. Thus the TP easily predicts the appearance of *dove*.

A contrary case is *hung*. Like *strike, struck* (see § 7), *hang, hung* (transitive) has been a strong verb throughout its history, but its finite past is not the one that is expected etymologically. To judge from the citations in the *OED, hung* might first have appeared in the 16th c.; on the other hand, some of the Middle English finite pasts spelled *hong(-)* might conceivably have contained /ʊ/ in the root. This case requires a thorough study in detail. However, it is already clear that the verb had several competing finite pasts in Middle English, and it would be reasonable to wager that the history of *hang* will turn out to be much like that of *strike* – not involving NLA nor the TP.

There is also a relatively recent phenomenon that seems to be different from both the phenomena dealt with in this paper. A finite past *snuck* (to *sneak*) first appears in the USA in comic writing in the mid-19th c.; William Faulkner, toward the end of his short story "Was", has a character say *you skun the hen-house one time too many*; Dizzy Dean is quoted as having said *he slud into third*;[22] a collection of oral histories of D-Day includes the statement *those that were wounded we drug up behind* (Astor 1994: 196); and a finite past *rutch* (to *reach*) was heard at a sporting event at Pennsylvania State University (Philip Baldi, p.c.; several of these forms are also adduced by Cheshire 1994: 122). It seems clear that /ʌ/ is spreading as a finite past tense vowel for verbs in North American English, but it is not clear exactly what is going on, nor why default-pattern verbs occasionally acquire such finite pasts. The discussion of Hogg (1988) and especially Cheshire (1994), with references, is highly relevant (and Cheshire's observations must be largely correct), but more information about the process is needed. That is a topic for future research.

10. Conclusion

A close examination of the 'irregularization' of verbs in Early Modern English confirms the usefulness of the Tolerance Principle in predicting which morphological rules will become productive. It also shows that the TP is not the whole story, as might have been expected. It appears that situations in which synonymous and similar but differently inflected verbs, such as *steke* and *stick*, are in competition

---

22. We are grateful to Patrick Stiles for calling this to our attention.

can lead to unexpected results; competition between multiple stems of a single verb, as in the case of *strike / strick*, can also yield unforeseen anomalies. It is very likely that adult native speakers are responsible for those innovative forms.

Acknowledgements

References

Albright, Adam & Bruce Hayes. 2003. Rules vs. analogy in English past tenses: A computational/experimental study. *Cognition* 90(2). 119–161. https://doi.org/10.1016/S0010-0277(03)00146-X

Astor, Gerald. 1994. *June 6, 1944: The voices of D-Day*. New York: St. Martin's.

Berko, Jean. 1958. The child's learning of English morphology. *Word* 14. 150–177. https://doi.org/10.1080/00437956.1958.11659661

Brunner, Karl. 1948. *Abriß der mittelenglischen Grammatik* (2nd edn.). Halle: Niemeyer.

Cheshire, Jenny. 1994. Standardization and the English irregular verbs. In Stein & Tieken-Boon van Ostade (eds.), *Towards a standard English 1600–1800*, 115–133. Berlin: Mouton de Gruyter.

Dobson, Eric John. 1968. *English pronunciation 1500–1700*. Oxford: Clarendon Press.

Ekwall, Eilert. 1956. *Studies on the population of medieval London*. Stockholm: Almqvist & Wiksell.

Fruehwald, Josef. 2013. Phonological influence on phonetic change. Philadelphia: University of Pennsylvania dissertation.

Hogg, Richard M. 1988. Snuck: The development of irregular preterite forms. In Graham Nixon & John Honey (eds.), *An historic tongue: Studies in English linguistics in memory of Barbara Strang*, 31–40. London: Routledge.

Kodner, Jordan. 2019. Estimating child linguistic experience from historical corpora. *Glossa* 4(1). 122. https://doi.org/10.5334/gjgl.926

Kodner, Jordan. 2020. Language acquisition in the past. Philadelphia: University of Pennsylvania dissertation.

Kroch, Anthony, Beatrice Santorini & Lauren Delfs. 2004. *The Penn-Helsinki parsed corpus of Early Modern English*. Philadelphia: Department of Linguistics, University of Pennsylvania.

Kytö, Merja, Jonathan Culpepper & Terry Walker. 2006. *A corpus of English dialogues 1560–1760*. Uppsala: Department of English, Uppsala Universitet.

Labov, William. 1994. *Principles of linguistic change. Volume 1. Internal factors*. Oxford: Blackwell.

Labov, William. 2020. The regularity of regular sound change. *Language* 96. 42–59. https://doi.org/10.1353/lan.2020.0001

Lass, Roger. 1994. Proliferation and option-cutting: The strong verb in the fifteenth to eighteenth centuries. In Dieter Stein & Ingrid Tieken-Boon van Ostade (eds.), *Towards a standard English 1600–1800*, 81–113. Berlin: Mouton de Gruyter.

MacWhinney, Brian. 2000. *The CHILDES project: Tools for analyzing talk*. 3rd ed. Mahwah, NJ: Lawrence Erlbaum Associates.

Marcus, Gary F., Steven Pinker, Michael Ullman, Michelle Hollander, T. John Rosen, Fei Xu & Harald Clahsen. 1992. *Overregularization in language acquisition*. Chicago: University of Chicago Press.

Milroy, James. 1994. The notion of "standard language" and its applicability to the study of Early Modern English pronunciation. In Dieter Stein & Ingrid Tieken-Boon van Ostade (eds.), *Towards a standard English 1600–1800*, 19–29. Berlin: Mouton de Gruyter.

Nevalainen, Terttu, Helena Raumolin-Brunberg, Jukka Keränen, Minna Nevala, Arja Nurmi, Minna Palander-Collin, Ann Taylor, Susan Pintzuk & Anthony Warner. 2006. *The parsed corpus of early English correspondence*. Text version. Helsinki: University of Helsinki.

Nida, Eugene. 1949. *Morphology: The descriptive analysis of words* (2nd edn.). Ann Arbor: University of Michigan Press.

Pinker, Steven. 1999. *Words and rules: The ingredients of language*. New York: Basic Books.

Seebold, Elmar. 1966. Die schwundstufigen Präsentien (Aoristpräsentien) der *ei*-Reihe. *Anglia* 84. 1–26. https://doi.org/10.1515/angl.1966.1966.84.1

Sneller, Elizabeth, Josef Fruehwald & Charles Yang. 2019. Using the Tolerance Principle to predict phonological change. *Language Variation and Change* 31(1). 1–20. https://doi.org/10.1017/S0954394519000061

Strik, Oscar. 2015. Modelling analogical change: A history of Swedish and Frisian verb inflection. Groningen: University of Groningen dissertation.

Xu, Fei & Steven Pinker. 1995. Weird past tense forms. *Journal of Child Language* 22(3). 531–556. https://doi.org/10.1017/S0305000900009946

Yang, Charles. 2016. *The price of linguistic productivity: How children learn to break the rules of language*. Cambridge, MA: MIT Press. https://doi.org/10.7551/mitpress/9780262035323.001.0001

Yip, Kenneth & Gerald J. Sussman. 1997. Sparse representations for fast, one-shot learning. *Proceedings of the National Conference on Artificial Intelligence* 521–527.

PART II

Syntax

CHAPTER 5

The reanalysis of VO in the history of English
Evidence for a language-internal account

Chiara De Bastiani

Università Ca' Foscari Venezia

In this paper, I present empirical data about selected Early Middle English (henceforth eME) texts which show how the reanalysis of VO as the unmarked word order constitutes a uniform development from Old English (henceforth OE). I employ an antisymmetric framework, in which information structural (henceforth IS) and weight interface conditions govern the pre- and post-verbal spell-out of constituents. The dialectal provenance of the texts and their transmission history is considered, to allow for a comparative analysis of the data.

Keywords: word order variation, Early Middle English, information structure, language change

1. Introduction

In this paper, I investigate eME subclauses from South-East Midlands and West Midlands texts, providing evidence for a language internal account in the transition from OE to eME.

Word order in OE has been the objective of a long-standing debate in the literature on language change and word order variation, starting from van Kemenade (1987) to more recent literature taking information structure and weight conditions as influencing word order to various degrees (Pintzuk & Taylor 2006, Taylor & Pintzuk 2011, 2012a–b, 2015; Cloutier 2009; Petrova & Speyer 2011; De Bastiani 2017, 2020; Struik & van Kemenade 2018). Whereas most attention has been devoted to word order variation in OE, eME texts have not received the same share of attention (but cf. Kroch & Taylor 2000; Taylor & Pintzuk 2011).

In the following, I build on De Bastiani (2017, 2020) and Struik & van Kemenade (2018) in proposing an investigation of the interaction between word order, information structure and weight of the constituents and focus on a selection

of eME texts. My aim is to show that there is a steady and consistent development throughout the history of English, as far as the mapping of constituents before and after the verb is concerned. Moreover, my investigation shows that not only is dialectal provenance of interest in the variation attested in eME, but also that the transmission history of the texts under consideration proves to be a factor in the variation observed. The paper is structured as follows: in § 2, I summarize the debate on word order variation in OE, and introduce the framework underlying the present research, whereas in § 3, I introduce the methodology and the sample selected. In § 4, I present the data and results; Section 5 concludes the paper.

2. The debate on OE word order and diachronic change

OE syntax presents variation in both the relative order of object and verb (henceforth O and V), as well as in the relative order of verb and auxiliary (henceforth V and Aux). The variation is resolved in the twelfth century, when the shift from OV to VO takes place (cf. Fischer et al. 2000).[1]

Different frameworks have been employed to analyze word order variation in OE. Van Kemenade (1987), for instance, proposes a base OVAux structure, and extraposition and V(P)R to derive surface VO orders. Pintzuk (1999), however, points out that, while the assumption of an OV grammar is needed to derive some of the word orders attested in OE, one can also find VO orders where the analysis of extraposition from an OV base structure is ruled out. In (1), in fact, a pronoun is extraposed:

---

1.   As an anonymous reviewer notes, there is still a significant number of OV orders in ME (Pintzuk & Taylor 2006); Fischer et al. (2000) report that, although 1200 is commonly analyzed as the point in time when VO is reanalyzed as the basic word order, OV orders presenting quantified and negated objects retain the pre-verbal position the longer. Negated and quantified elements are argued to be licensed by different syntactic requirements than referential objects. However, as the study by Foster & van der Wurff (1995) shows, surface OV with non-negated and non-quantified objects can be found. Notice however, that OV is about seven times more frequent in verse than in prose. Scrambling and leftward movement of objects could be linked to stylistic requirements. For some of the remaining OV orders analysed in Foster & van der Wurff, a scrambling analysis is compatible; this would entail that the object is moved for information structural reasons to a position outside the vP and involve hence a different type of derivation with respect to clauses which present clear evidence that the object is still in the vP domain (cf. De Bastiani 2020 for a proposal). It is therefore possible to assume that these cases represent marked structures, whereas surface VO order for non-quantified and non-negated objects was the norm. The exact analysis of these particular instances of OV order in ME would however go beyond the scope of the present paper.

(1)   *We wyllað secgan eow sum bigspell*
      We want tell   you a   parable
      "We want to tell you a parable"

(*ÆCHom* I, 14.1.212.6, in Fischer et al. 2000: 142)

Pintzuk (1999) concludes that OE word order variation was characterized by competing grammars, which differ in the headedness of both the IP and VP projections, giving rise to the word orders attested. These grammars compete throughout the whole OE period, with the head-initial grammar steadily gaining ground until the competition is resolved in the ME period. Other proposals involve the assumption of a uniform base grammar, and different leftward movement operations which derive the OV orders attested in OE. Such accounts were put forward by Roberts (1997), Fischer et al. (2000), Biberauer & Roberts (2005). In the accounts quoted, different triggers are assumed for the leftward movement operations yielding OV surface word order, cf., for instance, the checking of case features (Roberts 1997), the satisfactions of EPP features (Biberauer & Roberts 2005), and leftward licensing operations involving different types of objects (Fischer et al. 2000). For reasons of space, these proposals are not reviewed in the present paper.[2] An important contribution of the frameworks cited to the debate is the demonstration that OV surface word orders can be accounted for by assuming a uniform grammar, thus dispensing with the degree of optionality posed by an account of grammars in competition. However, the exact nature of the triggers offers some room for discussion; with the present paper, a revised form of the framework, proposed by Hinterhölzl (2014, 2015, 2017), is adopted, according to which both information structural and prosodic interface conditions regulate the spell-out of constituents after leftward licensing movement has taken place (cf. below). This accounts builds on previous research on OE and on the wealth of investigations on Early Germanic which show how information structure and weight of constituents influence word order (cf. Behaghel 1932; Bech 2001; van Kemenade & Los 2006; Cloutier 2009; Linde 2009; van Kemenade 2009; Petrova 2009; Hinterhölzl et al. 2005; Hinterhölzl & Petrova 2010, 2018; Petrova & Speyer 2011; van Kemenade & Westergaard 2012; Elenbaas & van Kemenade 2014; Milićev 2016; Struik & van Kemenade 2018, and Los & van Kemenade 2018, among many others).

Before introducing the framework, let us consider that Struik & van Kemenade (2018) present solid evidence to assume that a head initial base is indeed the unmarked option in OE. Struik & van Kemenade do not assume an a priori syntactic structure for their investigation and base their conclusions on the evidence

---

2. For a review of the different accounts on word order variation and the syntactic accounts proposed, cf. De Bastiani (2019, 2020); cf. moreover, Roberts (1997), Fischer et al. (2000), Fuß & Trips (2002), Haeberli (2008), Wallenberg (2009), among others.

provided by the data. Building on the literature on information structure and word order (Taylor & Pintzuk 2012a, b; Hróarsdóttir 2000, among others), they investigate the correlation of information structure and direct object position in a wide sample of OE non-translated texts. Summarizing their findings, they show that in both VAux and AuxV structures, it is the given status of the direct object that predicts word order; they moreover show that weight also significantly predicts word order, since longer objects prefer the post-verbal position.[3] Having seen that there is good evidence to postulate a uniform VO base for OE,[4] let us now introduce our framework. The framework employed in the present work is a revised version of Hinterhölzl's framework (2014, 2015, 2017), which builds on Biberauer & Roberts' (2005) antisymmetric account but considers IS and weight as relevant factors for the mapping of constituents in OE. In a similar way as to the proposal by Fischer et al. (2000), spell-out is decisive for the surface OV and VO order, and leftward licensing movement of different types of constituents is obligatory. According to the framework proposed, arguments of verbs undergo leftward licensing movement to a Case Phrase, whereas verb particles undergo leftward movement into the specifier of a low Aspect Phrase, and predicative elements undergo leftward movement to a Predicative Phrase. After these licensing movements have taken place, Hinterhölzl proposes that given elements must occupy a weak branch in prosodic structure ('Givenness-Transparency Condition'), whereas new elements must occupy a strong position ('Focus-Transparency Condition'). Finally, it is assumed that spell-out is also driven by syntactic weight. Syntactic weight is defined as follows:

(2)   A syntactic phrase XP counts as heavy if both its head X and the complement of X contain lexical material.   (Hinterhölzl 2014: 345)

The prosodic condition requires heavy elements to be spelled out in the post-verbal domain. Hinterhölzl defines the post-verbal position as a strong position, whereas the pre-verbal one is the weak position;[5] to satisfy the G-transparency condition, a given object must be spelled out in its checking position, while a new constituent

---

3.   For the complete investigation and details on the statistical tests, cf. Struik & van Kemenade (2018). The weight factor is examined also in Taylor & Pintzuk (2012a, b), who conclude that weight has a stronger impact than IS; it must be noticed, however, that their approach is slightly different than the one proposed in Struik & van Kemenade (2018), since Taylor & Pintzuk (2012a, b) examine whether post-verbal objects correlate with new information, which is not the case.

4.   For a discussion on the advantages of the antisymmetric approach, see also Roberts (1997) and Wallenberg (2009).

5.   Syntactic structure is equated in his framework to prosodic metric composition; in prosodic composition, it is the recursive branch which is stronger than the leftward branch (cf. Hinterhölzl 2015).

is spelled out in its base position inside the VP, resulting in VO word order. Heavy phrases, likewise, must be mapped on a strong branch. Notice that the interface conditions postulated in the framework only predict the spell-out sites for given, new or heavy objects; in other words, no straightforward prediction can be formulated about given and heavy objects, for instance;[6] one can postulate that the Givenness Transparency and the Weight Transparency conditions would conflict in the case of a given but heavy constituent. From the qualitative study of a restricted sample of OE sentences, De Bastiani (2017, 2020) shows that given elements in the post-verbal domain are either heavy or contrasted (i.e., they receive contrastive focus accent and are hence prosodically heavy). De Bastiani (2017, 2020) concludes that in the OE period, while there is a strong tendency for light and given objects to be spelled out in the pre-verbal domain, the weight condition can indeed prevail over the G-Transparency condition. This conclusion is also reached by Petrova & Speyer (2011), who claim that weight can indeed override the role of information structure. Moreover, what emerges from the investigation by Struik & van Kemenade (2018) and De Bastiani (2020) is that it is the pre-verbal domain that presents a restriction on the type of elements spelled out there, whereas the post-verbal domain hosts a heterogeneous set of elements. Given these considerations, only the G-Transparency and the prosodic mapping conditions are assumed in order to derive the word order variation in OE, given the discussion presented in De Bastiani (2020). The expected output resulting from the interaction of spell-out and interface conditions are given in the following:[7]

(3)   a.   [vP[CASEP O [CASE]][PREDP [PRED]][ASPP [ASP]][VP [V OCopy]]]

b.   [vP[CASEP Ogiven [CASE]][PREDP [PRED]][ASPP [ASP]][VP [V ~~Ogiven~~]]]

c.   [vP[CASEP ~~Oheavy~~ [CASE]][PREDP [PRED]][ASPP [ASP]][VP [V Oheavy]]]

In (3a), the object is licenced in [Spec,CaseP] and leaves a copy inside the VP, whereas in (3b) we find the pre-verbal spell-out of a given object, due to the G-Transparency condition, and in (3c) the object is spelled out in post-verbal position, due to the prosodic mapping condition.

According to Hinterhölzl, both the IS and the weight condition interact during the OE period, whereas the IS interface condition is lost in the transition from the OE to the eME period. This process is due to the grammaticalization of the definite

---

6.   I thank an anonymous reviewer for pointing this out.

7.   For a discussion about the metrical reasons for a definition of prosodic weight, and for the spell-out of the higher or lower copy, cf. Hinterhölzl (2014, 2015, 2017) and De Bastiani (2020, Chapter 1).

determiner, which would turn definite – and hence accessible and given DPs – into heavy phrases, which are progressively spelled out in the post-verbal domain, thus leading to the blurring of the G-Transparency condition. Before post-verbal spell-out becomes the unmarked option for each type of verbal object, a transitional stage is postulated, where only light elements are spelled out in pre-verbal position. This is both due to the prosodic mapping condition, since heavy objects are predicted to be spelled out in post-verbal position, but also to 'Preference for the Higher Copy' condition, which predicts that an element is spelled out in the checking position, unless other interface conditions intervene. The framework delineated by Hinterhölzl has the advantage of providing the necessary syntactic structure to derive Aux V O and Aux O V orders and furthermore provides triggers for the spell-out of the higher or lower copy of movement.[8]

There are numerous studies on the interaction of IS, prosody and syntax in Early Germanic (cf. Bech 2001; van Kemenade & Los 2006; Petrova 2009; Linde 2009; Hinterhölzl & Petrova 2018; Los & van Kemenade 2018, among others); therefore, an account in which both IS and prosody interface conditions intervene in the spell-out of constituents seems appropriate to the author of the present paper, especially because they fill the gap the aforementioned syntactic accounts leave as far as the apparent optionality in the choice of VO and OV orders is concerned.

In § 3, the research questions, the dataset and the methods are introduced.

3.   Research questions, methodology and sample

From the literature overview given in the previous section, it emerges that the different syntactic accounts have not provided a clear trigger for the expected outputs of their frameworks. Such triggers are provided in Hinterhölzl (2014, 2015, 2017), as commented on in § 2. Starting from Struik & van Kemenade's (2018) investigation, which shows that OV in OE was driven by the givenness of the object, I aim to show that the framework adopted accounts for both word order variation in the eME sample examined and for the loss of surface OV orders in the history of English. This approach postulates a language internal trigger for word order change; whether external factors accelerated the change in some dialect areas is left for future research (cf. Kroch & Taylor 2000; Fuß & Trips 2002; Emonds & Faarlund

---

8.   Hinterhölzl (2014, 2015, 2017), however, does not provide a derivation for VAux sentences. De Bastiani (2020: 28) adopts an analysis which combines Hinterhölzl's interface conditions with the pied-piping of the vP to [Spec,TP], as proposed by Biberauer & Roberts (2005). Following Milićev's (2016) analysis of VAux OE clauses, she proposes that the pied-piping of the vP is triggered by the backgrounding of the entire proposition.

2014).[9] According to Hinterhölzl, in fact, IS and weight were both influential for the spell-out of constituents in the OE period. De Bastiani (2017, 2020), moreover, shows that heavy elements are mostly mapped in post-verbal position already in OE. If Hinterhölzl's predictions are on the right track, we would expect eME sentences to display mostly light elements in pre-verbal position. Post-verbal elements are expected to be of a mixed nature, given the results provided in the literature quoted above. The present study concentrates on a sample of subordinate clauses (in order to minimize the impact on word order by finite verb movement; see Pintzuk & Taylor 2006, Taylor & Pintzuk 2011, 2012a, b; Struik & van Kemenade 2018) that contain a subject, an object and a non-finite verbal form. The corpus from which the sample is taken comprises the *Trinity Homilies, Vices and Virtues*, the *Lambeth Homilies, Life and Passion of Saint Juliana, The Guardianship of the Soul* and *Holy Maidenhood*.[10] The texts of the *Trinity Homilies* and *Vices and Virtues* come from the South East Midlands area;[11] the *Trinity Homilies* are dated around 1225,[12] and they share five sermons with the *Lambeth Homilies*, which, according to Morris (1969), are probably older. The *Trinity Homilies* are translations from Latin sermons, even though some seem to be original compositions (Morris 1873, 1969). They represent, however, a transcription and modernisation of probably older material. Likewise, the treatise of *Vices and Virtues* is dated around 1225 but was probably composed earlier. According to Hall, the language is conservative and is closer to the older text from which the present composition was copied.

As far as the *Lambeth Homilies* (henceforth LH) are concerned, Kroch & Taylor (2000) report that these homilies have been localized in the same West Midlands area as the Katherine Group (henceforth KG), however, Sisam (1951) remarks that the language constitutes the result of various manipulations of complex material. The *Homilies* are divided into two groups; the first group belongs to the PPCME2 MX1 period. They are dated before 1225 but are adaptations of older material from the 11th century. The second group belongs to period M1 and was probably composed later than the MX1 compositions. The texts of the KG are original

---

9.   The anonymous reviewers of this paper commented, in fact, that the internal account proposed here could be compatible with a scenario in which language contact might have accelerated the internal language change process postulated in the paper.

10.   The *Life and Passion of Saint Juliana, The Guardianship of the Soul*, and *Holy Maidenhood* belong to the Katherine Group.

11.   The philological information about the texts is taken from the PPCME2 corpus, texts arranged by dialect, Holthausen (1888), Hall (1972), Morris (1969), D'Ardenne (1961), Sisam (1951), Huber & Robertson (2016).

12.   According to Morris (1969), however, they were composed before 1200 AD.

compositions, dated around 1200–1225. While D'Ardenne (1961, 1977) strongly stresses the fact that this group of texts is genuinely English, with neither Norman nor Scandinavian influences, Trips (2002) remarks that they were composed in an area where a 10th century Scandinavian settlement was located. Huber & Robertson (2016), however, point out that the literary language of the KG is to be considered the product of a very composite language community, where influences of five languages can be traced, namely English, French, Scandinavian, Welsh and Latin.

Summarizing, we have the South East Midlands (henceforth SEM) texts on the one hand, located in an area which was on the border of the Danelaw. The texts are of a composite and stratified nature, representing manipulations of older material. On the other hand, we have texts from the West Midlands; the texts present a qualitative difference: in fact, the LH are adaptations of older material, whereas the texts of the KG were composed directly in the eME period.

Turning to the selection of the sample, a non-finite verbal form is needed in order to control for V-to-T movement. The data were collected by querying the PPCME2 corpus (Kroch & Taylor 2000) through CorpusStudio (Komen 2011). The search produces both sentences with AuxV and VAux order; however, since in the eME sample the number of VAux clauses is really low,[13] the present paper deals with AuxV clauses.

The editions of the texts were consulted to scrutinize the context of the results, since it is decisive in determining the IS value of the constituents.

I selected direct, indirect and PP objects of verbs surfacing between the inflected and the non-inflected verb and after the non-inflected verb, whereas nominal parts of complex predicates and arguments of predicates such as *be called, be named* were left out of the analysis, since they are inert for IS (cf. Petrova & Speyer 2011). Moreover, it is reasonable to assume that they occupy a different position in syntactic structure (as assumed by Hinterhölzl with his PredP).[14] Adjuncts modifying the reference of the event are excluded from the analysis, since it is reasonable to assume that they occupy a projection at the left of the vP, where their one-word adverbial counterparts are located (cf. Cinque 1999). Finally, personal object pronouns are also left out, since they have been analysed as being optional clitics in Pintzuk (1999); moreover, De Bastiani & Hinterhölzl (2020) show that the position of object pronouns in OE and eME is subject to slightly different information structural conditions; their

---

13. From the elicitation of clauses with a subject, an object and a non-finite verbal form in the texts selected for this study, only 8 subordinate clauses have VAux order in the South East Midlands texts, whereas the sample of subordinate clauses from the *Lambeth Homilies* presents only one, and there are no VAux clauses in the texts from the Katherine Group.

14. Petrova (2009), moreover, argues that nominal parts of complex predicates do not have referential status, and are thus inert for information structure.

progressive post-verbal spell-out follows however from the predictions laid out in Hinterhölzl's framework (2014, 2015, 2017).[15]

Pintzuk & Taylor (2006) also showed that quantified and negated objects behave differently from other types of objects. Given the analyses by Fischer et al. (2000), Moerenhout & van der Wurff (2000, 2005), it is clear that these objects undergo a different licensing to dedicated positions at the left of the VP; Fischer et al. (2000) propose that negated objects undergo movement to the Specifier of NegPhrase, which is located above the VP and that quantified objects undergo quantifier raising to the left of the VP. Finally, the fact that these elements are among the last items presenting surface OV order until late Middle English might be taken as evidence for their being subject to different constraints with respect to non-negated and non-quantified constituents. These items were only included in the investigation of prosodic weight, but were not analysed for IS.[16]

In order to determine whether an object is part of a complex predicate, the Middle English Dictionary (McSparran et al. 2001; henceforth MED) was consulted. An example of a complex predicate is given in (4):

(4) … *þat alle men sholden* **deað** *þolien.*
     … that all   men should   death undergo
     "That all men should undergo the penalty of death."

                                                      (CMTRINIT-MX1,143.1932)

The MED entry for the verb *þolien* is, in fact, the following:

> To be made to undergo (a penalty for misdeeds); ~ ded (deth), undergo the penalty of death   (Permanent link to this entry http://quod.lib.umich.edu/cgi/m/mec/med-idx?type=id&id=MED45378)

The MED was also used to establish whether a PP was part of the argument structure of a verb or an adjunct. In (5), an example of argument PP is given:

(5) *and ic ðe   wile scilden* **fram alle euele**, ….
     And I   you will shield from all   evils ….
     "And I will shield you from all harms"   (CMVICES1-M1,87.990)

---

15. For a discussion on the finer IS distinctions coinciding with the mapping of pronouns and how their post-verbal spell-out is integrated within the present framework, cf. De Bastiani & Hinterhölzl (2020).

16. Quantified and negated objects are usually defined as inert with respect to information structure, therefore their inclusion in the IS analysis would be misleading. However, they have been included in the weight analysis in order to have a complete picture of the prosodic composition of the pre- and post-verbal fields.

The MED entry for the verb *scilden* is in fact the following:

> To protect (sb., oneself, sth.) from harm, save, defend; (b) ~ from (with), to protect (sb.) against (sb. or sth.), save (sb.) from (sb. or sth.).
>
> (Permanent link to this entry: https://quod.lib.umich.edu/m/middle-english-dictionary/dictionary/MED39836/track?counter=2&search_id=2862230)

The weight of the constituents is assigned according to the definition provided by Hinterhölzl. Right branching constituents, as well as non-branching constituents with a PP or relative clause post-modifier are analysed as heavy. Moreover, also co-ordinated objects are analysed as heavy. Given the fact that by the beginning of the ME period a definite determiner is fully grammaticalized (cf. Allen 2016), definite DPs are coded as heavy, since a DP with a filled D head and a nominal complement is a right branching constituent. Nouns modified by a demonstrative, or a possessive pronoun, are analysed as left branching, and hence light.[17]

For the IS analysis, a binary distinction into given and new is used, based on Lambrecht (1994), Birner (2006), and Krifka (2007). A constituent is coded as given, if it has been mentioned in the immediate previous context, or if it is accessible through encyclopaedic knowledge such as (6):

(6) *God cweđ þurh   þes witeġan   muđ   Þet he walde **his gast***
     God said through the prophet's mouth that he wanted his spirit
     asenden ofer mennesc flesc.
     send   over human flesh
     "God said through the prophet's mouth, that he wanted to send his spirit over human flesh."   (CMLAMBX1-MX1,91.757)

Since my dataset consists mostly of religious texts, entities pertaining to the Christian Faith, such as *God, Jesus, the devil, paradise* as in (6) are labelled as given, since it is reasonable to assume that these were accessible for the intended audience of the works under investigation (cf. also Struik & van Kemenade 2018). Elements which are labelled as new are those which are mentioned for the first time in the context under discussion and are not part of inalienable possession or encyclopaedic knowledge.

---

17. Giusti (2015) shows in fact that demonstratives and possessive adjectives occupy a specifier position in the extended projection of the DP. Only the definite determiner is analysed as a head in her framework.

(7) *hit itimode   efter noes   flode þet eontas walden areran **ane buruh***
     It   happened after Noah's flood that giants wanted erect a   city
     ***and anne stepel** swa hehne; þet his Rof astiġe   up to heofene*
     and a   tower so   high   that his roof ascended up to heaven
     "It happened after Noah's flood, that giants wanted to erect a city and a tower
     so high, that its roof ascended up to heaven."   (CMLAMBX1-MX1,93.814)

Bridging inferables, such as (8), are analysed according to their mention in the previous context.[18] In Example (8), the reference of the term *son* is not an item of inalienable possession, and is analysed as a bridging inferable by its being anchored to the referent *Martianus* through the possessive adjective:

(8) *Martianus hæfde **his sunu** ær   befæst   to woruldlicre lare and to*
     Martianus had   his son   before committed to worldy   lore and to
     uđwitegunge ….
     philosophy ….
     "Martianus had entrusted his son to the study of worldly lore and to philosophy."
     (coaelive,ÆLS_[Julian_and_Basilissa]:184.1049)

In this example, the bridging inferable *his son* is mentioned at this point in the narration. Even though its being scrambled across an adverbial may point at its being analysed as a given entity by the author of the work, I refrained from such speculations in the assignment of the IS category, and assigned the label "new". The following section presents the results.

4. Results

Summarizing the debate presented in § 2, I am adopting an antisymmetric account, in which IS and weight interface conditions are responsible for the spell-out of constituents before and after the non-finite verb. Recall that Struik & van Kemenade (2018) point to the conclusion that in OV structures, the object is marked for its givenness, whereas in VO structures, the object presents heterogeneous features. Moreover, we expect light elements to prefer pre-verbal spell-out, whereas heavy elements are expected to surface post-verbally.

In the following, the distribution of direct, indirect and PP objects of verbs in the sample, coded according to the methodology given in § 3, is presented. As a

---

18. In the case of bridging inferables, the elements cannot be inferred without an explicit link to a preceding referent. These are opposed to, i.e., elaborating inferables, such as body parts, which are inferred without a necessary link to a previous referent (cf. also Taylor & Pintzuk 2012a, b and Struik & van Kemenade 2018).

reviewer points out, PP objects have never been included in similar investigations before (but cf. Cloutier 2009 for an account of directional objects). For reasons of space, it is not possible to present the percentages for each type of object separately; generally, PP objects favor the post-verbal position, independently of their IS status, even though pre-verbal PP objects can be found. As the detailed qualitative analyses presented in De Bastiani (2019, 2020) show, PP objects behave much like direct objects. For an overview of the distribution of direct objects only, cf. De Bastiani (2019).

Table 1. Distribution of constituents in the eME sample (De Bastiani 2020: 105, 106, 119)

		SEM texts	LH	KG
Total number of VP arguments		143	105	31
Pre-verbal		60	30	11
Analysed for weight		60	30	11
	Light	43; 71.7%	20; 66.7%	8; 66.7%
	Heavy	17; 28.3%	10; 33.3%	3; 33.3%
Analysed for IS		49	26	10
	Given	45; 91.8%	25; 96.2%	10; 100%
	New	4; 8.2%	1; 3.8%	0; 0%
Post-verbal		83	74	20
Analysed for weight		83	74	20
	Light	31; 37.4%	38; 51.4%	10; 50%
	Heavy	52; 62.6%	36; 48.6%	10; 50%
Analysed for IS		71	68	18
	Given	40; 56.3%	39; 57.4%	10; 55.5%
	New	31; 43.7%	29; 42.6%	8; 44.5%

As can be seen from Table 1, the pre-verbal constituents are in the majority of cases given, and about 70% of them consist of light constituents in the texts examined. The post-verbal domain hosts heterogenous constituents. In the post-verbal domain, in fact, one can find both given and new elements, with no marked preference for one IS category over the other, as well as light and heavy elements. Although we have to leave a statistical analysis to future research, these results are in line with Taylor & Pintzuk (2011, 2012a, b), Struik & van Kemenade (2018).

Similar conclusions can be drawn from the composition of the constituents in the pre-verbal domain of the LH. They are predominantly given, but there is also a 33% of them which consists of heavy constituents. The post-verbal domain presents a heterogeneous distribution of given, new, light and heavy objects. Even though the dataset from the KG is smaller, it shows a similar distribution of constituents with respect to the other texts presented in Table 1.

From the examination of non-pronominal objects in the selected texts, we can conclude that there is a continuity with respect to the OE period, as far as the mapping of pre-verbal elements is concerned. In fact, these are for the most part given. There are some new and pre-verbal constituents in the SEM texts and in the LH. These constitute either bridging inferables, or entities which are mentioned in the narration at the point in which they surface in our result sentence, but that can be linked to encyclopaedic knowledge possessed at the time, as in (9):

(9) *For ði   ðat tu   hauest **þine sennen and ðine euele þeawes** forlaten*
    Because that you have   your sins   and your evil habits left
    "Because you have relinquished your sins and your evil habits"
    (CMVICES1-M1,87.989)

This piece of direct speech is uttered by God; there is no mention in the previous context of sins and evil habits, and these constitute bridging inferables. Therefore, the referent is labelled as new. However, it can be argued that, given that this text is a treaty on sins and virtues, the fact that everyone has sins is part of the common ground shared by the author and the intended audience of the text.

In (10), the bridging inferable *his neighbour* is mentioned at this point in the narration and is accordingly labelled as new. However, one can argue that, from the perspective of the Christian faith, everyone has a neighbour, and this referent can be analysed as an elaborating inferable and hence given.

(10) *on þa oðre souen laȝe weren alswa sunderliche inna oðre stanine table*
    and the other seven laws were also   separately in   other stone table
    *brede   hu uwilc mon scal **his euenexta** beodan alswa he walde þet*
    educating how each man shall his neighbour act   as -so he wants that
    Me him bude.
    one him act
    "and the other seven laws were also written separately in the second stone table, instruction how each man shall act towards his neighbour, as he wants that people act towards him." (CMLAMBX1-MX1,13.119)

Table 1 shows that heavy elements can be mapped in pre-verbal position; these are, however, for the most part given. The distribution of constituents in the post-verbal domain shows that light elements are also found there, as are given ones. In order to have a clear picture of how given, new, light and heavy elements are distributed across the pre- and post-verbal domain, let us observe the following tables. Table 2 presents the distribution of constituents across the pre- and post-verbal domain in a pilot sample of OE subordinate clauses with AuxV order, as a standard for comparison with the eME texts; this sample of subordinate clauses was collected by querying the YCOE corpus (Taylor et al. 2003) for clauses with a subject, an

Table 2. Distribution of constituents in the OE sample (De Bastiani 2020: 71)

Arguments		101
Pre-verbal		38
Analysed for weight		38
	Light	29; 76.3%
	Heavy	9; 23.7%
Analysed for IS		29
	Given	28; 96.5%
	New	1; 3.5%
Post-verbal		63
Analysed for weight		63
	Light	30; 47.6%
	Heavy	33; 52.4%
Analysed for IS		49
	Given	26; 53.1%
	New	23; 46.9%

object and a complex verbal form. The results are sampled as to cover the whole OE period. The sample of AuxV clauses amounts to 101 sentences.

Table 2 suggests that the pre-verbal domain already hosts predominantly given and light elements in the OE period. The expectation is that there is a general decrease in the number of constituents in the pre-verbal domain in ME, so the following tables present the relative distribution of elements across the pre- and post-verbal domain, starting from the OE pilot sample reported in Table 2:

Table 3. Distribution of IS and Weight value in the OE sample (De Bastiani 2020: 73)

Arguments			Pre-verbal	Post-verbal
Weight value				
	Light elements	59	29; 49.2%	30; 50.8%
	Heavy elements	42	9; 21.4%	33; 78.6%
IS value				
	Given elements	55	28; 50.9%	27; 40.1%
	New elements	24	1; 4.2%	23; 95.8%

This table shows that given and light elements are distributed almost equally across the pre- and post-verbal domain, whereas the rates of heavy and new elements are much lower in the pre-verbal domain; such elements are spelled out in the post-verbal domain in the majority of cases. These data tell us that there is a preference for the pre-verbal domain to host light and given elements; this does not mean, however, that these elements are banned from the post-verbal domain.

Table 4. Distribution of IS and Weight value in the eME texts
(De Bastiani 2020: 108, 120)

Arguments		SEM		LH		KG	
		Pre-V	Post-V	Pre-V	Post-V	Pre-V	Post-V
Weight value							
	Light elements	43; 58.1%	31; 41.9%	20; 34.5%	38; 65.5%	8; 44.4%	10; 55.5%
	Heavy elements	17; 24.6%	52; 75.4%	10; 21.8%	36; 78.2%	3; 23.1%	10; 76.9%
IS value							
	Given elements	45; 52.9%	40; 47.1%	25; 39.1%	39; 60.9%	10; 50%	10; 50%
	New elements	4; 11.4%	31; 88.6%	1; 3.3%	29; 96.7%	0; 0%	8; 100%

Let us now turn to the eME data in Table 4. The results from the SEM texts pattern with the OE sample. The distribution of new and heavy elements in the post-verbal domain is more restricted, but light and given elements are distributed almost equally in pre- and post-verbal position. The LH present a distribution that can be interpreted as a further path along the reanalysis of VO as the basic word order. Whereas the frequencies of given and light constituents in the pre-verbal domain are higher than new and heavy ones, note that their percentage in the pre-verbal domain is lower both with respect to the OE sample, as well as with respect to the SEM texts. The distribution of constituents from the KG patterns with the SEM texts and the OE sample. However, the differences in percentages between the three texts may be due to the different size of the sample. We can conclude that there is a tendency across time to progressively spell out constituents in the post-verbal position, even though there are differences across the dialectal areas.

The data examined so far show that there is a continuity with the OE period, and that the postulation of the G-Transparency condition accounts for the cases of pre-verbal constituents. Moreover, the data show that the majority of heavy elements are found in post-verbal position.

The data examined so far include non-pronominal objects. When the distribution of personal object pronouns is considered, interesting differences arise between the data.

The data in Table 5 were collected by recording the number of object pronouns in the sentences of the dataset. As can be seen, in the SEM texts and in the LH the number of post-verbal pronouns is around 10%. In other words, object pronouns favour the pre-verbal position or the position to the immediate left of the auxiliary.

Table 5. Distribution of object pronouns in the dataset
(De Bastiani & Hinterhölzl 2020: 19–20)

	SEM	LH	KG
Pre-T	65/123; 52.8%	20/43; 46.5%	9/32; 28.1%
Pre-V	42/123; 34.2%	20/43; 46.5%	9/32; 28.1%
Post-V	16/123; 13%	3/43; 7%	14/32; 43.8%

If the data from Table 5 are combined with the results from Table 1, it can be concluded that the SEM texts and the LH represent the intermediate stage sketched in § 2, namely a stage in which the elements mapped in pre-verbal position are predominantly light. The LH represent probably a more advanced stage along the reanalysis of VO, since the non-pronominal elements mapped in pre-verbal position exhibit a lower frequency compared to the OE sample and the SEM texts, leaving object pronouns as the typical pre-verbal elements. The next stage in the reanalysis of VO is represented by the KG, where object pronouns are progressively being spelled out in post-verbal position. Finally, from the analysis conducted so far, it emerges that there is a uniform development from the OE period and that the language change scenario sketched in § 2 can account for the differences in mapping of constituents in the different texts. It should be borne in mind, however, that the more conservative character of the SEM texts and the LH can also be ascribed to their being a manipulation of older material. In fact, since the LH and the texts of the KG come from the same dialectal area, one would expect a similar distribution of constituents. While the distribution of non-pronominal constituents is similar, the mapping of object pronouns yields a different picture. The syntax of the texts from the KG shows that these texts were further along the path of reanalysing VO as the basic word order. The fact that these texts, directly composed in the eME period, present a more innovative syntax, is not surprising when one considers that the other texts in the sample constitute copies of older material, a factor which has not received due attention in previous investigations.

5.   Conclusion

In this paper, I have presented empirical data from eME subordinate clauses about the interaction of IS, weight and mapping of the constituents. The aim was twofold: (1) to show that there is a uniform development with respect to the OE period, and (2) to show that transmission history of the texts examined can in part account for the different distributions of constituents in the dataset. In fact, the texts of the LH and of the KG, although localized in the same area, show differences in the mapping of constituents when object pronouns are added to the picture. On the other hand,

the mapping of object pronouns in the SEM texts and the LH, both texts being manipulations of older materials but coming from two different areas, is very similar.

As far as the theoretical framework employed is concerned, I suggested in § 2 that the influence of the F-Transparency condition, originally assumed by Hinterhölzl (2014, 2015, 2017), needs to be reconsidered, in light of the data presented by Struik & van Kemenade (2018) and De Bastiani (2017, 2020). In fact, what emerges from the data is that pre-verbal elements are clearly marked for their given status, but post-verbal elements are not marked either for their new IS status, or for their weight. It is therefore reasonable to assume that it is the G-Transparency condition that requires the spell-out of the higher copy of movement; moreover, light elements are also spelled out by default in the pre-verbal domain, given that the prosodic condition applies to branching objects. The analysis of the eME data points to the fact that these conditions were still in effect for the pre-verbal spell-out of constituents, but that the G-Transparency condition was probably subject to blurring, since in a text such as the LH, the vast majority of non-pronominal objects is mapped in post-verbal position. The condition requiring light elements to be spelled out in the pre-verbal domain is the one responsible for the spell-out of object pronouns, a condition which is ultimately lost with the progressive spell-out of all types of objects in the post-verbal domain, as the data from the KG suggest.

One last point that needs to be addressed is the persistence of OV orders throughout the late Middle English period; as was suggested in §§ 2–3, the last elements presenting surface OV order are quantified and negated objects. For these, a syntactic account assuming leftward movement to dedicated positions was put forth by Fischer et al. (2000). As Fischer et al. (2000) point out, the shift from surface OV to surface VO order takes place around the twelfth century, but it is reasonable to assume that the reanalysis of post-verbal spell-out as the unmarked option for all types of objects might have proceeded at different rates. The analysis presented here shows that non-pronominal objects in eME present post-verbal spell-out in the majority of cases, but as Table 5 shows, pronouns are mostly spelled out in pre-verbal position in the texts composed earlier. This piece of evidence is interpreted in the framework adopted as an indication for the blurring of the G-Transparency condition, since the distribution of objects (both pronominal and non-pronominal) is regulated by their syntactic weight. With the progressive spell-out of also pronominal objects in the post-verbal domain, evidence for the prosodic mapping of constituents is lost as well, driving the reanalysis of post-verbal spell-out as the unmarked option. The reanalysis of post-verbal spell-out as the unmarked option applies first to non-pronominal objects, and later on to pronominal objects. It is therefore reasonable to assume that post-verbal spell-out for different types of objects undergoes reanalysis at different rates, depending on the nature of the triggers prompting leftward movement and pre-verbal spell-out.

The studies of Foster & van der Wurff (1995) and Moerenhout & van der Wurff (2000, 2005) show that OV orders could indeed be found in late ME, but the instances in poetic texts outnumber by far those in prose, and by the end of the ME period, they are restricted to specific syntactic requirements. There is therefore little evidence for leftward movement of non-negated and non-quantified objects, which are probably subject to stylistic reasons and/or restricted to marked syntactic contexts. In other words, the contexts in which leftward movement is still productive by the end of the ME period are driven by clear syntactic requirements, involving the checking of features, whereas evidence for leftward movement and pre-verbal spell-out of other types of object had probably become scant. If spell-out is indeed reanalysed as the unmarked option, it follows that checking is realised covertly. As already stated above, there is very little evidence for leftward movement and pre-verbal spell-out, so that it can be assumed that negated and quantified objects also undergo post-verbal spell-out in analogy with the other types of objects, albeit later, since the trigger for leftward movement is probably strong enough to resist reanalysis the longer. These assumptions open up interesting avenues for future research on the interaction of pre- and post-verbal spell-out in Late ME with respect to different types of objects.

Acknowledgements

Sources

D'Ardenne, Simonne T. R. O. (ed.) 1961. *Þe liflade and te passiun of Seinte Juliene*. Oxford: Oxford University Press.

D'Ardenne, Simonne T. R. O. 1977. *The Katherine Group edited from ms. Bodley 34*. Bibliothèque de la Faculté de philosophie et lettres de l'Université de Liège fasc. 215. Paris: Société d'Edition Les Belles Lettres.

Hall, Joseph (ed.) 1972. *Selections from Early Middle English 1130–1250*, 2nd edn. (1st edn. 1920). Oxford: Clarendon.

Holthausen, Ferdinand. 1888. *Vices and virtues. Part 1*. EETS O.S. 89. London: Trübner.

Huber, Emily R. & Elizabeth Robertson (eds.). 2016. *The Katherine Group: MS Bodley 34*. Kalamazoo, Michigan: Medieval Institute Publications (Robbins Library Digital Projects). http://d.lib.rochester.edu/teams/publication/huber-and-robertson-the-katherine-group

Komen, Erwin. 2011. *CorpusStudio*, Nijmegen: Radboud University. http://erwinkomen.ruhosting.nl/software/CorpusStudio/

Kroch, Anthony & Ann Taylor. 2000. *The Penn-Helsinki parsed corpus of Middle English (PPCME2)*. Department of Linguistics, University of Pennsylvania. CD-ROM, 2nd edn. http://www.ling.upenn.edu/hist-corpora/

Mc Sparran, Frances, Paul Schaffner, John Latta, Alan Pagliere, Chris Powell & Matt Stoeffler (eds.). 2001. *Middle English dictionary*, University of Michigan Digital Library Production Service, https://quod.lib.umich.edu/m/med/

Morris, Richard (ed.). 1969. *Old English homilies and homiletic treatises. Part I*. EETS O.S. 29, 34. New York: Greenwood Press. Originally published by Trübner (London, 1868).

Morris, Richard (ed.). 1873. *Old English homilies of the twelfth century. Second series*. EETS O.S. 53. London: Trübner.

Taylor, Ann, Anthony Warner, Susan Pintzuk & Frank Beths. 2003. *The York–Toronto–Helsinki parsed corpus of Old English prose*. York, UK: Department of Language and Linguistic Science, University of York. Available through the Oxford Text Archive.

References

Allen, Cynthia. 2016. The definite determiner in Early Middle English: What happened with þe?. In Sten Vikner, Henrik Jørgensen & Elly van Gelderen (eds.), *Let us have articles betwixt us – Papers in historical and comparative linguistics in honour of Johanna L. Wood*, 43–82. Aarhus: Department of English, School of Communication & Culture, Aarhus University.

Bech, Kristin. 2001. Word order patterns in Old and Middle English: A syntactic and pragmatic study. Bergen: University of Bergen PhD dissertation. https://bora.uib.no/handle/1956/3850 (26 February, 2021).

Behaghel, Otto. 1932. *Deutsche Syntax*, Band 4. Heidelberg: Carl Winters Universitätsbuchhandlung.

Biberauer, Theresa & Ian Roberts. 2005. Changing EPP parameters in the history of English: Accounting for variation and change. *English Language and Linguistics* 9(1). 5–46. https://doi.org/10.1017/S1360674305001528

Birner, Betty J. 2006. Inferential relations and noncanonical word order. In Betty J. Birner & Gregory L. Ward (eds.), *Drawing the boundaries of meaning: Neo Gricean studies in pragmatics and semantics in honor of Laurence R. Horn*, 31–51. Amsterdam: John Benjamins. https://doi.org/10.1075/slcs.80.04bir

Cinque, Guglielmo. 1999. *Adverbs and functional heads: A cross-linguistic perspective*. Oxford: Oxford University Press.

Cloutier, Robert A. 2009. *West Germanic OV and VO: The status of exceptions*. Utrecht: LOT

De Bastiani, Chiara. 2017. Prosody and information structure in Old English: Testing the framework. *Linguistica Atlantica* 36(2). 57–97

De Bastiani, Chiara. 2019. An interface based account for the grammaticalization of the VO word order in the history of English. Venice: Ca' Foscari University of Venice PhD dissertation.

De Bastiani, Chiara. 2020. *Verb and object order in the history of English: A language-internal account*. Newcastle upon Tyne: Cambridge Scholars Publishing.

De Bastiani, Chiara & Roland Hinterhölzl. 2020. On the syntax of object pronouns in Old English and Early Middle English. *Glossa: A Journal of General Linguistics* 5(1). 43. https://doi.org/10.5334/gjgl.890

Elenbaas, Marion & Ans van Kemenade. 2014. Verb particles and OV/VO in the history of English. *Studia Linguistica* 68(1). 140–167. https://doi.org/10.1111/stul.12018

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

Fischer, Olga, Ans van Kemenade, Willem Koopman & Wim van der Wurff. 2000. *The syntax of early English*. Cambridge: Cambridge University Press.

Foster, Tony & Wim van der Wurff. 1995. The survival of object-verb order in Middle English: Some data, *Neophilologus* 79(2). 309–327. https://doi.org/10.1007/BF00999787

Fuß, Eric & Carola Trips. 2002. On the validity of the Double Base Hypothesis. *Journal of Comparative Germanic Linguistics* 4. 171–224. https://doi.org/10.1023/A:1016588502117

Giusti, Giuliana. 2015. *Nominal syntax at the interfaces: A comparative study of languages with articles*. Newcastle upon Tyne: Cambridge Scholars Publishing.

Haeberli, Eric. 2008. Review of Roberts, Ian. 2007. *Diachronic syntax*. Oxford: Oxford University Press. *Lingua* 118. 1998–2007.

Hinterhölzl, Roland. 2014. On the interaction between syntax, prosody and information structure: An interface approach to word order developments in Germanic. In Kristin Bech & Kristine Gunn Eide (eds.), *Information structure and syntactic change in Germanic and Romance languages*, 341–376. Amsterdam: John Benjamins. https://doi.org/10.1075/la.213.13hin

Hinterhölzl, Roland. 2015. An interface account of word-order variation in Old High German. In Theresa Biberauer & George Walkden (eds.), *Syntax over time*, 299–318. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199687923.003.0018

Hinterhölzl, Roland. 2017. From OV to VO in English: How to Kroch the nut. In Bettelou Los & Pieter de Haan (eds.), *Word order change in acquisition and language contact: Essays in honour of Ans van Kemenade*, 9–34. Amsterdam: John Benjamins.
https://doi.org/10.1075/la.243.02hin

Hinterhölzl, Roland, Svetlana Petrova & Michael Solf. 2005. Diskurspragmatische Faktoren für die Topikalität und Verbstellung in der althochdeutschen Tatianübersetzung. In Ishihara, Shinichiro, Michaela Schmitz & Anne Schwarz (eds.), *Approaches and findings in oral, written and gestural language* (Interdisciplinary Studies on Information Structure 3), 143–182. Potsdam: Universitätsverlag Potsdam.

Hinterhölzl, Roland & Svetlana Petrova. 2010. From V1 to V2 in West Germanic. *Lingua* 120. 315–328. https://doi.org/10.1016/j.lingua.2008.10.007

Hinterhölzl, Roland & Svetlana Petrova. 2018. Prosodic and information structural factors. In Agnes Jäger, Gisella Ferraresi & Helmut Weiß (eds.), *Clause structure and word order in the history of German*, 277–288. Oxford: Oxford University Press.

Hróarsdóttir, Þorbjörg. 2000. *Word order change in Icelandic: From OV to VO*. Amsterdam: John Benjamins. https://doi.org/10.1075/la.35

van Kemenade, Ans. 1987. *Syntactic case and morphological case in the history of English*. Dordrecht: Foris. https://doi.org/10.1515/9783110882308

van Kemenade, Ans. 2009. Discourse relations and word order change. In Roland Hinterhölzl & Svetlana Petrova (eds.), *Information structure and language change*, 91–118. Berlin: Mouton de Gruyter.

van Kemenade, Ans & Bettelou Los. 2006. Discourse adverbs and clausal syntax in Old and Middle English. In Ans van Kemenade & Bettelou Los (eds.), *The handbook of the history of English*, 224–278. Oxford: Blackwell. https://doi.org/10.1002/9780470757048.ch10

van Kemenade, Ans & Marit Westergaard. 2012. Syntax and information structure: Verb second variation in Middle English. In Anneli Meurman-Solin, María José López-Couso & Bettelou Los (eds.), *Information structure and syntactic change in the history of English*, 87–118. New York: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199860210.003.0005

Krifka, Manfred. 2007. Basic notions of information structure. In Caroline Féry, Gisbert Fanselow & Manfred Krifka (eds.), *Interdisciplinary studies on information structure*, Volume 6, 13–55. Potsdam: Potsdam Universitätsverlag.

Kroch, Anthony & Ann Taylor. 2000. Verb-object order in Middle English. In Susan Pintzuk, George Tsoulas & Anthony Warner (eds.), *Diachronic syntax: Models and mechanisms*, 132–163. Oxford: Oxford University Press.

Lambrecht, Knud. 1994. *Information structure and sentence form: Topic, focus and the mental representations of discourse referents*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511620607

Linde, Sonja. 2009. Aspects of word order and information structure in Old Saxon. In Roland Hinterhölzl & Svetlana Petrova (eds.), *Information structure and language change*, 367–389. Berlin: Mouton de Gruyter.

Los, Bettelou & Ans van Kemenade. 2018. Syntax and the morphology of deixis: The loss of demonstratives and paratactic clause linking. In Marco Coniglio, Andrew Murphy, Eva Schlachter & Tonjes Veenstra (eds.), *Atypical demonstratives*, 127–157. Berlin: Mouton de Gruyter.

Milićev, Tanya. 2016. Syntax and information structure of the Old English VP. Novi Sad: University of Novi Sad PhD dissertation.

Moerenhout, Mike & Wim van der Wurff. 2000. Remnants of the old order: OV in the Paston Letters. *English Studies* 81(6). 513–530. https://doi.org/10.1076/enst.81.6.513.9184

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. https://doi.org/10.1017/S1360674305001553

Petrova, Svetlana. 2009. Information structure and word order variation in the Old High German Tatian. In Roland Hinterhölzl & Svetlana Petrova (eds.), *Information structure and language change*, 251–279. Berlin: Mouton de Gruyter.

Petrova, Svetlana & Augustin Speyer. 2011. Focus movement and focus interpretation in Old English. In Edgar Onea & Malte Zimmermann (eds.), *Focus marking strategies and focus interpretation, Lingua* Special Issue 121 (11). 1751–1765. https://doi.org/10.1016/j.lingua.2011.06.004

Pintzuk, Susan. 1999. *Phrase structures in competition: Variation and change in Old English word order*. New York: Garland.

Pintzuk, Susan & Ann Taylor. 2006. The loss of OV Order in the history of English. In Ans van Kemenade & Bettelou Los (eds.), *The Handbook of the history of English*, 249–278. Oxford: Blackwell. https://doi.org/10.1002/9780470757048.ch11

Roberts, Ian. 1997. Directionality and word order change in the history of English. In Ans van Kemenade & Nigel Vincent (eds.), *Parameters of morphosyntactic change*, 397–426. Cambridge: Cambridge University Press.

Sisam, Celia. 1951. The scribal tradition of the Lambeth Homilies. *The Review of English Studies* 2(6). 105–113. https://doi.org/10.1093/res/II.5.105

Struik, Tara & Ans van Kemenade. 2018. On the givenness of OV word order: A (re)examination of OV/VO variation in Old English. *English Language and Linguistics* 24. 1–22. https://doi.org/10.1017/S1360674318000187

Trips, Carola. 2002. *From OV to VO in Early Middle English*. Amsterdam: John Benjamins. https://doi.org/10.1075/la.60

Taylor, Ann & Susan Pintzuk. 2011. The interaction of syntactic change and information status effects in the change from OV to VO in English. *Catalan Journal of Linguistics* 10. 71–94. https://doi.org/10.5565/rev/catjl.61

Taylor, Ann & Susan Pintzuk. 2012a. Verb order, object position and information status in Old English. *York Papers in Linguistics Series* 2. 29–52.

Taylor, Ann & Susan Pintzuk. 2012b. The effect of information structure on object position in Old English: A pilot study. In Anneli Meurman-Solin, María José López-Couso & Bettelou Los (eds.), *Information structure and syntactic change in the history of English*, 47–65. New York: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199860210.003.0003

Taylor, Ann & Susan Pintzuk. 2015. Verb order, object position and information status in Old English. In Theresa Biberauer & George Walkden (eds.), *Syntax over time: Lexical, morphological, and information-structural interactions*, 318–335. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199687923.003.0019

Wallenberg, Joel C. 2009. Antisymmetry and the conservation of C-Command: Scrambling and phrase structure in synchronic and diachronic perspective. Philadelphia: University of Pennsylvania PhD dissertation. (Publicly Accessible Penn Dissertations 77). https://repository.upenn.edu/edissertations/77

CHAPTER 6

The role of (the avoidance of) centre embedding in the change from OV to VO in English

Rodrigo Pérez Lorido
University of Oviedo

Authors from Colman (1988) to Ogura (2001, 2004) have discussed the role of Relative Clause Extraposition as a means to avoid centre-embedded structures in Old English, and its relevance as a factor in the change from OV to VO word order, but to date no study has addressed the full extent to which clausal embedding is allowed in Old English, the contexts for its application or its quantitative limits. Based on an extensive corpus analysis, this paper provides evidence about the factors (both language internal and external) that interact in the generation of centre-embedded structures in Old English while providing a reassessment of the role the avoidance of this kind of structures played in the change from OV to VO.

Keywords: Change OV to VO, language processing, clausal embedding, centre-embedded relative clauses, Relative Clause Extraposition, dependency length minimisation theories, Dependency Locality Theory, Early Immediate Constituents, Old English, German

1. Introduction

This paper is about the role centre-embedded structures, like the sentence in (1), played in the change from OV to VO word order in English.

(1) *gif* **se** *þegen* <u>þone</u> <u>þræl</u>[1] *[þe he ær ahte] fullice* **afille** …
      if the thane the   slave that he before owned shamefully kills
      "if the thane shamefully kills the slave that he had previously owned … " [1]
                                                                (cowulf,WHom_20:2.98.1706)

---

1. Throughout this paper preverbal head objects of the superordinate clauses are underlined, and their modifying relative clauses (whether embedded or extraposed) placed inside square brackets. The subjects and verbs of the superordinate clauses appear in bold type.

Structures like these, in which a preverbal head object is modified by a relative clause, are difficult to process for several reasons. Firstly, the inherent 'heaviness' (in terms of 'size', length …) of the clausal object generates long, disrupted syntactic dependencies between the subject and the verb (and probably between the object itself and the verb), and secondly, this configuration creates discontinuous relationships between arguments and predicates going from both edges inward, which cause the memory load at the most embedded noun to be greater that the capacity of most people's working memory. Considering the first point, clearly as objects become longer in SOV structures, the subject-verb dependency will become longer, requiring increased effort from short-time memory, and straining the mental processor. Therefore, dependency length minimisation theories such as Gibson's (1998, 2000) Dependency Locality Theory or Hawkins's (1994, 2007) Early Immediate Constituents – which postulate that words are ordered to make dependencies between them as short as possible – will predict that 'heavy' objects will be postponed, giving way to VO structures.

To avoid the problems long, disrupted dependencies cause to the human mental processor, two solutions stand out cross-linguistically. The first one is postposing part or the whole of the heavy preverbal argument to a position after the verb, either as Relative Clause Extraposition (2a) or moving the entire object (head and RC) to the right (Heavy-NP Shift) (2b)). This way, syntactic dependencies (subject-verb/ object-verb) are reduced and nested dependencies are transformed into concatenated, consecutive ones, rendering more processing-optimal strings:

(2) a. #*… gif se þegen þone þræl [ ] fullice afille [þe he ær ahte]*
                                                              (RC Extraposition)
     b. #*… gif se þegen fullice afille þone þræl [þe he ær ahte]]*   (Heavy-NP Shift)

This is a pervasive phenomenon, which is attested in most SOV languages with postnominal relatives, such as Dutch or German (see the examples in (3) and (4)):

(3) a. (ʔ) *dat **Jan** het meisje [dat hij kuste] **kent**.*
             that Jan the girl   that he kissed knows
             " *… that Jan knows the girl that he kissed*"   (RC Embedding)
     b. *dat **Jan** het meisje [ ] **kent** [dat hij kuste].*
             that Jan the girl   knows that he kissed   (RC Extraposition)

(4) a. *dass **Peter** das Buch, [das Maria gekauft hat,] bald **liest**.*
             that Peter the book that Maria bought   has soon reads
             " *… that Peter will soon read the book that Maria bought*"
     b. *dass **Peter** das Buch [ ] bald **liest**, [das Maria gekauft hat].*

Another grammatical resource which reduces the long, disrupted S-V/O-V dependencies created by centre-embedded object relatives is the use of recapitulative elements,[2] whose presence is particularly striking in Old English. Different degrees of recapitulation (from less to more explicit) can be found in OE, including recapitulative subjects (*hi(e)* in (5a) below), combinations of recapitulative subject + object (*ic hi* in (5b)) and even recapitulative subordinator + subject + object (*ðæt we ða* in (5c)), whenever complex nesting of clauses is involved:

(5) a. *ðæt* **hie** … *ðæt geswinc ðæra costnunga*   [*ðe hi   ðrowað*] *hi(e)*
          that they   the labour of-the-temptations that they endure   they
          forsion &   geðolien.
          despise and suffer
          "that they … despise and and suffer the labour of the temptations they endure"
          (cocuraCP: 52.407.27.2807)

     b. *þæt* **ic** *his sceal her fela   oferhebban, &   þa spell* [*þe ic secge*]
          that I of-it shall here much omit   and the story that I   tell
          ic hi **sceal gescyrtan.**
          I   it shall shorten
          "that I must here pass over much, and must shorten the story which I tell"
          (coorosiu, Or_1:8.27.22.536)

     c. *ðæt* **we** *eac sumæ bec,   [ða ðe niedbeðearfosta sien eallum monnum*
          that we also certain books which needful   are for.all men
          to wiotonne], ðæt we ða   on ðæt geðiode **wenden**   *ðe   we*
          to know   that we those in the language translate-SBJV that we
          ealle gecnawan mægen,
          all   know   may
          "that we also translate certain books which are most needful for all men to know in the language which we can all understand"
          (coprefcura,CPLetWærf:49.29)

Returning to our central topic, the role of processing (associated to the avoidance of 'heavy' preverbal arguments) has a long pedigree in the literature on word order change in English and has often been put forward as a major causal factor in the change from OV to VO (Strang 1970; Stockwell 1977; Traugott 1992). On some occasions the focus has been placed on the avoidance of centre embedding itself (Colman 1988; Ogura 2001, 2004), hinting sometimes at a certain impairment in the cognitive capacities of the speakers of Old English and at their inability to

---

2.   Some theorists might argue, on the contrary, that including the recapitulative pronoun in fact increases the overall dependency length of the sentence (cf. Tily 2010: 62–63). My view, however, is that recapitulation effectively does away with unresolved dependencies, and that examples containing resumptive pronouns do not qualify as examples of genuine centre embedding.

process complex syntactic structures (O'Neil 1976: 53; Mitchell 1985: 616). A radical position in this respect is Ogura (2001, 2004), for whom processing issues and – specifically – the avoidance of centre embedding was not just one more factor but the trigger of the change OV > VO in English. Ogura (2001: 238) suggests that relative clauses (the commonest type of centre-embedded clause cross-linguistically) originated in English as separated from their antecedents ((6a) below) and, in a later stage, the head object noun was pulled to a position before the relative clause, producing the VO pattern (6b). The head noun and the relative could not be united (according to Ogura) by moving the relative back after the head noun (6c), because that would produce perceptually problematic centre-embedded structures.

(6)   a.   *&   hi   þær   manega beboda setton [þe   to Christendome belimpað]*
                             O   V   RC[3]
           and they there many   canons set   which to Christianity   concern
           "and they there laid down many canons relating to Christianity" [3]
                                       (ID cochronE,ChronE_[Plummer]:1102.15.3380))
      b.   *& hi þær setton manega beboda [þe to Christendome belimpað]*
                         V   O   RC
      c.   *& hi þær manega beboda [þe to Christendome belimpað] setton*
                     O   RC   V

                                          [Examples from Ogura 2001: 245]

According to Ogura, therefore, RCs in the history of English were not extraposed from a position within the sentence brace to a position outside it but *originated* outside the sentence brace, with the integration of the RC and the preverbal object never – or very rarely – occurring during the Old English period. This hypothesis has never really been seriously challenged to date,[4] but I see a number of weak points in it:

–   As Cloutier (2009: 103) suggests, it seems unlikely that a small subset of preverbal ('heavy') objects (those with attached RCs) was enough to instantiate the OV to VO change alone.

---

3.   Henceforth RC stands for 'relative clause'.

4.   Ogura's position is occasionally mentioned in the literature on language change in English (e.g. Hogg & Denison 2006: 187), but no critical voices have been raised against it, with the exception perhaps of Cloutier (2009) and Fischer et al. (2017). Cloutier (2009: 103) points out that centre-embedded structures represent too specific a context to have been the trigger of the change from OV to VO in English, and Fischer et al. (2017) – while admitting that the postposition of heavy objects obviously favoured the change OV > VO – also indicate that OV languages always have alternatives for arranging sequences other than centre embedding, so they consider it doubtful " … that processing difficulties necessitated the OV-to-VO shift in English" (p. 196).

–   The existence of complex, embedded relative clauses is attested in some of the earliest Germanic languages. (See Hock 1988, 1991 for Gothic).
–   Ogura's conclusions stem from a poorly designed database:
    –   Her corpus comprises only one text (the *Peterborough Chronicle*). If the dataset were extended (specially covering earlier periods of OE) different conclusions might be arrived at.
    –   She includes examples of recapitulation in her database (in fact, her only examples of centre-embedded relatives are examples of recapitulation).

Therefore, to validate Ogura's hypothesis and – above all – to try to shed some light on the role of performance in the change from OV to VO in English, I decided to carry out a systematic analysis of object relative clauses in Old English, inasmuch as they may appear in centre-embedded position within the clause, according to the following research plan: (i) Check the incidence of centre-embedded object relatives against extraposed ones in a large corpus of OE texts. (ii) Map the results against different variables of both a syntactic and a pragmatic nature, including variation over time, and style. (iii) Examine the interaction of the different structural factors that trigger centre-embedded structures in OE and other Germanic languages, insofar as they may provide evidence about underlying SOV/SVO order. (iv) Check up the limits of central recursion in Old English. (v) Provide new empirical evidence about the true role of language processing in the change from OV to VO in English.

2.   Data, corpus and data retrieval

2.1   Database

For the definition of 'object' I restricted my search to NP direct objects, regardless of their case-marking characteristics, though the vast majority of them are in the accusative case. As for verbal position, since my analysis focuses on centre-embedded structures, I concentrated on v-final clauses, that is, the only ones capable of producing this type of configuration. In the cases of extraposition, verbal position is defined with regard to the position of the object head. I considered structures with one or two verbal forms (v = finite verb; V = non-finite verb) in the patterns indicated below, which represent the only possible arrangements in my corpus when objects are modified by relative clauses[5] (see patterns and illustrative examples below):

---

5.   Other patterns of embedding, though not present in my corpus, are actually attested in OE, such as SvORCV. Pintzuk & Taylor (2006: 254) report 13 instances of this pattern in their study, whose corpus comprises the entire YCOE for the Old English period.

Verbal Position: <u>1 verb</u>   <u>2 verbs</u>
                    SORCv   SORCVv / SORCvV [embedding]
                    SOvRC   SOVvRC / SOvVRC [extraposition]
                    SvOVRC

Embedding
<u>1 verb</u> SORCv

(7) *butan <u>þu</u> … <u>þone soþan God</u> [þe þe   gesceop]* **oncnawest***, ne miht*
     unless you   the   true   God who you created   recognise   not might
     þu   hæle   habban
     you healing have
     "unless you acknowledge the true God who created you, you cannot have heal-
     ing"   (coaelive,ÆLS_[Sebastian]:227.1348)

<u>2 verbs</u> SORCVv

(8) *Hwa is þæt þe   <u>eall ða</u> <u>yfel</u> [þe hi   donde wæron]* **asecgean mæge**
     who is that who all the evil that they doing were   tell   may
     oððe areccean?
     or   narrate
     "who is that who may tell or narrate all the evil things that they were doing?"
                                                               (coorosiu,Or_1:8.27.26.537)

SORCvV

(9) *&* **hio** *<u>þæt ylce</u> <u>gewin</u> [þe hio hine on bespon   mid manigfealdon*
     and she the same war   that she him on persuaded with many
     firenlustum]   twa &   feowertig wintra **wæs dreogende**
     wicked.desires two and forty   years   was carrying.on
     "and for forty-two years she carried on the same war, which she brought upon
     him by her many wicked desires"   (coorosiu,Or_1:2.22.9.438)

Extraposition
<u>1 verb</u> SOvRC

(10) *þæt* **he** *<u>þone</u> <u>dæg</u>* **forlure**   *[þe he noht   to gode on ne gedyde];*
      that he the   day   lose-SBJV that he nothing of good on not did
      "that he lost the day on which he did no good"
      (cochronA-1,ChronA_[Plummer]:81.1.86)

<u>2 verbs</u> SOVvRC

(11) *ðæt* **mon** *lustlice   <u>ðone</u> <u>lareow</u>* **gehieran wille**   *[ðe mon ne lufað].*
      that one willingly the   teacher listen to wish-SBJV that one not love
      " … that one (may) willingly listen to the teacher that one does not love"
                                                               (cocura,CP:19.147.12.995)

SOvVRC

(12) *swelce* **hie** … **ða** *scylda* **willen forgietan** [ðe hie wið
as.if they the sins wish ignore that they against
hiene geworhton].
him committed
"as if they wished to ignore the sins that they had committed against him"
(cocura,CP:45.343.14.2309)

SvOVRC

(13) *þæt* **he sylf sceolde** *ða* *swaran wita* **onfon** [þe he þam
that he self should the heavy punishment receive that he to.the
preoste gemynte]
priest meant
"that he himself should receive the heavy punishment which he had meant for the priest" (coaelive,ÆLS_[Alban]:45.4024)

Interesting as they may be, I have not included in my analysis any examples of top-icalised objects with attached relative clauses, either with OSV or OVS word order (see (14a,b) below), as my emphasis is on the dependencies that are established between subject and verb and the processing problems that the clausal objects that 'intervene' between them cause. Therefore, in my database all objects follow their respective subjects.

(14) a. *þætte* [*þa seolfan moldan* [*þær his lichoma gefeol]]obj* **monige**
that the same clay where his body fell many many
mensubj **neomende wæron;**
men carrying were
"that the very clay where his body fell was carried away by many men"
(cobede,Bede_3:7.178.5.1739)

b. *ic ongite þætte* [*eall* [*ðæt þu me ær reahtes]]obj me* **reahte**
I perceive that all that you me before said me said
Godsubj *þurh ðe,*
God through you
"I perceive that all that you said to me before, God said to me through you"
(coboeth,Bo:36.103.28.2012)

Finally, regarding clause type, I decided initially to restrict my search to subordi-nate clauses to avoid the effects of V2, but soon realised that coordinate clauses, especially those with elliptical subjects (see Stockwell & Minkova 1990) and those whose first conjunct is v-final (see Pintzuk 1999) would also be good candidates for displaying v-finality, as proven by search results. See the graph below for the distribution of examples across clause types:
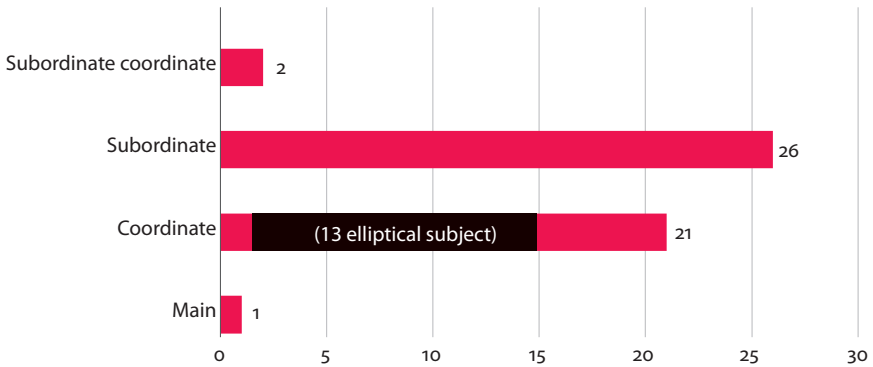
Figure 1. Number of examples of centre-embedded object relative clauses in the corpus according to clause type

In this chart we can observe that – as expected – the highest rate of objects with embedded relative clauses in the corpus is found in subordinate clauses (28 out of 50, including subordinate coordinate ones), followed by coordinate clauses, with 21 relevant examples, 13 of which have elliptical subjects. Unsurprisingly, only one example of a genuine embedded object relative appears in a main clause (15):[6]

(15) *Ðære tide eac swylce* **Eastseaxan** *þone geleafan*, [*þe hy   geara*
      at.this time also likewise East.Saxons the   faith   that they formerly
      awurpon, þa   heo Mellitum þone biscop ut   ascufon], *mid gyrnfulnesse*
      rejected   when they Mellitum the   bishop out drove   with diligence
      Oswiæs   þæs cyninges eft   onfengon.
      of.Oswio the king   again obtained
      "At this time also, through the efforts of king Oswio, the East Saxons returned to
      the faith, which they had formerly rejected when they expelled bishop Mellitus"
                                                    (cobede,Bede_3:16.224.7.2293)

2.2   Methodology

Concerning the methodology of the research, I have systematically compared the incidence of centre-embedded object relative clauses in the corpus with that of extraposed ones (patterns (16a) and (16b) below respectively), dismissing altogether examples of postposed head + relative clause (pattern (16c)). Therefore, in the measurements I carried out in §§ 3.1–3.1.1, the overall incidence of centre-embedded structures is generally expressed against the sum of centre embedding and

---

6.   Note, however, that the relative clause in this example is clearly non-defining. Hawkins (2007) points out that appositive relatives have "a more complete clausal structure and greater independent processability" (p. 134).

extraposition examples. The results were then mapped against different variables both language internal and external, trying to discover significant correlations. A comparison of the structural parameters that interact in the generation of embedded or extraposed structures in Old English and Modern German was carried out in §§ 3.1.2–3.2, trying to find out whether the processing strategies in both languages were similar and how this related to aspects of (changing) syntactic typology.

(16)   a.   O[RC]V (Embedding)
             *gyf **ðu** **þas** <u>wyrte</u> [ðe man ricinum nemneð] on þinre*
             if   you this plant that man ricunum call   in your
             *æhte   **hafast***
             possession have
             "if you have in your possession this plant that people call 'ricinum' … "
                                          (coherbar,Lch_I_[Herb]:176.1.2541)

       b.   OV[RC] (Extraposition)
             *gyf **hwa**   **þas** <u>wyrte</u> mid him **hafað** [þe we peristereon*
             if   whoever this plant with him has   that we peristereon
             nemdon] …
             called
             "if whoever has with him this plant that we called 'peristereon' … "
                                          (coherbar,Lch_I_[Herb]:67.1.1164)

       c.   VO[RC] (Heavy NP-Shift)
             *gif **hi**   **hæfdon** <u>hine</u> [þe wæs ærur   heafod to ðam unræde] …*
             if   they had   him who was formely leader to the plot
             "if they had him who had been before the leader of the plot"
                                          (cochronE,ChronE_[Plummer]:1087.43.3014)

2.3   Corpus and data retrieval

As regards corpus and data retrieval, the database for this study comprises 10 texts from the early and late OE periods (the boundary set about the year 950 approximately), which tot up around 540,000 words. They are:

Early OE (850–950) [264,040 words]: Preface to the *Cura Pastoralis* (CPLetW); *Cura Pastoralis* (CP); *Orosius* (Or); *Boethius* (Bo), Bede (Bede); *Anglo-Saxon Chronicle*, MS A (ChronA).

Late OE (950–1050) [275,439 words]: *Anglo-Saxon Chronicle*, MS E (ChronE); Ælfric's *Lives of Saints* (ÆLS); Old English Herbarium (Lch 1 (Herb)); Wulfstan's *Homilies* (WHom).

The texts belong to different genres and styles, ranging from the elaborate prose of Wulfstan to the simpler narrative-annalistic style of the *Anglo-Saxon Chronicle*. A special point is made in the study of whether texts were originally written in OE

or if they were translations of a Latin original, as I thought the influence of the Latin rhetorical tradition, understood as imitation of the 'periodic' style (whose trademark is precisely clausal embedding) might have some bearing on the overall incidence of clausal embedding in the corpus.

For my data source I used the *York-Toronto-Helsinki Parsed Corpus of Old English Prose* (Taylor et al. 2003). The data were retrieved using Corpus Search2 (Randall 2005–2007).

3.   Analysis

3.1   Single centre embedding

As you can observe in the table below, single centre embedding of object relatives occurs with moderate frequency in the corpus (50 tokens), which represents an incidence of 19.9% out of the sum of all examples of centre-embedded and extraposed object relatives:

Table 1. Percentage of embedded object relative clauses in the corpus

	Embedded	Extraposed	Total	% Embed
Or	2	24	26	7.7
CPLetW	1	–	1	100
CP	7	43	50	14
Bede	26	61	87	29.9
Bo	1	17	18	5.6
ChronA	3	5	8	37.5
ChronE	0	10	10	0
ÆLS	3	28	31	9.7
Lch I (Herb)	3	3	6	50
Wulfstan	4	10	14	28.6
Total	**50**	**201**	**251**	**19.9**

Considering the distribution of centre-embedded object relative structures across the different texts and text types in the corpus, some striking facts immediately catch the eye, such as the very low impact of relative clause embedding in argumentative texts like the OE *Boethius*[7] (5.6% of the relevant contexts), and – conversely – its

---

7.   Some, like Otten (1964), Godden (1992) or Los (p.c.), have suggested, however, that the OE *Boethius* did not always succeed in reproducing the subtle complexity of the Latin original, but – rather – reflected the difficulties the author had in translating it, pointing even to a certain clumsiness or awkwardness in the final rendering (Godden 1992: 525).

remarkably high incidence in texts of a less elaborate nature such as MS A of the *Anglo-Saxon Chronicle* (actually, 2nd highest in the list with 37.5% of the relevant contexts) or the OE *Herbarium* (50%).[8] This may suggest that the use of centre-embedded structures in OE reflected the practice of individual writers and their internal grammars, overriding any general considerations about genre, style or level of formality of a text. A closer look at the data, however, weakens this claim somehow: for example, two of the three examples of direct objects with embedded relative clauses in MS A of the *Anglo-Saxon Chronicle* involve quantifiers ((17a,b) below), and it is a well-known fact that quantified and negated elements keep OV order longer than any other kind of object in the history of English (see, among others, Moerenhout & van der Wurff (2005), or Pintzuk & Taylor (2006)):

(17) a. *& Ø þa   ealle [þe hie alǣdan   ne mehton] tobrǣcon*,
        and them all   that they take.away not could   broke.up
        "and (they) broke up all those they could not take away,"
                                                    (cochronA-2b,ChronA_[Plummer]:896.16.1114)
     b. *& Ø eal [þæt se arcebiscop at him crafede]   eadmedlice gefylde*
        and   all that the archbishop at him demanded humbly   fulfilled
        "and (he) humbly fulfilled all that the archbishop demanded from him"
                                                    (cochronA-10,ChronA_[Plummer]:1070.25.1479)

Similarly, one of the only two examples of centre-embedded relative clauses in the OE *Orosius* (another text commonly regarded as stylistically less elaborate)[9] also involves a quantified object (18):

(18) *Hwa is þæt þe   eall ða yfel [þe hi   donde wæron] asecgean mǣge*
     who is that who all the evil that they doing were   tell   may
     oððe areccean?
     or   narrate
     "who is that who may tell or narrate all the evil things that they were doing?"
                                                    (coorosiu,Or_1:8.27.26.537)

This is a remarkable example, in which the presence of the quantifier *eall* probably promoted the cohesion of the complex NP object, preventing the relative clause from becoming extraposed, even at the expense of creating a massive verb cluster with four verbs.

---

8.   Obviously, in these texts the number of examples is too low to be of statistical significance.

9.   The *Orosius* has often been put on a par with the *Anglo-Saxon Chronicle* in representing "a colloquial narrative style" (Stockwell & Minkova 1991: 375). Clark (1971: 217), however, remarks that the style of the *Orosius*, while being "… simple, with little subordination other than temporal and relative clauses …, is not as terse as that of the early *Chronicle*".

3.1.1   *External factors*

To have a wider picture of the interaction of the external factors with purely structural ones in the generation of centre-embedded structures in Old English, I analysed the global figures for relative clause centre embedding in the corpus against three non-structural variables: origin of the text (latinate or not), degree of stylistic elaboration, and date. The results are as follows:

Table 2. Percentage of embedded object relative clauses in Latin-based
and non-Latin-based texts in the corpus

	Embedding examples	Total	% embedding
Latin-based texts	39	187	20.9%
Non Latin-based texts	11	64	17.2%

We can see in Table 2 that the difference in the incidence of embedded object relatives between texts written originally in Old English and those translated from Latin is not great (20.9% vs 17.2% in favour of texts of Latin origin). However, I would like to point out that the text unanimously regarded as the most latinate of all the major OE prose texts (Bede's *Historia Ecclesiastica*, see Cichosz et al. 2016) presents a remarkable incidence of object centre-embedding in proportion to relative clause extraposition (29.9%), and is the highest in absolute incidence, with 26 tokens overall (or a 4.81 mean rate in normalised frequency per 100,000 words), a long, long distance from the second one, the *Cura Pastoralis*, with a 1.29 mean rate (7 tokens).

Concerning the level of stylistic development of the texts, I grouped them into two categories: the more elaborate ones (comprising the *Cura Pastoralis*, Bede, Boethius and Wulfstan) and the less elaborate ones (Orosius, ChronA, ChronE, the Herbarium and Ælfric's work).[10] The results of the analysis show a greater correlation between the level of formality of the text and the use of embedded relativisation, with almost 10 percentage points in favour of the more crafted prose:

Table 3. Percentage of embedded object relative clauses in stylistically
+/− developed texts in the corpus

	Embedding examples	Total	% embedding
+ elaborate prose	39	170	22.9%
− elaborate prose	11	81	13.6%

---

10. This classification, though obviously debatable, reflects current opinions on the stylistic characteristics of Old English texts, such as Godden (1992: 523–535) or Bernárdez & Tejada (1995: 235–236). The position of Ælfric's *Lives of Saints* – as pointed out to me by B. Los – is controversial, since this work might also fit in the 'elaborate prose' category on the basis of the findings in Lipp (1969), Hurst (1972), or Waterhouse (1983).

Therefore, I argue that we can assume that, however ingrained centre embedding was in the grammar of the OE speakers, the influence of style on it was considerable, and that it clearly interacted with the purely grammatical conditions that constrained it.

Finally, considering the diachronic axis, the corpus analysis indicates that embedded object relatives were used consistently more often in early OE (21.1%) that in late OE (16.4%) in the relevant contexts. Importantly, if we assume that the presence of embedded object relative clauses is an indicator of underlying OV order (since there should be no reason for relative clauses to occur inside the sentence brace/pre-verbally in SVO languages), then, the decrease in the frequency of embedded object relatives from early to late Old English could be said to reflect the change in progress from OV to VO proposed in Pintzuk's (1999) or Pintzuk & Taylor (2006) within the canvas of the Double Base Hypothesis. In fact, this study confirms the data in Pintzuk & Taylor (2006), as the proportion of preverbal objects with relative clauses decreases over time in almost the same proportion (about 6 percent) as general objects do in their study (see Table 4):

Table 4. Percentage of preverbal object relative clauses and preverbal general objects in early and late Old English compared

	Proportion of embedded object relative clauses			Proportion of preverbal objects in early/late Old English (Pintzuk & Taylor (2006))		
	Embedding examples	Total	% embedding	Preverbal objects	Total objects	% preverbal
Early OE (before 950)	40	190	21.1	803	1,416	56.7
Late OE (after 950)	10	61	16.4	1,165	2,310	50.4

3.1.2   *Internal factors*

Leaving aside the external factors that interact in the use of single clausal embedding in OE, I turn for the rest of this section to the internal, syntactic parameters that control the use of embedded object relativisation in Old English, to find out if the OE facts respond to similar restrictions on processing as other modern SOV languages.

Distance-based theories of linguistic complexity such as Gibson's (1998, 2000) Dependency Locality Theory or Hawkins's (1994, 2007) Early Immediate Constituents select two components that interact heavily in the choice between embedding or extraposition of object relative clauses when it comes to minimising processing domains in SOV languages (using modern German as a model): the so-called Extraposition Distance and the Relative Clause Size. I define each concept below

highlighting it (by underlining) in the examples given, which are variations from
(4) above:

Extraposition distance

<u>In centre embedding</u>:
*The number of remaining word forms of the superordinate clause after the end of the
subordinate clause*

(19) *Ich hoffe, dass **Peter** [*VP* [*NP* das Buch, [das Maria gekauft hat,] <u>bald</u> <u>**liest**</u>].*
I   hope that Peter   the book that Maria bought has   soon read
"I hope that Peter will soon read the book that Maria has bought"

(Extraposition distance = 2 words)

<u>In extraposition</u>:
The number of word forms between the end of the antecedent NP and the relative clause

(20) *Ich hoffe, dass **Peter**[*VP*[*NP*das Buch <u>bald</u> <u>**liest**</u>, [das Maria gekauft hat]].*

(Extraposition distance = 2 words)

Relative clause size

Length (in number of words) of embedded or extraposed relative clause

(21) a.   *Ich hoffe, dass **Peter**[*VP*[*NP*das Buch, [<u>das Maria gekauft hat,</u>] bald **liest**].*
b.   *Ich hoffe, dass **Peter**[*VP*[*NP*das Buch bald **liest**, [<u>das Maria gekauft hat</u>]].*

(Relative clause size = 4 words)

Both experimental work (Uszkoreit et al. 1998; Konieczny 2000) and extensive
corpus studies (Shannon 1992; Strunk 2010; Bader & Häussler 2012) have been
carried out to assess the overall efficiency of one structure over the other taking into
account the interaction of these two parameters but – to cut a long story short –
Hawkins's (1994, 2007) predictions are that (i) shorter centre-embedded relative
clauses should be more acceptable than longer ones, and (ii) extrapositions should
be more acceptable if their extraposition distance is small.

The results of the corpus analysis for Old English corroborate Hawkins's find-
ings: as we can see in Table 5, the average size of the RCs in the examples of embed-
ding in the corpus (4.9 words) is considerably shorter than the size of the RCs in the
examples of extraposition (6.7 words), and –importantly – the average size of the

Table 5. Average length of relative clause and extraposition distance in embedding
and extraposition examples in the corpus (in number of words)

	RC length	Extraposition distance
In embedding	4.9	4.8
In extraposition	6.7	1.9

extraposition distance is *much* shorter on average in the examples of extraposition (1.9 words) than in the examples of embedding (4.8 words).

We can say, therefore, that – apparently – the structural characteristics that govern the processing of single centre embedding and extraposition of object relative clauses in Old English and ultimately determine the choice between one and the other do not significantly differ from those for other modern SOV languages. To corroborate this, I compared the incidence of object relative clause embedding and object relative clause extraposition in Old English (corpus data) and Modern German, using the size of the relative clause and the extraposition distance as determinant variables. For the German data, I used Uszkoreit et al.'s (1998) study, based on a corpus of 12,000 sentences taken from the *Frankfurter Rundschau* newspaper and an expanded corpus tagged for parts of speech, as it appears in Hawkins (2007: 145–146).

As shown in Figure 2, which represents the extraposition rates in Old English and Modern German according to the size of the relative clause, the lines run fairly parallel, the rate of extraposition increasing progressively in both languages as the RC increases in size (from 3 words to 15 words and over). It must be noted, however, that the overall rate for the use of RC extraposition is systematically higher in OE than in Modern German at each reference point.
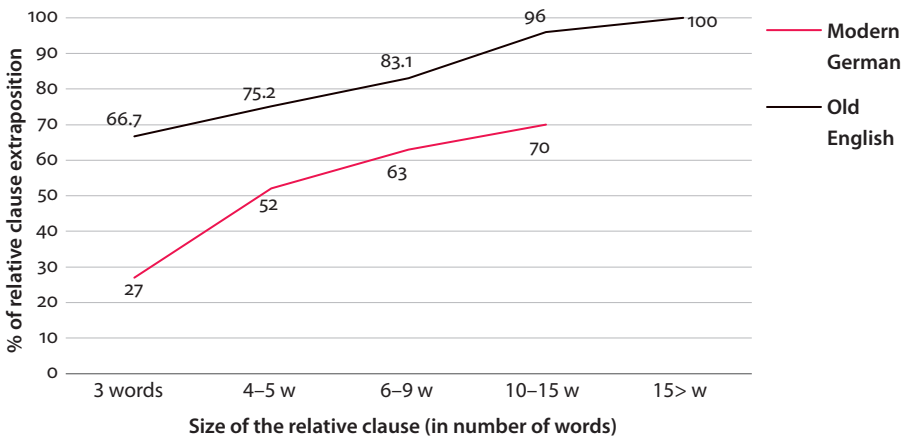
Figure 2. Percentage of extraposition according to the size of the relative clause in Modern German and Old English compared

Regarding the extraposition distance (Figure 3), lines run fairly close to one another for distances between 1 and 2 words (which represent the majority of the examples), with Modern German then showing a preference for non-extraposition as the extraposition distance becomes 3 words long or more:
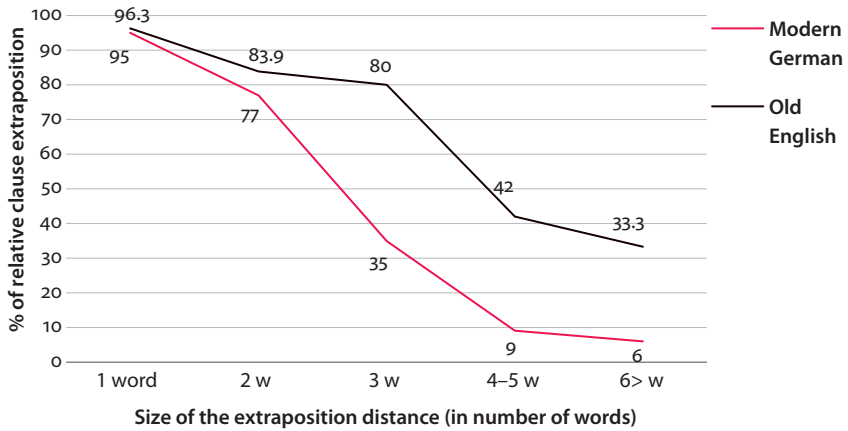
Figure 3. Percentage of extraposition according to the size of the extraposition distance in Modern German and Old English compared

This is interesting, but – as pointed out in Hawkins (2007: 142–146) – the real test for performance strategies in terms of optimal constituent recognition domains comes from the interaction of these two parameters (RC size and extraposition distance) and not from taking each one of them in isolation. Therefore, I carried out a study of such interaction, the results of which are given in Table 6 (the RC size is represented in the horizontal axis and the extraposition distance in the vertical axis):

Table 6. Percentage of object relative clause extraposition in the corpus according to the interaction of RC size and extraposition distance

	Relative clause size														
	2–3 words			4–5			6–9			10–15			15 >		
Extraposition distance	Emb	Extr	% extr	Emb	Extr	% extr	Emb	Extr	% extr	Emb	Extr	% extr	Emb	Extr	% extr
V	1	13	92.8	3	42	93.3	0	31	100	0	11	100	0	7	100
XP$_1$> +V	3	1	25	3	19	86.3	2	21	91.3	1	6	85.7	0	1	100
XP$_2$> +V	0	2	100	3	14	82.3	4	10	71.4	0	2	100	0	1	100
XP$_{3–4}$> +V	1	0	0	9	3	25	5	2	28.5	0	3	100	–	–	–
XP$_{5–9}$> +V	1	0	0	6	1	14.2	2	5	71.4	1	2	66.6	–	–	–
XP$_{10}$> +V	2	0	0	2	0	0	1	0	0	–	–	–	–	–	–
% extraposition	66.7%			75.2%			83.1%			96%			100%		

The results of the analysis confirm the predictions in Hawkins (2007) and other studies of locality and dependency minimisation carried out from a psycholinguistic perspective like Bader (manuscript), Bader & Häussler (2012) and Trotzke & Bader (2013). Namely:

a. with very short extraposition distances the percentage of extraposition is always very high, no matter the size of the relative clause. See the table above and the Old English examples in (22) in which the extraposition distance is just one word – the verb of the superordinate clause, *onscunie* "reject" in (22a), *doð* "do" in (22b), *gedwellen* "obscure" in (22c) and *bebyrigdon* "buried" in (22d) –, and the size of the extraposed relative clauses ranges from 3 to 10 words:

(22) a. *swa þæt þu … þa ðincg* **onscunie** *[þe God onscunað].*
           so   that you   the things reject-SBJV that God rejects
           "so that you reject the things that God rejects"
                                                    (coaelive,ÆLS[Ash_Wed]:125.2767)

      b. *ðæt* **hie** *ðæt ilce* **doð** *[ðæt hie on ðæm oðrum tældon].*
           that they the same do   that they on the   others blamed
           "that they do the same that they blamed in the others"
                                                    (cocura,CP:44.333.20.2256)

      c. *ðæt* **hie** *ðæt* **gedwellen**   *[ðæt oðre   men ryhtlice &*
           that they that obscure-SBJV that other men rightly and
           gesceadwislice ongieten   habbað],
           sagaciously   perceived have
           "that they obscure that which others have rightly and sagaciously understood"   (cocura,CP:48.365.23.2470)

      d. *hwi* **hi** *þa*   **bebyrigdon** *[þe his beboda   forsawon and for*
           why they those buried   who his commands despised and for
           heora scyldum ofslagene lagon];
           their sins   slain   lay
           "why they buried those who despised his commands and lay slain because of their guilt"   (coaelive,ÆLS_[Cecilia]:203.7237)

b. as the size of the extraposition distance increases (especially from 3–4 words onwards), extraposition rates decrease and embedding becomes the preferred configuration, unless relative clauses are exceptionally large; see the following examples of embedding from the *Herbarium* and the OE *Orosius*, which display extraposition distances ranging from 3 to 6 words: *mid him hafað* "with him have" in (23a), *on þinre æhte hafast* "in your possession have" in (23b), and *twa & feowertig wintra wæs dreogende* "forty-two years was carrying on" in (23c).

(23) a. *gyf* **hwa**   *þysse wyrte wyrttruman [þe man crision &*
          if   whoever of-this-herb root   that one crision and
          *oðrum   naman clæfre nemneð] mid him **hafað**…*
          by.another name clover name   with him has
          "if whoever has with him the root of this herb that people call 'crision'
          and by another name 'clover' … " (coherbar,Lch_I_[Herb]:70.1.1182)

     b. *gyf **ðu** þas wyrte [ðe man ricinum &   oðrum   naman*
          if   you this herb that one ricunum and by.another name
          *+++[11] nemneð] on þinre æhte   **hafast** …*
          +++ call   in your possession have
          "if you have in your possession this herb that people call 'ricinum' and,
          by another name, '+++' … "   (coherbar,Lch_I_[Herb]:176.1.2541)

     c. *&   **hio** þæt ylce gewin [þe hio hine on bespon   mid*
          and she the same war   that she him on persuaded with
          manigfealdon firenlustum]   twa &   feowertig wintra
          many   wicked.desires two and forty   years
          ***wæs dreogende**.*
          was carrying.on
          "and, for forty-two years she carried on the same war, which she brought
          upon him by her many wicked desires"   (coorosiu,Or_1:2.22.9.438)

c. when RC's are exceptionally large (15 words or more), this factor completely
    outweighs the extraposition distance, and extraposition becomes the norm.

In conclusion, there is nothing that apparently sets Old English substantially apart
from other modern SOV languages like Modern German from the point of view of
the structural constraints that govern the optimal processing of sentences contain-
ing single embedded preverbal objects. When the interaction of these constraints
naturally leads to centre embedding, that is usually the case also in Old English,
and when they interact to trigger extraposition of the relative clause, extraposition
is also the norm in OE. This is so at least as far as single centre embedding is con-
cerned. Some important differences, however, arise when we talk about multiple
clausal embedding, which I discuss in the following section.

3.2   Multiple centre embedding

It is well-known (Yngve 1960; Chomsky & Miller 1963; Reich 1969; Reich & Schütze
1991) that in actual linguistic performance multiple centre embedding normally
leads to processing breakdown when the degree of syntactic embedding exceeds a

---

11. Illegible in the MS.

rather small limit, usually two levels of syntactic depth (Karlsson 2007: 24; Trotzke et al. 2013: 18). Therefore, sentences like (24c) and (25c) in the examples below, containing doubly centre-embedded relative clauses, are usually considered as unacceptable by most English speakers:

(24) a. The man left
      b. The man [the boy heard] left
      c. *The man [the boy [the woman saw] heard] left

(25) a. The reporter disliked the editor
      b. The reporter [who the senator attacked] disliked the editor
      c. *The reporter [who the senator [who John met] attacked] disliked the editor
                                              [Examples from Gibson 1998: 105]

In spite of this, double (and sometimes triple) centre embedding is found in SOV languages, like Modern German, as in the following example from Trotzke & Bader (2013):

(26) *Probleme gab es,   weil   einige Kurse, [die   in der schönen*
      problems were there because some courses which in the nice
      Broschüre [die   man vorher zugeschickt bekommt], aufgelistet waren],
      brochure which one before sent.out   got   listed   were
      gestrichen worden waren
      cancelled became were
      "There were problems because some courses which were listed in the nice brochure which was sent out in advance were cancelled"

Of course, the processing problems mentioned before for simple centre embedding are there for double centre embedding too. Therefore, instead of having two embedded relative clauses, most German speakers prefer to rearrange the complex and difficult-to-parse structure in (26) using RC extraposition to different degrees and in different combinations. This renders the following additional attachment positions for two dependent relative clauses in Modern German, running from the most difficult (27a) to the least difficult to process (27c):[12]

(27) a. **Higher RC embedded, lower RC extraposed**
          Problem gab es, weil einige Kurse [die in der schönen Broschüre aufgeliest waren, [die man vorher zugeschickt bekommt]], gestrichen worden waren.

---

12. Evidence for the processing difficulty of one arrangement over the other has massively been produced, both from experimental work in psycholinguistics using eye-tracking and self-paced reading experiments, and from textual studies of the overall incidence of each pattern in different corpora, including Korthals (2001), Bader (manuscript) and Trotzke et al. (2013).

b. **Higher RC extraposed, lower RC embedded**
Problem gab es, weil einige Kurse gestrichen worden waren [die in der schönen Broschüre, [die man vorher zugeschickt bekommt], aufgeliest waren]].

c. **Both RCs extraposed**
Problem gab es, weil einige Kurse gestrichen worden waren [die in der schönen Broschüre aufgeliest waren, [die man vorher zugeschickt bekommt]]

As far as Old English is concerned, no genuine examples of double relative centre embedding have been found in the corpus, with all the contexts suitable for multiple recursion of this type presenting extraposition of one or both relative clauses. So, for instance, constructs like (28a) below, with double centre embedding (in this case combining subject and object relativisation) are apparently never found in OE. The actual rendering in the corpus for (28a) is (28b), with extraposition of both relative clauses (the verbs in (28)–(30) are in bold):

(28) a. ⁺*forðan þe   seo sawul [ðe þonne scyppend [þe hi and hire*
because that the soul   that the   lord   that it and its
geferan   gesceop] lufað] is gesælig.
companions created loves   is blessed

b. *forðan þe seo sawul **is** gesælig [ðe þonne scyppend **lufað** [þe hi **gesceop** and hire gefæran]].*
"because the soul that loves the God that created it and its fellow-pilgrims is blessed"   (coaelive,ÆLS_[Christmas]:169.130)

Double relative clause extraposition is not, however, the only available option in Old English for the arrangement of complex sentences involving two relative clauses: of the three patterns mentioned in (27) for Modern German, Old English also displays (27a), with the higher RC centre embedded and the lower RC extraposed. We can see an example of it in (29):

(29) *þæt heretogan   &   domeras, &   þa maðmhirdas [þe þæt fioh*
that commanders and judges   and the treasurers   who the money
hioldon** [þe   mon þam ferdmonnum on   geare **sellan sceolde]] **hæfdon
kept   which one   to the soldiers   every year give   should   had
mæstne weorðscipe.
greatest honour
"that commanders, and judges, and the treasurers who kept the money which they were every year to give to the soldiers, had the greatest honour"
(coboeth,Bo:27.64.11.1194)

Finally, examples of pattern (27b) above (extraposed higher RC and embedded lower RC) are also found in the textual record:

(30) *From þæm he <u>monig þing</u> **gehyrde**, [þa   ðe   ge   him sylfum ge*
from them he many things heard   those which both to   himself and
*eallum þæm,   [þe hit **gehyran woldon**], swiðe halwende **wæron***
to all   of those who it   hear   wanted   very   beneficial were
***to weotonne**].*
to be known
"From them he heard many things which were very beneficial to be known, both for himself and for those who would hear them" (cobede,Bede_3:14.214.22.2188)

One explanation for the dearth of examples of multiple centre embedding in Old English might be – as Karlsson (2007: 2, 27) points out – that multiple centre embedding is simply rare cross-linguistically, and that its use in most European languages was normally confined to very formal communicative contexts or to sophisticated literary styles heavily influenced by the Latin rhetorical tradition. However, multiple centre embedding is not that exceptional in Modern German, nor confined to specific registers or styles, as Korthals (2001) and Bader (manuscript), among others, have demonstrated.[13] Therefore, if OE behaved in a similar way to Modern German with regards to the processing of multiple central recursion, an explanation beyond style for the absence of such structures in OE should probably be desirable.

I suggest (tentatively) that the absence of double centre embedding in OE may have to do with the interaction between language processing and the shaping of grammars, and specifically with the possibility that such interaction may not necessarily be unidirectional (with parsing mechanisms being able to shape grammars but not the other way around).

If we assume a position on processing and language parsing like that in Vasishth & Lewis (2006), Vasishth et al. (2010) or Frank et al. (2015), according to which the human parser is an adaptive system that changes depending on the kinds of actions it has to perform and on the linguistic input it receives, strict SOV languages, like Modern German, would be better adapted to making predictions about upcoming verbs – and subsequently to processing multiple centre-embedded structures – than SVO ones because of the very robust memory representations for left branching configurations German speakers have as a result of OV complementation. However, if we assume that Old English is shifting slowly – on independent grounds – from OV to VO (as suggested in Pintzuk 1999 or Pintzuk & Taylor 2006) or that it

---

13. Korthals (2001) and Bader (manuscript) have studied the incidence of double centre embedding in Modern German written texts using large computerised corpora, suggesting an average ratio of 17 instances per million words. So, taking into account that the size of my corpus is 539,479 words, the expected number of doubly centre-embedded relative clauses in it should be 9.1.

actually *is* essentially VO (as recently pointed out by Struik & van Kemenade 2018), then Old English would lack the robust left-branching configurations which would help model the processing capacities of its speakers, making them capable of processing double centre-embedded structures easily. To put it rather bluntly: Old English would be OV enough to produce single centre-embedded structures, but not OV enough to produce double centre-embedded ones. This hypothesis, tentative and debatable as it is, would only be possible in a scenario of gradualism and syntactic variation in the base like the one proposed in Pintzuk's (1999) Double Base Hypothesis and – if proved valid – would represent clear evidence of it.

4.   Conclusions

The results of the analysis presented here suggest (contra Ogura 2001, 2004) that single centre embedding was a viable grammatical option in Old English, with a considerable impact in the textual record, and that – while its avoidance undoubtedly contributed to the change from OV to VO in English – it was certainly not the trigger of it. Style, register and imitation of Latin models probably played a role in the choice for centre-embedded relative objects in OE, overlapping with the purely grammatical constraints operating on them, but this is difficult to assess given the limited range of text types in Old English. On the other hand, the interaction of the structural parameters that combine in search of optimal processing for single-embedded relative objects in Old English does not seem to greatly differ from other modern SOV languages, like Modern German, suggesting similar parsing strategies for that particular configuration for both groups of speakers. Moreover, assuming that preverbal relative objects are indicative of OV order, the decrease in their frequency from early to late Old English could be considered as an indicator of the change in progress from OV to VO, as suggested in Pintzuk (1999), Fuss & Trips (2001) or Pintzuk & Taylor (2006).

Finally, the analysis reveals that the maximum number of levels of relative object centre embedding in Old English is one. The lack of examples of double centre embedding in the corpus could be explained – apart from the overall cross-linguistic exceptionality of those structures – by considerations affecting the typology of OE: given the fact that multiple centre embedding is strongly associated with OV complementation, its absence in OE could be the consequence of the language having largely shifted in the classical period – on independent grounds – from OV to VO.

Acknowledgements

References

Bader, Marcus. Manuscript. Complex center embedded relative clauses in German. 2012. Johan Wolfgang Goethe-Universität. Frankfurt am Main.

Bader, Marcus & Jana Häussler. 2012. Constraints on center-embedding as revealed by corpus data. Paper presented at the Formal Linguistics and the Measurement of Grammatical Complexity Symposium, Seattle, WA, March 23–24.

Bernárdez, Enrique & Paloma Tejada. 1995. Pragmatic constraints to word order and word-order change in English. In Andreas H. Jucker (ed.), *Historical pragmatics*, 217–241. Amsterdam: John Benjamins. https://doi.org/10.1075/pbns.35.15ber

Chomsky, Noam & George A. Miller. 1963. Introduction to the formal analysis of natural languages. In R. Duncan Luce, Robert R. Bush & Eugene Galanter (eds.), *Handbook of mathematical psychology*, Volume 2, 269–321. New York: Wiley.

Cichosz, Anna, Jerzy Gaszewski & Piotr Pęzik. 2016. *Element order in Old English and Old High German translations*. Amsterdam: John Benjamins. https://doi.org/10.1075/nss.28

Clark, Cecily. 1971. The narrative mode of *The Anglo-Saxon Chronicle* before the Conquest. In Peter Clemoes & Kathleen Hughes (eds.), *England before the conquest: Studies in primary sources presented to Dorothy Whitelock*, 215–235. Cambridge: Cambridge University Press.

Cloutier, Robert A. 2009. *West Germanic OV and VO. The status of exceptions*. Utrecht: LOT.

Colman, Fran. 1988. Heavy arguments in Old English. In John M. Anderson & Norman Macleod (eds.), *Edinburgh studies in the English language* 1, 33–89. Edinburgh: John Donald.

Fischer, Olga, Hendrik de Smet & Wim van der Wurff. 2017. *A brief history of English syntax*. Cambridge: Cambridge University Press. https://doi.org/10.1017/9781139049559

Frank, Stefan L., Thijs Trompenaars & Shravan Vasishth. 2015. Cross-linguistic differences in processing double-embedded relative clauses: Working-memory constraints or language statistics? *Cognitive Science* 40(3). 554–578. https://doi.org/10.1111/cogs.12247

Fuss, Eric & Carola Trips. 2001. Variation and change in Old and Middle English: On the validity of the Double Base Hypothesis. *Journal of Comparative Germanic Linguistics* 4(3). 171–224. https://doi.org/10.1023/A:1016588502117

Gibson, Edward. 1998. Linguistic complexity: Locality of syntactic dependencies. *Cognition* 68. 1–76. https://doi.org/10.1016/S0010-0277(98)00034-1

Gibson, Edward. 2000. The dependency locality theory: A distance-based theory of linguistic complexity. In Alec P. Marantz, Yasushi Miyashita & Wayne O'Neil (eds.), *Image, language, brain. Papers from the first Mind Articulation Project Symposium*, 95–126. Cambridge, MA: MIT Press.

Godden, Malcolm. 1992. Literary language. In Richard M. Hogg (ed.), *The Cambridge history of the English language. Volume 1: The beginnings to 1066*, 490–535. Cambridge: Cambridge University Press. https://doi.org/10.1017/CHOL9780521264747.009

Hawkins, John A. 1994. *A performance theory of order and constituency*. Cambridge: Cambridge University Press.

Hawkins, John A. 2007. *Efficiency and complexity in grammars*. Oxford: Oxford University Press.

Hock, Hans Henrich. 1988. Relative clauses and rebracketing in Old English. In Elmer H. Antonsen & Hans Henrich Hock (eds.), *Germanic linguistics II: Papers from the Second Symposium on Germanic Linguistics*, 35–55. University of Illinois at Urbana-Champaign: Indiana University Linguistics Club.

Hock, Hans Henrich. 1991. On the origin and development of relative clauses in early Germanic, with special emphasis on *Beowulf*. In Elmer H. Antonsen & Hans Henrich Hock (eds.), *Stæfcræft: Studies in Germanic linguistics*, 55–89. Amsterdam: John Benjamins. https://doi.org/10.1075/cilt.79.05hoc

Hogg, Richard & David Denison. 2006. *A history of the English language*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511791154

Hurst, James. 1972. *Ælfric*. New York: Twayne.

Karlsson, Fred. 2007. Constraints on multiple center-embedding of clauses. *Journal of Linguistics* 43. 365–392. https://doi.org/10.1017/S0022226707004616

Konieczny, Lars. 2000. Locality and parsing complexity. *Psycholinguistic Research* 29. 627–645. https://doi.org/10.1023/A:1026528912821

Korthals, Christian. 2001. Self embedded relative clauses in a corpus of German newspaper texts. In Kristina Striegnitz (ed.), *6th ESSLLI student session, August 13–14, 2001*, 179–190. Helsinki: University of Helsinki.

Lipp, Frances R. 1969. Ælfric's Old English prose style. *Studies in Philology* 66. 689–718.

Mitchell, Bruce. 1985. *Old English syntax*. Oxford: Clarendon Press. https://doi.org/10.1093/acprof:oso/9780198119357.001.0001

Moerenhout, Mike & Wim van der Wurff. 2005. Object-Verb order in early sixteenth-century English prose: An exploratory study. *English Language and Linguistics* 9. 83–114. https://doi.org/10.1017/S1360674305001553

Ogura, Mieko. 2001. Perceptual factors and word order change in English. *Folia Linguistica Historica* 22(1–2). 233–253.

Ogura, Mieko. 2004. Evolution of word order. *Folia Linguistica Historica* 25(1–2). 21–39. https://doi.org/10.1515/flih.2004.25.1-2.21

O'Neil, Wayne. 1976. Clause adjunction in Old English. *General Linguistics* 17. 199–211.

Otten, Kurt. 1964. *König Alfreds Boethius* (Studien zur englischen Philologie, n.f. 3). Tübingen: Max Niemeyer. https://doi.org/10.1515/9783111385990

Pintzuk, Susan. 1999. *Phrase structures in competition: Variation and change in Old English word order*. New York: Garland.

Pintzuk, Susan & Ann Taylor. 2006. The loss of OV order in the history of English. In Ans van Kemenade & Bettelou Los (eds.), *The handbook of the history of English*, 249–278. Oxford: Blackwell. https://doi.org/10.1002/9780470757048.ch11

Randall, Beth. 2005–2007. CorpusSearch 2. http://corpussearch.sourceforge.net/

Reich, Peter A. 1969. The finiteness of natural language. *Language* 45. 831–843. https://doi.org/10.2307/412337

Reich, Peter A. & Carson T. Schütze. 1991. Syntactic embedding: What can people really do? In Barbara Brunston & Tom Wilson (eds.), *Toronto Working Papers in Linguistics* 11(1). 91–97.

Shannon, Thomas F. 1992. Towards an adequate characterization of relative clause extraposition in Modern German. In Irmengard Rauch, Gerald F. Carr & Robert L. Kyes (eds.), *On Germanic linguistics: Issues and methods* 253–281. Berlin: De Gruyter. https://doi.org/10.1515/9783110856446.253

Stockwell, Robert P. 1977. Motivations for exbraciation in Old English. In Charles Li (ed.), *Mechanisms of syntactic change*, 291–314. Austin: University of Texas Press.

Stockwell, Robert P. & Donka Minkova. 1990. Verb phrase conjunction in Old English. In Henning Andersen & Konrad Koerner (eds.), *Historical linguistics 1987*, 498–515. Amsterdam: John Benjamins. https://doi.org/10.1075/cilt.66.34sto

Stockwell, Robert P. & Donka Minkova. 1991. Subordination and word order change in the history of English. In Dieter Kastovsky (ed.), *Historical English syntax*, 367–408. The Hague: Mouton. https://doi.org/10.1515/9783110863314.367

Strang, Barbara. 1970. *A history of English*. London: Methuen.

Struik, Tara & Ans van Kemenade. 2018. On the givenness of OV word order: A reexamination of OV/VO variation in Old English. *English Language and Linguistics* 24. 1–22. https://doi.org/10.1017/S1360674318000187

Strunk, Jan. 2010. Enriching a treebank to investigate relative clause extraposition in German. In Nicoletta Calzolari, Khalid Choukri, Bente Maegaard, Joseph Mariani, Jan Odijk, Stelios Piperidis, Mike Rosner & Daniel Tapias (eds.), *Proceedings of LREC 2010*, 1911–1917. Paris: ELRA.

Taylor, Ann, Anthony Warner, Susan Pintzuk & Frank Beths. 2003. *York-Toronto-Helsinki Parsed Corpus of Old English Prose. Electronic texts and manuals available from the Oxford Text Archive*.

Tily, Harry. 2010. The role of processing complexity in word order variation and change. Stanford: Stanford University PhD dissertation.

Traugott, Elizabeth C. 1992. Syntax. In Ricahrd M. Hogg (ed.), *The Cambridge history of the English language. Volume 1, The beginnings to 1066*, 168–289 Cambridge: Cambridge University Press. https://doi.org/10.1017/CHOL9780521264747.005

Trotzke, Andreas & Marcus Bader. 2013. Against usage-based approaches to recursion: The grammar-performance distinction in a biolinguistic perspective. Paper presented at the 36 GLOW conference, Lund, April 2–6.

Trotzke, Andreas, Marcus Bader & Lyn Frazier. 2013. Third factors and the performance interface in language design. *Biolinguistics* 7. 1–34.

Uszkoreit, Hans, Thorsten Brants, Denys Duchier, Brigitte Krenn, Lars Konieczny, Stephan Oepen & Wojciech Skut. 1998. Studien zur performanz-orientierten Linguistik: Aspekte der Relativsatz-extraposition im Deutschen. *Kognitionwissenschaft* 7. 129–133. https://doi.org/10.1007/s001970050065

Vasishth, Shravan & Richard L. Lewis. 2006. Argument-head distance and processing complexity: Explaining both locality and antilocality effects. *Language* 82. 767–794. https://doi.org/10.1353/lan.2006.0236

Vasishth, Shravan, Katja Suckow, Richard L. Lewis & Sabine Kern. 2010. Short-term forgetting in sentence comprehension: Crosslinguistic evidence from verb-final structures. *Language and Cognitive Processes* 25. 533–567. https://doi.org/10.1080/01690960903310587

Waterhouse, Ruth. 1983. *The triangular clause relationship in Ælfric's Lives of Saints and in other works*. New York: Peter Lang.

Yngve, Victor H. 1960. A model and an hypothesis for language structure. *Proceedings of the American Philosophical Society* 104. 444–466.

Syntactic changes in verbal clauses and noun phrases from 1500 onwards

Gerold Schneider

University of Zurich

Can the promise of data-driven methods hold in historical linguistics? Can they detect salient syntactic changes and open new research avenues? I first use data-driven measures to detect patterns in the ARCHER corpus. Secondly, I qualitatively interpret the differences and build hypotheses. Thirdly, I validate these, with the Penn Corpora, investigating frequency and creativity. We observe several trends – among them: verbal ('Doric') style is decreasing, nominal ('Attic') style is increasing. A cascade from full to non-finite clause, and from paratactic to hypotactic style unfolds. Furthermore, constituent order is increasingly becoming fixed, strengthening the principle of dependency length minimisation. While data-driven approaches entail a complex interpretation step, their holistic perspective goes beyond well-trodden envelopes of variation to more global language models.

Keywords: data-driven methods, syntax, ARCHER corpus, Penn Parsed Corpora, Penn Treebank, nominal style, entropy, word order, information structure

1. Introduction

The development of English has been intensively studied, covering the entire period from the first known manuscripts to PDE. While there are very many detailed studies describing individual phenomena, and the broad changes have been summarised, there is a risk of oversights, and the assessment of the quantitative effects of each of the changes is still incomplete. While corpus-based research has given us a high level of precision, we cannot be sure about recall – have we missed any important developments and changes?

Both oversight and quantitative effect can be addressed by using data-driven approaches. Data-driven, also called corpus-driven or bottom-up, approaches systematically investigate correlation patterns. This study uses statistical metrics based on relative frequencies to exhaustively assess the most salient differences between

frequencies in historical corpora, at the level of part-of-speech-tags and syntactic relations.

Ananiadou et al. (2006: 572) argue that any type of statistically significant difference, connection or association between not clearly connected entities and facts could be meaningful and thus merit being investigated. For this, data-driven methods offer the advantages of making fewer theoretical assumptions, and being able to bring previously undetected patterns to the surface.

> We should strive to be open to the patterns observable in language … and rebuild a picture of language and meaning which is not only consistent with the evidence but also exploits it to the full … the first stage should be an attempt to inspect the data with as little attention as possible to theory.   (Sinclair & Carter 2004: 10)

Their disadvantage is that they are very sensitive to data: " … since the information provided by the corpus is placed centrally and accounted for exhaustively, then there is a risk of error if the corpus turns out to be unrepresentative" (Tognini-Bonelli 2001: 88). It is thus unclear how many of the salient features that data-driven methods deliver correspond to actual diachronic developments, or if they largely report coincidences from the corpora used and errors from the processing steps. Automatic annotation, which is part of this processing, on the one hand has the advantage that it provides large amounts of data from a different perspective. Word class annotation allows us to see morphosyntactic patterns, and syntactic annotation allows us to detect word order phenomena. On the other hand, automatic processing is also error-prone, and because higher level annotation builds on lower levels (for example, parsing uses part-of-speech tagging) there is a serious risk of error propagation.

Data-driven approaches have not been used much in historical linguistics, as for example Hilpert & Gries (2017: 44–45) point out. They also describe the characteristic of data-driven methods that they often reverse the roles of quantitative and qualitative analysis. There is no initial hypothesis: the patterns in the data (detected in a first step) then need to be analysed qualitatively (in a second step), and the arising hypothesis can then be tested (in a third step).

The current paper explores whether the promise of data-driven methods is going to hold. My orientation is thus of a methodological nature. I am addressing the following research questions.

1. Can data-driven methods detect salient syntactic changes?
2. What operations are needed to be successful in this task?
3. Can data-driven approaches bring up new avenues of research?
4. Up to which point does the problem of reliance on the data affect data-driven approaches?

The current study uses fully automatic data-driven measures to detect patterns in a first step, which we refer to as the diagnostic step. This involves automatic annotation for part-of-speech and syntactic dependency relation, and overuse metrics, i.e., statistical measures based on frequencies which deliver the most strongly overused patterns for the investigated periods. The most overused patterns are then treated as candidates for diachronic change, which need to be qualitatively interpreted in a second step. In a third step I then test, in classical corpus-based fashion, whether the patterns and my interpretation correspond to measurable and well-known differences in historical data, or if they can even deliver hitherto undescribed changes. I refer to this as the validation step.

I further address the question of which levels of representation are needed to automatically detect patterns: in particular, I test part-of-speech tags and grammatical dependency relations, and use historical data whose spelling variants have been mapped to PDE, also using automatic procedures. The automatic parser and tagger that I am using have been designed for PDE. The language before about 1500 can be expected to be too different from PDE to allow these tools to work reliably.

It is good practice in data science to keep the learning corpus and application corpus separate, in order not to over-adapt and reduce the risk of reporting idiosyncrasies in the data; this also minimises the risk of the findings being unrepresentative (see again Tognini-Bonelli's (2001: 88) comment quoted above). I learn the patterns from one corpus, the ARCHER corpus (Biber et al. 1994), and test their validity on the manually annotated Penn Parsed Corpora (Kroch et al. 2000; Kroch et al. 2004; Kroch et al. 2016), and the Penn Treebank (Marcus et al. 1993).

The paper is structured as follows. After describing the preparation of the data and introducing my methods in section 2, section 3 shows which features my data-driven step delivers. The following sections then take up the five strongest suggested trends, and compare them against the literature. Sections 4–8 then discuss the developments of participial clauses, verb-noun ratios, noun complexity, use of non-finite clause complementation, and information structure.

2.   Preparing the corpora

In this section, I first describe the corpora that I use, and then how I have prepared my data.

2.1   Data

As data, I use an automatically parsed, part-of-speech tagged and syntactically analysed version (Lehmann & Schneider 2012) of the 3.2 million word ARCHER corpus version 3.2 (Biber et al. 1994) for the data-driven step, and compare it to the manually annotated Penn Parsed Corpora (Kroch et al. 2000; Kroch et al. 2004; Kroch et al. 2016), to which I refer as *Penn historical* in the following, and the Penn Treebank (Marcus et al. 1993).

The manually annotated Penn historical corpora are an invaluable source for the description of linguistic change in English. They cover the period from 1100 to 1914. As I focus on a timeframe starting in 1500, I largely exclude the texts from before 1500. But as I wanted to include PDE, I was looking for comparable material for the 20th century. The University of Pennsylvania has also compiled and manually annotated the Penn Treebank (Marcus et al. 1993),[1] henceforth PTB, which consists of texts from the Wall Street Journal (WSJ) newspaper and the BROWN Corpus. PTB was the best option, given the fact that the Penn Treebank is the only available manually annotated PDE source with comparable annotation. Nevertheless, the comparison faces several challenges. First, the annotation is similar, but not identical, as I will demonstrate in the methods sections. Second, the genre of newspaper language in WSJ is so different from the genres encountered in the Penn historical corpora that we need to be very careful when interpreting the results. The genre-balanced BROWN Corpus seems a slightly better choice, but as the genre mix is still very different from that of the Penn historical corpora, the same caveat applies here too. On the positive side, the comparison between WSJ and BROWN can reveal how strong the influence of genre is, thus addressing RQ4.

I have split our data into seven time periods, largely following the splits of the Penn historical distribution: 1100–1500, 1500–1569, 1570–1639, 1640–1710, 1700–1800, 1800–1914, BROWN 1961 and WSJ 1989. I will only partly include the first period and the last two periods.

2.2   Methods of data preparation

I apply a three-step approach: first a data-driven diagnostic step, second my qualitative interpretation, and third a classical, data-based validation step. I start with a data-driven step, which aims to deliver the most salient changes between the early and the late periods of ARCHER.

---

1.   https://catalog.ldc.upenn.edu/LDC99T42

For this, I rely on overuse metrics, centrally O/E, Observed divided by Expected, where the Observed frequency comes from the subcorpus of interest (e.g., a certain period), and the Expected frequency is the frequency that we would find if the counts were homogeneously distributed across the entire corpus. If, for example, a subcorpus is half the size of the entire corpus, the subcorpus has 9 occurrences of a given phenomenon and the entire corpus 12, then O/E = (9/(12/2)) = 9*2/12 = 1.5. In other words, as the subcorpus is half the size of the entire corpus, we 'expect' to find 12/2 = 6 occurrences if the distribution is homogeneous. However, we 'observe' 9. O/E is thus 9/6 = 1.5, i.e., the phenomenon occurs 1.5 times as often as expected, it is overused by 50%.

I rely on a range of features, from lexical item to POS tag (using the Treetagger version of the Penn tagset, see Schmid 1994)[2] and POS frequency, and syntactic dependency label and direction (Schneider 2008). Those features which are strongly overused are good candidates for actual differences in the linguistic data. As most of my features are based rather on form than linguistic function, the form-function mapping is not direct. Each feature needs to be critically assessed in my second step. Sentences containing the features are inspected in the light of the research literature, and working hypotheses emerge. However, not all features lead to a clear hypothesis. Some overused features send ambiguous signals, or may stem from errors in the automatic analysis. For example, high frequency of the tag *VBG*(verb in the *-ing* form) can occur for several reasons: more progressive forms, but also more present participles can be an explanation. Manually inspecting a large set of sentences and other partly ambiguous signals often gives rise to clearer hypotheses. In order to keep interpretation manageable and overcome sparse data issues, I differentiate only two periods: an early period, consisting of the ARCHER texts from 1600 to 1800, and a late period, consisting of the ARCHER texts from 1900 to 1999. In order to compare the poles of the earliest and the latest data, and to compare two periods for which the size of the textual data is comparable, I have discarded the texts from 1800 to 1899.

In the third step, I validate the suggested hypotheses using the manually annotated Penn corpora. I have split the data into 7 periods, as seen in § 2.1, which allows us to follow increases, decreases and peaks with considerably finer granularity than in the binary division used for our diagnostic first step. The validity of step 1 partly depends on the accuracy of the pre-processing steps: the spelling normalisation tool and the automatic tagger and parser that I use. I give a brief evaluation of these tools in the following subsection.

---

2.   An overview of the tagset is also available at https://www.sketchengine.eu/english-treetagger-pipeline-2/

2.3   Evaluation of spelling normalisation, tagger and parser

Reliable results for the O/E method are only possible if the same spelling variants are used throughout ARCHER. I have therefore applied spelling normalisation, using the tool VARD2 (Baron & Rayson 2008), which maps historical variants to PDE variants. A detailed evaluation of VARD2 is given in Schneider et al. (2017) and I give a summary of the most important metrics in Table 1.

Table 1. Evaluation of VARD

	17thC	18thC	19thC
Precision	88.3%	83.7%	90.6%
Recall	74.6%	72.3%	47.5%
Unnormalised	94.8%	99.2%	99.4%
Normalised	98.1%	99.7%	99.6%

Table 1 shows that in the 17th century 94.8% of the word tokens are spelled as in PDE (row "Unnormalised"), while 5.2% are spelled differently and should thus be normalised. After the normalisation step, 98.1% of all tokens are spelled as in PDE, while 1.9% remain spelled differently or were mapped incorrectly to PDE. In terms of precision and recall, 88.3% of the changed variants are mapped correctly to PDE (precision), and 74.6% of the variants that should be mapped to a different PDE form are correctly mapped (recall). Precision measures how many of the instances reported by the system are true positives, i.e., correct. Recall measures how many of the true positives are actually reported by the system, i.e., actually found. The recall on the 19th century is very low, but this should be considered against the fact that only few word forms need to be changed (0.6%). The percentage of correctly normalised words still increases, from 99.4% to 99.6%.

Even after accurate spelling normalisation, the accuracy of tagging and parsing a historical corpus can still be affected by other phenomena, such as changed word order, word class probabilities that may change for individual words, or historical words that are not known in the tagger lexicon. Schneider et al. (2016) find that tagging accuracy drops from 95% on PDE to about 88% in the 17th century. Due to changes in capitalisation, a major source of error is the failure of the tagger to distinguish between proper name and other nouns. If the distinction between proper names and other nouns is left underspecified, the accuracy of the tagging rises to 90% for the 17th century, and higher in later centuries. Note that the parsing step will not be affected by the error of confusing proper names and other nouns – both types of nouns end up as NPs at the level of the phrase, which is the input to the parsing step.

The parser (Schneider 2008), too, performs less accurately on historical texts than on PDE texts. As parsing depends on spelling normalisation and POS tagging, due to error propagation, I expect slightly higher error rates than on the previous, lower-level annotation steps. As no new constructions have entered the grammar of English since 1600, the differences I would expect to find between the subperiods will mainly be differences of frequency, with all the major constructions in place. Rissanen (1999: 187) confirms that by about 1700 "the structure of the language was gradually established so that eighteenth-century standard written English closely resembles the present-day language". Indeed, the parser's grammar is able to recognise many of the syntactic phenomena that have become rare in PDE.
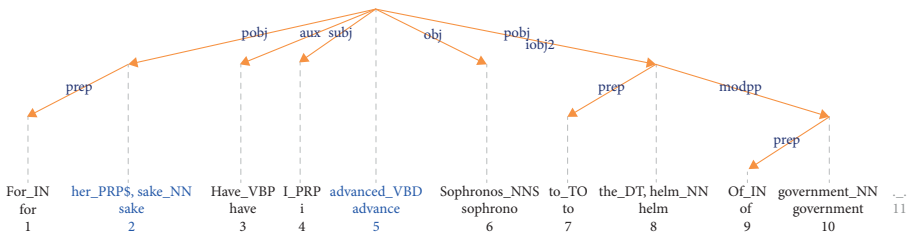
Figure 1. *For her sake Have I advanced Sophronos to the helm of government* (ARCHER 1628ford_d1b)

Figure 1 shows a syntactic analysis where the PP has been fronted and the auxiliary verb *have* has been moved outside the verb group, leading to subject-auxiliary inversion. Nevalainen (2006: 113) describes this non-canonical word order, which particularly often leads to a split clause, as follows:

> In the sixteenth century, inversion often took place after initial adverbial elements … . It was particularly common when the verb phrase consisted of an auxiliary and a main verb … . This pattern of inversion triggered by sentence-initial adverbials largely disappeared in the Early Modern English period.

For a full evaluation of the parser, see Schneider et al. (2015), who assess subject, object, PP-attachment (to which verb or noun, irrespective of argument or adjunct status) and sentence subordination relations, and note that F-score, a measure of parser accuracy, drops from 79% on PDE to 73% in the 17th century. This means that, in the earlier period, on average about every fourth instance of the evaluated relations is incorrect.

While such a high level of noise is problematic, two important characteristics need to be borne in mind. First, as long as errors are unsystematic, my aim of assessing the diachronic differences is not affected. Error analyses show that the vast majority of errors are unsystematic (Schneider 2008; Schneider et al. 2015). We could

see some systematic errors, though: most obviously, some relations have much lower accuracy than others. However, this is inconsequential for diachronic comparison: as long as the levels of accuracy do not change considerably, the level of noise stays similar and differences in frequencies also reflect differences in the signal.

Second, and this is in fact also a systematic error that I noticed, sentences with non-canonical word order are much more likely to be parsed incorrectly, which is of course unsurprising: if analyses involving canonical word order are possible, the parser will usually prefer them. This has the effect that the parser systematically underestimates the frequency of non-canonical word-order. In Figure 2, the PPs *with dance* and *with music* are attached incorrectly to preceding constituents (*with dance* is attached to *were; with music* is attached to *dance*), as the parser prefers to interpret *served* as a participle modifying the noun *music*. The effect of this tendency is that the reported frequencies of non-canonical word order (from the PDE perspective) are too low, the estimates thus too conservative. This increases the risk of low recall, and of type II errors in significance testing, but reduces precision and the risk of type I errors.
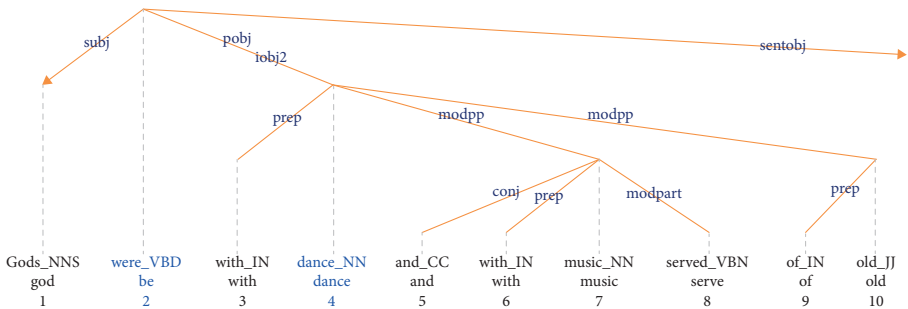
Figure 2. *Gods were with dance and with music served of old … (ARCHER 1613camp_d1b)*

The manually annotated Penn corpora, on which we test the diagnosed differences in ARCHER, have the advantage that they contain almost no parsing mistakes. I describe the method of accessing them in the following.

2.4   Querying the manually annotated corpora

The syntactic annotation of the Penn parsed corpora can be accessed via query languages that have been specifically developed for this purpose, with the aim of delivering the power of regular expression searches to syntactic data. I am using *Tregex* (Levy & Andrew 2006). It can be accessed equally via a text-based UNIX command line and a user-friendly graphical interface. Figure 3 shows a session with the graphical interface.

Figure 3. Query session with Tregex

While the query language is straightforward, there are a number of complications that need to be considered. I will discuss two.

First, there are often several ways in which a structure can be searched for, often with a considerable difference in reported hits. For the example of verb-attached PPs (§ 5.2) we could for example query for an IP that is modified by a PP; in tregex the query is */IP/ < /PP/*. This query returns 41555 hits in the period from 1800–1914. An alternative query is to find a main verb which has a PP sister node, in tregex */^V/ $ /PP/*. This query returns 35634 hits in the same period. The differences stem from the fact that the first query includes verbless clauses (IP-SMC, inflectional phrase – small clause), such as the one in Figure 4.
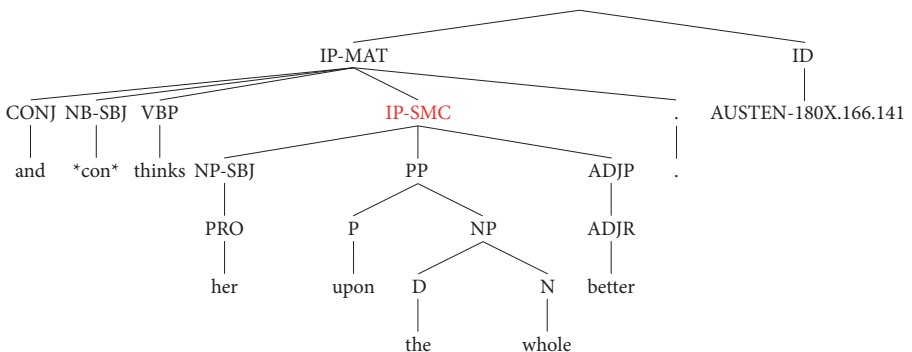
Figure 4. The verbless clause … *her upon the whole better* (AUSTEN-180X, 166.141. 1800–1914)

The second query also misses instances with BE and HAVE as main verb. This is because BE and HAVE do not have a tag starting with *V* (but with *B* and *H* instead). But including them in the query would lead to doublets: in Figure 5, the PP *at this time* is a sister both of the auxiliary (BEP) and the main verb (VAG).
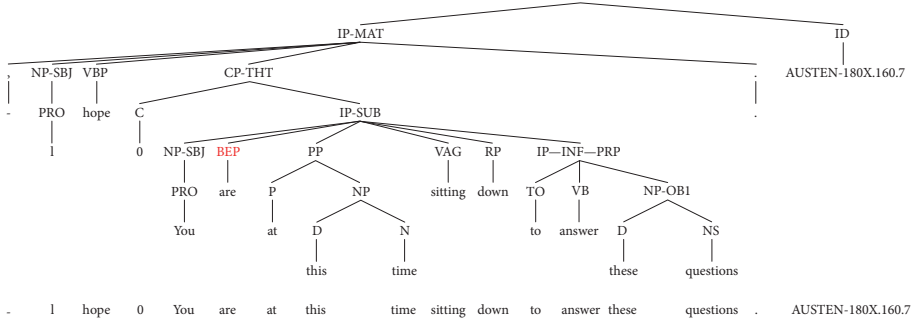
Figure 5. PP *at this time* which is both a sister of the auxiliary and the main verb (AUSTEN-180X.160.7, 1800–1914)

While the second query thus misses true positives, it also returns instances that the first query misses: for example, those in which the verb is a past participle heading an ADJP, as in the example in Figure 6.
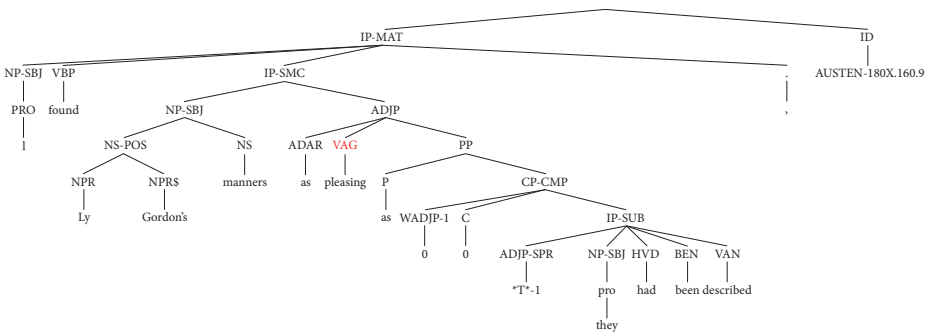
Figure 6. ADJP in *I found Ly Gordon's manners as pleasing as they had been described* (AUSTEN-180X,160.9, 1800–1914)

The second complication is that the PTB, BROWN and WSJ corpora are annotated with a slightly different annotation scheme from the one used in the Penn historical corpora. Some queries are equally applicable to all corpora, while others lead to different or no results. In our particular example, the PTB does not use an IP-style annotation, which is why the first of the suggested queries fails to return any hits. This is why I decided to use the comparable, second query for all corpora.

3.   Findings from the diagnostic step on ARCHER

I apply data-driven overuse measures to identify the quantitatively most salient differences between ARCHER early (1600–1800) and late (1900–1999) in order to find out if the differences correspond to reported changes in the literature (RQ1), or if they are due to corpus idiosyncrasies (RQ4) or pre-processing errors (RQ2). At the POS tag level, we obtain the differences given in Table 2. The data is sorted by decreasing O/E of the early period, which has the effect that the quantitatively strongest and thus characteristic features of the early period appear at the top of the table: they 'bubble up'. We give a manual comment in the last column.

For example, rows 1 and 2 show that participial adjunct clauses, and particularly free participial adjunct clauses, are more frequent in the early period: rows 4 and 19, following inspection of the examples, can be explained by the fact that full relative clauses are more frequent in the early period, as opposed to reduced relative clauses (Hundt et al. 2012). Rows 5, 8, and 16 indicate that the earlier period made more use of longer sentences and full subordinate clauses. Rows 3 and 7 are partly due to titles and polite addresses (*your highness, your most humble servant*) but this explanation probably does not explain the full effect. One further reason is that possessive pronouns and determiners were not mutually exclusive until EModE (Ibaraki 2009: 78).

In the early ARCHER material, we more frequently find the following article-pronoun combination, which is highly marked in PDE.

(1)   The cause of **this_DT our_PRP$** undertaking was your information.
$$\text{(ARCHER 1636heyl\_p1b)}$$

Row 6 (PDT) points to the related trend that, as multiple determiners are increasingly avoided, predeterminers like *all* and *such* have also reduced in frequency (Ibaraki 2009).

Rows 3 and 13 are due to elatives and the frequent use of *most* and *best* as adverbs, for example in the following sentence.

(2)   by virtue of which excellent airs, the sky got a **most** clear complexion, looked smug and smooth, and had not so much as a wart sticking on her face.
$$\text{(ARCHER 1603dekk\_p1b)}$$

That verbs seem to be more frequent (9, 10, 11) is partly explained by the fact that nouns are more frequent in the late period, as the same data, sorted by O/E of the late period, reveals. This is shown in Table 3, particularly in rows 10 and 12. Table 3 thus shows the opposite end of the list from Table 2: the quantitatively strongest features of the late period appear at the top of the table.

Table 2. Most salient tag-based features of the early ARCHER period

#	TAG	F(early)	F(late)	E	O/E(early)	O/E(late)	T(early)	T(late)	Example	Comment
1	VHG	1001	353	677	1.479	0.521	12.452	−12.452	**having** shewed	perfect progressive participial
2	VBG	2002	812	1407	1.423	0.577	15.862	−15.862	**being** attired	progressive passive participial
3	RBS	1263	644	953.5	1.325	0.675	10.023	−10.023	Most	elative, adverb
4	WDT	10094	5760	7927	1.273	0.727	24.339	−24.339	Which	full relative clause
5	:	21873	12701	17287	1.265	0.735	34.880	−34.880	:	complex clauses
6	PDT	2620	1559	2089.5	1.254	0.746	11.606	−11.606	All	predeterminers
7	PRP$	30792	18495	24643.5	1.249	0.751	39.167	−39.167	Your	possessive pronouns
8	,	102740	62145	82442.5	1.246	0.754	70.691	−70.691	,	complex clauses
9	VB	9430	6486	7958	1.185	0.815	16.501	−16.501	Be	Verb frequency
10	MD	19484	14165	16824.5	1.158	0.842	20.504	−20.504	Shall	Verb frequency, modality
11	VHN	286	208	247	1.158	0.842	2.482	−2.482	Had	Verb frequency
12	CC	52380	39847	46113.5	1.136	0.864	29.182	−29.182	And	paratactic style
13	JJS	2071	1621	1846	1.122	0.878	5.237	−5.237	Best	elatives, adverb
14	LS	1239	984	1111.5	1.115	0.885	3.824	−3.824	(list item)	
15	NNPS	2176	1749	1962.5	1.109	0.891	4.819	−4.819	(proper name)	(Capitalisation)
16	IN/that	9494	7661	8577.5	1.107	0.893	9.896	−9.896	That	full subordinate clauses
17	FW	341	279	310	1.100	0.900	1.761	−1.761	a priori	Latin, French
18	TO	35161	28827	31994	1.099	0.901	17.706	−17.706	To	Verb frequency
19	WP	6272	5185	5728.5	1.095	0.905	7.181	−7.181	Who	full relative clause
20	VVN	39437	32677	36057	1.094	0.906	17.800	−17.800	Eaten	Verb frequwncy

Table 3. Most salient tag-based features of the late ARCHER period

#	TAG	F(early)	F(late)	E	O/E(early)	O/E(late)	T(early)	T(late)	Example	Comment
1	UH	1328	2933	2130.5	0.6233	1.3767	−17.3862	17.3862	wow!	spoken-like
2	)	1769	3557	2663	0.6643	1.3357	−17.3242	17.3242	)	scientific
3	(	1763	3450	2606.5	0.6764	1.3236	−16.5218	16.5218	(	scientific
4	SENT	56682	87462	72072	0.7865	1.2135	−57.3265	57.3265	.	shorter sentences
5	RP	4342	6657	5499.5	0.7895	1.2105	−15.6084	15.6084	give up	verbal particle
6	VVZ	8104	12262	10183	0.7958	1.2042	−20.6023	20.6023	eats	not 'eateth'
7	EX	1673	2414	2043.5	0.8187	1.1813	−8.1960	8.1960	there	existentials
8	VVG	13439	18984	16211.5	0.8290	1.1710	−21.7751	21.7751	eating	progressive forms
9	CD	15743	21588	18665.5	0.8434	1.1566	−21.3912	21.3912	123	scientific
10	NNS	43871	54589	49203	0.8905	1.1095	−24.2812	24.2812	houses	noun compounds
11	VBN	2337	2825	2581	0.9055	1.0945	−4.8028	4.8028	been	perfect
12	NN	147474	176123	161798.5	0.9115	1.0885	−35.6117	35.6117	house	noun compounds

Table 3 also reveals shorter sentences (row 4), more spoken-language features (rows 1 and 5), increase of scientific style (rows 2, 3, 9), and more –*ing* forms (row 8, largely due to progressives). Row 6 is a side-effect of the fact that spelling normalisation sometimes fails. Verb forms ending in –*eth*, which are frequent in the early period, are not always correctly mapped to PDE, and thus not recognised as verbs in the third person singular (tag VVZ) by the tagger.

Summarising, we can discern the following five trends, which are well known in historical linguistics:

– **Participial clauses** seem to have been more frequent in the early period, according to ARCHER. Fonteyn & Cuyckens (2013) show that, in particular, free participial adjunct clauses were most frequent in the 17th century. We illustrate this trend using the Penn parsed corpora in § 4.
– **From verbal to nominal**: most changes testify to an enormous increase in nouns and complex NPs. The change from a more verbal to a more nominal style has been observed by Halliday (2004: 102–108), who compares the more scientific 'Attic' style to the plain, more verbal, 'Doric' style – "where the Attic is a nominal, the Doric is a clausal style" (Halliday 2004: 108). Leech et al. (2009) have shown that the trend towards nominal style continues, comparing LOB from 1960 to FLOB from 1990. I illustrate the development and measure the loss of creativity of a verbal construction in § 5.
– **Noun compression**: Attic style entails both complex NPs and more noun compounds. This style spreads from scientific discourse (Hundt et al. 2012) and press (Biber 2003) to other genres. In § 6, I show that compressing information forms a cascade from full to reduced relative clauses or PPs, and noun-noun constructions. I also trace individual lexicalisations.
– **From full clause to non-finite clause:** a part of this cascade is the move from full to non-finite clauses introduced by *to* and finally -*ing* forms, a trend described by Givón (1979) and tested on the Penn parsed corpora by Green (2017). I have added the PTB to check recent developments, and automatically annotated data to test them as a method, in § 7.

The suggested trends are also corroborated, and further extended, when considering the syntactic annotation. Table 4 shows the most overused dependency labels of the early period (at the top of the list) and the late period (at the bottom of the list). The syntactic overuse indicates trends that were not visible in our tag-based approach of Table 2. The high frequency of conjunctions (Table 2 row 12) could have been caused by nominal coordination, while the *compl* relation, which marks subordinated clauses, and *modrel*, which marks relative clauses, both point to clausal complexity. The higher frequency of *obj2* (second nominal object of ditransitive

verb) directly indicates changes in verb complementation that cannot be seen from part-of-speech tags, not even from sequences of tags.

Any errors at the tagging level are not only propagated at the level of syntactic annotation, but also scaled up (see the comments in brackets in the comments column of Table 4 below). The *strandprep* relation (stranded preposition) should find stranded prepositions, for example in relative clauses without pied-piping. But the relation massively overgenerates on some of the sentences from EModE. Many prepositions that cannot be integrated into the tree are attached to this label when the structure of the sentence gets too complex.

Table 4. Most overused dependency labels for the early period (top) and the late period (bottom) of ARCHER

Relation	Early	Late	E	O/E(early)	O/E(late)	Comment
(strandprep)	1648	974	1311	1.257	0.743	(error propagation)
comma	54548	33310	43929	1.242	0.758	complex long sentences
bridge	47529	33344	40436.5	1.175	0.825	fragments
obj2	1419	1026	1222.5	1.161	0.839	more ditransitive double noun constructions (§§ 5.2 & 5.3)
modrel	15548	11278	13413	1.159	0.841	relative clauses (§ 7)
aux	2758	2086	2422	1.139	0.861	verb group order (§ 8)
compl	19684	15924	17804	1.106	0.894	full subordinating clauses
conj	93226	78610	85918	1.085	0.915	rambling, paratactic style
…						
modpart	3316	3525	3420.5	0.969	1.031	reduced relative clauses (§ 7)
adj	2866	3154	3010	0.952	1.048	temporal adjuncts
(gen)	9	10	9.5	0.947	1.053	(error propagation)
(rquant)	71	84	77.5	0.916	1.084	(error propagation)
adjtrans	632	754	693	0.912	1.088	complex NPs (§ 6)
vpart	295	613	454	0.65	1.35	phrasal verbs
hyph	461	2041	1251	0.369	1.631	hyphenated words

Table 5 lists the top 4 overused relation+direction tuples, i.e., the combination of dependency label and direction.

Table 5. Most overused dependency label + direction tuples

Relation	Direction	Early	Late	E	O/E(early)	Comment
obj	←	12213	8129	10171	1.502	Bananas I like
sentobj	←	2085	1458	1772	1.43	Fronted subordinate clauses
pobj	←	6681	5781	6231	1.156	Fronted verb-attached PP
pobj	→	61419	54457	57938	1.128	Verb-attached PP

These findings indicate a further trend:

– **Fixation of word order**: the increasing fixation of word order turns out to be a further strong change. After Verb-Second (V2) is lost as a canonical word order, Verb-Second-like structures are used to mark various discourse functions like episode boundaries or introducing a conclusion in EModE. These relic structures mimic rather than copy the original construction (Los 2013) and retreat further. Fronting decreases, VO order at the expense of V2 increases, and fewer verbs are used ditransitively. In my annotation, verb-attached PPs refer to both arguments and adjuncts that are governed by a verb. The strong reduction in instances with non-canonical word order may also be due to these retreating V2-like structures. I will discuss these changes in § 8.

In order to verify whether the reported overuses stem from known changes in the development of English, I now consider further evidence. I will use a different corpus, the manually annotated Penn corpora, in the following sections. In data science terms we have learned features in step 1, and manually interpreted them in step 2, where we have seen that most of the trends correspond to actual and well-known developments. Now we need to evaluate them against a corpus which is less error-prone in step 3. In order to avoid overfitting, the learning corpus and the evaluation corpus should be different.

4.   Participial clauses

Lexical verbs in the *-ing* form are a strong feature of the late periods, where they are mostly used as progressives (Table 3 row 6). Furthermore, participial clauses on the whole increase in frequency or stay stable, as Table 6 shows. These trends seem to contradict our data-driven expectations (Table 2), where the tags VHG (*having*) and VBG (*being*) strongly decrease between the early and late ARCHER period, but agree with Fonteyn & Cuyckens (2013), who show that present participial free adjuncts were particularly frequent in EModE. My queries on the Penn corpora in Table 6 show that the data-driven method reported a strong trend for VHG and VBG because participials with *have* (*having seen*) and with *be* (*being*) were fashionable around 1600–1800, and then reduced again. The construction is also less frequent in earlier centuries. Table 6 gives the frequencies per million words (the particularly high values are in boldface).

Table 6. Frequencies of participial clauses (per million words)
(high values are in boldface)

Description	1100–1500	1500–1569	1570–1639	1640–1710	1700–1800	1800–1914	BROWN 1961	WSJ 1989
Participial Clause	1929.1	5425.5	**6934.4**	**9243.4**	**9204.0**	**9248.9**	4405.5	2125.5
Participial Clause with ING	1546.1	4001.2	4696.6	6098.7	6168.9	6434.5	3327.2	1659.2
Fronted Participial	327.5	1568.8	**2409.0**	**2136.4**	1702.6	913.9	871.0	383.5
Fronted Participial with 'be' or 'have'	15.7	65.4	133.8	165.3	150.3	97.0	36.9	12.8
Present Participial with 'have'	41.6	168.6	**401.5**	**526.8**	**572.0**	**455.1**	119.8	59.2
Present Participial with 'be'	9.3	445.5	**992.9**	**991.9**	**904.0**	561.4	46.1	8.9

An example of a participial with *have* is given in Figure 7, and a participial with *be* in Figure 8.

```
                              IP-MAT                                     ID
   ┌────┬──────┬─────────┬──────────┬───┬─────────────┬─────────┐
NP-SBJ VBD   NP-OB1    NP-MSR      ,      IP-PPL        .    BOSWELL-1776,51.562
  │     │    ┌───┐    ┌────┐            ┌────┬─────┬────────┐
*pro* Kept  D    N    Q    N     ,    HAG   VBN   NP-OB1      ;
            │    │    │    │          │     │    ┌─────┐
           the house all  day       having taken ADJ    N
                                                  │      │
                                                strong physic
```
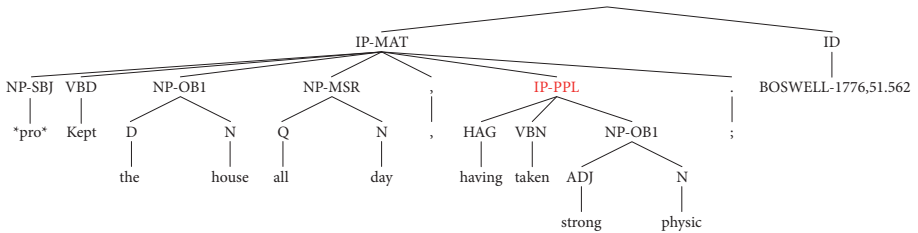
Figure 7. *Kept the house all day, having taken strong physic* (BOSWELL-1776, 51.562, 1700–1800)

```
                           IP-MAT                              ID
   ┌───────┬──────────────┬────┬───┬──────────┬─────┬──┐
NP-LFD-1 ,   IP-PPL       ,   RP   NP-SBJ-RSP=1 VBP  .   ARMIN-E2-H,14.226
   │        ┌────┬──────┐      │       │        │
  NPR   ,  BAG   ADJP    ,    up     PRO       comes  ;
   │        │     │             │
  Jack    bing   ADJ            he
                  │
                 dry
```
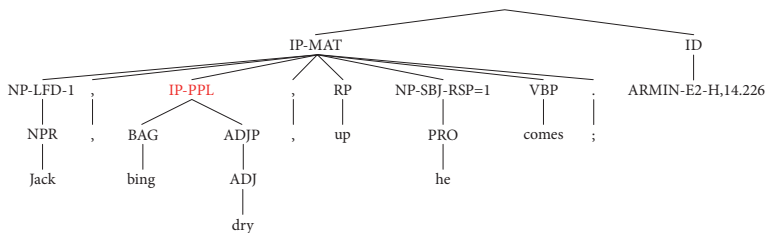
Figure 8. *Jack, being dry, up he comes* (ARMIN-E2_H, 14.226, 1570–1639)

In summary, the data-driven approach has shown us not only that progressives increase, but also that one specific *-ing* form, namely present participial free adjuncts, was particularly frequent in EModE, as argued by Fonteyn & Cuyckens (2013). While these findings are well-known, data-driven methods reliably detect them. I should add that only the extended version of the Penn Treebank tagset, which the Treetagger uses, manages to pick up this difference: in the standard Penn tagset, all verbs in the *-ing* form have the tag VBG (verb in the *-ing* form). Data-driven methods clearly depend on the granularity and quality of annotation.

5.   From verbal to nominal

My data from the parsed ARCHER corpus show a general decrease of verbs, and an increase of nouns, as Table 2 illustrates. This change has been described as being particularly evident in the 20th century (Mair et al. 2002; Leech et al. 2009), while Säily et al. (2011) report relative stability in noun frequency between 1415 and 1681. Let us consider the changes in more detail, using the Penn corpora.

5.1   Verb/noun ratio

That there is a relative decrease in the number of verbs, and an increase in the number of nouns, is also suggested by the data from the Penn corpora. Table 7 lists the frequencies, normalised per million words.

Table 7. Verb and Noun tags per million words, and their ratio

	1570–1639	1640–1710	1700–1800	1800–1914	BROWN 1961	WSJ 1989
Ratio of verbs to nouns	57.01%	57.81%	55.70%	50.66%	74.45%	42.70%
Verbs	119304	118542	114510	109338	165377	136046
Nouns	209277	205054	205567	215832	222146	318578

The linguistic change from verbal to more nominal is generally confirmed. The relative stability of the earlier periods (Säily et al. 2011) until about 1700 is also confirmed. Further, there seems to be enormous genre variation (BROWN is a balanced corpus, WSJ a newspaper corpus, see § 2.1), which partly explains why BROWN has many more verbs. A further problem is that there are major annotation differences between PENN historical parsed and PTB (in spite of their having been compiled at the same University). While nominalised *-ing* forms tend to be analysed as nouns in the latter, they are verbs in the former.

5.2   Verb and noun modification

A major reason for the increase in noun frequency is that there are more complex NPs (§ 6), which may also entail that verbs become less frequent, and that verb modification becomes less frequent. We test this hypothesis now. I checked adverbs, clausal modification (no major change in frequency, but see § 7), and phrasal modification (PP-attachment), which reduces in frequency, and verbal complementation (objects). Adverbs show no clear trend and no strong change.

Let us first discuss PP-attachment. PPs can modify either nouns or verbs.[3] The parsed ARCHER corpus reports more noun-PP, and fewer verb-PP in later periods. This is also confirmed in the PTB, but genre influence is also very strong: the largest difference here is between BROWN (balanced corpus) and WSJ (newspaper corpus). The general trend, however, is the same, and reveals a reduction in the ratio of verb-attached PPs to noun-attached PPs from almost 3 to 1. In our core period of investigation, from 1500 to 1914, we still find a reduction from 2.9 to 1.8, or from 1.68 to 1.46 in PPs attached in the canonical direction. The frequency of verb-attached PPs decreases, while that of noun-attached PPs increases. Table 8 gives the frequencies (per million words) and the ratio of verb-PP / noun-PP in the third and second last rows. The last row lists the percentage of verb-attached PPs that follow the verb, i.e., are in the canonical direction, a point to which we will return in § 8.

Table 8. Verb-attached PPs and noun-attached PPs (per million words); rows 1 to 4 are the absolute numbers (all verb attached PPs, verb-attached PPs to the right, noun attached PPs, noun-attached PPs with the preposition of); rows 5–7 are ratios

	1100– 1500	1500– 1569	1570– 1639	1640– 1710	1700– 1800	1800– 1914	BROWN 1961	WSJ 1989
verb-PP	78576	70226	72600	68285	73997	66460	56703	49040
verb-PP (where PP to the right)	59424	52115	53911	52599	58319	53815	49576	42151
noun-PP	27104	31036	29817	29460	33485	36953	36320	47629
noun-PP(with of)	23897	27021	24495	23198	28428	30766	21514	22837
vPP(to the right) / nPP	2.192	1.679	1.808	1.785	1.742	1.456	1.365	0.885
vPP / nPP	2.899	2.263	2.435	2.318	2.210	1.799	1.561	1.030
vPP(to the right) / vPP	0.756	0.742	0.743	0.770	0.788	0.810	0.874	0.860

---

3. I do not distinguish between arguments and adjuncts.

We next turn to verbal complementation. While V+obj remains quite stable, the frequency of V+obj+obj2 (ditransitive verb with double object) is reducing, as Table 9 shows. The strongest reduction in the frequency of the double object construction happens just after the Middle English period; frequencies pick up again in the course of Early Modern English (see also Zehentner 2019: 152), but drop after 1700. We can also observe that, contrary to the trend of a reduction in verb-attached PPs, the ditransitive pattern with Object and a PP introduced by *to*, given in the last line, remains frequent. The new ditransitive option thus is a successful pattern.

Table 9. Verb modification: Verb+Obj and Verb+double Obj (ditransitive) constructions, per million words

	1100–1500	1500–1569	1570–1639	1640–1710	1700–1800	1800–1914	BROWN 1961	WSJ 1989
V+obj	77858	55336	62404	60382	58277	53511	50857	49065
V+obj +obj2	3345	1958	2392	2287	1631	1294	1304	742
V+obj +to-PP	2879	2085	2464	3032	2887	2716	2018	2445

On the nominal side, I checked adjective modification (which increases, as Table 10 shows), noun compounds (Section 6), and relative clauses (Section 7).

Table 10. Noun modification: Adjective+noun constructions, per million words

	1100–1500	1500–1569	1570–1639	1640–1710	1700–1800	1800–1914	BROWN 1961	WSJ 1989
adj+n	30414	33098	30515	36092	36607	43823	44930	67892

5.3 Changes in productivity

We have seen that the ditransitive construction has reduced in frequency (see Table 9). The creativity of a construction can also be measured by its productivity: to how many types can it be applied? Structures that are no longer canonical often retreat to a few lexical items or fixed phrases, if they do not disappear altogether. The distribution of verb types in a corpus, as a consequence, only contains few types if creativity is low, and is often dominated by even fewer types. In terms of information theory, low creativity leads to a high degree of order, and low information.

For assessing creativity, entropy (Shannon 1951) is an ideal measure. Shannon Entropy (H) measures the degree of absence of order, that is, how unskewed a distribution is:

$$H = -\sum_{i=1}^{N} p_i * log_2(p_i)$$

The most easily interpretable measure is relative entropy = H / H_{max}, where H_{max} is the entropy if there is random distribution, i.e., $log_2(N)$.

$$H_{rel} = \frac{H}{H_{max}}, \; where \; H_{max} = log_2(N)$$

A perfectly random distribution has a relative entropy of 1; a distribution that is dominated by one variant has (almost) 0, and a Zipfian distribution about 0.7.

Relative entropy depends on the randomness of the distribution of types (if each type has equal token frequency, entropy is highest) but also the number of types (more types leads to higher entropy). We need to take into consideration the simple fact that the presence of more tokens typically also leads to more types and consequently an inflated measure of relative entropy. We have taken the same number of tokens for each period for the calculation of relative entropy, namely the top N for each period, where N is the frequency in the period with the lowest count: in our case, BROWN, with 566 occurrences of ditransitive verb constructions.

Retreating constructions have low relative entropy in their lexical slots, i.e., the set of lexical material that can be used in them is subject to strong constraints. The fact that this is the case here, with the verb types showing increasingly lower entropy, is confirmed in Table 11, which indicates a marked reduction in creativity. The ditransitive construction is dominated by *give* in PDE, and has retreated to few core types, while we find a wide range of verbs, for example *bear, cry, bid, forbid, put, set, bequeath, strike, enjoin, appoint* and *advise* in earlier periods.

Table 11. Relative entropy of heads of the ditransitive construction

	Types of verb	Tokens	TTR	Rel. Entropy	Rel. Entropy (first 566)	F per million	Words
1100–1500	1307	3310	2.533	0.8999	0.9420	3062.5	1080815
1500–1569	307	1075	3.502	0.8398	0.8622	1849.2	581325
1570–1639	288	1449	5.031	0.7565	0.8109	2254.9	642592
1640–1710	224	1184	5.286	0.7294	0.7777	2151.0	550446
1700–1800	137	699	5.102	0.7486	0.7493	1568.0	445786
1800–1914	128	661	5.164	0.7412	0.7537	1232.8	536173
BROWN 1961	201	**566**	2.816	0.7358	0.7412	1304.1	434002
WSJ 1989	144	1319	9.160	0.7509	0.7628	1300.3	1014344

However, a caveat must be added: some of the high levels of entropy in the early periods must be due to non-standardised spellings. Even so, higher frequencies of non-standard spellings cannot solely account for all of it, and manually sifting the lists reveals that the early periods do actually contain many more verb types. Some examples of verbs that occur as ditransitives in the earlier periods but do not in PDE are given below.

The verb *bear* has retreated to the fixed pattern *bear someone company*, while it used to be more productive, as in the example given in Figure 9.
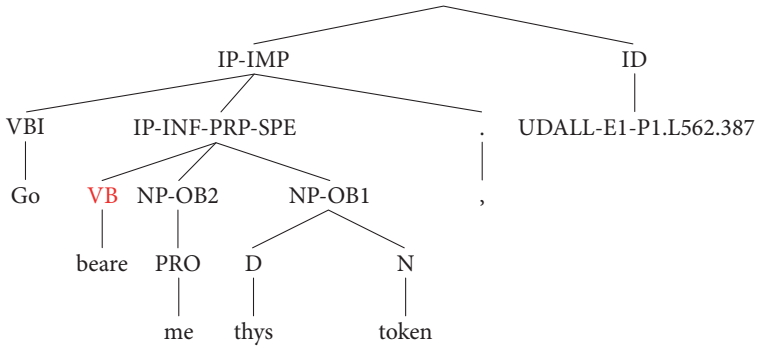
Figure 9. *Go beare me thys token* (UDALL-E1-P1,L562.387, 1500–1569)

A ditransitive use of *cry* is given in Figure 10.
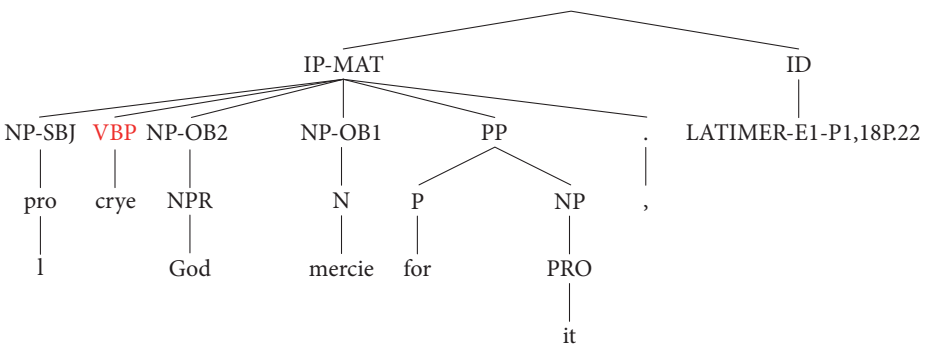
Figure 10. *I crye God mercie for it* (LATIMER-E1_P1,18P.22,1500–1569)

A ditransitive form of *advise* can be seen in Figure 11. The example illustrates full creative use of the form, first in a relative clause, in which the direct object is relativized, and secondly in a form in which the direct object is the free relative *That which hee aduiseth him*.
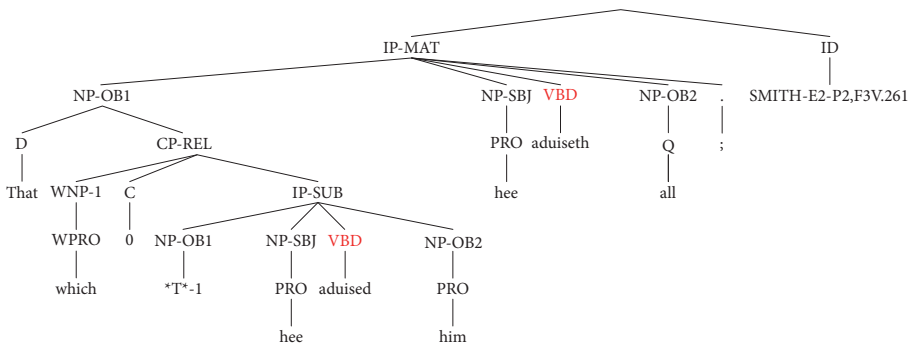
Figure 11. *That which hee aduised him he aduiseth all* (SMITH-E2-P2,F3V.261, 1570–1639)

While the verb construction of ditransitives has thus retreated to core types and fixed expressions, one relatively rare nominal construction has increased considerably, as we discuss in the following section.

6. Noun compression

The increase of the construction of noun compounds has been one of the quantitatively strongest changes in the 20th century, as described, for instance, in Leech et al. (2009). The change from a more verbal to a more nominal style (Halliday 2004: 102–108), comparing the earlier 'Doric' plain style and the later 'Attic' nominal style, entails more NPs as well as more complex NPs. His example of the Attic style is *Investment in a rail facility implies a long-term commitment* (Halliday 2004: 105). A more Doric version of such a sentence would be: *If you invest in a new facility for the railways you will be committing [funds] for a long term* (2004: 105). His example also includes a change from complex NP *facility for the railways* to the noun compound *rail facility*. Biber (2003) explains the change by the increasing demand for compression of information. The change is in agreement with information theory which says that, particularly in scientific and press genres, there is a constant trend towards shortening and abolishing redundant information. Once the thematic relations between the nouns in a noun-PP construction are clear, the preposition is redundant. This development is particularly spearheaded by one-off formations used as referring expressions (Halliday 2001: 185, Kastovsky 2006: 207), some of which make a career as neologisms. For the COHA corpus, the envelope of variation from PP (*Noun1 preposition Noun2*) to compound noun (*Noun2 Noun1*) for the most frequent types is shown in Schneider et al. (2017). In spoken language, there are equally strong cognitive trends militating against compression, as real-time planning and processing of very compressed NPs is taxing. This is why

spoken registers show a balance between areas of high and low information load, while scientific registers are dominated by compression. As an example, we show in Figure 12 the plots of probability for a word in its context (so-called *surprisal*, Levy & Jaeger 2007) from Schneider & Grigonyte (2018). The horizontal axis of this histogram gives the bigram surprisal (a higher value expresses a lower probability), the vertical axis the frequency of bigrams which are in the interval.
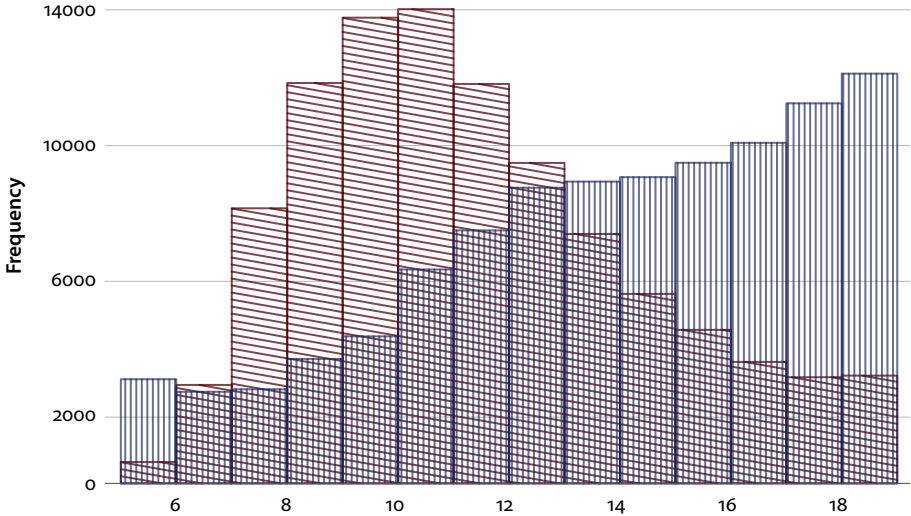
Figure 12. Surprisal of BNC spoken demographic (diagonal bars) compared to BNC pure science (vertical bars)

We have seen that the frequency of complex NPs increases according to my data-driven steps. The frequency of noun-noun sequences also increases markedly. In the period from 1950–1999, the relative frequency of two nouns in a sequence is four times higher than in the period from 1700–1750 (the numbers also show a slight increase before 1700, but this is mainly due to mistagging, particularly of capitalised verbs, which were quite frequent in that period, and never corrected by the spelling normalisation tool that I have used, and are hence mistagged as nouns). Thus, we need to compare these results to more reliable data, such as the manually annotated Penn corpora. The frequency of noun-noun modifications (per million word) is given in the third column of Table 12, confirming the strong increase (doubling from 1700–1800 to 1800–1914), but also a strong influence of genre: the balanced BROWN 1961 corpus shows an 5-fold increase compared to 1700–1750, while WSJ 1989, which contains newspaper language, a genre which can be expected to compress the most, shows a 15-fold increase.

The high frequency of noun compounds could be due to a few extremely frequent types. In order to assess whether the results are due to the construction itself changing, rather than to a few isolated lexical changes, we can measure creativity, for which I again use entropy. The relative entropy is given in the last but one column of Table 12. Relative entropy is calculated as explained in § 5.3. We can see that relative entropy increases considerably, particularly from 1700 onwards. It is tempting to rely solely on TTR (token per type), a well-known measure of vocabulary richness, which is given in column 5. These figures are misleading, however, as TTR heavily depends on the number of tokens, so the first requirement is that the comparison should be restricted to the same number of tokens (as I have done for relative entropy, see § 5.3, here 744 in the period from 1700–1800). But even then, TTR only considers the number of types, not patterns of dominance, i.e., whether just one, or very few types dominate the distribution. Entropy is therefore a much more informative measure. Relative entropy increases (the lowest point is also at about 1700) which indicates that the noun compound structure has become more creative. The surprisingly low entropy in WSJ is largely due to a high degree of repetition of technical terms from the financial section, headed by *rates, spending*, and *funds*. This is a reminder that WSJ contains only one genre, unlike the other corpora.

Table 12. Frequency and entropy of noun compounds; TTR = Token per type ratio

Source	Types of Noun2	Per million	Tokens	TTR	Rel. Entropy (first 744)	Words
1500–1569	573	2330.9	1355	2.365	0.889	581325
1570–1639	553	1736.7	1116	2.018	0.888	642592
1640–1710	461	1796.7	989	2.145	0.892	550446
1700–1800	427	1669.0	**744**	1.742	0.930	445786
1800–1914	897	3168.8	1699	1.894	0.960	536173
Brown 1961	2037	9449.3	4101	2.013	0.977	434002
WSJ 1989	3884	27069.7	27458	7.070	0.954	1014344

In summary, while noun compression is a well-known phenomenon, a data-driven approach showed us that it is one of the strongest changes quantitatively. Noun compression is a phenomenon known from scientific and press genres which has spread to other genres, thus also bringing shifts in register to the surface. The condensed style developed specifically for one text type (scientific/academic writing), in tandem with scientific discoveries, and a greater need to communicate in writing rather than in speech; this means that we have identified a very specific driver for language change in these periods. Entropy showed us that the complex NP construction not only became more frequent, but also more creative.

7.   From full clause to non-finite clause

To-infinitival clauses as verb complement came to be in competition with full clauses in Middle English (Los 2005) and continued to be so until the Late Modern English period (Rohdenburg 1995). This trend also shows up in our data, as in row 16 of Table 2, and the overused *compl* relation in Table 4. The rise of the gerund added a new type of non-finite clause to the existing complementation types, with some verbs – perception verbs like *see* and retrospective verbs like *remember* – increasingly complemented by gerunds (Fanego 2010; De Smet 2013). Mair (2006) shows that *remember* + gerund is increasing in Late Modern English. I illustrate the case of *remember* in Figure 13. Due to sparse data, I had to resort to COHA, which unfortunately only starts in 1800. The initial impression is that the change seems to consist of a two-stage cascade, in which the first stage (complementation by full clause) decreases, while the last stage (*-ing* complementation) increases, and the intermediate stage seems stable at first sight.
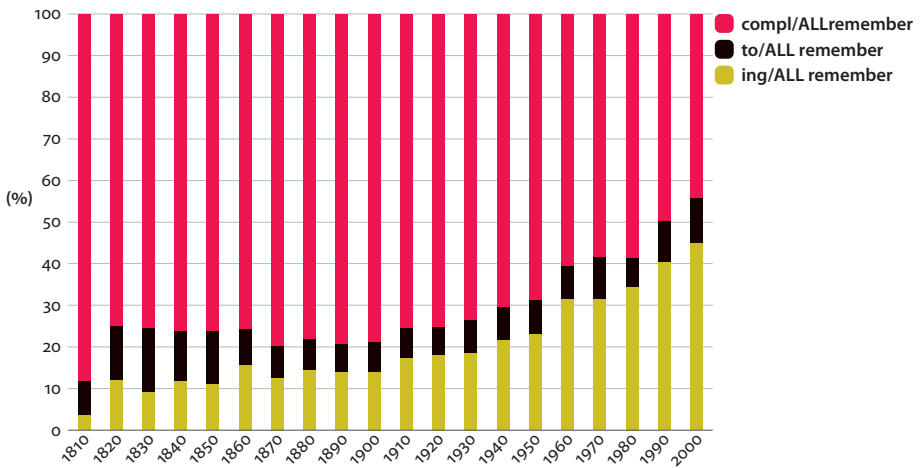
Figure 13.   Complementation of *remember* with full clause, *to*-infinitive and *-ing* form in COHA

On closer inspection, however, it turns out that the process observed with *remember* does not extend to other verbs – it may be an isolated lexical rather than a structural change. Extending our inspection to other verbs allows us to obtain enough data also from smaller corpora such as ARCHER, but the picture which emerges is different, or at least more complex.

Figure 14 shows a comparison of the surface frequencies of verb + *that*, verb + *-ing*-form, and verb + *to* in the ARCHER corpus. While the expected increase
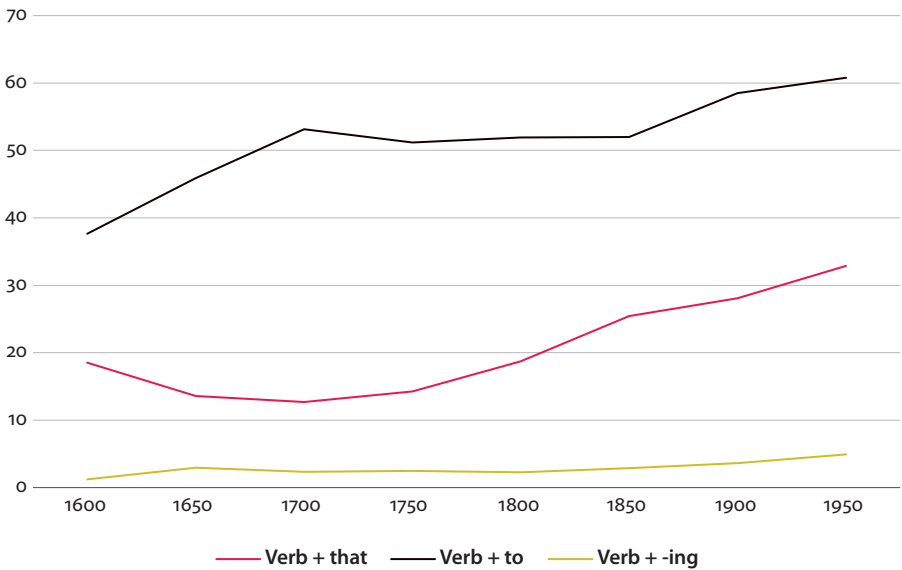
Figure 14. Development of the frequencies of complementation with full clause, non-finite clause with *to*, and *-ing* form, with a search at the level of surface sequence

of verb + *-ing*-form (which is 4 times as frequent in the latest than in the earliest period), and verb + *to* (which increases by 60%) can be observed, the frequency of verb + *that* only decreases until about 1700, and rises again afterwards. This suggests that the increase of non-finite forms does not occur at the expense of the finite forms in any direct way.

Different reasons can be found for the increase of verb + *that*. One is anticipated in Table 5: fronted subordinate clauses were far more frequent in earlier periods, which means that our query searching for surface sequences of verb followed by *that* cannot find them. Indeed, the frequency of the fronted subordinate clauses which are introduced by *that* reduces, from about 1.7 per 10,000 words in the 17th century to below 0.3 in the 20th century. A second reason is given in line 12 of Table 2: the frequency of the tag CC, coordinating conjunction, is one of the strongest features of the early period. While the earlier texts often have very long sentences, they are marked by a paratactic style with many clausal coordinations, while later texts often have a more hypotactic, subordinating style. Verbal conjunction and clausal subordination are compared at the level of the dependency syntax analysis in Figure 15. Verbal coordination reduces in frequency as the paratactic style decreases. In contrast, clausal subordination remains at high levels, as potential losses due to increases in nonfinite complementation are more or less compensated by a more hypotactic style.
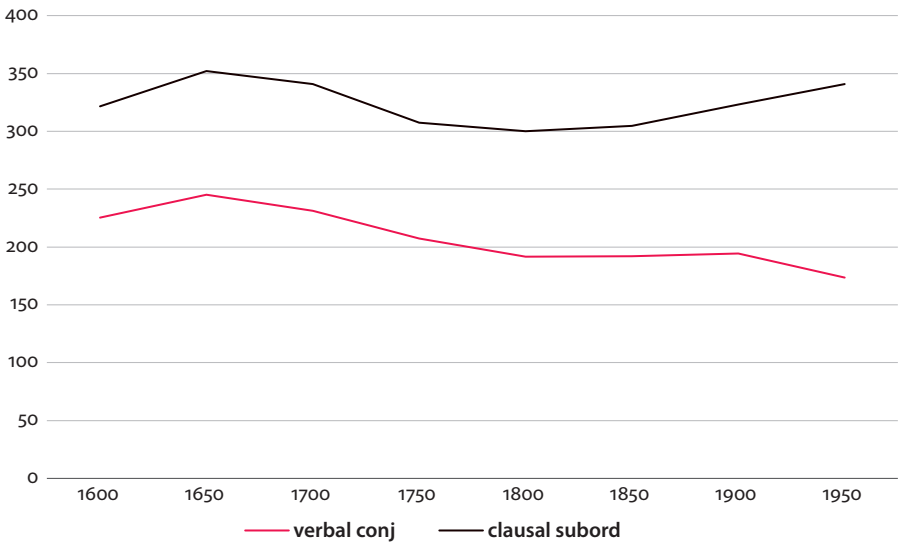
Figure 15. Verbal conjunction compared to clausal subordination, at the syntax level

At a more abstract level, we can see that trends in the data can lead to a situation in which the signals cancel each other out. Most forms can reflect more than one construction, so there is no 1:1 form-to-function mapping. Subordinated full clauses are, on the one hand, in competition with subordinated non-finite clauses, so their frequency is expected to decrease; on the other hand, they are also in competition with the paratactic, coordinating style, which leads to an increase. The two developments cancel each other out to some extent. Fortunately for our purposes, most changes also lead to several effects at the form level, which confirm some of the trends. My hypothesis that the paratactic style is showing a reduction is validated by the fact that coordination (tag CC) is a marked feature of the early ARCHER period. To see the bigger picture, then, we need to understand how the various features interact, and much depends on our qualitative interpretation of the results. In an optimistic interpretation, data-driven methods help us to deconstruct overly binary envelopes of variation. I will return to this point in § 8.

The relative increase we found of infinitival clauses and of present participle clauses in the Penn historical corpora is also confirmed in Green (2017: 97–98). Green (2017: 93) also reports findings suggesting a related trend, from full clause to a reduced clause or a phrasal construction. This may well be the same phenomenon that is reflected in rows 4 (*which*) and 16 (*that*) of Table 2. I give the frequencies of full compared to reduced relative clauses in Table 13. While there is a reduction of full relative clauses and an increase in reduced relative clauses in the ARCHER period, the picture is more complicated. First, the two earliest periods seem to pattern differently. Second, genre remains an important factor, as the marked difference

between the BROWN corpus and WSJ indicates. Hundt et al. (2012) have shown that the shift from full to the more packaged relative clauses is particularly characteristic of the scientific genre.

Table 13. Full relative clauses vs. reduced relative clauses, per million words

	1100–1500	1500–1569	1570–1639	1640–1710	1700–1800	1800–1914	BROWN 1961	WSJ 1989
Full relative clauses	15729.8	11231.2	11126.8	14224.8	11222.9	11278.1	9087.5	8433.0
Reduced relative (RRC)	1510.9	4174.9	3541.9	3722.4	3896.5	4330.7	1617.5	3575.7
RRC / Full relative	**0.096**	**0.372**	**0.318**	**0.262**	**0.347**	**0.384**	**0.178**	**0.424**

8.   Information structure

We have seen in § 5.2 (Table 8, last row) that the frequency of verb-PP not only reduces, but the constituent order is increasingly being fixed. This trend can already be observed in the ME period, see, e.g., Zehentner (2019: 175). Fronted PPs, i.e., PPs preceding their verbal attachment site, are growing rarer, reducing from 25% in the earliest period to 14% in the 20th century. The dramatic reduction of fronted objects and PPs leaves a strong signal in our data-driven investigation (Table 5, rows 1 and 3). Fixation of word and constituent order is a gradual process. We list several variable word-order phenomena in Table 14. Row 2 (subject inversion) in assertive clauses has retreated to a closed class of verbs, mainly utterances (e.g., X *said she*). The split form (row 3), in which the auxiliary verb appears before the subject while the main verb follows it (see § 2.3 and Nevalainen 2006: 113), is also reducing in frequency. At a higher level, we can also observe (last line) that there are fewer long-distance dependencies (i.e., trace-antecedent structures) overall in recent periods.

Table 14. Phenomena of variable constituent order, per million words

F per 1 million words	1100–1500	1500–1569	1570–1639	1640–1710	1700–1800	1800–1914	Latest/Earliest
subj-V	95461	73663	79724	72986	75579	70093	0.95
V before subj	6919	3881	2877	1324	1137	1238	0.32
split: AUX before SUBJ before V, not Q	729	506	302	189	195	168	0.33
Long-Distance Dependencies	39380	33126	31102	33473	28877	25283	0.76

These findings raise the question of what underlying forces there are behind these changes. One attractive explanation is that after English lost verb-second order, the function of the first position in the clause is gradually changing: while it could accommodate a wide variety of topicalised constituents, it is increasingly reduced to serving as a subject position (Los 2012, Los & Dreschler 2012).

> The crucial feature of verb-second syntax is not, in our view, the obligatory second position of the finite verb in the main clause, but the special character of the first position of that clause … . The movement of the finite verb merely serves to demarcate this special initial position … it offers a dedicated position for links to the immediately preceding discourse.   (Los 2012: 41)

Until about the fifteenth century, English was a verb-second language: the finite verb moved to second position while the first position accommodated topicalised constituents, which are often links to the previous discourse. After these links to given information are no longer available to be coded by PPs, there are fewer possibilities – only PPs expressing time-adjuncts or place-adjuncts can be used in unmarked constructions.

Information structure typically proceeds from given information to new information. After the first position is no longer available due to the shift from verb-second to SVO order, other devices are needed to order structures in such a way that given discourse participants can be placed before new ones. The prepositional variant of the dative shift is one such option. We have seen in § 5.2 and Table 9 that the prepositional variant, i.e., the option to use a ditransitive verb with an object and a to-PP, is quite successful. Unfortunately, the ditransitive is not directly marked, neither in the Penn Parsed corpora nor in the output of the Dependency parser. Verbs with object and a *to*-PP retrieve both proper ditransitives, e.g., *give it to John* and Small Clause constructions, e.g., *drive her to London*. This makes a direct comparison between object+*to*-PP and a double object construction difficult. Arppe et al. (2010) even doubt if the structures can be directly compared. Only a minority of the forms can be exchanged (*give a book to John/give John a book* is fine, but not *donate a painting to the museum*/*donate the museum a painting, or even *give John it), and a rich set of factors almost deterministically predict which form is used (Bresnan et al. 2007). Alternations such as the dative shift and noun compound vs. prepositional modification (§ 6) have come under criticism.

> Our focus on alternations is the result of theoretical heritage from generative syntax and a matter of methodological convenience. Most linguistic decisions that speakers make are more complex than binary choices … alternations are as simplistic and reductionistic as the theories of language that originally studied them.
>   (Arppe et al. 2010: 12)

Speakers, writers, listeners and readers do not only, or mainly, take decisions between alternable variants; it seems that the decisions are taken elsewhere. In fact, speakers take decisions at every word, or due to the use of prefabricated structures, at every couple of words. For this, entrenchment, information structure and further ordering principles play an important role.

While given-before-new is an important ordering principle, there are other principles, and they are related. Consider for instance the ordering principle of end-weight as a factor, which is also known as short-before-long (Behaghel 1909). This principle says that longer constituents tend to be placed towards the end of the clause. As given information is often realised with a pronoun, it will be difficult to separate out the effect of given-before-new and short-before-long. In the dative alternation, short-before-long is a driving factor (in addition to given-before-new): if the indirect object is short, the double object variant, which places the indirect object first, is preferred, while long indirect objects are usually expressed with the PP-variant placing them at the end.

Short-before-long can be reduced to a tendency that complements should be closer to their governor than adjuncts, and generally every constituent should be as close to its governor as possible (Futrell et al. 2015). This principle is also known by the name Dependency Length Minimisation,[4] a principle for minimising ambiguity, in line with Hawkins (1994, 2004). If minimal attachment is generally preferred, there are fewer attachment options and thus less scope for ambiguity: the less intervening material between a governor and the head of its dependent, the less scope for ambiguity. If every constituent is as close to the governor as possible, short-before-long emerges automatically. Additionally, complex constituents offer more attachment sites: if they are last in the clause (or first in a head-right language), their potential attachment sites cannot be used, which leads to less ambiguity. While these principles generally apply in all periods of English, Dependency Length Minimisation has been claimed to increase diachronically (Gulordava & Merlo 2015).

Let us consider the situation in our data. A comparison reveals that the average length of the dependency (measured in chunks) is indeed shorter in the late ARCHER period, supporting Gulordava & Merlo (2015). It looks as if the available ordering options were used in a way that more closely approaches the Dependency Length Minimisation Principle.

---

4.   As in head-right languages like Japanese a trend for long before short has been observed (Yamashita & Chang 2001), Dependency Length Minimisation is a more accurate description of the underlying forces.

Table 15. Average length of dependencies and chunks

ARCHER Period	Dependency Length (in chunks)	
	Incl. *bridge*	Excl. *bridge*
Early (1600–1799)	2.637	2.616
Late (1900–1999)	2.311	2.264

Dependency Length Minimisation seems to be observed better by PDE, which may lead one to the interpretation that English word ordering has become more efficient. In particular, the dependency between the auxiliary in V2 and the finite verb in VL could seem uneconomical, and add a further, direct argument for Dependency Length Minimisation. Gildea & Temperley (2010) point out that German, which has V2 syntax with the main verb in VL (akin to Old English), observes Dependency Length Minimisation to a much lesser degree than English. But an alternative explanation of the longer dependency length in earlier periods is that PDE grammars are written for SVO order, in which other orders, such as relic V2 structures, can only be expressed with longer dependencies and often movements, i.e., long-distance dependencies. This would also explain the last row of Table 14. Further research is needed to answer these questions.

9.   Conclusions

This study has presented and applied a three-step method to detect salient diachronic changes: first a data-driven diagnostic step on the automatically-parsed ARCHER corpus, second an interpretation of patterns in the light of previous research, and third a detailed confirmation step using the manually annotated Penn corpora. It is shown that the data-driven approach largely delivered reliable trends, and automatic parsing delivers reliable signals despite the higher error rate on historical data. Five salient diachronic changes could be confirmed and explored.

Taking up RQ1 (Can data-driven methods detect salient syntactic changes?) the answer is a clear yes. In Table 2, 19 out of the 20 top features are due to well-known diachronic developments. In Table 3, 11 out of the top 12 features equally point to actual changes. In Table 4, 12 out of the 15 strongest overuse features indicate linguistic developments, while 3 are due to systematic errors of the automatic analysis. The top four entries of Table 5 are all true positives.

In answer to RQ2 (What operations are needed to be successful in this task?) we can conclude the following. Spelling normalisation is essential, obviously for lexical changes, but also for morphosyntactic changes, as the quality of POS tagging

depends on the normalisation. The parsing step itself builds on POS tagging, thus normalisation is again crucial. It is also important to apply the data-driven methods at more than one level. For example, verb complementation (§ 5) and word order phenomena (§ 8) could only be detected by using syntactically annotated data. Also, the label set can be crucial: the distinction between participials and progressive forms only appears when using the more fine-grained Tree-Tagger tagset, and fixation of word order and information structures (§ 8) needs syntactic information which also includes dependency direction.

Concerning RQ3 (Can data-driven approaches bring up new avenues of research?) the answer is more debatable. On the one hand, all the phenomena mirrored in the data-driven features are well-known in diachronic linguistics, so the benefit may seem limited to confirming that previous research has not missed any important change. On the other hand, we could also quantify various phenomena against each other and obtain a more holistic view. We could for example collect further evidence that the loss of V2 has far-reaching consequences (§ 7), and that compression by noun compounds has spread from one small register to becoming one of the quantitatively strongest linguistic changes of the 20th century (§ 8). Data-driven methods thus offer an opportunity to re-assess and deconstruct interrelated changes (for example word order, fixation and information structure in § 8), to step from segmented individual changes to broader perspectives, and to pay tribute to the patterns in the spirit of Sinclair & Carter (2004). This can only work, however, if careful qualitative interpretation is used, checking candidates with traditional data-based methods and critically considering related issues and alternative explanations. The upshot of the highly quantitative data-driven method is that the qualitative step is at least equally important. The two have just reversed their sequence (Hilpert & Gries 2017): in corpus-based, hypothesis-driven research the qualitative step triggers the quantitative one, while in corpus-driven approaches the quantitative step comes first, and the patterns then need to be interpreted qualitatively. Furthermore, the fact that form-function mappings are not 1:1 further complicates the analysis. A further avenue of research is that data-driven methods involve probabilistic language models. Language models, which include word and tag-sequence models, taggers, surprisal, parsers, etc. allow one to make predictions beyond individual constructions and can thus also pave the way leading from envelopes of variation to a global perspective.

As a next point, let us address RQ4 (Up to which point does the problem of reliance on the data affect data-driven approaches?). First, the quality of automatic annotation affects results. In my answer to RQ1 I argued that this influence is relatively small, but depends on data preparation, such as spelling normalisation. Second, the data itself may reflect side effects, for example genre rather than diachronic change.

This is indeed a serious problem. However, data-driven and data-based approaches are equally affected, and I think that the important interpretative step actually gives one a good opportunity to detect and assess side-effects. While Tognini-Bonelli (2001: 88) points out the risks and argues that they may be more affected, we also see an opportunity: the global perspective which data-driven methods offer, and also force us to take in order to interpret the complex feature lists, help us to detect skews in the data and correlated features.

Further, we should note that the reliance on a particular annotation set may turn out to be as big a problem as genre differences. For instance, only the extended version of the Penn Treebank tagset, which the Treetagger uses, picked up participials (§ 4). Data-driven methods profit from, but equally depend on, the granularity and quality of annotation. A practical answer is that it is vital to include several levels of annotation: I have used POS tags and syntax.

Finally, the mapping difficulties between slightly differing annotations, in our case Penn historical versus Penn Treebank, leads to serious challenges. This issue needs to be addressed in further research. As an additional outlook, I conclude that automatically parsed large data can get one further due to the reduction of the sparse data problem. I will apply these methods to larger corpora, such as the COHA corpus and earlier material such as the CLMET corpus.

Possible shortcomings of this study include error rates and error propagation at any level. More advanced annotation tools of the future may reduce these error rates. Furthermore, I have relied on the information theoretic measure O/E and on entropy. Additional metrics and methods can probably reveal further differences.

References

Ananiadou, Sophia, Douglas B. Kell & Jun-ichi Tsujii. 2006. Text mining and its potential applications in systems biology. *Trends in Biotechnology* 24(12). 571–579. https://doi.org/10.1016/j.tibtech.2006.10.002

Arppe, Antti, Gaëtanelle Gilquin, Dylan Glynn, Martin Hilpert & Arne Zeschel. 2010. Cognitive corpus linguistics: Five points of debate on current theory and methodology. *Corpora* 5(1). 1–27. https://doi.org/10.3366/cor.2010.0001

Baron, Alistair & Paul Rayson. 2008. VARD 2: A tool for dealing with spelling variation in historical corpora. *Proceedings of the Postgraduate Conference in Corpus Linguistics*. Birmingham, UK: Aston University.

Behaghel, Otto. 1909. Beziehungen zwischen Umfang und Reihenfolge von Satzgliedern. *Indogermanische Forschungen* 25. 110–142.

Biber, Douglas. 2003. Compressed noun-phrase structures in newspaper discourse: The competing demands of popularization vs. economy. In Jean Aitchison & Diana Lewis (eds.), *New media language*, 169–181. London: Routledge.

Biber, Douglas, Edward Finegan & Dwight Atkinson. 1994. ARCHER and its challenges: Compiling and exploring a representative corpus of historical English registers. In Udo Fries, Peter Schneider & Gunnel Tottie (eds.), *Creating and using English language corpora: Papers from the 14th International Conference on English Language Research on Computerized Corpora, Zurich 1993*, 1–13. Amsterdam: Rodopi.

Bresnan, Joan, Anna Cueni, Tatiana Nikitina & Harald Baayen. 2007. Predicting the dative alternation. In Gerlof Bouma, Irene Krämer & Joost Zwarts (eds.), *Cognitive foundations of interpretation*, 69–94. Amsterdam: Royal Netherlands Academy of Science.

De Smet, Hendrik. 2013. *Spreading patterns: Diffusional change in the English system of complementation* (Oxford Studies in the History of English). New York: Oxford University Press.

Fanego, Teresa. 2010. Variation in sentential complements in eighteenth- and nineteenth-century English: A processing-based explanation. In Raymond Hickey (ed.), *Eighteenth-century English: Ideology and change*, 200–220. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511781643.012

Fonteyn, Lauren & Hubert Cuyckens. 2013. The development of free adjuncts in English and Dutch. *Leuven Working Papers in Linguistics* 2. 160–195.

Futrell, Richard, Kyle Mahowald & Edward Gibson. 2015. Large-scale evidence of dependency length minimization in 37 languages. *Proceedings of the National Academy of Sciences of the USA* 112(33). 10336. https://doi.org/10.1073/pnas.1502134112

Gildea, Daniel & David Temperley. 2010. Do grammars minimize dependency length? *Cognitive Science* 34(2). 286–310. https://doi.org/10.1111/j.1551-6709.2009.01073.x

Givón, Talmy. 1979. *On understanding grammar*. New York: Academic Press.

Green, Clarence. 2017. *Patterns and developments in the English clause system*. Singapore: Springer Nature. https://doi.org/10.1007/978-981-10-2881-6

Gulordava, Kristina & Paola Merlo. 2015. Diachronic trends in word order freedom and dependency length in dependency-annotated corpora of Latin and Ancient Greek. In *Proceedings of the 3rd International Conference on Dependency Linguistics (Depling 2015)*, 121–130. Uppsala: Uppsala University.

Halliday, Michael A. K. 2001. Literacy and linguistics: Relationships between spoken and written language. In Anne Burns & Caroline Coffin (eds.), *Analyzing English in a global context: A reader*, 181–193. New York: Routledge.

Halliday, Michael A. K. 2004. *The language of science*. Collected works of M. A. K. Halliday, Volume 5 (Jonathan J. Webster, ed.). London: Continuum.

Hawkins, John A. 1994. *A performance theory of order and constituency*. Cambridge: Cambridge University Press.

Hawkins, John A. 2004. *Efficiency and complexity in grammars*. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199252695.001.0001

Hilpert, Martin & Stefan Th. Gries. 2017. Quantitative approaches to diachronic corpus linguistics. In Merja Kytö & Päivi Pahta (eds.), *The Cambridge handbook of English historical linguistics*, 36–53. Cambridge: Cambridge University Press.

Hundt, Marianne, David Denison & Gerold Schneider. 2012. Retrieving relatives from historical data. *Literary and Linguistic Computing* 27(1). 3–16. https://doi.org/10.1093/llc/fqr049

Ibaraki, Seishirou. 2009. The development of the determiner system in the history of English. *English Linguistics* 26. 67–95. https://doi.org/10.9793/elsj.26.1_67

Kastovsky, Dieter. 2006. Vocabulary. In Richard Hogg & David Denison (eds.), *A history of the English language*, 199–270. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511791154.005

Kroch, Anthony, Ann Taylor & Beatrice Santorini. 2000. *The Penn-Helsinki Parsed Corpus of Middle English (PPCME2)*. Department of Linguistics, University of Pennsylvania. CD-ROM, 2nd edn., release 4.

Kroch, Anthony, Beatrice Santorini & Lauren Delfs. 2004. *The Penn-Helsinki Parsed Corpus of Early Modern English (PPCEME)*. Department of Linguistics, University of Pennsylvania. CD-ROM, 1st edn., release 3.

Kroch, Anthony, Beatrice Santorini & Ariel Diertani. 2016. *The Penn Parsed Corpus of Modern British English (PPCMBE2)*. Department of Linguistics, University of Pennsylvania. CD-ROM, 2nd edn., release 1.

Leech, Geoffrey, Marianne Hundt, Christian Mair & Nicholas Smith. 2009. *Change in contemporary English: A grammatical study*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511642210

Lehmann, Hans Martin & Gerold Schneider. 2012. BNC Dependency Bank 1.0. In Signe Oksefjell Ebeling, Jarle Ebeling & Hilde Hasselgård (eds.), *Studies in variation, contacts and change in English, Volume 12: Aspects of corpus linguistics: compilation, annotation, analysis*. Helsinki: VARIENG.

Levy, Roger & Galen Andrew. 2006. Tregex and Tsurgeon: Tools for querying and manipulating tree data structures. In *Proceedings of the 5th International Conference on Language Resources and Evaluation (LREC 2006)*. Genoa, Italy: European Language Resources Association.

Levy, Roger & Florian T. Jaeger. 2007. Speakers optimize information density through syntactic reduction. In Bernhard Schölkopf, John C. Platt & Thomas Hofmann (eds.), *Proceedings of the Twentieth Annual Conference on Neural Information Processing Systems*, 849–856. Cambridge, MA: MIT Press.

Los, Bettelou. 2005. *The rise of the to-infinitive*. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199274765.001.0001

Los, Bettelou. 2012. The Loss of verb-second and the switch from bounded to unbounded systems. In Anneli Meurman-Solin, Maria José López-Couso & Bettelou Los (eds.), *Information structure and syntactic change in the history of English* (Oxford Studies in the History of English), 21–46. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199860210.003.0002

Los, Bettelou. 2013. Recycling "junk": A case for exaptation as a response to breakdown. In Ritsuko Kikusawa & Lawrence A. Reid (eds.), *Historical linguistics 2011: Selected papers from the 20th International Conference on Historical Linguistics, Osaka, 25–30 July 2011*, 267–288. Amsterdam: John Benjamins. https://doi.org/10.1075/cilt.326.18los

Los, Bettelou & Gea Dreschler. 2012. The loss of local anchoring: From adverbial local anchors to permissive subjects. In Terttu Nevalainen & Elizabeth C. Traugott (eds.), *The Oxford handbook of the history of English* (Oxford Handbooks in Linguistics), 859–871. New York: Oxford University Press.

Mair, Christian. 2006. The 19th century as a critical period in the formation of the modern English system of nonfinite complement clauses: The case of remember. In Merja Kytö & Erik Smitterberg (eds.), *Nineteenth-century English: Stability and change*, 215–228. Cambridge: Cambridge University Press.

Mair, Christian, Marianne Hundt, Geoffrey N. Leech & Nicholas Smith. 2002. Short term diachronic shifts in part-of-speech frequencies: A comparison of the tagged LOB and F-LOB corpora. *International Journal of Corpus Linguistics* 7(2). 245–264. https://doi.org/10.1075/ijcl.7.2.05mai

Marcus, Mitchell P., Beatrice Santorini & Mary Ann Marcinkiewicz. 1993. Building a large annotated corpus of English: The Penn Treebank. *Computational Linguistics* 19. 313–330.

Nevalainen, Terttu. 2006. *An introduction to Early Modern English*. Edinburgh: Edinburgh University Press.

Rissanen, Matti. 1999. Syntax. In Roger Lass (ed.), *The Cambridge history of the English language, Volume 3: Early Modern English 1476–1776*, 187–331. Cambridge: Cambridge University Press.

Rohdenburg, Günter. 1995. On the replacement of finite complement clauses by infinitives in English. *English Studies* 76. 367–388. https://doi.org/10.1080/00138389508598980

Säily, Tanja, Terttu Nevalainen & Harri Siirtola. 2011. Variation in noun and pronoun frequencies in a sociohistorical corpus of English. *Literary and Linguistic Computing* 26. 167–188. https://doi.org/10.1093/llc/fqr004

Schmid, Helmut. 1994. Probabilistic part-of-speech tagging using decision trees. In *Proceedings of the International Conference on New Methods in Language Processing, Manchester, UK*, https://www.cis.uni-muenchen.de/~schmid/tools/TreeTagger/data/tree-tagger1.pdf

Schneider, Gerold. 2008. Hybrid long-distance functional dependency parsing. Zurich: University of Zurich Doctoral thesis.

Schneider, Gerold & Gintare Grigonyte. 2018. From lexical bundles to surprisal and language models: Measuring the idiom principle on native and learner language. In Joanna Kopaczyk & Jukka Tyrkkö (eds.), *Applications of pattern-driven methods in corpus linguistics*, 15–55. Amsterdam: John Benjamins. https://doi.org/10.1075/scl.82.02sch

Schneider, Gerold, Marianne Hundt & Rahel Oppliger. 2016. Part-of-speech in historical corpora: Tagger evaluation and ensemble systems on ARCHER. In *Proceedings of KONVENS 2016*. Bochum, Germany: Bochumer Linguistische Arbeitsberichte.

Schneider, Gerold, Hans Martin Lehmann & Peter Schneider. 2015. Parsing Early Modern English corpora. *Digital Scholarship in the Humanities* 30(3). 423–439. https://doi.org/10.1093/llc/fqu001

Schneider, Gerold, Eva Pettersson & Michael Percillier. 2017. Comparing rule-based and SMT-based spelling normalisation for English historical texts. In Gerlof Bouma & Yvonne Adesam (eds.), *Proceedings of the NoDaLiDa 2017 Workshop on Processing Historical Language*, 40–46. Linköping: Linköping University Electronic Press.

Shannon, Claude E. 1951. Prediction and entropy of printed English. *The Bell System Technical Journal* 30. 50–64. https://doi.org/10.1002/j.1538-7305.1951.tb01366.x

Sinclair, John McHardy & Ronald Carter. 2004. *Trust the text: Language, corpus and discourse*. Routledge, London. https://doi.org/10.4324/9780203594070

Tognini-Bonelli, Elena. 2001. *Corpus linguistics at work*. Amsterdam: John Benjamins. https://doi.org/10.1075/scl.6

Yamashita, Hiroko & Franklin Chang. 2001. 'Long before short' preference in the production of a head-final language. *Cognition* 81. B45–B55. https://doi.org/10.1016/S0010-0277(01)00121-4

Zehentner, Eva Maria. 2019. *Competition in language change: The rise of the English dative alternation*. Berlin: De Gruyter. https://doi.org/10.1515/9783110633856

CHAPTER 8

Prepositions in Early Modern English argument structure and beyond

Eva Zehentner and Marianne Hundt
University of Zurich

This paper is the first to use a bottom-up, corpus-based, exploratory approach
to the full range of prepositions in Early Modern English argument structure.
Contrary to what previous research leads us to expect, the overall token fre-
quency of prepositions during this period decreases, and they are not always suc-
cessful against the older NP-variants. Similarly, our case study challenges earlier
suggestions that PP-complements are particularly frequent in second-language
varieties of English. With respect to the functions taken on by PPs in the clause,
however, we provide preliminary evidence that more complement-like uses in-
crease at the expense of more adjunct-like arguments, i.e., that PPs become more
important as core elements of the clause.

Keywords: prepositions, verbal argument structure, bottom-up approach, Early
Modern English, World Englishes

1. Introduction

Prepositions are an integral part of Present-Day English (PDE). Among other
things, they play an important role in the system of verbal complementation: not
only are they used to express adverbials of instrument, location or time (*John wrote
the letter **with a pencil in Rome on Monday***), but PPs also frequently mark the direct
or indirect objects of verbs (*John relied **on his mother** to give the book **to Mary***).
Accordingly, their semantic and syntactic features have received much attention
(e.g., Hoffmann 2005, 2011).

Diachronically, prepositions are interesting as they are claimed to feature
prominently in the typological shift from a more synthetic to a largely analytic
language, which English has undergone. According to previous research, prep-
ositional phrases (PPs) have over time greatly increased in terms of their token
frequency and with respect to the different functions they serve (e.g., Mustanoja

1960: 348 or Traugott 1972: 127; also Lundskær-Nielsen 1993: 113–115). Typically, this functional extension is said to put the new PP-patterns in competition with resident NP-patterns, as illustrated in (1) and (2).

(1) a. *þæs geares*_{NP} *þe   Crist   <u>acenned</u> <u>wæs</u>*
       the year   REL Christ born   was
       "[in] the year that Christ was born"
                                                    (ÆCHom i.80.30; Mitchell 1985: 586)

    b. <u>*Com*</u> *on wanre   niht*_{PP} *scriðan sceadugenga*
       came on gloomy night   gliding shadowgoer
       "there came gliding on a gloomy night the one who walks in darkness"
                                    (Beo. Th. 1410; B. 703; Bosworth-Toller 2010, s.v. *sceadu-genga*)

(2) a. *he lytle werede*_{NP} *unieþelice   æfter wudum <u>for</u>*
       he little troop   with difficulty after woods   went
       "he travelled [with] a small troop with difficulty through woods"
                                                    (ChronA 74.28 (878); Sato 2009: 32)

    b. *Ða   bebead   se fæder ðæm consule ðæt he mid his fierde*_{PP}
       then commanded the   father the   consul that he with his army
       angean <u>fore</u>
       back   go
       "then the father commanded the consul that he should go back with his army"   (Ors. 3, 10; Bosworth-Toller 2010, s.v. *mid*)

As shown in (1a)–(b), adverbials of time, but also place or manner, as well as relations such as accompaniment (2a)–(b) could be encoded either by an NP or a PP in Old English (OE) (Traugott 1992: 207). Preliminary evidence on PDE has been taken to suggest that PP-adverbials greatly increased at the expense of the nominal patterns over time, eventually taking over almost entirely – today, NP-adverbials are exceptional. Similarly, more complement-like uses of PPs already existed in OE, but were still less frequent and more restricted than other (mostly nominal) options (Mitchell 1985: 512; Allen 1995: 217–219). For instance, although a number of transitive (3) and ditransitive (4) verbs could already occur with PP- instead of NP-objects, this possibility was limited to individual verbs or specific verb classes: In the case of ditransitives, it was only verbs of caused motion such as *send* or *bring* or verbs of communication (*talk, say*) which appeared in the now well established *to*-prepositional pattern (see, e.g., De Cuypere 2015b: 3–7).

(3) a. *Ne   <u>gefeah</u>   he þære fæhðe*_{NP}
       NEG rejoiced he the   feud
       "He did not rejoice in this feud"
                                                    (Beowulf i, 109; Bosworth-Toller 2010, s.v. *fæhþ*)

    b.   *Ðonne motan we **in ðære engellican blisse**$_{PP}$ gefeon*
         Then   must   we in the   angelic   bliss   rejoice
         "then may we in this angelic bliss rejoice"
                  (Blickl. Homl. 83, 3; Bosworth-Toller 2010, s.v. *gefēon*)

(4) a.   <u>*Sende ic **Wylfingum**$_{NP}$*</u> *… ealde madmas*
         sent I Wyflings   old   treasures
         "I sent the Wylfings ancient treasures"
               (Beo. Th. 946; B. 471.; Bosworth-Toller 2010, s.v. *sendan*)
    b.   <u>*sende*</u> *his gewrit **to þam wælhreowan casere**$_{PP}$*
         sent his letter to the   cruel   emperor
         "sent his letter to the cruel emperor"
         (coaelive,ÆLS_[Julian_and_Basilissa]: 249.1090; De Cuypere 2015b: 17)

Many constraints on complementation with PPs were increasingly lost from Middle English (ME) onwards, and PP-complements became more frequent. In some cases, both NP- and PP-variants survive. A prime example of this is the PDE dative alternation featuring variation between NP (*John gave **Mary** a book*) and PP (*John gave a book **to Mary***) in the recipient role. Similarly, there is still variation with transitive verbs (e.g., the co-existence of *believe*+NP and *believe*+PP in PDE). On the other hand, PPs often came to eventually oust the non-prepositional patterns (e.g., *rejoice+NP vs. *rejoice*+PP; *wonder+NP vs. *wonder*+PP). This is especially visible in instances of tight verb-preposition combinations such as *rely on* or *consist of*, typically referred to as 'prepositional verbs'. These patterns first appeared in the 14th century and have gained in frequency since (Denison 1981; Claridge 2000). In PDE, the PP is obligatory, and intransitive uses are ungrammatical (*He relied Ø).

   The emergence of new prepositional verbs in (non-)standard varieties of English such as (5) indicates that the process of extension is still ongoing. Likewise, PP-complements are said to be particularly productive in institutionalised second-language varieties (ESL) of English, as reported by, e.g., Nesselhauf (2009) and illustrated in (6).

(5) a.   the beds were so narrow that one would <u>fall</u> **off of them**$_{PP}$…
                                    (2015; COCA, FICTION)
    b.   On Friday around 6 p.m., Mark Sturgis, 22, told his father that he was going to <u>meet</u> **up with his longtime friend and co-worker**$_{PP}$, Tom Meo, 21.
                                    (2017; COCA, NEWS)

(6) a.   First, I would be <u>explaining</u> **about the gender inequality**$_{PP}$ … which is what I would be <u>discussing</u> **about**$_{PP}$ in the second part of this essay.
                              (ICE-FJ: W1A-016; Schneider & Zipp 2013: 4)
    b.   Whenever I <u>enter</u> **into the class**$_{PP}$ they laugh at me   (ICE-IND, S1a-001)

While specific aspects of the diachronicity of English PPs (such as the preposi-
tional passive, or preposition stranding) have received detailed attention (see § 2),
the overall development of PP-complements, particularly in relation to structural
alternatives, is still under-researched. The present paper aims to take a first step
towards addressing these research gaps by reporting on a pilot investigation of
PP-patterns in the Early Modern English (EModE) period as well as selected va-
rieties of World English (WE), giving special attention to the competition be-
tween PP-patterns and NP-uses and allowing for cross-fertilisation of diachronic
and World Englishes research. Making the connection between World Englishes
research and language history is also important because patterns found in institu-
tionalised ESL varieties are not necessarily recent innovations (see also Schneider
et al. 2020: 6 on 'extended' pluralisation of non-counts in historical varieties of
English). Corpus data from the Early Modern English period are therefore ex-
pected to provide important background for the study of language contact in the
more recent history of English.

   We make use of a syntactically annotated corpus, aiming to retrieve the full
range of prepositions in verbal argument structure in the period in question. Our
findings on the overall distribution of PPs and the different functions they exhibit
in the data are reported on in § 3. We then discuss the competition between PP-
and NP-patterns, zooming in on one verb class, namely verbs of communication
(§ 4). We complement this case study with a closer look at the same set of verbs in
WEs, to test the hypothesis that ESL varieties show a higher use of PP-patterns and
potential overlap in the use of verb-specific patterns. Section 5 concludes the paper
and points out directions for further research.

2.   Prepositions: Previous research on variation and change

The extant literature on prepositional complementation patterns in the history of
English falls into three broader groups: the first is concerned with changes in the se-
mantics and syntax of (individual) prepositions in earlier periods, often within the
context of grammaticalisation theory. The most extensive survey of PP-adverbials
is Iglesias-Rábade's (2011) study of Middle English (ME) location-prepositions,
indicating that the semantics of the prepositions widened considerably over time.
Similarly, Sato's (2009) investigation of prepositions in Old English (OE) shows in-
creasing semantic bleaching in prepositional adverbials. Lundskær-Nielsen (1993)
provides a comprehensive overview of various issues involved in the use of preposi-
tions in OE and ME, including their semantics. Molencki (2005, 2007, 2011, among
others), Cziskek-Kiliszewska (2014, 2015) as well as De Cuypere (2013) each deal
with the semantic range of individual prepositions in early English.

The second group of studies has tended to focus on selected phenomena in the history of English PPs. Yáñez-Bouza (2015), for instance, offers an account of preposition stranding (as well as pied-piping) from Early to Late Modern English (see also Bergh & Seppänen 2000). PP-ellipsis, a further aspect relevant for the diachrony of PPs in English, is treated in various publications by Nykiel (e.g., 2014, 2015), while preposition reduplication is discussed in Bergh (1998), inter alia. Denison (e.g., 1993) links the history of the prepositional passive with the development of prepositional verbs, whereas Dreschler (2015) connects it with word order changes. Claridge's (2000) in-depth investigation of prepositional verbs considers their relation to a number of other multi-word verbs in EModE. PPs have furthermore featured prominently in WEs research, particularly in ESL contexts where prepositional constructions appear to be used more extensively than in English as a native language (ENL) varieties. Among others, Schneider & Zipp (2013) provide evidence of new verb-preposition combinations such as *discuss about* or *demand for* in a number of WEs; Siew (2016) and Nelson & Hongtao (2012) report on innovative prepositional verbs in Singapore and African Englishes, respectively.

The third research strand looks at specific prepositional constructions which have come to stand in an alternation relationship with nominal patterns in PDE: these include, e.g., the prepositional member of the dative alternation (see, among many others, McFadden 2002; De Cuypere 2015a, 2015b; Zehentner 2019).

In sum, there is a substantial body of research on PPs in the history of English and WEs. However, previous diachronic studies mostly dealt with narrowly defined aspects of their development only, instead of the PP-system as a whole, or are problematic due to being based on insufficient empirical evidence (either using very small databases or being greatly restricted in scope). Similarly, previous WEs research has had to be selective with respect to the range of patterns under investigation because available corpora were not syntactically annotated.

Finally, possible connections between developments in different varieties of English today and changes observed in the history of English have – surprisingly – so far not been investigated at all. That is, synchronic variation in PDE and ESL varieties of English as well as ongoing or recent changes in PP-usage have rarely been linked to change in the more distant past. On the contrary, most research has focussed on individual periods and/or varieties. WEs research into PP complementation has largely neglected the connection with previous stages of the language, in particular the potential of 'innovative' patterns being present in the historical input varieties. What is therefore clearly missing is a systematic, comprehensive, unified, longitudinal account of the development of verb complementation in English, approached from the perspective of prepositions.

While addressing this lack in a more encompassing manner is clearly outside the scope of the present paper, we here report on a pilot study on EModE and

selected WEs, carried out in the context of a larger project on the diachrony of English PPs. Our main aim is to provide first results with the view to illustrating the novel perspective provided by a bottom-up approach across both time and space and the research questions it opens up for further investigation.

3.   Prepositional patterns in EModE

3.1   Methodology

The main data for our study come from the *Penn-Helsinki Parsed Corpus of Early Modern English* (PPCEME), which covers the time span of 1500–1720 and includes approximately 1.8 million words (Kroch et al. 2004). In a very first step, we used the POS and syntactically annotated versions of the corpus to determine the overall frequency of prepositions in the entire corpus (as compared to other POS). In a second step, we retrieved all clauses featuring verbs in combination with PPs from the corpus by means of *CorpusSearch*2 (Randall 2009); more specifically, we extracted instances of verbs either directly governing PPs or sequences of NPs and PPs, resulting in a total of over 30,000 clauses. Our data include both tokens such as (7a), where a verb is immediately followed by a prepositional phrase, as well as (7b), which features an NP-(direct) object plus a PP-(indirect) object.

(7)   a.   *The Chancelor saide that after long debating they <u>departyd</u> **for that tyme***
            "the chancellor said that after long debate they departed for that time"
            (AMBASS-E1-P2,3.2,25.20)
      b.   *Sir, in my most hartiest wyse I <u>recommennd</u> me **unto you***
            "Sir, I commend myself to you in the most heartfelt manner"
            (APLUMPT-E1-H,167.4)

For the present study, we limited our analysis to a subset of the data by taking a random sample of 200 tokens for each of the three sub-periods of the corpus (E1-E3, each comprising data from about 70 years, respectively). This random subset of verb-attached PPs was analysed according to their function adapting Hoffmann's (2005, 2011) classification scheme for PPs in PDE.

In his work, Hoffmann challenges the traditional distinction between complements (as obligatory dependents of the verb) and adjuncts, with the latter seen as optional and more loosely attached to the clause (see, e.g., Huddleston & Pullum 2002: 224–228). Although many PPs can indeed be distinguished into complements and adjuncts based on a range of semantic and syntactic criteria, it has also been shown that a strict complement-adjunct dichotomy often cannot be upheld: "In practice it is hard to make an absolute distinction between free combinations and

fixed multi-verb verbs” (Biber et al. 1999: 403; see also Quirk et al. 1985: 501–511, 1162–1167). Hoffmann instead proposes a more fine-grained classification with regard to factors such as optionality or degree of syntactic freedom, ranging from optional adjuncts (‘sentence adjuncts’ in Hoffmann’s terminology) such as *John ate lasagne **in Rome on Monday*** to fully obligatory complements in tightly connected verb-preposition combinations (subcategorising for specific prepositions, as in *rely on, consist of, give to*), with intermediate types such as ‘goal/source’ PPs (*John ran **to the church***), ‘accompaniment’ (*John committed the crime **with his dad***), ‘affected location’ (*John slept **in a bed***) or optional complements like *John worked **at the job***.

In our analysis, we distinguish between ‘adjunct’ and ‘complement’ PPs, which we divide into smaller groups, inspired by Hoffmann’s more fine-grained taxonomy. Adjuncts are thus further sub-divided into prototypical adjuncts, i.e., adverbials expressing time, manner, or location (illustrated in (8a)), ‘mixed’ types including sources or goals (8b), as well as ‘complement-like’ instances (8c).

(8)   a.   *And Haran <u>dyed</u> … **in the londe where he was borne***
            “and Haran died in the land where he was born” ‘adverbial’
                                                                            (TYNDOLD-E1-P1,XI,20G.341)

      b.   *and this day she charged Jug Altham to <u>goe</u> **to Hattfeeld***
            “and on this day she charged Jug Altham to go to Hatfield” ‘mixed’
                                                                            (MASHAM-E2-H,78.10)

      c.   *gave her may=tie= cause to <u>looke</u> **about her***
            “gave her majesty reason to look around her” ‘c-like’
                                                                            (RCECIL-E2-H,304.14)

Complements in our analysis are divided into fixed expressions and prototypical complements. The former set (labelled ‘idiom’) includes instances of patterns combining a preposition with a ‘bare’ noun which, in combination with a verb, expresses a prototypical activity connected with the (institutional) noun: for example, *go to sea* in (9a) does not denote motion towards a goal, but refers to seafaring, just as *go to university* often means “to study towards a university degree” rather than physical movement towards a place.[1] Prototypical complements (‘compl(obj)’) comprise instances of PPs functioning as direct (9b) or indirect objects (9c) of transitive or ditransitive verbs. Finally, subject complements such as (9d) are also included among the PP-complements.

---

1.   On the special semantics of such bare noun constructions, see, e.g., de Swart & Zwarts (2009), Goldberg (2013) or Hundt (2018). V-P-bare-N constructions of this type are likely to have had special status in earlier English already; their distribution and development will therefore need to be considered in greater detail in follow-up work to this pilot study.

(9) a. *he doth not believe that the Duke of York will <u>go</u> **to sea** again* 'idiom'
         "he does not believe that the Duke of York will go to sea again"
                                                          (PEPYS-E3-H,7,411.53)

     b. *he persuaded the Earl rather to flee into Wales … than to <u>trust</u> **to the City***
         "he persuaded the Earl rather to flee to Wales than to trust the city"
                                                          (ESSEXSTATE-E2-P2,202.84)

     c. *he <u>restored</u> **unto the Lord Roch** all his Cattell* 'compl(obj)'
         "he restored all his cattle to the Lord Roch"   (PERROTT-E2-P1,52.46)

     d. *the Amorites <u>dwell</u> **in the mountaines***: 'subjC'
         "the Amorites dwell in the mountains" (AUTHOLD-E2-H,XIII,20N.903)

Before moving on to the results of this part of the study, the following section briefly introduces the main hypotheses we aim to test.

3.2   Hypotheses

Building on observations made in earlier literature about the functions taken on by PPs in different stages in the history of English, we anticipate that the data will show changes in the inventory of PP-complementation patterns that can be approached in terms of grammaticalisation (see, e.g., Hundt 2001; Tyler & Evans 2003 or Rostila 2007 for earlier discussions on PPs in the context of grammaticalisation).[2] More precisely, we predict that the development of PP-constructions in the history of English may have proceeded along the following pathway:

Initially expressing spatial or temporal meaning (*John died **in Rome on Monday***), PPs extended to cover additional adverbial functions such as 'accompaniment' (*John cooked dinner **with his dad***) or 'instrument' (*John killed the cat **with a knife***), both subsumed under 'mixed' in our analysis. While in these cases the PP was still outside the VP, it nevertheless added a participant role to the event. Subsequently, the relations between the elements in such constructions became increasingly tight, with the PP moving inside the VP construction (e.g., *John slept **in a bed*** ['affected location'] or *John ran **to the church*** ['goal/sourcePP']). In some cases, strong links then formed between specific verb and preposition types. For example, *talk* is used exclusively with a PP-complement in Present-Day ENL varieties outside of highly idiomatic constructions like *to talk shop*; other fixed combinations include *work at something* or *keep something in a place* (see the sub-type of 'c-like').

---

2.   Note that the change could also be analysed as an instance of grammatical 'constructionalisation' as laid out in, e.g., Traugott & Trousdale (2013) or Smirnova (2015). While this paper is sympathetic to such accounts, a more explicitly constructional approach would require additional data and a more fine-grained analysis of the patterns extracted in a bottom-up approach than is possible in this pilot study.

Importantly, however, these complement-types are optional (consider also verbs of cognition such as *believe* [*in, on*] or *think* [*about, of*]). The process of increasing VP-integration, ultimately yielding prepositional verbs like *rely on* (reflected in the results on group 'compl(obj)'), indicates that the elements in the clause have come to be perceived as an indivisible chunk rather than separate constituents.

We expect our findings to provide us with first evidence for these different stages in the development of prepositional complementation. Specifically, a general rise in PP-frequency (as is characteristic for grammaticalising elements) should be seen, as well as differences in the frequency distribution of the individual patterns reflecting a cline according to degree of grammaticalisation. Further investigation into the precise distribution and features of these patterns will give more insights into the processes at play.

3.3   Results

In order to test whether prepositions in general, i.e., in any function, have increased in the course of the EModE period, their frequency per 10,000 words was determined for the three sub-periods of the corpus. As shown in Figure 1, the data – perhaps surprisingly – do not confirm this initial hypothesis: rather than growing, the overall token frequency of prepositions significantly decreases over time ($p < 0.0001$), dropping from almost 1,800 tokens in E1 to about 1,650 in E3 (per 10,000 words).
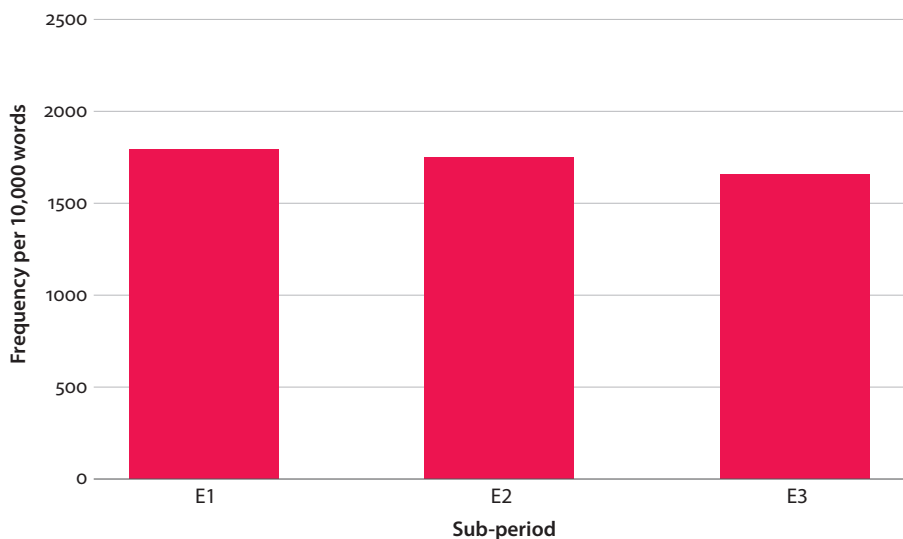
Figure 1. Normalised token frequency (per 10,000 words) of prepositions in the PPCEME, E1-E3

This observation is noteworthy, particularly as it corresponds (at least to some extent) with Szmrecsanyi's (2012: 658) conclusions about syntheticity and analyticity in the history of English, showing that the situation is by no means as straightforward as often assumed. It is, however, equally clear that the variability seen in Figure 1 cannot simply be attributed to their role in verb complementation and related to typological change across time. For example, the definition of 'preposition' applied in the annotation of the Penn-Helsinki corpora is relatively broad, including, e.g., subordinating conjunctions as exemplified in (10). This is likely to have had an impact on the distribution shown in the data.

(10)   *And **after**_–P þat þis   Donewal was dede, his sones … departede þe   lande*
       and after   that this Donweal was dead his sons   divided   the land
       "And after this Donewal was dead, his sons divided the land"
                                                     (PPCME2; CMBRUT3,24.708)

Moving on to the types of prepositional patterns, Figure 2 presents the proportional distribution of adjuncts and complements in the random sample of instances collected from the corpus ($N = 600$). It can be seen that there is a gradual but persistent increase of more complement-like structures at the expense of adjunctival types within EModE (E1/E3: $\chi2 = 4.5944$, $p = 0.03$). While the latter still account for more than 50% of the tokens in the sub-period of E1, the exact opposite holds for the latest sub-period, E3. As discussed above, we take this change to be indicative of PPs undergoing further grammaticalisation processes – while adjuncts represent the
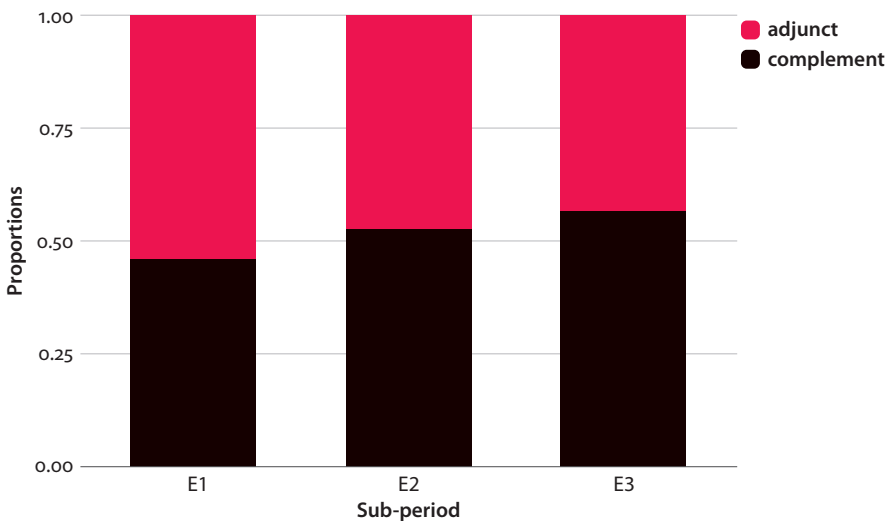
Figure 2. Token frequency (proportion) of PPs as adjunct: Complement in the PPCEME, E1-E3

older, less grammaticalised variants, complements are arguably more procedural and are more tightly integrated into the VP. However, the dataset analysed here is still rather small. In the future, we will use more data from the corpus in order to move beyond the periodisation provided by the corpus compilers. Applying *Variability-based Neighbour Clustering* (Gries & Hilpert 2008) will allow us to derive the relevant stages for diachronic change in a more bottom-up fashion and thus uncover developments that might be masked in an a priori approach to periodisation.

Since a binary distinction between adjuncts and complements is difficult to uphold, Figure 3 provides results of a more fine-grained analysis of PPs in verbal argument structure. First, it is noticeable that all sub-types proposed on the basis of Hoffmann's (2005, 2011) PDE model can also be found in the EModE dataset, indicating that the initial extension of PPs to an increasing number of functions took place before the investigated timeframe. Second, no major changes can be observed between the earliest and latest sub-period. Comparing the proportional distribution of all sub-types against each other, we find that adverbial, mixed and c-like uses slightly decrease over time (E1: 25/17/12% to E3: 20/15/9%), while 'idiom' uses remain more or less steady. Both 'compl' and 'subjC' increase over time (from about 37 to 44, and 0 to 5%, respectively). However, these changes are non-significant. Nevertheless, the distribution supports the assumption that the oldest (and least grammaticalised) functions are among the more frequent patterns, with the innovative extended patterns only slowly diffusing through the population.
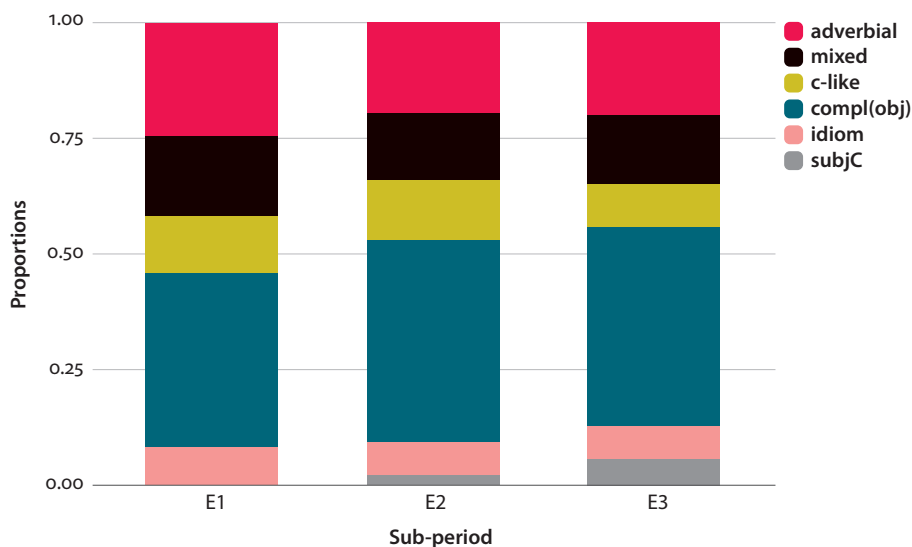
Figure 3. Token frequency (proportion) of subtypes of PPs in the PPCEME, E1-E3

In sum, the data presented in this section provide tentative support for the hypotheses put forward above, in that the frequency patterns correspond to the assumed cline of grammaticalisation.

4.   Competition between NP- and PP-complementation in EModE and WEs

4.1   Methodology

In order to get some preliminary insights into the question of competition between prepositional and nominal complementation patterns in both the history of (standard) English and World Englishes, we concentrated on one particular set of verbs, namely those expressing communicative events. We took the most frequently attested verbs in our random PPCEME subsets as a starting point. Specifically, we investigated occurrences of the verbs *answer, say, talk, enquire, repeat* and *whisper* in both the PPCEME as well as various components of the *International Corpus of English* (ICE; Greenbaum 1996). For EModE, we extracted all occurrences of the specified communication verbs from the PPCEME. Overall frequencies for *say* were so high that a random subset of 300 instances was used for further processing.

Our WEs data come from five ENL varieties (American, British, Canadian, New Zealand and Irish English) and eight ESL varieties (Hong Kong, Singapore, Indian, Philippine, Sri Lankan, Jamaican, Ghanaian and Fiji English). For American, Fiji, Ghanaian and Sri Lankan English, only written data (approximately 400,000 words each) are available; for the other WEs, the ICE components amount to approximately 1 million words each. Here, we made use of corpus' dependency parsing (Schneider 2008) and lemmatisation, i.e., we retrieved all instances of the verbs with PP-complements and NP-objects, resulting in a total of 10,031 concordance entries. Since the manual annotation of the ICE components is not fully reliable, and PP-attachment, in particular, is problematic, we manually annotated a random subset of 1,000 instances of the frequent verbs plus all occurrences for *enquire* and *whisper*, excluding all false positives, leaving a total number of 790 relevant hits.

The resulting concordances for all corpora and components were manually coded for complementation type according to whether they appear with NP-arguments (11a) or with prepositional arguments (11b); 'other' patterns such as intransitives (11c) or uses with clausal objects were excluded from the present study. Ditransitives with a prepositional object were included among the instances with PP- rather than NP-arguments.

(11) a. *He only shewed how improbable it was, that in a public house he should <u>talk</u> such things*_{NP}

"He only showed how improbably it was that he should talk about such things in a pub"   (BURNETCHA-E3-P1,2,172.285)

 b. *to <u>talk</u> of business*_{PP} *… is a Theme not quite so entertaining to you as that of the Ladies*

"to talk of business is a topic not quite as entertaining to you as that of ladies"   (VANBR-E3-P1,48.291)

 c. *Otherwise he and I never <u>talk'd</u>*

"otherwise he and I never talked"   (THOWARD2-E2-P1,1,94.436)

Note that this analysis presents a rough, preliminary picture only – in many cases, the function of the NP is in fact not directly comparable with that of the PP, meaning that a closer investigation is definitely necessary. Furthermore, the choice of verbs included in this study grew out of the bottom-up approach we initially took. And while they form a coherent group semantically, they only represent a subset of the class of communication verbs. Thus, they can only serve as a starting point for further work and are taken here mainly to illustrate how a bottom-up approach with subsequent top-down analyses of semantically related lexical items can provide complementary data sets. We believe that these data, while only representing the very tip of a substantial ice-berg, can yield a first idea as to whether PPs actually increase at the expense of NPs over time and whether prepositional patterns are indeed more prevalent in ESL than in ENL varieties, as laid out in the following section.

4.2   Hypotheses

The aim of this part of our study is to gain insights into the different ways of competition resolution between nominal and prepositional patterns. That is, we assume that when expanding in uses, PPs entered competition with resident patterns expressing the same functions, mainly NPs. As the main possible outcomes of such competition, we expect to see either loss of one of the variants, or eventual co-existence of both variants, often with functional differentiation of the patterns.

Loss of one variant should hold in a large majority of cases. We hypothesise that the prepositional pattern typically ousted the nominal construction. Examples of this development are manifold: as already mentioned above, PP-constructions appear to have almost completely replaced NP-variants of adverbials of time, place and manner (e.g., Traugott 1992: 207); the same is found with accompaniment- and instrument-adjuncts (e.g., *John killed his cat* *[*with*] *a knife*), as well as a set

of originally transitive verbs (e.g., *wonder about/at* vs. **wonder*+NP, *rejoice in* vs. **rejoice*+NP) and prepositional verbs such as *rely on* (Claridge 2000). The study of 'novel' prepositional verbs in ESL varieties should provide us with the opportunity to observe emerging competition as well as competition resolution in progress, since these innovations alternate with the standard nominal patterns (e.g., PDE *explain*+NP vs. FijE *explain+about; talk*+NP vs. *talk+about*).

Although the overall history of prepositions seems to be characterised by expansion and successful competition against nominal alternatives, there are also counter-developments. As Claridge (2000: 98) shows, instances where the NP-construction has prevailed can easily be found. For example, verbs such as *congratulate* see NP-complements being reinstated at the expense of the PP-patterns – collocations with *with*, possible in EModE, are not available anymore today (12).

(12) a. And they <u>congratulated</u> **with her**_{PP}
                        (1609, *Bible* (Douay) I. Luke i. 58; OED, s.v. *congratulate*)
       b. All the world…did <u>congratulate</u> **me**_{NP}
                        (1668, S. Pepys *Diary* 5 Mar. (1976) IX. 103; OED, s.v. *congratulate*)

This also holds for ESL varieties: while these WEs are generally described as more 'preposition-friendly', nominal patterns can occasionally be shown to compete with the now standard prepositional patterns. In (13a)–(b), for example, *talk* is used with NP-objects, most likely as an analogical extension from schematic patterns such as (13c), which also feature an NP-object and are attested in (informal, spoken use) in metropolitan varieties (see further Mukherjee & Schilk 2012: 196).

Importantly, transitive uses of *talk* with an NP-complement are also attested in EModE (see (11a) above), so seemingly 'novel' uses in ESL varieties might in fact go back to the historical input.

(13) a. From this time onward you must try not to lose your temper and try to <u>talk</u> **pleasant things**_{NP} all the time.   (ICE-SL, W2b-026)
       b. From now onwards don't <u>talk</u> **dirty jokes**_{NP} in front of me.
                                                                                  (ICE-Sing, S1a-031)
       c. So we're <u>talking</u> **magnitude earthquakes**_{NP} of six point five seven closing to seven point five.   (ICE-GB, S2a-025)

Moreover, there is variation among national standard varieties of English in this respect, too. American English (AmE), for instance, currently prefers to use *protest* and *appeal* with NP-complements (14), whereas the preferred complementation in British English (BrE) is with a PP (15), and varieties like New Zealand English use both complementation patterns (Hundt 1998: 108–112).

(14) a. The amazing outpouring of people at Women's Marches around the country and the world to <u>protest</u> **the evil**$_{NP}$ that the Trump Administration represents can and must be the start of something great and powerful.

(2017; COCA, NEWS)

    b. He <u>appealed</u> **the verdict**$_{NP}$ and was granted a new trial in 2005, …

(2016; COCA, NEWS)

(15) a. Thousands of people turned out at the Royal Showground today to <u>protest</u> **against a private members bill**$_{PP}$ which aims to abolish blood sports.

(BNC, K1F)

    b. The Sun was granted leave to <u>appeal</u> **against his order**$_{PP}$.   (BNC, CBM)

And while the pattern with the NP-complement for *protest* is also attested in EModE (16) and would thus have formed part of the input in the evolution of AmE, a cursory glance at historical American data shows that the NP-complement now preferred in the US is not an instance of colonial lag but one of post-colonial revival (Hundt 2009) of a conservative complementation pattern.[3]

(16)   Protest, (protestari)… is by way of complaint to <u>protest</u> **a mans bill**$_{NP}$

(1607, J. Cowell Interpreter sig. Ggg1; OED, s.v. *protest*)

By contrast to the instances just outlined, which reflect ongoing competition, stable (conditioned) co-existence seems to have been reached in other cases – verbs of cognition such as *think* and *believe* have preserved both the prepositional and non-prepositional option in PDE. These options were presumably interchangeable at some point, but have come to diversify functionally: PDE *believe+in* means "have confidence, faith in", whereas the nominal construction *believe+NP* expresses a sense of "to give intellectual assent to, accept the truth or accuracy of (a statement, doctrine, etc.)" (OED, s.v. *believe*). Other, more systematic and schematic correspondences include, e.g., the PDE dative alternation (*John gave Mary a cake ~ John gave a cake to Mary*).

    In sum, the interaction between PPs and other structural options is clearly much more complex than increasing replacement. This challenges the assumption of a straightforward development towards more analytic (specifically prepositional) means of expressions in the history of English, which is at least implicitly present in many textbook analyses (cf. Szmrecsanyi 2012 for a critical evaluation). It furthermore confirms that a quantitative investigation into the precise dynamics – with a particular focus on connecting developments in the more distant and recent

---

3. A quick-and-ready search the *Corpus of Historical American English* (COHA) shows that PP-complements are preferred throughout the 19th century and that bare NP complements only increase from the 1920s.

past and across varieties of English – is needed. The following section begins to address this lack by testing the basic assumption that PPs generally succeed against NP-complements on one set of verbs.

4.3   Results

In this study, we take a first step towards untangling the complex interactions just discussed by looking at a subset of relevant verbs, namely verbs of communication such as *answer* or *talk*. Figure 4 shows that again, our initial hypothesis is not confirmed: PP-complements significantly decrease in favour of NP-complements during the course of the EModE period, at least within this subset ($N = 396$; E1/E3: $\chi^2 = 5.707$, $p = 0.01689$).[4] While PP-patterns account for 69% of tokens in E1, this number drops to just above 50% in E3, meaning that NPs come to take up slightly less than half of the instances at the end of EModE.
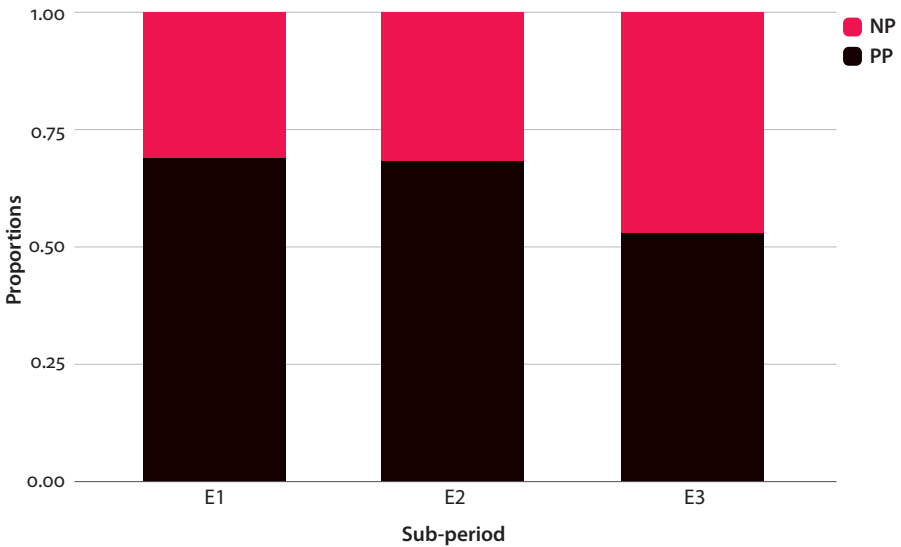
Figure 4. Token frequency (proportion) of complementation patterns (NP : PP) for communication verbs in the PPCEME, E1-E3

---

4.   The instances are distributed across the individual verbs in the following way: *answer*: $N = 113$, *say*: $N = 166$, *talk* = 79, *enquire* = 20, *repeat* = 13, *whisper* = 5.

We then model our data with a non-parametric regression analysis (applicable to non-normally distributed data) by fitting a conditional inference tree (ctree). The algorithm makes use of recursive partitioning and predicts outcomes on binary splits of the data (Hothorn et al. 2006). Starting off with the most important predictor variable, the data are split into homogenous sets (with a minimum significance level of $\alpha = 0.05$) until further splits no longer increase the homogeneity of the data. The response variable in our analysis is the complementation type (with the levels 'PP' vs. 'NP'); predictor variables are the sub-periods in PPCEME and the individual verbs. Figure 5 shows that 'period' is not selected as a predictor variable in the model, indicating that change across the subsections of the PPCEME is not significant.[5] Instead, the lexical verbs are significant predictor variables, with *answer* and *repeat* strongly favouring NP-complements and the remaining verbs in the set tending towards PP-complements, particularly *talk*.
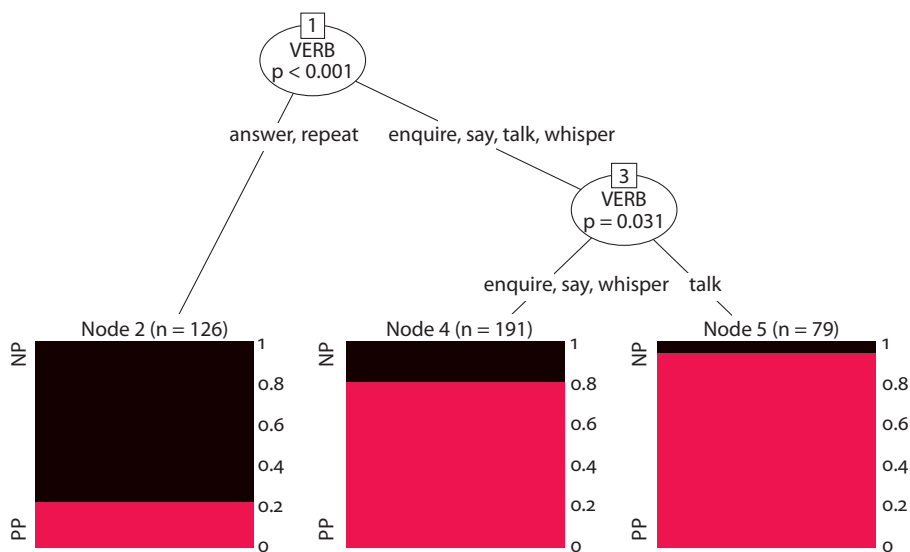
Figure 5. Complementation of communication verbs in the PPCEME, E1-E3 (ctree)

---

5. Again, a Variability-based Neighbour Clustering approach based on a larger dataset might reveal diachronic developments that are potentially masked in a dataset that uses an 'external' division of the data into sub-periods.

We conclude our case study with a closer look at the same verb set in WEs, to test the assumption that ESL varieties are particularly PP-friendly. Of the 790 instances, 574 (72.7%) are followed by a PP-complement.[6] Contrary to our expectations, there is no significant difference between ENL and ESL varieties in their overall complementation patterns: 276 or 73% take a PP-complement in our ENL data compared with 298 or 72.2% in our ESL evidence. A conditional inference tree (Figure 6) again reveals that individual communication verbs differ significantly in their complementation patterns, with *answer* and *repeat* strongly preferring NP complements and *whisper* and *talk* overwhelmingly taking PP-complements. Unlike our EModE data, the complementation profiles of *enquire* and *say* in ICE indicate a less marked preference for PP-complementation in PDE (see node 4 in Figure 6). In other words, we do not see the expected shift in complementation patterns between EModE and PDE from NP-complementation towards PP-complementation with these verbs.
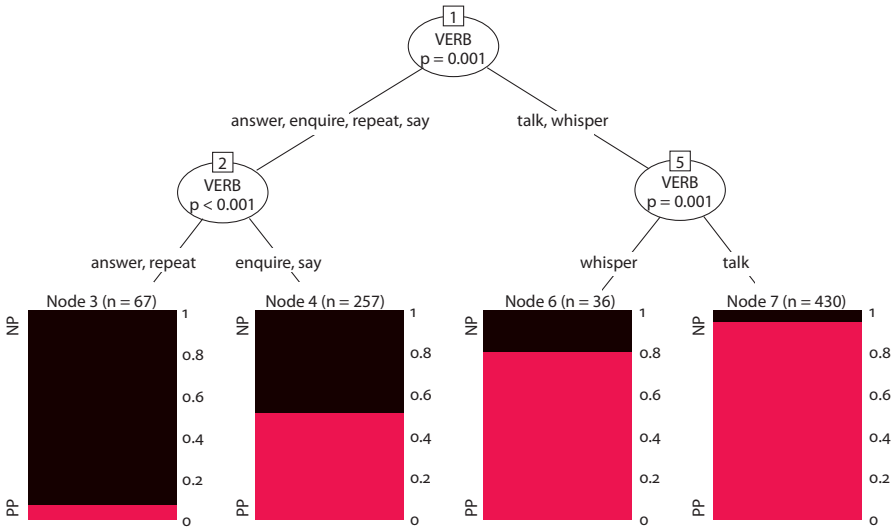
Figure 6. Complementation of communication verbs in ICE (ctree)

Note that 'period' and 'variety' not being selected as significant predictors in the ctrees just presented is perhaps not too surprising – as mentioned before, the chosen verbs cannot necessarily be assumed to form a clearly coherent group, and it is therefore to some extent expected that the individual verb preferences will greatly outweigh other factors. While we have used only a small set of communication verbs in our study, the lexical preferences for one or the other complementation

---

6.   With respect to individual verbs, our ICE dataset is comprised of *answer* ($N = 54$), *enquire* ($N = 6$), *repeat* ($N = 13$), *say* ($N = 251$), *talk* ($N = 430$) and *whisper* ($N = 36$).

pattern are furthermore likely to be a robust finding in a follow-up study, too, seeing that previous research into the dative alternation (e.g., Zehentner 2019: 200–202) also provided evidence of collocational 'attraction' of different verbs to complementation patterns.

Finally, despite there not being any statistically significant differences between ENL and ESL varieties, we can still find some qualitative evidence of divergent complementation patterns in our ESL data, such as the following:

(17) a. Can we <u>talk</u> **something**$_{NP}$ **about the situation**$_{PP}$ in these days.
<div align="right">(ICE-Ind, S1a-039)</div>

  b. I can <u>talk</u> **many things**$_{NP}$ with uh **with my brothers**$_{PP}$ but sometimes I don't know uh what to say with my sisters uhm uhm. (ICE-HK, S1a-027)

  c. You want to <u>talk</u> **anything**$_{NP}$ more interesting.   (ICE-Sing, S1A-015)

  d. You yourself did not <u>enquire</u> **it**$_{NP}$ **anybody**$_{NP}$ as to what was the cause of the quarrel till till none informed him   (ICE-Ind, S1b-070)

  e. He would sit down along with his bag and baggage he will talk to the person <u>enquires</u> **him**$_{NP}$ **about the expenditure**$_{PP}$ (ICE-Ind, S1b-018)

It is notable that (17a)–(c) are ESL constructions that are also attested in EModE (see (11a) above), and are thus difficult to classify as 'extended' or 'innovative' use in the absence of diachronic evidence for ESL varieties. Similarly, transitive uses of *enquire* are also present in EModE, as (18) shows:

(18)   he <u>enquired</u> earnestly **the true cause of the next dayes execution**$_{NP}$
<div align="right">(JOTAYLOR-E2-P1,3,79.C2.53)</div>

In sum, the competition between PP- and NP-patterns is clearly highly intricate, as is the precise relation between recent developments in ESL and non-standard varieties of English, and the history of the standard.

5.   Conclusion

Although changes in argument structure tend to be very complex and thus defy simple narratives along the lines of 'EModE has become more PP friendly', our case study has provided evidence of some interesting diachronic developments in the distribution of the patterns. The analysis of a randomised subset of the data shows that the proportion of constructions on the cline from complements to adjuncts changes over time, with complement-like types increasing at the expense of more adjunctival patterns. We have argued that this change may most fruitfully be approached in terms of grammaticalisation, as the changes reflect an increasing expansion of the PPs towards more 'core' (and arguably more grammatical) functions.

With regard to competition between structural variants, we have shown that PPs cannot be assumed to straightforwardly and progressively oust (older) nominal patterns, but that a more likely scenario is that a range of different competition resolution processes will hold for different (groups of) verbs. Within the rather narrow set of communication verbs focussed on in our case study, there are striking differences in the preferences of individual verbs, and NPs seem to hold their ground (or increase) at least with some of them within EModE. 'Lexical verb' also proves to be a significant predictor variable in the choice of NP- vs. PP-complements in WEs, whereas variety type (ENL vs. ESL) does not. On a quantitative level, our evidence thus contests the claim that ESL varieties are generally more prone to PP-uses. Infrequent (and thus salient) 'divergent' patterns of use emerge from our qualitative analyses. These, in turn, cannot all simply be interpreted as straightforward innovation in the ESL data, seeing that the PP-complementation they show is often also attested from earlier stages of English (in this case, EModE).

In future work, we aim to approach both the functional expansion of PPs and their competition with other patterns in constructionist terms, modelling them as changes to the network of English argument structure constructions. We will also pay attention to the emergence and regional spread of patterns with increased schematicity such as V-P-N, where the nominal slot is filled with a bare noun and the meaning is non-compositional (e.g., *go to sea/church, be at work/in hospital*).

To conclude, our study has allowed us to revisit and re-assess some common and well-known hypotheses about the development of prepositions in verbal argument structure in the history of English. We have also begun to show how a bottom-up approach can be a useful starting point for further hypotheses and complemented with top-down analyses (both quantitative and qualitative). In addition, and very importantly, our case study on a subset of communication verbs shows that historical data provide valuable depth to studies on complementation patterns in WEs (and vice versa).

Sources

Bosworth–Toller = Bosworth, Joseph. 2010. *An Anglo–Saxon dictionary online*. (edited by T. Northcote Toller & others; compiled by Sean Christ & Ondrej Tichý). https://bosworthtoller.com/ (6 January, 2021)

COCA = Davies, Mark. 2008–. *The corpus of contemporary American English (COCA): 560 million words, 1990-present*. https://corpus.byu.edu/coca/ (6 January, 2021)

COHA = Davies, Mark. 2010–. *The corpus of historical American English (COHA): 400 million words, 1810–2009*. https://corpus.byu.edu/coha/ (6 January, 2021)

ICE = *International corpus of English*. http://www.ice-corpora.uzh.ch (6 January, 2021)

OED = *Oxford English Dictionary Online*. 2017. Oxford: OUP. http://www.oed.com/ (6 January, 2021)

PPCEME = Kroch, Anthony, Beatrice Santorini & Lauren Delfs. 2004. *The Penn-Helsinki parsed corpus of Early Modern English* (PPCEME), 1st edn., release 3. Philadelphia, PA: University of Pennsylvania. http://www.ling.upenn.edu/ppche-release-2016/PPCEME-RELEASE-3 (6 January, 2021)

PPCME2 = Kroch, Anthony & Ann Taylor. 2000. *The Penn-Helsinki parsed corpus of Middle English* (PPCME2), 2nd edn., release 4. Philadelphia, PA: University of Pennsylvania. http://www.ling.upenn.edu/ppche-release-2016/PPCME2-RELEASE-4 (6 January, 2021)

Randall, Beth. 2009. *CorpusSearch 2: A tool for linguistic research*. Philadelphia, PA: University of Pennsylvania. http://corpussearch.sourceforge.net/ (6 January, 2021)

References

Allen, Cynthia. 1995. *Case marking and reanalysis: Grammatical relations from Old to Early Modern English*. Oxford: Oxford University Press.

Bergh, Gunnar. 1998. Double prepositions in English. In Jacek Fisiak & Marcin Krygier (eds.), *English historical linguistics 1996*, 1–13. Berlin: Mouton. https://doi.org/10.1515/9783110804072.1

Bergh, Gunnar & Aimo Seppänen. 2000. Preposition stranding with *wh*-relatives: A historical survey. *English Language and Linguistics* 4(2). 295–316. https://doi.org/10.1017/S1360674300000265

Biber, Douglas, Gray, Stig Johansson, Geoffrey Leech, Susan Conrad & Edward Finegan. 1999. *Longman grammar of spoken and written English*. Harlow: Pearson.

Ciszek-Kiliszewska, Ewa. 2014. Middle English preposition *twēn(e)*. *Studia Anglica Posnaniensia* 49(3). 91–111. https://doi.org/10.1515/stap-2015-0008

Ciszek-Kiliszewska, Ewa. 2015. The Middle English preposition and adverb *atwēn*. In Brian Lowrey & Fabienne Toupin (eds.), *Studies in linguistic variation and change: From Old to Middle English*, 41–63. Cambridge: Cambridge Scholars.

Claridge, Claudia. 2000. *Multi-word verbs in Early Modern English. A corpus-based study*. Amsterdam: Rodopi. https://doi.org/10.1163/9789004333840

De Cuypere, Ludovic. 2013. Debiasing semantic analysis: The case of the English preposition *to*. *Language Sciences* 37. 122–135. https://doi.org/10.1016/j.langsci.2012.12.002

De Cuypere, Ludovic. 2015a. A multivariate analysis of the Old English ACC+DAT double object alternation. *Corpus Linguistics and Linguistic Theory* 11(2). 225–254. https://doi.org/10.1515/cllt-2014-0011

De Cuypere, Ludovic. 2015b. The Old English *to*-dative construction. *English Language and Linguistics* 19(1). 1–26. https://doi.org/10.1017/S1360674314000276

Denison, David. 1981. Aspects of the history of English group-verbs. With particular attention to the syntax of the Ormulum. Oxford: Oxford University PhD thesis.

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

Dreschler, Gea. 2015. *Passives and the loss of verb second: A study of syntactic and information-structural factors*. Utrecht: LOT.

Goldberg, Adele. 2013. Constructionist approaches. In Thomas Hoffmann & Graeme Trousdale (eds.), *The Oxford handbook of construction grammar*, 15–31. Oxford: Oxford University Press.

Greenbaum, Sidney (ed.). 1996. *Comparing English worldwide: The International Corpus of English*. Oxford: Clarendon.

Gries, Stefan & Martin Hilpert. 2008. The identification of stages in diachronic data: Variability-based neighbor clustering. *Corpora* 3(1). 59–81. https://doi.org/10.3366/E1749503208000075

Hoffmann, Thomas. 2005. Variable vs. categorical effects: Preposition pied piping and stranding in British English relative clauses. *Journal of English Linguistics* 33(3). 257–297. https://doi.org/10.1177/0075424205282891

Hoffmann, Thomas. 2011. *Preposition placement in English: A usage-based approach*. Cambridge: Cambridge University Press.

Hothorn, Torsten, Kurt Hornik & Achim Zeileis. 2006. Unbiased recursive partitioning: A conditional inference framework. *Journal of Computational and Graphical Statistics* 15. 651–674. https://doi.org/10.1198/106186006X133933

Huddleston, Rodney & Geoffrey Pullum. 2002. *The Cambridge grammar of the English language*. Cambridge: Cambridge University Press. https://doi.org/10.1017/9781316423530

Hundt, Marianne. 1998. *New Zealand English grammar. Fact or fiction?* Amsterdam: John Benjamins. https://doi.org/10.1075/veaw.g23

Hundt, Marianne. 2009. Colonial lag, colonial innovation, or simply language change? In Günter Rohdenburg & Julia Schlüter (eds.), *One language, two grammars: Morphosyntactic differences between British and American English*, 13–37. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511551970.002

Hundt, Marianne. 2018. Variable article usages with institutional nouns. An 'oddment' of English? In Alex Ho-Cheong Leung & Wim van der Wurff (eds.), *The noun phrase in English: Past and present*, 113–142. Amsterdam: John Benjamins. https://doi.org/10.1075/la.246.05hun

Hundt, Markus. 2001. Grammatikalisierungsphänomene bei Präpositionalobjekten in der deutschen Sprache. *Zeitschrift für germanistische Linguistik* 29(2). 167–191. https://doi.org/10.1515/zfgl.2001.012

Iglesias-Rábade, Luis. 2011. *Semantic erosion of Middle English prepositions*. Frankfurt: Lang.

Lundskær-Nielsen, Tom. 1993. *Prepositions in Old and Middle English*. Odense: Odense University Press. https://doi.org/10.1075/nss.9

McFadden, Thomas. 2002. The rise of the to-dative in Middle English. In David Lightfoot (ed.), *Syntactic effects of morphological change*, 107–123. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199250691.003.0006

Mitchell, Bruce. 1985. *Old English syntax*, Volume 1. Oxford: Clarendon. https://doi.org/10.1093/acprof:oso/9780198119357.001.0001

Molencki, Rafał. 2005. On the syntactic and semantic development of *after* in medieval English. In Marcin Krygier & Liliana Sikorska (eds.), *Naked wordes in Englissh*, 47–67. Frankfurt: Lang.

Molencki, Rafał. 2007. On the rise of the temporal preposition/conjunction *before*. In Marcin Krygier & Liliana Sikorska (eds.), *To make his Englissh sweete upon his tonge*, 37–54. Frankfurt: Lang.

Molencki, Rafał. 2011. New prepositions and subordinating conjunctions of Romance origin in Middle English. In Jacek Fisiak & Magdalena Bator (eds.), *Foreign influences on Medieval English*, 9–24. Frankfurt: Lang.

Mukherjee, Joybrato & Marco Schilk. 2012. Exploring variation and change in New Englishes: Looking into the International Corpus of English (ICE) and beyond. In Terttu Nevalainen & Elizabeth C. Traugott (eds.), *The Oxford handbook of the history of English*, 189–199. Oxford: Oxford University Press.

Mustanoja, Tauno F. 1960. *A Middle English syntax*, Volume 1. Helsinki: Société Néophilologique.

Nelson, Gerald & Ren Hongtao. 2012. Particle verbs in African Englishes: Nativization and innovation. In Marianne Hundt & Ulrike Gut (eds.), *Mapping unity and diversity world-wide: Corpus-based studies of New Englishes*, 197–213. Amsterdam: John Benjamins. https://doi.org/10.1075/veaw.g43.08nel

Nesselhauf, Nadja. 2009. Co-selection phenomena across New Englishes. Parallels (and differences) to foreign learner varieties. *English World-Wide* 30(1). 1–26. https://doi.org/10.1075/eww.30.1.02nes

Nykiel, Joanna. 2014. Semantic dependencies and the history of ellipsis alternation. In Michael Adams, Laurel Brinton & Richard D. Fulk (eds.), *Studies in the history of the English language VI: Evidence and method in histories of English*, 51–70. Berlin: de Gruyter.

Nykiel, Joanna. 2015. Constraints on ellipsis alternation: A view from the history of English. *Language Variation and Change* 27(2). 227–254. https://doi.org/10.1017/S0954394515000071

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

Rostila, Jouni. 2007. *Konstruktionsansaetze zur Argumentmarkierung im Deutschen*. Tampere: Juvenes.

Sato, Kiriko. 2009. *The development from case-forms to prepositional constructions in Old English prose*. Bern: Lang.

Schneider, Gerold. 2008. Hybrid long-distance functional dependency parsing. Zurich: University of Zurich PhD dissertation. https://www.zora.uzh.ch/id/eprint/7188/ (8 October, 2018)

Schneider, Gerold & Lena Zipp. 2013. Discovering new verb-preposition combinations in New Englishes. *Studies in Variation, Contacts and Change in English* 13. [Online journal]

Schneider, Gerold, Marianne Hundt & Daniel Schreier. 2020. Pluralized non-counts nouns across Englishes: A corpus-linguistic approach to variety types. *Corpus Linguistics and Linguistic Theory* 16(3). 515–546. https://doi.org/10.1515/cllt-2018-0068

Siew, Tan. 2016. Charting the endonormative stabilization of Singapore English. In Gerhard Leitner, Azirah Hashim & Hans-Georg Wolf (eds.), *Communicating with Asia: The future of English as a global language*, 69–84. Cambridge: Cambridge University Press.

Smirnova, Elena. 2015. Constructionalization and constructional change: The role of context in the development of constructions. In Jóhanna Barðdal, Spike Gildea, Elena Smirnova & Lotte Sommerer (eds.), *Diachronic construction grammar*, 81–106. Amsterdam: John Benjamins. https://doi.org/10.1075/cal.18.03smi

de Swart, Henriëtte & Joost Zwarts. 2009. Less form – more meaning: Why bare singular nouns are special. *Lingua* 119(2). 280–295. https://doi.org/10.1016/j.lingua.2007.10.015

Szmrecsanyi, Benedikt. 2012. Analyticity and syntheticity in the history of English. In Terttu Nevalainen & Elizabeth C. Traugott (eds.), *The Oxford handbook of the history of English*, 654–665. New York: Oxford University Press.

Traugott, Elizabeth C. 1972. *A history of English syntax*. New York: Holt, Rinehart, and Winston.

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

Traugott, Elizabeth C. & Trousdale, Graeme. 2013. *Constructionalization and constructional changes*. Oxford: Oxford University Press.
https://doi.org/10.1093/acprof:oso/9780199679898.001.0001

Tyler, Andrea & Vyvyan Evans. 2003. *The semantics of English prepositions: Spatial scenes, embodied meaning, and cognition*. Cambridge: Cambridge University Press.
https://doi.org/10.1017/CBO9780511486517

Yáñez-Bouza, Nuria. 2015. *Grammar, rhetoric and usage in English: Preposition placement 1500–1900*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511732522

Zehentner, Eva. 2019. *Competition in language change: The rise of the English dative alternation*. Berlin: De Gruyter Mouton. https://doi.org/10.1515/9783110633856

Should with non-past reference

A corpus-based diachronic study

Lilo Moessner

RWTH Aachen University

This chapter combines a quantitative and a qualitative analysis of the use of *should* with non-past reference in the historical periods Old English, Middle English (ME), and Early Modern English (EModE). It is based on data from the *Helsinki Corpus of English Texts* and the *Middle English Dictionary*. The quantitative analysis detects a few late ME examples, but establishes EModE as the period when *should* with non-past reference became firmly established, and from the qualitative analysis the conclusion is derived that *should* as a competitor of the present subjunctive in mandative constructions developed as a result of its meaning extension from obligation to other shades of non-fact modality.

Keywords: quantitative analysis, qualitative interpretation, mandative construction, Old English (OE), Middle English (ME), Early Modern English (EModE), meaning extension, main clause, relative clause, noun clause

1.　Introduction: Setting the scene

More than twenty years ago Marianne Hundt published a paper with the title "It is important that this study (*should*) be based on the analysis of parallel corpora: On the use of the mandative subjunctive in four major varieties of English" (1998). *Should* in this title is in parentheses and thus marked as an optional element. The subjunctive *be* and the verbal syntagm *should* + infinitive are variants in this *that*-clause. The results of Hundt's analysis of the BROWN Corpus (the Brown University Standard Corpus of Present-Day American English, Francis & Kucera 1979) and the LOB Corpus (the British-English forerunner of BROWN, Johansson et al. 1978) support the claims of PDE grammars that in *that*-clauses depending on matrix clauses with verbs, adjectives or nouns expressing emotions, demands, requests, or wishes – so-called 'mandative' contexts – the subjunctive competes with putative *should* + infinitive and that the *should*-construction is typical of British

English. Among her 111 examples from the LOB-corpus 97, i.e., almost 90%, attest the *should*-construction, only 14 the subjunctive, while the situation in her data from the BROWN Corpus shows the reverse. From a synchronic point of view this result is not surprising, because it is well known that subjunctive frequency is low in present-day British English. What is surprising from a diachronic point of view, however, is that a present tense subjunctive should alternate with a modal construction containing the morphological past of present tense *shall*.

The issue touches on one of the characteristics of modal auxiliaries that sets them apart from lexical verbs: tense relationships – as between *shall* and *should* – do not parallel those of lexical verbs; *must* and *need* have no 'preterite forms' at all, and those modals that do, like *shall*, have acquired meanings and distributions of their own, appearing "in a range of tentative, hypothetical or polite expressions. … [R]eference to past time is uncommon and typically restricted" (Warner 1993: 9).

This PDE situation raises two questions, which I will try to answer in this paper. These are:

1. When did the *should*-construction become a competitor of the present subjunctive in so-called mandative constructions?
2. Which developments made this competition possible?

2. Structure of the paper and research method

In § 3 I review previous publications on the use of *should* in the historical periods of English, focussing on explanations of the transition of *should* as a past tense marker to *should* as a modality marker in present tense contexts. Against this background it becomes appropriate to carry out a quantitative analysis of the occurrences of *should*, and my data come from the Old English (OE), Middle English (ME), and Early Modern English (EModE) parts of the *Helsinki Corpus of English Texts* (HC). The ME data are compared to the relevant entries in the *Middle English Dictionary* (MED). The searches for the variants of *should* in the HC are based on the antconc-tools 'list' and 'concordance'. From their results I derive an answer to my first research question (§ 4). After the presentation of the quantitative results, I proceed to their qualitative interpretation, and there I offer an answer to my second research question (§ 5). The last section of this paper summarizes my findings.

3.   Uses of *should* in the historical periods of English

Together with *can/cuþe, mæg/meahte* and some others the pair *sceal/sceolde* belongs to the class of preterit-present verbs in OE. Joined by the anomalous verb *will/ wolde* they underwent a number of interesting changes, e.g., their status changed from lexical verbs to modal auxiliaries and they acquired the feature abnormal time reference: "Not only the present forms, but the past forms of the modal auxiliaries can be used to refer to present or future time" (Quirk et al. 1985: § 3.30). When and how these changes happened is an unsettled question, but there are signs that they started already in OE.

Goossens (1987) claims that OE *cunnan* and *sculan* represent the extremes on the scales of auxiliarisation and grammaticalisation. Based on his analysis of 200 occurrences of each of these premodals, he concludes that *sculan* is "not a full verb (any more) in Old English" (1987: 127), since it does not occur as a non-finite form in his corpus. He also argues that the present tense forms of *sculan* should be treated separately from its past tense forms because the latter show a higher degree of grammaticalisation than the former. He analyses all his past tense examples in clauses of purpose or result and 50% of those in object clauses after verbs of voli- tion as "a marker of a specific type of subclause", but he does not provide a single example with a present tense verb form in the corresponding matrix clause. This combination would have been convincing proof that *should* was fully grammati- calised at this early stage.

This is also the assessment of the OE situation by Anderson (1991: 28), who warns his readers that they should not overinterpret the few signs of a dissolution of the past/non-past relationship between *sceolde* and *sceal* in OE. What he would accept as clear proof of this dissolution "would be examples with a *sceolde* in a clause dependent on a mandative non-past verb".

This is exactly what Traugott (1972: 150) seems to offer. She refers to the *should*-construction as a "segmentalised" subjunctive, and she assumes that it "orig- inated in such OE sentences as the following, where it expresses obligation, and is embedded in a main clause with the past tense of reported speech". In her OE example the *should*-construction depends on a past tense form of *beodan*, which she translates as "commanded". She argues that in ME this construction was extended to non-past contexts and lost its former sense of obligation. As an illustration of this hypothesis she takes over the following example from the *South English Legendary* cited by Visser (1963–1973: § 1546):

(1)   *Many gon naked and bidde þat sum man   heom **scholde** beweue*
      Many go naked and ask   that some person them should clothe
      "Many go naked and ask that someone should clothe them"

                                             (*South English Legendary* 420, 7)

The same example is also quoted in the OED.[1] Yet a closer look at the context reveals that *gon* and *bidde* are not present tense forms of the corresponding verbs, but infinitives. The verb of the matrix clause *saiȝh* "saw" governs several accusative with infinitive (a.c.i.) constructions.

(2)   *him þouȝte, he saiȝh manie men : liggen in beddes of golde, And manie*
      him thought he saw   many men lie   in beds   of gold and many
      sitte at heyȝe borde : and habbe al þat huy wolde; and many gon naked
      sit   at high table   and have all that they wanted and many go naked
      *and bidde þat sum   man   heom **scholde** beweue*
      and ask   that some person them should   clothe
      "he thought he saw many people lie in beds of gold, and many sit at high table
      and have all they wished, and many go naked and ask that somebody should
      clothe them"   (*South English Legendary* 420, 5–7)

The same example is quoted by Fischer in her treatment of ME object clauses, where she points out that after verbs expressing a wish, command or exhortation the subjunctive alternates with modal constructions, especially with *should* + infinitive. Like Traugott, she places the first occurrences of this *should* after present-tense matrix verbs in ME, "from about 1300 onwards" (1992: 315).

   López-Couso & Mendez Naya (1996, 2006) follow Traugott's line of argumentation. Unlike Traugott, however, they place the origin of *should* in non-past contexts in OE, but they restrict their claim to complements of the high frequency verbs *beodan* and *biddan* in the OE and ME parts of the HC. Their assumption that this process had started already in OE is nourished by what they call sporadic OE examples. One of their examples is of *should*, from Alfred's translation of Pope Gregory's *Cura Pastoralis*; the others are of *motan* "must".

(3)   *Đa sahlas is beboden þæt **sceoldon** bion mid golde befangne*
      The poles is ordered that should   be   with gold encased
      "The poles are ordered to be encased in gold"   (*Cura Pastoralis* 170/14)

The Latin original of this passage is:

(4)   *Qui auro   quoque jubentur   operiri*
      They with gold also   are.ordered to.be.covered   (Bramley 1874: 126)

---

1.   S.v. *shall* 22. This is the section where the use of *shall* in noun clauses is described. The quotations are accompanied by the following note: "In dependence on expressions of will, desire, command, advice, request … [t]he substitution of *should* for the earlier *shall* (itself a periphrastic substitute for the more primitive use of the present subjunctive …) may have arisen from instances in which the governing vb. was in the modal past tense."

In this complex Latin example two constructions are combined, namely *jubere aliquid* "to order something" and *operire vectes auro* "to cover the poles with gold", such that the second construction, shown in square brackets below, is embedded in the first as its accusative object (= a.c.i.): *jubere* [*operire vectes* (acc.) *auro*] "to order [to cover the poles with gold]". Both constructions appear as passive constructions, i.e., the a.c.i.-construction becomes a nominative and infinitive (n.c.i.)-construction, with the clause in square brackets as subject of *juberi*: *juberi* [*vectes* (nom.) *operiri auro*] "the poles to be ordered to be covered with gold". As a consequence of the embedding of the n.c.i.-construction into the first (passive) construction, the nominative *vectes* becomes the subject of *juberi* and number concord appears on the verb: *jubentur* "be ordered", 3pl (see (4)). The noun *vectes*, which is used in the preceding context to express a part of the construction of the ark, is pronominalised by *qui*:[2] *qui jubentur operiri auro*. This is the construction which appears in the Latin text in (4).

The translator was faced with two problems: one was that neither the person who receives the order nor the person who covers the poles with gold are specified in the Latin text. Problem two was that *bebeodan* 'to order', the OE equivalent of Latin *jubere*, does not accept a passive infinitive construction as object.[3] This precarious situation was probably the reason why King Alfred produced a translation which he himself describes as "sometimes word by word and sometimes according to the sense" (Sweet 1872: 7). He first tried to reproduce the linguistic structure of the Latin original and at the same time to capture its meaning by replacing the plural subject pronoun *qui* by *ða sahlas*, the substantival syntagm of the context that it takes up. But already when it came to the translation of the verb form, he deviated from his original strategy and used a singular instead of a plural verb form probably to avoid the reader's interpretation of the *subject ða sahlas* as the receiver of the order. After deciding on translating Latin *jubere* by OE *bebeodan* the translator was obliged to use a *that*-clause instead of an infinitive construction, but the issue he faced was that no element of the original text was left, which could serve as a subject of the *that*-clause. As a consequence King Alfred produced a clumsy translation which an OE reader who was not familiar with the Latin original had probably a hard time understanding.

Henry Sweet (1872: 170/171), whose edition of King Alfred's translation (= Example (3)) was quoted, provided a PDE version of the text, which captures the meaning of the Latin original, but not its form:

---

2. This feature is known as 'Relativische Anknüpfung' or 'Relativischer Anschluß', cf. Moessner 1992: 338f.).

3. See BT, s.v. *bebeodan*.

(5)   The poles were to be cased in gold

Bramley's nearly cotemporaneous edition of Pope Gregory's Latin original made use of a linguistic pattern for his PDE translation which was not available or was at least very rare in King Alfred's lifetime.[4]

(6)   And they are commanded to be covered with gold   (Bramley 1874: 127)

In their own corpus study López Couso & Mendez Naya (1996: 416) detected the following two late ME examples.

(7)   *siþ   Crist   byddiþ men of his suyt   þat þei **schulde** not haue*
      since Christ asks   men of his retinue that they should   not have
      two cotis.
      two coats
      "since Christ asks the people of his retinue that they should not have two coats"
                                           (c.1400, *Wycliffite Sermons*, I 480)

(8)   *Ffor moyses doth bydde in oure lawe that Euery Advowterere we*
      for   Moses does ask   in our law   that every adulterer   we
      xuld   *qwelle.*
      should kill
      "For Moses asks in our law that we should kill every adulterer"
                                           (?a 1475, Ludus Coventriae, 203)

They interpret them as indications of the "dissolution of the past/non-past relationship between *shall/should*, thus anticipating the Present-day English … non-past obligational use found in *should* …" (1996: 416).

   The assumption of an OE origin of the use of *should* "with reference to present or future time or in timeless contexts … as a variant of subjunctive forms" and that this use was grammaticalised in Early Modern English is also shared by Rissanen (1999: 235). Yet his claim does not primarily target noun clauses; it is illustrated by this main clause from Shakespeare's works:

(9)   So **should** a murtherer looke, so dead, so grimme.
                                           (*Midsummer Night's Dream* III.2)

Ogawa (1989) – who analysed the relation between inflected verb forms (indicative and subjunctive) and constructions with the modal verbs *agan, cunnan*, *durran, magan*, *motan, *sculan, þurfan, willan* and *wuton* in dependent and independent sentences in a poetic corpus and in an earlier and a later OE prose corpus – comes

---

4.   Los (2005: 179) assumes that *to*-infinitive constructions started to replace *that*-clauses "between 950 and 1050".

to a different conclusion. He rejects the so-called substitution theory, which claims that the growing numbers of ambiguous forms resulting from the simplification of the verbal paradigm were the reason for their replacement by modal constructions already in OE. He argues that "this 'substitution theory' does not derive from any full documentation of the processes it assumes" (1989: 230). In his corpus analysis he identifies format, construction type, genre, and meaning of the matrix verb as the factors, which in addition to chronology influence the distribution of inflected verb forms vs. modal constructions already in OE, and from these findings he concludes that the development of modal constructions as a subjunctive equivalent postdates OE at least (Ogawa 1989: 232).

The lesson to be learnt from this overview of earlier studies on *should* is that in synchronic PDE studies the focus is on its use in object clauses depending on suasive expressions, whereas in historical studies other syntactic environments are considered as well. In these environments *should* alternates with the subjunctive. This was the motivation for including all environments in my study where the subjunctive is attested, namely main clauses and several types of subordinate clauses: adjectival relative clauses, noun clauses, comparative clauses, and adverbial clauses.

4.   Quantitative analysis

4.1   Old English

The search of the variants of *should* in the 70 files of the OE part of the HC produced 380 hits (= 9.20/10,000 words), but only 56 of them contained *should* with non-past reference (= 14.74%). The texts of the earliest sub-period (= O1) contain no occurrence at all, then the absolute frequency rises from 9 in O2, to 15 in O3, and 32 in O4. In relative terms this corresponds to 1.00/10,000 words in O2, to 0.60/10,000 words in O3, and to 4.75/10,000 words in O4. The dialect analysis revealed that examples of *should* with non-past reference are attested only in texts from the West Saxon dialect area, unsurprisingly, as the great bulk of extant texts are from that area.

The examples belong to the following construction types: main clause (12 examples), relative clause (23 examples), comparative clause (10 examples), adverbial clause (7 examples). Among the four remaining examples there are two in which *should* occurs in a clause which functions as a noun postmodifier (= appositive clause in Quirk et al.'s terminology (1985: 1049–1053); types such as *the commandment that we should not lie* and *the reason why we should not lie*). Only two examples attest *should* in noun clauses:

(10)   *hit is cuð,   þæt seo forestihtung **sceolde** beon mid benum gefylled*
       it is known   that the prophecy   should be   mid prayers fulfilled
       "it is known that the prophecy should be fulfilled through prayers"

(COGREGD3, p. 55)

(11)   *Eala, eala, fela is þæra,   þe sacerdhades on unriht   gyrnað, swa*
       alas alas many is of those-GEN who priesthood in injustice desire as
       hit þincan mæg, swyðost for idelum gylpe   and for gitsunge
       it   seem may mostly for vain   arrogance and for desire
       *woruldgestreona,   and ne cunnon na, þæt hy   cunnon **sceoldan**.*
       of wordly goods-GEN and not can   not what they can   should
       "Alas, alas, there are many, who aim at priesthood unrightfully, as it may seem,
       mostly because of vain arrogance and because of their desire of worldly riches,
       and they are not capable of what they should be capable of."

(COINSPOL, p. 97)

Example (10) contains the same construction as the title of Hundt's 1998 article, where the *that*-clause functions as subject. Traugott (1992: 234) prefers to analyse similar OE examples as "complements of predicates", because they do not occur in initial position. Whatever analysis one chooses, the adjective *cuð* does not belong to the class of emotive adjectives or to those which express concepts concerned with modality or volition, which according to Quirk et al. (1985: 1224) are usually followed by *that*-clauses with putative *should*. Example (11) illustrates an object clause with putative *should*, but it differs from PDE mandative constructions in that the matrix clause does not contain an expression of command, request, or wish. It is in fact a free relative clause: the introductory element *þæt* of the dependent clause is not a conjunction, but a nominal relative pronoun.

The conclusions which can so far be drawn from my OE data are that they contain only very few examples of *should* with non-past reference and that among those few there are only two which qualify as noun clauses, which, however, are not completely parallel to PDE dependent mandative constructions with putative *should*. So the hypothesis of an OE origin of putative *should* is not supported by my data.

4.2   Middle English

The ME part of the HC comprises 91 files, and they contain 600 occurrences of *should* (= 9.86/10,000 words), but only 149 have non-past reference (= 24.83%). Of these three are attested in sub-period M1 (= 0.27/10,000 words), 14 in sub-period M2 (= 1.44/10,000 words), 66 in M3 (= 3.58/10,000 words), and another 66 in M4 (= 3.09/10,000 words).

Most occurrences of *should* with non-past reference are attested in texts from the East Midland dialect area (75 = 50.33%), followed by those from the Northern dialect area (38 = 25.50%), by those from the Southern dialect area (20 = 13.42%), and those from the West Midland dialect area (16 = 10.73%). When we compare these figures to the distribution of texts from the different dialect areas represented in the ME part of the HC, *should* with non-past reference is overrepresented in texts from the East Midland and Northern dialect areas.[5]

The different construction types occur with the following frequencies: main clause (78 examples), adjectival relative clause (13 examples), comparative clause (6 examples), adverbial clause (23 examples), noun postmodifier (12 examples), noun clause (17 examples).[6]

When mandative constructions, i.e., *that*-clauses in which present tense subjunctives and *should*-constructions alternate, are defined as depending on matrix clauses which contain an expression of wish, command, or exhortation, only four of the 17 examples of noun clauses of my ME corpus qualify as prototypical mandative constructions.[7] They function as objects. Two were presented above as (7)–(8); the remaining two, (12) and (13), are from the early 14th century:

(12) *drightin for-bede þat þou* **suld**   *thinc   to do suilk a dede*
      lord   forbid   that you should consider to do such a deed
      "God forbid that you should consider doing such a deed"

                                                        (CMCURSOR, p. 446)

(13) *Vr   lagh will, bath þat es and was, þat in þe muth   o tua or thre*
      Our law will both that is and was that in the mouth of two or three
      Suld   *alkin soth   stand and be.*
      should always truth stand and be
      "our law, both the one that is and the one that was, requires that in the mouth of [every] two or three truth should always be and remain"

                                                        (CMCURSOR, p. 1018)

In descriptions of subjunctive use in earlier periods of English the list of elements governing noun clauses in which the subjunctive alternates with modal constructions also include "expressions of emotion … and other mental activities" and "verbs of saying, declaring, lying and denying" (Visser 1963–1973: §§ 872–875). The other 13 instances of *should* in my ME corpus occur in these environments,

---

5.   The shares of texts from the individual dialect areas are: East Midland 48.35%, Northern 10.99%, West Midland 16.48%, Southern 20.88%, Kentish 3.30%.

6.   The 17 noun clauses represent 11.41% of the 149 examples containing *should* with non-past reference.

7.   They represent 23.53% of the 17 noun clauses containing *should* with non-past reference.

10 as objects, three as subjects; most of them date from the last two ME sub-periods (cf. Examples (14) and (15)):

(14)   *For yif I sende him unto, J wene men* **sholde** *him shame do*
       for if I sent   him onto I fear men should him shame do
       "For if I send [someone else] to him, I am afraid people might put him to shame"   (CMHAVELO, p. 53)

(15)   *me thynketh for your worship that ye   shold   not procede ayenst*
       me seems   for your worship that you should not proceed against
       hym ouer hastely
       him over hastily
       "it seems to me that your worship should not proceed against him too hastily"   (CMREYNAR, p. 54)

These figures show that the relative frequency of *should* is slightly larger in the ME than in the OE corpus, that the percentage share of *should* with non-past reference is nearly twice as large as in the OE corpus, and that ME is the first period in which there are some instances of mandative constructions with *should* which fulfill the requirement that their matrix clauses contain an expression of command, wish, or exhortation.

A similar picture evolves from an analysis of the 1,416 examples of *should* in 35 'senses' quoted in the MED. Thirty-three senses are distinguished for *should* as a modal auxiliary, and *should* as 'past form with present meaning' is illustrated by 253 examples in 26 of the 33 senses (= 17.86% of all examples). They occur in clauses with the following syntactic functions: main clause (110 examples), noun clause (62 examples), adverbial clause (39 examples), adjectival relative clause (23 examples), and comparative clause (19 examples). The 62 noun clauses represent 24.51% of all examples of *should* with non-past reference. Only 12 of them (= 19.35%) qualify as prototypical mandative constructions, i.e., they depend on a matrix clause containing an expression of command, request, or wish. They are comparable to Examples (7)–(8) and (12)–(13) from the HC:

(16)   *The kyng comaundeth his constable anon… That he ne*   **sholde**   *suffren…*
       the king commands his constable anon   that he not should suffer
       Custance   in with his regne for tabyde   Thre   dayes.
       Constance in with his realm for to abide three days
       "The king commands his constable immediately … that he should not suffer … Constance to remain in his kingdom for three days"
                                        ((c1390) Chaucer CT.ML. (Manly-Rickert) B.796)

(17)   *I   am   not   so   hardy   …   For   to   desire   that   ye*   **shulde**   *love me.*
I   am   not   so   bold   for   to   desire   that   you   should   love   me
"I am not so bold to desire that you should love me"
<div align="right">((a1500 (c1370) Chaucer Comp. L. (Benson-Robinson) 85)</div>

(18)   *God   forbede   that   al   a   compaignye*   **Sholde**   *rewe   a   synguler   mannes   folye.*
God   forbid   that   all   a   company   should   rue   a   single   man's   folly
"God forbid that an entire fraternity should pay for the folly of a single indi-
vidual"   ((c1395) Chaucer CT.CY. (Manly-Rickert) G.997)

It is quite striking that ten of these 12 examples have the form *God forbede/shilde/
deffende þat* …, which Iyeiri (2017: 210–213) proposes to analyse as a set phrase
functioning either as a main clause or as a sentential adverb. If this analysis is ac-
cepted for the ten examples from the MED, it should also be applied to Example (12)
from the HC. In the present context I prefer the analysis as noun clauses for all
of them.

In five examples from the MED the verb of the matrix clause belongs to the
categories verbs of saying and expressions of emotion and other mental activities
(cf. Example (19)):

(19)   *I   hope   that   men*   **schuldyn**   *se   This lond   amende*
I   hope   that   men   should   see   this country improve
"I hope that people see this country improve"
<div align="right">((a1393) Gower CA (Frf 3) 8.3004)</div>

A comparison of the relevant constructions in the ME part of the HC and in the
MED reveals that the share of mandative constructions with *should* depending
on present tense matrix clauses amounts to 11.41% in the HC and to 27.42% in
the MED.[8]

4.3   Early Modern English

The 81 files of the EModE part of the HC contain 1,106 occurrences of the variants
of *should* (= 20.07/10,000 words); of these 295 (= 26.67%) have non-past reference.
Sub-period E1 contains 68 examples (= 3.57/10,000 words), E2 contains 88 exam-
ples (= 4.64/10,000 words), and E3 139 examples (= 8.13/10,000 words). According
to these figures the relative frequency of *should* is more than twice as large in the
EModE than in the ME corpus, but the percentage share of *should* with non-past
reference increases by only 2% between ME and EModE.

---

8.   The much larger share of examples in the MED must be interpreted against the fact that they
were collected from the entry *shulen* and not from whole texts as those in the HC.

The examples are distributed across the construction types as follows: main clause (78), noun clause (105), relative clause (18), comparative clause (12), adverbial clause (54), noun postmodifier (28). The noun clauses realise the functions subject (32), subject complement (16), and object (57).

Out of the 16 subject complement clauses of my data 15 come from the same text; it describes man's anatomy, and the relevant sentences have the structure '*the method/the cause/why*-clause + *is* + *that*-clause'. Descriptions of mandative constructions usually do not include this type, and I will follow the same strategy.

(20)   *The seconde <u>cause</u> is, that the matter that commeth of woundes made in the head pearsing the skul, **shoulde** by it the better be defended and kepte from Piamater, and hurting of the brayne.*   (CESCIE1A, p. 29)

The 32 subject clauses belong to the construction types '*it* + impersonal verb + (experiencer object) + *that*-clause' and '*it/this* + *is/becomes/seems* + adjective + (*that*)-clause':

(21)   *it <u>behoueth</u> that they that haue done such offences **should** be broughte (not with angre, but rather with merciful accusers) vnto iudgemente*
(CEBOETH1, p. 102)

(22)   *Thou dost rightly judg: but he who thinks it hard to assent to a Conclusion, it is <u>fit</u> he **should** demonstrate that the Premises are untrue*   (CEBOETH3, p. 178f.)

The adjectives which are complemented by a *that*-clause with *should* share the semantic feature evaluative, and elements of this semantic class figure also in the studies on PDE mandative constructions.

The 57 object clauses of my data depend on verbs of saying (e.g., *mention, tell*), of mental activity (e.g., *fear, doubt, suppose, imagine*), and of volition (e.g., *desire, endeavour, forbid, command, wish*). The last group, which is represented by 21 examples, meets best the requirement of mandative constructions.

(23)   *The worde of God doeth shew plainlie that there be witches, & commaundeth they **should** be put to death.*   (CEHAND2A, p.B2V)

(24)   *if that you doe desire that I **should** then come unto you, I will doe it with all possible speed.*   (CEPRIV2, p. 29)

Generalising on my EModE data, I conclude that this is the first period in which the construction type which in studies on PDE is described as the mandative construction is robustly attested (89 examples). Consequently the answer to my first research question is that it was only in the EModE period that the combination of *should* + infinitive became a competitor of the mandative subjunctive.

5. Interpretation of the quantitative results

When we assume that the HC is a representative corpus and its data allow general-
isations about language use in the periods OE, ME, and EModE, my data show that
the relative frequency of *should* increases from period to period and that the largest
frequency change per 10,000 words occurred between ME and EModE (Table 1).

Table 1. *Should* in the HC

	OE	ME	EModE
Size of corpus in words	413,250	608,570	551,000
Tokens of *should*	380	600	1,106
	9.20/10,000	9.86/10,000	20.07/10,000

How can we explain this frequency increase? One possibility is that *should* started to
occur in contexts where non-fact modality was expressed by other linguistic means
before. The subjunctive immediately comes to mind as one of those other linguistic
means. It is a widely held belief that it was replaced by modal constructions. Fischer
(1992: 262) claims that at the end of the ME period modal constructions far out-
weighed the subjunctive, and Gorrell (1895: 117) assumes that this development
started already in the OE period. As a matter of fact, OE past tense matrix clauses,
especially those with verbs of demanding, ordering, wishing, etc., often governed
that-clauses with a so-called hortative subjunctive (matrix verb underlined, sub-
junctive in bold):

(25) *Ond se casere þa* <u>*bebead*</u> *þæt hine man* **forlete** *unbyrgedne*
And the emperor then ordered that him one left-SBJV unburied
"and then the emperor ordered that one should leave him unburied"
(COMARTYR, p. 23)

In past tense the indicative vs. subjunctive contrast was even less formally marked
than in present tense, so that particularly in the sub-periods O3 and O4 modal
constructions with obligational *should* + infinitive replaced the subjunctive, and
they are at least partly responsible for the frequency increase of *should* in the later
periods (matrix verb underlined, *should* in bold).

(26) *Ure drihten* <u>*bebead*</u> *his discipulum þæt hi* **sceoldon** *læran. and*
Our Lord commanded his disciples that they should instruct and
tæcan eallum þeodum ða ðing þe he sylf him tæhte
teach all nations the things that he self them taught
"our Lord commanded his disciples that they should instruct and teach all
nations those things which he himself had taught them" (COAEPREF, p. 6)

As a second reason for the frequency rise of *should* I consider the possibility that *should* acquired features which made it fit for new contexts. Already in OE – in my data from sub-period O2 onwards, but more often in sub-periods O3 and O4 – *should* is also attested in present tense contexts, and here most frequently in main clauses and in relative clauses (Table 2).

Table 2. Tokens of *should* with non-past reference in different construction types

Tokens of *should* with non-past reference	OE	ME	EModE
	56	149	295
	1.36/10,000	2.45/10,000	5.35/10,000
	14.74%	24.83%	26.67%
in main clauses	12	78	78
in relative clauses	23	13	18
in comparative clauses	10	6	12
in adverbial clauses	7	23	54
in noun postmodifiers	2	12	28
in noun clauses	2	17	105

In these examples *should* does not express obligation, but advisability, suitability, or appropriateness.

(27) *Hit is earfoðe eall to gesecganne, þæt se   beðencan sceal, ðe   scire*
It is difficult all to say   that that-M think   shall who shire
*healt. Ne **sceolde** he nan ðing forgyman, ðe   æfre to note mehte:*
holds. Not should he no   thing neglect   that ever to use might
ne   forða   musfellan ne, þæt git læsse is, to hæpsan þinn.
neither therefore mousetrap nor what yet less is for hasp   pin
"It is difficult to list everything that he should consider who governs a shire.
He should neglect nothing that ever might be useful: neither a mousetrap nor
what is even smaller, the peg for a hasp."   (COLAW4, p. 455)

(28) *Nu her þara   banena byre nathwylces frætwum   hremig on flet*
Now here of.those slayers son some   trappings-INS boasting on floor
gæð, morðres   gylpeð, ond þone maðþum byreð, þone þe ðu   mid
goes murder-GEN boasts and that treasure wears which you with
*rihte rædan **sceoldest.***
right possess should
"Now some son or other of those slayers walks in the hall, exulting in the trap-
pings, boasts of the killing and wears the treasure which by rights you should
possess!"   (transl. Swanton) (COBEOWUL, p. 32)

The percentage share of *should* in present tense contexts is small, but it rises from 14.74% in OE to 26.67% in EModE with the largest frequency leap between OE and ME. The frequency increase of *should* in present tense contexts is another explanation for the overall frequency increase of *should* from period to period, and it is most conspicuous in noun clauses.

The figures in Table 2 suggest that *should* with non-past reference entered the stage via the construction types relative clause and main clause, but from Table 3 we learn that we have to wait until the EModE period before we find a more than negligible number of occurrences of *should* in contexts which meets the requirement of a mandative construction. Table 3 also shows that *should* with non-past reference in EModE noun clauses is more frequent after matrix verbs of saying and thinking than after matrix verbs expressing a command or wish, which is in line with Ogawa's (1989) finding that the meaning of the matrix verb influences the form of the verbal syntagm in the dependent clause.

Table 3. Absolute frequencies and percentage shares of *should* with non-past reference in HC-OE, HC-ME, HC-EModE; and MED

	OE	ME	EModE	MED
should non-past reference	56	149	295	253
	14.74%	24.83%	26.67%	17.87%
in noun clauses	2	17	105	62
	3.57%	11.41%	35.59%	24.51%
in mandative constructions	0	17	89	17
	0.00%	11.41%	84.78%	27.42%
in mandative constructions type *command*	0	4	21	12
	0.00%	23.53%	23.60%	70.59%
in mandative constructions type *say/think*	0	13	68	5
	0.00%	76.47%	76.40%	29.41%

Can we find an explanation for the sequence in which *should* intruded into these construction types?

I suggest that the explanation lies in the contribution of *should* to the meaning in each of them. In main clauses and in relative clauses it is due to the presence of *should* that we understand that a certain action or object is suitable or appropriate. In Example (27) *should* makes us understand that mousetraps and other little things are appropriate tools for a landowner, and in Example (28) *should* makes us understand that a particular treasure would be more suitable for somebody else to wear. In mandative constructions we have a different meaning constellation. This becomes obvious when we consider the elements which bring about the mandativeness. In Example (23) which comes from a handbook on witches and witchcraft,

the *that*-clause contains the command to kill witches. The matrix clause verb form carries this meaning, but it is enhanced by the presence of *should*, which preserves its original obligational meaning. In Example (24) the action desired by the writer of the letter is that the addressee comes to him/her. When we ask which elements in the sentence contribute to this meaning, it is obvious that the decisive element is the matrix verb *desire*. What does *should* contribute to the meaning of the *that*-clause? Would the meaning change if *should* was left out, if the sentence read *if you do desire that I then come unto you*? As a non-native speaker, I argue that *should* adds a note of uncertainty to the *that*-clause; the writer of the letter is not sure if the addressee really desires his/her visit.

(29)   *But Nature gives to every thing that which is agreeable to, and convenient for them, and endeavours that they **should** not perish before their time.*

(CEBOETH3, p. 144)

Example (29) expresses the intention of the agent of the matrix clause. This meaning is conveyed by the combination of the meaning of the matrix clause verb and the modality of futurity which is added by *should*.

These examples show – and this holds for all examples of *should* in mandative constructions of my EModE corpus – that *should* can occur with its original obligational meaning, but that it can carry other meanings as well. In my EModE data the mandativeness is brought about by the combination of the meaning of the matrix verb and several shades of non-fact modality expressed by *should*. The answer to my second research question is therefore that *should* became only fit as a competitor of the subjunctive in mandative constructions when its original modality of obligation had been extended to include other shades of non-fact modality as well. The meaning extension started in OE relative clauses and made its way via main clauses into noun clauses.

6.   Summary and conclusion

Starting from the present-day British English situation where the present subjunctive and putative *should* + infinitive alternate in mandative constructions and where the second variant is much more frequent than the first, I investigated *should* in all contexts where the subjunctive is attested in the historical periods OE, ME, and EModE. It was the aim of the paper to identify the period when the *should*-construction became a competitor of the present subjunctive in mandative constructions and to trace the developments which led to this competition.

The quantitative analysis of my data from the HC yielded the result that *should* with non-past reference is attested already in OE, but that its percentage share

amounted only to 14.74% of all occurrences of *should*. In OE *should* with non-past reference occurred most frequently in relative clauses, in ME in main clauses. The analysis of the relevant data from the ME part of the HC and from the MED showed that the mandative construction with putative *should* can be traced back to late ME, but that it became firmly established only in EModE. I interpret this as a result of the meaning extension of *should* from obligation to other shades of non-fact modality.

References

Anderson, John. 1991. Should. In Dieter Kastovsky (ed.), *Historical English syntax*, 11–30. Berlin: Mouton de Gruyter. https://doi.org/10.1515/9783110863314.11

Bramley, H. R. 1874. *S. Gregorii Magni regulæ patoralis liber. S. Gregory on the pastoral charge: The Benedictine text*, with an English Translation. Oxford: James Parker and Co.

BT: Bosworth, Joseph & T. Northcote Toller. 1898. *An Anglo-Saxon dictionary*. London: Oxford University Press.

Fischer, Olga. 1992. Syntax. In Norman Blake (ed.), *The Cambridge history of the English language, Volume 2. 1066–1476*, 207–408. Cambridge: Cambridge University Press. https://doi.org/10.1017/CHOL9780521264754.005

Francis, W. Nelson & Henry Kucera. 1979. *Brown Corpus manual: Manual of information to accompany a standard corpus of Present-Day edited American English for use with digital computers*. http://icame.uib.no/brown/bcm.html

Goossens, Louis. 1987. The auxiliarization of the English modals: A functional grammar view. In Martin Harris & Paolo Ramat (eds.), *Historical development of auxiliaries*, 111–143. Berlin: Mouton de Gruyter. https://doi.org/10.1515/9783110856910.111

Gorrell, Joseph Hendren. 1895. *Indirect discourse in Anglo-Saxon*. Baltimore: The Modern Language Association of America.

HC: *The Helsinki Corpus of English Texts*. 1991. *Department of Modern Languages*, University of Helsinki.

Hundt, Marianne. 1998. It is important that this study (*should*) be based on the analysis of parallel corpora: On the use of the mandative subjunctive in four major varieties of English. In Hans Lindquist, Staffan Klintborg, Magnus Levin & Maria Estling (eds.), *The major varieties of English. Papers from Maven 97, Växjö 20–22 November 1997* (Acta Wexionensia Humaniora 1), 159–175. Växjö University.

Iyeiri, Yoko. 2017. Recent changes in the use of the verb 'forbid'. *Memoirs of the Faculty of Letters, Kyoto University* 56. 195–218.

Johansson, Stig, Geoffrey Leech & Helen Goodluck. 1978. *Manual of information to accompany the Lancaster-Oslo/Bergen Corpus of British English, for use with digital computers*. Oslo: Department of English, University of Oslo.

López-Couso, María José & Belén Mendez Naya. 1996. On the use of the subjunctive and modals in Old and Middle English dependent commands and requests. Evidence from the Helsinki Corpus. *Neuphilologische Mitteilungen* 97. 411–421.

López-Couso, María José & Belén Mendez Naya. 2006. Complement selection in Early English dependent desires: A look at commands and requests. *Estudios Ingleses de la Universidad Complutense* 14. 33–53.

Los, Bettelou. 2005. *The rise of the to-infinitive*. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199274765.001.0001

MED: Middle English Dictionary. https://quod.lib.umich.edu/m/middle-english-dictionary/dictionary

Moessner, Lilo. 1992. Relative constructions and functional amalgamation in Early Modern English. In Matti Rissanen, Ossi Ihalainen, Terttu Nevalainen & Irma Taavitsainen (eds.), *History of Englishes. New methods and interpretations in historical linguistics*, 336–351. Berlin: Mouton de Gruyter. https://doi.org/10.1515/9783110877007.336

OED. Oxford English Dictionary. https://www.oed.com/

Ogawa, Hiroshi. 1989. *Old English modal verbs: A syntactical study*. Copenhagen: Rosenkilde and Bagger.

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. Volume 3. 1066–1476*, 187–331. Cambridge: Cambridge University Press.

Sweet, Henry. 1872. *King Alfred's West-Saxon version of Gregory's pastoral care* (EETS 45/50). London: Trübner.

Traugott, Elizabeth Closs. 1972. *A history of English syntax*. New York: Holt, Rinehart and Winston.

Traugott, Elizabeth Closs. 1992. Syntax. In Richard M. Hogg (ed.), *The Cambridge history of the English language, Volume 1. The Beginnings to 1066*, 168–289. Cambridge: Cambridge University Press. https://doi.org/10.1017/CHOL9780521264747.005

Visser, F. Th. 1963–1973. *An historical syntax of the English language*. Three Volumes. Leiden: Brill.

Warner, Anthony. 1993. *English auxiliaries: Structure and history*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511752995

PART III

Semantics and pragmatics

Shifting responsibility in passing information

Stance-taking in Sir Thomas Bodley's diplomatic correspondence

Gabriella Mazzon

University of Innsbruck

Within the subfield of historical pragmatics, this paper focuses on the linguistic means used for transmitting knowledge and signalling stance in a sample of Early Modern diplomatic correspondence, a text-type that, while partly codified and formulaic, shows individual variation according to topic, correspondents, and other contextual factors. Sir Thomas Bodley's diplomatic missions in several countries produced a substantial corpus of correspondence between him and the Court. The paper examines Bodley's strategies for the reporting and assessment of information, while conveying different levels of involvement vs. distancing. The results show that Bodley is especially cautious about the reliability of sources of the information he gathers, as is shown by the vast array of linguistic means he employs to convey his stance towards the information and its sources.

Keywords: historical pragmatics, diplomatic correspondence, Early Modern English, stance, evidentiality, epistemic scales

1.   Introduction and theoretical background

This contribution moves within historical pragmatics (Jucker & Taavitsainen 2010), a relatively young branch of research that investigates how pragmatically relevant elements were encoded in the past. The sub-branch that is of interest here is the one that investigates expressions established as part of text-types produced by specific communities and/or social groups, as well as any variation in these expressions to encode modulations of affect, emotion and intersubjectivity tied to the micro-context and to the management of face and relational needs. All these elements can be subsumed under the notion of stance, employed here as a superordinate term (Du Bois & Kärkkäinen 2012: 438); this notion refers to the ways in which identity can be signalled using language and other semiotic systems.

Stance represents the intersection between general socio-cultural categories and reaction to immediate contexts, altering the relation between language items and extra-linguistic worlds and introducing linguistic variation related to the subjective attitudes of the speaker in any given interaction.

Letters are an ideal field for pragmatic studies; Early and Late Modern English collections of correspondence have recently been the object of intensive investigation, especially because the epistolary genre seems to have developed its still current form in the Early Modern English period (Del Lungo Camiciotti 2014: 18), and because the boundary between public and private letters was less clear-cut, from both the socio-cultural and the stylistic points of view, than in modern times (Marcus & Evans 2019: 67–68). Even if correspondence is partly codified and formulaic, it shows individual variation according to topic, correspondents and other variables, and is particularly rich in stance markers.[1] Work on diplomatic correspondence (Fitzmaurice 2006; Okulska 2006), has concentrated on the narration of events, the linguistic means used for the conveying of knowledge and the subjective stance taken in regard to that knowledge. Okulska also discusses the reporting of discourse in Early Modern English correspondence, as does Evans (2017). Okulska, in particular, emphasises that diplomatic correspondence has a certain "potential" to include fixed as well as optional linguistic patterns (2006: 48), but does not show abundant evaluative marking, with a preference for neutral stance (2006: 67) – as we will see, this is not consistent with my findings.

A subset of Sir Thomas Bodley's correspondence was the object of a previous study that considered several pragmatically relevant items such as terms of address and the Politeness system. Within that paper, one section was also devoted to stance-taking and evidentials, as expressed both by Bodley and by his correspondents (Mazzon 2014). For the present paper, I take into account a different and more substantial subset of letters (i.e., two hundred letters, and only letters written by Bodley himself), in order to look more closely at the ways in which Bodley introduces new information, refers to shared/unshared information, takes subjective, epistemic and evaluative stance[2] towards the information, indicates its relevance, or acknowledges the limitations of his knowledge.

---

1.   For a review of studies see Palander-Collin (2010) and contributions in Culpeper (2011); on non-private correspondence in particular, also from a socio-cultural perspective, see contributions in Dossena & Fitzmaurice (2006), especially Del Lungo Camiciotti (2006); see also Gray et al. (2011: 231), Williams (2013), and Del Lungo Camiciotti (2014).

2.   See Nuyts (2001) for the definition of subjectivity adopted here. For definitions and discussions of stance see Du Bois (2007) and, in a more strictly sociolinguistic perspective, Kiesling (2009).

A first example (1) shows how complex the expression of subjective stance can be in such correspondence – here, evidentials (in a wider sense, as explained below) and epistemic markers (in bold), as well as performatives (in roman), counterfactuals and modality markers (underlined), co-occur and contribute to the careful weighing of the information:

(1)   *I <u>would yt were</u> not so **apparant** as I do expresse yt, but <u>for mine owne parte I stand assured</u>, & yt is a comfort unto me that your L. <u>shall</u> **parceave** yt that matters have bin caried worse yn these countreyes then by waie of writing from mee <u>could</u> conveniently be signified.* [0929]³

Here, Bodley employs several means of hedging and qualification of stance, and shows (in spite of the use of the verb *assure*) his uncertainty. What makes this example particularly interesting and representative of Bodley's strategies is the facework that accompanies the expression of stance towards content, which involves boosting the interlocutor's face and damaging his own face.

The study takes into account widely researched evidential and epistemic markers which convey the text-producers' authority over knowledge and their level of responsibility towards the content of their report (Mushin 2013: 628), while also hedging or boosting the importance of the propositional content reported (Hyland 2005: 175). The items employed to this aim, in particular, are verbs and phrases of reporting and cognition, modal markers and other elements indicating stance such as markers of evaluation, quantification and approximation (Chafe 1986; Nuyts 2001; Clift 2006; Gray & Biber 2014; Marin-Arrese 2015). In accordance with an established view (see, e.g., Papafragou 2000; Palmer 2003), a wider perspective on modality, not limited to modal verbs, is adopted here. I also adopt an inclusive view of evidentiality, which sees linguistic elements conveying epistemicity, reliability, and source-giving as devices that cooperate in the construction of the expression of speaker stance, in particular when it comes to legitimisation and (inter)subjective validation.⁴ Epistemic and evaluative stance concur to present the writer's point of view (whether in terms of personal attitudes or of evaluation of an epistemic state) and often reinforce each other, and this is why they are treated together here. The analysis will be mostly qualitative, as stance can only be interpreted within a specific context of production, although reference to the high frequency of some items will be made (Du Bois 2007: 78; Gray & Biber 2014: 220). Table 1 provides

---

3. The number following each example is the ID number of the letter the passage is extracted from, as assigned in the database.

4. Some current literature keeps epistemic stance and evaluative stance separate, but see Hyland (2005), Du Bois & Kärkkäinen (2012: 438), Mushin (2013: 633), Fetzer & Oishi (2014: 322) for different categorisations that suggest a more integrated view.

Table 1. A classification of the forms found. Items in SMALL CAPS refer to lexemes; items in *italics* are not analysed in other literature on Early Modern English stance

Stance type	Category	Forms/examples
Subjective	Involvement	I, my (private) opinion, *for mine owne parte*
		We, friend, BE PERSUADED
	Distancing	Impersonal constructions: it is thought, etc.
		Passive forms: *I AM ASSURED* that…
		He, they
Epistemic	Reporting	Representative performatives: SIGNIFY, ADVERTISE, INFORM, CONFIRM…
		Verbs of saying: SAY, TELL, WRITE
		REPORT (n.), SPEECH
	Reinforcement	Direct speech
		Appeal to shared knowledge (exx. 36, 37)
		Appeal to direct witness (ex. 25)
		PROOF, TRUTH, WARRANT (n.), NO DOUBT, FOR CERTAIN, WITH CERTAINTY
		CERTIFY, *ASSURE, BE MANIFEST*
	Downtoning	THINK, SUPPOSE, PERCEIVE, DOUBT (n./v.), PRESUME, CONJECTURE, (n./v.) ALLEGE, *IMAGINE, SUSPECT*, SEEM, APPEAR, BE UNCERTAIN, BRUIT (n./v.)
		may, can
Manner – Evaluation	Generalisation, quantification	*GENERALLY, ABROAD*, COMMON OPINION/CONJECTURE, SOME/ALL/EVERY MAN
		Impersonal forms: it is feared…
	Deferring	Man of credit/experience/understanding/judgement
	Contradicting	Exx. 30–33

the classification used in this investigation. The example items have been taken from Gray et al. (2011: 221–222, 230, 247–248), with additional forms (in italics) that were not investigated by previous research on Early Modern English stance.

After a brief illustration of the sample and of its relevance for this type of investigation (§ 2), the study looks at how new knowledge is framed or introduced by specific practices that immediately establish the degree of directness and/or reliability and certainty of the transmitted piece of knowledge (§ 3). Section 4 discusses in detail the ways in which Bodley refines his stance by moving along epistemic scales or by qualifying his sources, while section 5, which precedes a short conclusion, focusses on two discourse variables, having to do with the polarities 'appearance vs. truth' and 'shared vs. new information', which play an important role in conveying stance in these texts.

2. The sample

The database employed for this study is a collection of Sir Thomas Bodley's diplomatic correspondence, which has recently been made available in searchable text format (see Source).

Before retiring and founding the well-known library that bears his name, Thomas Bodley (1545–1613) was sent by Queen Elizabeth I on several missions to the United Provinces of the Netherlands, as well as to several locations in Germany, Denmark and France. In his capacity as a representative of the Council of State, he conducted various negotiations in relation to Protestant revolts and to the mercantile issues descending from the related political conflicts. The difficulties of these negotiations increasingly cast doubts on the efficacy of his actions. At the same time, he was dissatisfied by the fact that instructions and responses to his queries seem to have been scant and slow to come. He became increasingly plaintive and rancorous, and repeatedly asked to be recalled, so that he did not get further employment at the end of his missions and subsequently retired from political life.

The database of his diplomatic correspondence includes nearly 700 letters written by Bodley, several in two or three copies, coming from different archives, as well as many letters addressed to Bodley by different people, which will not be considered in this paper. A subset of two hundred letters written by Bodley was selected, respecting a proportional ratio in terms of year of compilation (the database ranges from 1585 to 1596, with the highest numbers clustering from 1589 to 1592), and addressee (with Bodley's direct superiors Burghley and Walsingham representing the large majority).

Bodley's two main correspondents held very important roles at court: William Cecil, 1st Baron Burghley, was at the time Lord Treasurer and had previously been Scretary of State, leader of the Privy Council, and creator of the network of intelligence that Francis Walsingham subsequently led for many years. Burghley died in 1592 and was succeeded in office by his son Robert Cecil, 1st Earl of Salisbury. Thus, some of the later letters in my sample addressed to *my Lord Treasurer*, or even *Lord Burghley*, are intended for Robert. Sir Francis Walsingham was principal secretary to Queen Elizabeth, as well as her 'spymaster', from 1573 to his death in 1590, and he supported the Protestant cause in the Netherlands even more strongly than Lord Burghley, who took over again as secretary when Walsingham died.

Further correspondents, marginally present in the corpus and thus also in our sub-sample, include Robert Devereux (who was later the protagonist, as Earl of Essex, of the famous rebellion), as well as various representatives of the Company of the Merchant Adventurers, who were in conflict with the Hanseatic League. All these correspondents grew increasingly dissatisfied with Bodley's results over the years, and therefore the diplomatic correspondence is extremely interesting

in terms of face-work and the expression of power hierarchies, but the content focusses on the transmission and reporting of knowledge. The reliability and relevance of Bodley's reports became crucial in his correspondence, as he increasingly needed to maintain his credibility and usefulness. At the same time, he needed not to sound overconfident, in order to avoid being contradicted and risk losing face in case he was disproved. This delicate balance can be seen already in the way transmitted pieces of information are introduced, which is the focus of the next section.

3.   Framing reliability of information

As mentioned, I take into consideration the quite extensive range of expressions used to report knowledge, as well as an equally ample range of expressions indicating stance along the evaluative and epistemic ranges. Most of the examples found can be identified as pertaining to three types of stance commonly labelled as epistemic, affect, and manner (Clift 2006: 570). The items involved are, first and foremost, verbs of perception and cognition and performative verbs, but nominals and adjectivals are also frequent (Okulska 2006: 53–54, 65–67: the importance of stance nouns is increasingly taken into account; Gray & Biber 2014: 240). Apart from the conveying of positionings on the 'involvement > distancing' scale (see Hyland 2005), Bodley modulates language forms in his letters also to refer to the shared/unshared nature of information, and to the quality and reliability of his sources.

If we look at how pieces of information are introduced, we can first see the presence of personal pronouns, not only the *I* of involvement and responsibility-taking,[5] but also the occasional *we*, as in (2); in such cases, it is not always clear who exactly is included in this plural reference – more than anything else, this appears to be a strategy to reinforce Bodley's own point of view or authority in reporting something. A converse strategy can be noted when Bodley uses self-reference for delimitation or contrast, as will be briefly mentioned in § 5.

(2)   *It may please your L. to bee advertized that since the writing of my last which **I** sent to your L. the 9 of this moneth, **wee** have had no intelligence from the Camp*
[0924]

Even more noticeable in the sample, however, is the frequency of impersonal constructions (such as *it is said that …*), which indicate stance by 'de-personalising'

---

5.   On the importance of self-mention as a stance-marking strategy see Hyland (2005: 181). On identity construction through self-reference in early correspondence see Palander-Collin (2011). Variation in self-reference is not separately analysed here, but integrated in the review of epistemic devices. Several elements which are mentioned in my classification (see the Table 1) are included in Hyland's notion of engagement (here referred to as involvement).

or distancing claims or statements, while the co-occurring evidentials, whether in personal or impersonal form, signal the degree of commitment of the speaker to the knowledge reported. It may seem counter-intuitive to assign a value of evidentiality and speaker commitment to third-person or impersonal expressions (Nuyts 2001: 385): in such cases, stance in reporting can be argued to be 'one step removed' or transferred, as in (3), in a form of indirect speech, while direct speech, following Mushin (2013: 638), can in itself be considered as "an alternative strategy to a grammaticalised evidential system".

(3) *Master Buzenval **complaineth**, …, that wheras **he had notice** from me of the pleasure of her Highnes for soliciting the states, … he <u>should oppose</u> himself against it, wherein **he saieth** he was mistaken: … But yet howsoever[nevertheless], that <u>he was willing</u> to advance it, and **had spoken** about it to the states*   [1136]

A crucial element in the expression of stance-taking is represented by the verbs of cognition and communication (i.e., verbs of saying and representative performatives) that introduce a piece of information, mostly adding epistemic-stance marking, or follow it, mostly with evaluative-stance import.[6] The most common verb in our sample is THINK,[7] in about half of the cases in its impersonal form *it is thought*, or with third-person reference.[8] Also quite common is SIGNIFY, in transitive constructions, with the now obsolete meanings of "inform, notify", limited to Early Modern English (*OED*). Other verbs employed for representative speech acts of transmitting knowledge are ADVERTISE, also in the meaning "apprise, notify, warn" followed by a *that*-clause, considered now rare/obsolete by the *OED*, and INFORM. The former is mostly used when Bodley is reporting to his superiors, often in letter openings, while the latter is more frequently found when Bodley reports that he has received information, and thus in the passive voice. Other frequent verbs are PERCEAVE, SUPPOSE, *BE PERSUADED* (the active form is not used as an epistemic evidential), and DOUBT. Verbs of saying like SAY and TELL or WRITE are not extremely frequent except in the third person, when dialogues are reported in a form of 'indirect speech', as was seen in Example (3) above. Most of the uses are therefore not cases of 'personal meta-communication' such as those studied in dialogic texts by

---

6. A comparable analysis of this type of performative was carried out on a different and slightly later letter sample by Williams (2013: 113ff.).

7. Verbs in small caps refer to lexemes; those also in italics are not included in other studies on stance in Early Modern correspondence quoted in this paper.

8. This result is consistent with findings in Evans (2017: 49) on reporting discourse in sixteenth-century correspondence, in which it is considered a central verb in the expression of evidentiality, with an increase in the use of the passive form in the seventeenth century (Palander-Collin 1999: 135). The semi-lexicalised form *methinks* appears very rarely in the sample, again consistent with other findings (López Couso 1996; Palander-Collin 1999: 137).

Landert (2017), but rather of 'speech description' of the type investigated by Grund (2017). In the sample, however, remarks on the way in which reported speech was delivered in terms of tone or structure, of the type that Grund describes for other co-temporary texts, are not frequent, which is why we do not focus on this typology here, except for the few comments below.

The verbs of saying and of cognition found in the sample often encode a degree of certainty or reliability of the information, as conveyed, for instance, by the difference between PRESUME and CERTIFY. In between these poles stand verbs like *ASSURE*, CONJECTURE, ALLEGE, *IMAGINE* and *SUSPECT*; FEAR can also be included in this list, although it often introduces news that can be considered negative for the recipient, and therefore could convey facework, in addition to epistemic stance. Most of the examples found refer however to the latter: *it may be feared/it is to be feared/it is generally feared* all indicate possible events, rather than announcements, as the modalisation indicates. While CONJECTURE and PRESUME appear mostly in the first person singular, indicating the level of certainty that Bodley expresses about the interpretation of events, ALLEGE only appears in the third person, and thus in cases of 'knowledge reported through indirect speech'. As mentioned earlier, many of the reporting expressions appear in the passive form and in impersonal constructions – this creates a distancing effect, and on some occasions also obscures, or lends increased vagueness to, the sources of the information reported. This concerns *ASSURE* and CERTIFY, and partly *SUSPECT*. *ASSURE* often appears in the passive (*I am assured that …*), indicating that the responsibility for the reliability of the information lies with the original source and not primarily with Bodley himself.

Sources are often named with added details (as will be mentioned in the next section), but are sometimes left unspecified and only gain authority through adverbs like GENERALLY or ABROAD in the original meaning of 'at large, widely', as in *it is generally said that ….* These constructions can be placed on a scale of 'reporting > assessing', which shows Bodley's evaluative stance on the quality of the information transmitted.

There is, finally, a strengthening of the evaluative meanings into evidentials for verbs of appearance, a development studied by Gisborne & Holmes (2007), particularly for SEEM and APPEAR, in which the turning point seems to be precisely in the first part of Early Modern English. Not unexpectedly, these verbs occur mostly in impersonal forms in our sample, often in the raised version that became current at the same time. Their frequency in the subset of letters analysed is however not high, particularly since the 'appearance' meaning detracts from the reliability of the information given through general implicature, and this conveys the suggestion of 'non-truth', which Bodley was of course keen to avoid (see also § 5).[9]

---

9. Conversely, these meanings, and the related hedging and mitigation functions, are central to the later scientific correspondence analysed by Gotti (2006: 41–42).

4.   Modulating epistemic scales and specifying source authority

All the reporting verbs and evidentials exemplified in the previous section are further qualified and modulated through the use of epistemic modals, placing the statement on a scale of certainty. As already mentioned, Bodley employs epistemic stance variation with reference to his own reports more often than when reporting others' speech, as in (4)–(6) below:

(4)   *And besides* **I may conjecture by sundrie special tokens**, *that Co. Maurice underhande hath bin the autor of this course.*   [0438]

(5)   *What my pleading will prevaile,* **I can not certainly say**, *because the mater may happen to be resumed by them*   [0438]

(6)   *for which* **I could not write with certainty, nor specify** *so much as came after to my knowledge*   [1060]

Further qualifications, including hedging expressions, are employed to convey doubt and uncertainty, adding to the subtle modulation of Bodley's carefulness about his responsibilities, especially when there is a contrast between different sources of evidence. This phenomenon is quite frequent, in harmony with the notion that evidentials tend to highlight counterclaims and counterarguments rather than direct claims (Barton 1993: 752). These hedging expressions are also epistemically scalar, for instance FOR CERTAIN as opposed to BE UNCERTAIN, as well as BE MANIFEST, or NO DOUBT vs. SOME DOUBT.

There is frequent 'stacking' of forms which all contribute to a particular stance position. This happens when Bodley chooses to delimit and qualify the reliability of the information he sends (7)–(8), often by downgrading its relevance, or by adding disclaimers (9)–(10), especially as concerns the type of source (*brute, privat letters*) or the lack of confirmation through other or further sources (*for certaine, warrant*).

(7)   *The truthe* **heerof I knowe not otherwise but by his owne report,** *who* **as far as I can parceave** *hath* **no other proofe nor conjecture but by the speeches of John Damport**; *which also* **I knowe not** *how truly* **he avoucheth**   [1410]

(8)   *We are alwaies heer* **so ill provided of intelligence** *of the Ennemy, as* **we know no other certainty** *of his going for France, then* **the brute of the people**, *which is the same that it hath bin that he goeth undoutedly.*   [1079]

(9)   **I have enquired very harde to knowe** *the autor of that Frenche discorse which I sent to my L. of Essex,* **but no bodie heere** *can tell* *it* *for certaine*   [1264]

(10) **I am nether informed** *in time /convenient/ of suche alterations nor in particular maner, and* **by privat letters onely, without warrant sufficient to proceede upon it.**   [1198]

Bodley is occasionally led to confess a total lack of knowledge on certain topics (11)–(14), often highlighting his inability to acquire information due to secrecy or other impediments, in spite of his efforts – several face-saving strategies are therefore employed in such contexts. One exception is (13), where there is no mitigation of the threat to Bodley's own face performed through confessing his lack of information.

(11)   *And what reason they might have, not to sue unto her Highnes for continuance of her succors,* **I** <u>can not</u> **otherwise conjecture** *…. For these are* **maters debated in their secret assemblies, and not imparted unto me**   [1199]

(12)   *but that they hold an other course, then they have done heere tofore,* **I** <u>can not</u> **learne it yet of any**.   [1136]

(13)   *But of any resolution either taken or intended to be taken about it,* **I am utterly ignorant**.   [1264]

(14)   *If Count Hollocke have any suche intelligence with those of Groninghen, as her Highnes is informed,* **it hath bin caried so secretly, that I have not heard it reported**.   [0221]

Of special interest if we are looking at evidentiality is, of course, the way in which Bodley's sources are referred to. In the sample, sources can be characterised by quantity or by quality: starting with the former, one first strategy is indicating how many people share the same information and consider it reliable (15)–(16), with several recurring expressions, such as MY (PRIVATE) OPINION in contrast with THE COMMON OPINION/CONJECTURE or AN OPINION OF MANY. Other common ways of quantitative specification are SOME/MOST/EVERY MAN, ALL MEN (followed by verbs of cognition, opinion or saying), while THE BRUIT (in the now rare or marginal sense of "talk amongst people in general, report, rumour" used in English as from Caxton's time, according to the *OED*) is also often quantitatively modified. The related verb *BRUIT*, found in (16) in the past participle, is reported as still current by the *OED* in the sense of "spread as a report or rumour", and frequently in the passive and in impersonal constructions, precisely to highlight the underspecification of sources, but possibly also their quantitative relevance, on the scale "some/many/all say that …".

(15)   *and there is* **no man heere that** <u>douteth</u>, *but that uppon the sight of her Majesties letter, they will presently condescend to a good composition.*   [0242]

(16)   **Some** *of those in these quarters that have* **best intelligence** *out of Spaine,* **give it out** <u>for certaine</u>, *that the* **bruited** *preparations are nothing so great, as* **heere they did imagine**.   [0736]

A second strategy of quantification is the joint mention of different sources of both the 'quotative' (17)–(18) and 'hearsay' (19) types, i.e., referring to specific texts and interactions or to more generic events and communicative exchanges respectively (Fetzer & Oishi 2014: 323), in which the sources are named in a more indirect way or not named at all.

(17) *But **the bruite is very constant** of his going for France, … **Whiche is also con-*
      firmed, by letters from Brussels   [0331]

(18) *Heerof **we are daily more** <u>assured</u>, not onely **by those cowrses** which he hath*
      *taken all this sommer, but **by divers lettres intercepted, & chiefly nowe of late***
      ***by certain neutrall parsons, men of very good jugement**, which have conversed*
      with the chiefest of the enemies   [1079]

(19) *The same **occurence /advertisement/ ys come not onely from the Councel & the***
      ***Count at the campe, but from divers others in divers /other/ places**, to the states*
      that are heer.   [1166]

When the authority of the source is reinforced not by quantity, but by quality, reliability is emphasized by referring to the specific features of the source itself, e.g., a direct witness of the events reported, or a person of generally established authority, expressed by use of nominals such as *credit, experience, understanding*, as in (20)–(24). Levels of authority can be indicated through certain relational terms used for referents, see, e.g., the use of *friend* in (21) (on the use of this referent in later correspondence, see Nevala 2011: 67–70).

(20) ***I have signified** at lenght unto my L. Tresurer, and have prayed his L. to imparte*
      it unto you, but yf youer L. bee desirous to bee enformed of the same, and of all
      *the dealinge of this peopele in particular maner, **the bearer hereof which was***
      ***suitour for the parties and hath bin heare present at all that hath passed** <u>can</u>*
      <u>report it verye thorowly</u> …   [1188]

(21) *… nevertheles upon this occasyon I have had some conference heer **with a speciall***
      ***freind & one well acquainted with the affaires of those quarters**, who doth hold*
      *yt <u>very likely</u>, to take good effect, **& will give me farder informacion***   [0864]

(22) *Heare I send yow the copies of two intercepted lettres …. They were sent from*
      *Brussels for Rome, and **written by one William Creytton** a Scottishe Jesuit, **a***
      ***man of credit and experience, and the chiefest plotter, as I have heard**, of these*
      troubles in Scotland.   [0672]

(23) *This is th'effect of such talk as I have had with such as ar heer **men of experience***
      ***& good understanding.** With whose opinions I thought yt my duty to acquaynt*
      your L.   [0207]

(24) *But I have very good intelligence, and suche as I doe thinke I may boldly write <u>for certaine</u>, that heeretofore in time of peace, the contribution brought in to the Gouvernor of that towne, hath amounted every moneth to ten thou- sand poundes sterling*   [0438]

Bodley often delegates the report to an (eye)witness or competent third person, distancing himself further or avoiding taking any responsibility at all for the content that is being conveyed. Notice the difference between (25)–(27), where Bodley avoids reporting altogether, and (28), in which a different person is entrusted with the responsibility of confirming (*certifie*) the piece of information given by Bodley himself.

Finally, Bodley entrusts the transmission of information to another type of source entirely, i.e., to other texts, e.g., letters that he sends (translated, deciphered or in the original as it may be) enclosed in his messages. These are either intercepted letters or copies of letters he received. The interpretation and evaluation of the new information is completely delegated to the recipient in (29).

(25) *I will not trouble yow with relating, what we heare of their proceedinges, that are before Gertrudenbergh, for that **Sir Francis being there, <u>can</u> deliver it more directly**.*   [0438]

(26) *They have bin many wayes diminished by the service of this sommer, **as will be written to your L. by Sir Francis himself***   [1189]

(27) *What hath passed betweene the states general and me about the articles of my Instructions, and other like buisnes, **I doe not <u>dout</u> but Master Secretary will informe your L.***   [0802]

(28) *… the prisoner Vasseur is at length delivered, but with very muche troble and sute, **as the bearer herof Sir William Russels servaunt, who hath folowed it long, with exceeding great paines, <u>can</u> certifie your H.***   [0880]

(29) ***I have written the wordes, as they lie in the letter, leaving the construction to your L. conjecture***.   [0236]

Bodley is equally cautious in expression when he is trying to distance himself from the reported news with disclaimers or other forms indicating non-involvement, as already mentioned, or to stress the relative unreliability of the sources, either because the source enjoys low credit (30), or because of contradictory reports. He assesses and balances the reliability of different sources (31)–(32), or a piece of information received from someone else but contradicted by him (33).

(30) *Albeyt wee have sent abroad very many to learne his dessignes & the state of his army yet we finde theyr informations to differ so much, that **wee <u>can</u> build upon nothing**, but …*   [1058]

(31)   *Heere goeth a speeche* that the D. of Parma is returned from the frontiers of France to Brussels: **as some men report**, to conferre with the Emperors Ambassadors: **but as other men affirme, upon intelligence had** that his government was bestowed upon the old Count Mansfeld.   [0361]

(32)   *Of Schincke* **it hath bin heere reported, albeit most men hold it** <u>untrue</u>, that for 40000 dallers to himself, …, he hath yelded up his fort to the D. of Cleve [0117]

(33)   **I do not** <u>doubt</u> but the report of Parmaes death is come to your L which **is almost assured** is heer by Lettres from Andwerp & Bruxels, **but the Authors credyt is not suche that we can** <u>hold yt for certayne</u>, **& yet the advertisement of his deadly sicknes hath bin often continued by sondry mens lettres.**   [0831]

These examples lead us to a brief examination of further strategies that surround the acquired knowledge conveyed, i.e., those involving ascertaining the 'truth' and referring to common ground.

5.   Further stance-related strategies: Appearance vs. truth, shared vs. new information

In the previous section, we saw how Bodley makes use of a range of devices to scale epistemic stance. Here we examine how Bodley occasionally stresses the contrast between appearance and truth, carefully distinguishing between what is 'only' reported and what is considered factual (at least by him), as in (34). He may even explicitly comment on his own caution, claiming his epistemic stance as the deciding factor in reporting news, thereby reinforcing his own reliability (35). Notice the use of *mine owne parte* as a self-reference in both examples to distinguish between Bodley's own stance and other evidence, thus emphasizing the subjectivity and partiality of his views.

(34)   *It* <u>seemeth</u> **by his report** that an effectuall letter from her Majestie will accomplishe all, **though for mine owne part,** I see them all this contrey over possessed with suche a carelesnes of gratifieng her Highnes **as I** <u>dout</u> **the successe**.   [0142]

(35)   *I do assure your L.* **for mine owne parte that I** <u>cannot</u> **bee more carefull** then I have bin & will bee, to **advertise nothing** <u>for certain</u> **unles I know yt** <u>to bee certain</u>, **& whatsoever ys** <u>doutfull</u> **to report yt thereafter**.   [0924]

The appeal to shared knowledge is considered a further device to signal speaker engagement (Hyland 2005: 182), and we do indeed find several such cases, especially when Bodley refers to previous correspondence and other forms of shared knowledge; both of these are quite frequent in the sample, but a couple of examples may suffice (36)–(37).

(36)   *and **as by former letters I have certified your L.** unles they may be obeied, I am very muche afraide, it will breede an incon-venience.*   [1047]

(37)   *May it please your good L. **by Sir John Norreis report, and by suche former letters as I have written to your L.** of which one was of the 18 and an other of the 23 of the last, I trust yow are advertised of the states resolution …*   [1026]

Finally, I draw attention to the way in which Bodley adjusts the immediacy of transmission in the reporting of knowledge. This can obviously be very relevant to the construction of Bodley's self-image as an efficient agent; therefore, there is frequent highlighting of the contrast between 'hot' news versus follow-up of previously known information. This is achieved through intensified time deictics (*right nowe, at this very instant*), as in (38)–(41).

(38)   *Postscript: **Heere is come intelligence right nowe**, that the companies in Guelders, which were sent, as in my last I advertised your L. to surprise the towne of Straelen, have taken the repulse. [0297];*

(39)   *and will not remove till ether the Enemie be passed the Rhine, **wherof ther is som speeche at this very instant**, or the castel of Coevorden sufficiently fortified,*   [1182]

(40)   *It may please your L. to be advertised, that **by Letters whiche I have right nowe receaved from Embden, and out of Friseland, it is signified unto me,** that …*   [0242]

(41)   ***Heere** is **at this instant** a troublesome advertisement come from the Campe, …*   [0715]

6.   Conclusion

The sample investigated has proved very helpful for the study of stance in Early Modern English,[10] and the consistent patterns it shows can further specify the features of diplomatic correspondence as outlined by previous research. Analyses like the present one highlight the advantages of an integrated perspective on stance, incorporating epistemic and evaluative items, which concur to the same communicative aims. Studies on later stages of English show that the degree of 'personalisation' in the expression of stance seems to increase over the Early Modern English period (Marcus & Evans 2019: 81), yet our earlier sample shows abundant stance

---

10. This is also due to the fact that the dating of our documents coincides with a period of innovation in the use of evidentials (Gisborne & Holmes 2007: 14, 19). Note the similar strategies in reporting sources in letters from about a century later than in our sample, in the field of scientific correspondence, in which epistemic stance is also important (Gotti 2006: 29–32).

markers. There is a wider inventory of epistemic stance markers, especially in verbs, although affective stance markers are still frequent, mostly within self-reference or source evaluation.

Much remains to be done, for instance looking at this correspondence from the other end, i.e., analysing the stance of Bodley's correspondents in accepting or questioning the credibility of Bodley's reports. A further step could be the specific analysis of responses to the transmission of knowledge, possibly adding a social network or communities of practice perspective to look at the system of evidentials and stance-taking elements, which situates them in terms of accommodation and of convergence. The exploitation of politeness systems to modulate stance, as illustrated in our first example, needs further investigation, given the typically asymmetrical relations that can be found in diplomatic correspondence.

Acknowledgments

Source

Diplomatic Correspondence of Thomas Bodley, 1585–1597, Centre for Editing Lives and Letters, based at University College London. Chief editor Robyn Adams. Last accessed April 2019. URL: http://www.livesandletters.ac.uk/bodley/bodley.html

References

Barton, Ellen. L. 1993. Evidentials, argumentation, and epistemological stance. *College English* 55(7). 745–769. https://doi.org/10.2307/378428

Chafe, Wallace. 1986. Evidentiality in English conversation and writing. In Wallace Chafe & Johanna Nichols (eds.), *Evidentiality: The linguistic coding of epistemology*, 261–272. Norwood, NJ: Ablex.

Clift, Rebecca. 2006. Indexing stance: Reported speech as an interactional evidential. *Journal of Sociolinguistics* 10(5). 569–595. https://doi.org/10.1111/j.1467-9841.2006.00296.x

Culpeper, Jonathan. (ed.). 2011. *Historical sociopragmatics*. Amsterdam: John Benjamins. https://doi.org/10.1075/bct.31

Del Lungo Camiciotti, Gabriella. 2006. "Conduct yourself towards all persons on every occasion with civility and in a wise and prudent manner; this will render you esteemed": Stance features in nineteenth-century business letters. In Marina Dossena & Susan Fitzmaurice (eds.), *Business and official correspondence: Historical investigations*, 153–174. Bern: Peter Lang.

Del Lungo Camiciotti, Gabriella. 2014. Letters and letter writing in Early Modern culture: An introduction. *Journal of Early Modern Studies* 3. 17–35.

Dossena, Marina & Susan Fitzmaurice (eds.). 2006. *Business and official correspondence: Historical investigations*. Bern: Peter Lang.

Du Bois, John W. 2007. The stance triangle. In Robert Englebretson (ed.), *Stancetaking in discourse. Subjectivity, evaluation, interaction*, 139–182. Amsterdam: John Benjamins. https://doi.org/10.1075/pbns.164.07du

Du Bois, John W. & Elise Kärkkäinen. 2012. Taking a stance on emotion: Affect, sequence, and intersubjectivity in dialogic interaction. *Text & Talk* 32(4). 433–451. https://doi.org/10.1515/text-2012-0021

Evans, Mel. 2017. Royal language and reported discourse in sixteenth-century correspondence. *Journal of Historical Pragmatics* 18(1). 30–57. https://doi.org/10.1075/jhp.18.1.02eva

Fetzer, Anita & Etsuko Oishi. 2014. Evidentiality in discourse. *Intercultural Pragmatics* 11(3). 321–332. https://doi.org/10.1515/ip-2014-0015

Fitzmaurice, Susan M. 2006. Diplomatic business: Information, power, and persuasion in Late Modern English diplomatic correspondence. In Marina Dossena & Susan Fitzmaurice (eds.). *Business and official correspondence: Historical investigations*, 77–106. Bern: Peter Lang.

Gisborne, Nikolas & Jasper Holmes. 2007. A history of English evidential verbs of appearance. *English Language and Linguistics* 11(1). 1–29. https://doi.org/10.1017/S1360674306002097

Gotti, Maurizio. 2006. Communal Correspondence in Early Modern English: The Philosophical Transactions Network. In Marina Dossena & Susan Fitzmaurice (eds.), *Business and official correspondence: Historical investigations*, 17–46. Bern: Peter Lang.

Gray, Bethany & Douglas Biber. 2014. Stance markers. In Karin Aijmer & Christoph Rühlemann (eds.), *Corpus pragmatics: A handbook*, 219–248. Cambridge: Cambridge University Press. https://doi.org/10.1017/cbo9781139057493.012

Gray, Bethany, Douglas Biber, & Turo Hiltunen. 2011. The expression of stance in early (1665–1712) publications of the Philosophical Transactions and other contemporary medical prose: Innovations in a pioneering discourse. In Irma Taavitsainen & Päivi Pahta (eds.), *Medical writing in Early Modern English*, 221–248. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511921193.013

Grund, Peter. 2017. Description, evaluation and stance: Exploring the forms and functions of speech descriptors in Early Modern English. *Nordic Journal of English Studies* 16(1). 41–73. https://doi.org/10.35360/njes.394

Hyland, Ken. 2005. Stance and engagement: A model of interaction in academic discourse. *Discourse Studies* 7(2). 173–192. https://doi.org/10.1177/1461445605050365

Jucker, Andreas H. & Irma Taavitsainen (eds.) 2010. *Historical Pragmatics*. Berlin: de Gruyter. https://doi.org/10.1515/9783110214284

Kiesling, Scott F. 2009. Style as stance: Stance as the explanation for patterns of sociolinguistics variation. In Alexandra Jaffe (ed.), *Stance: Sociolinguistic perspectives*, 171–193. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195331646.003.0008

Landert, Daniela. 2017. Meta-communicative expressions and situational variation of stance marking: *I say* and *I tell (you)* in Early Modern English dialogues. *Nordic Journal of English Studies* 16(1). 120–144. https://doi.org/10.35360/njes.396

López Couso, Maria José. 1996. On the history of *methinks*: From impersonal construction to fossilized expression. *Folia Linguistica Historica* 17(1–2). 153–169. https://doi.org/10.1515/flih.1996.17.1-2.153

Marcus, Imogen & Mel Evans. 2019. "Right trusty and well-beloved": The socio-pragmatics of gender, power and stance in sixteenth-century English letters. In Ursula Lutzky & Minna Nevala (eds.), *Reference and identity in public discourses*, 67–98. Amsterdam: John Benjamins. https://doi.org/10.1075/pbns.306.03mar

Marín-Arrese, Juana I. 2015. Epistemic legitimisation and inter/subjectivity in the discourse of parliamentary and public inquiries. *Critical Discourse Studies* 12(3). 261–278. https://doi.org/10.1080/17405904.2015.1013484

Mazzon, Gabriella. 2014. The pragmatics of Sir Thomas Bodley's diplomatic correspondence. *Journal of Early Modern Studies* 3. 117–131.

Mushin, Ilana. 2013. Making knowledge visible in discourse: Implications for the study of linguistic evidentiality. *Discourse Studies* 15(5). 627–645. https://doi.org/10.1177/1461445613501447

Nevala, Minna. 2011. Altering distance and defining authority: Person reference in Late Modern English. In Jonathan Culpeper (ed.), *Historical Sociopragmatics*, 61–82. https://doi.org/10.1075/bct.31.04nev

Nuyts, Jan. 2001. Subjectivity as an evidential dimension in epistemic modal expressions. *Journal of Pragmatics* 33. 383–400. https://doi.org/10.1016/S0378-2166(00)00009-6

Okulska, Urszula. 2006. Textual strategies in the diplomatic correspondence of the Middle and Early Modern English periods: The narrative report letter as a genre. In Marina Dossena & Susan Fitzmaurice (eds.), *Business and official correspondence: Historical investigations*, 47–76. Bern: Peter Lang.

Palander-Collin, Minna. 1999. Male and female styles in 17th correspondence: I THINK. *Language Variation and Change* 11. 123–141. https://doi.org/10.1017/S0954394599112018

Palander-Collin, Minna. 2010. Correspondence. In Andreas H. Jucker & Irma Taavitsainen (eds.), *Handbook of pragmatics: Historical pragmatics*, 651–677. Berlin: Mouton de Gruyter. https://doi.org/10.1515/9783110214284.7.651

Palander-Collin, Minna. 2011. Variation and change in patterns of self-reference in early English correspondence. In Jonathan Culpeper (ed.), *Historical sociopragmatics*, 83–108. https://doi.org/10.1075/bct.31.05pal

Palmer, Frank. 2003. Modality in English: Theoretical, descriptive and typological issues. In Roberta Facchinetti, Manfred Krug & Frank Palmer (eds.), *Modality in Contemporary English*, 1–17. Berlin, Mouton de Gruyter. https://doi.org/10.1515/9783110895339.1

Papafragou, Anna. 2000. *Modality: Issues in the semantics-pragmatics interface* (Current Research in the Semantic/Pragmatic Interface 6). Amsterdam: Elsevier Science. https://doi.org/10.1163/9780585474199

Willliams, Graham T. 2013. *Women's epistolary utterance: A study of the letters of Joan and Maria Thynne*, 1575–1611. Amsterdam: John Benjamins. https://doi.org/10.1075/pbns.233

Theatrical practices and grammatical standardization in eighteenth-century Britain

YOU WAS and YOU WERE

James Hyett and Carol Percy

University of Toronto

This chapter extends the discussion of second-person pronouns and of eighteenth-century language norms by examining three playwrights' use of YOU WAS in the period before its proscription by grammarians like Robert Lowth (1762). Our analysis corroborates historical sociolinguistic surveys by Tieken-Boon van Ostade (2002) and Laitinen (2009) that identify YOU WAS as informal/familiar before 1762. We also show that YOU WAS was salient and associated with social lowness in comedies well before it was stigmatized by prescriptivists. Finally, while acknowledging the distinctness of theatrical language, we have identified one useful general property of comedy and especially romantic comedy. The stylized opposition between comic characters (particularly lovers) can signal sociolinguistic salience, here of YOU WAS/WERE by 1747.

Keywords: eighteenth-century British English, second-person pronoun, prescriptivism, normative linguistics, historical sociolinguistics, qualitative methods, dramatic literature, comedy, style-shifting, gender

1. Introduction

The stories and lives presented on the English stage of the eighteenth century were primarily concerned with two interconnected pursuits: the accumulation of wealth and the acquisition of social status. The heightened emotions and interactions of dramatic characters means that they present a prime opportunity to gather insight into salient historical linguistic variants used in everyday life. Eighteenth-century English comedies, written almost entirely in prose and full of references to contemporary life, can reasonably be used as clues to actual contemporary usage, since playwrights used (and occasionally abused) common and recognizable linguistic patterns in order to differentiate characters. Charting the use of linguistic variants

across texts written by the same playwrights throughout the century can be used as a tool for investigating diachronic language variation and change. Our investigation illuminates the development of second-person pronouns and of eighteenth-century language norms. In this paper we present evidence for the social indexing of the linguistic variant YOU WAS in mid-eighteenth-century English through analysis of the use of YOU WAS in comedies primarily published in the 1750s. YOU WAS was explicitly proscribed in the early 1760s by influential grammarians such as Robert Lowth (1762). But on what grounds? Laitinen (2009: 209) has observed that by the 1750s, YOU WAS was already receding from some men's letters. And Tieken-Boon van Ostade observed that in the 1750s, Lowth used YOU WAS only with his wife (2002: 90). Our analysis of contemporary comedies casts more light on the 1750s and suggests the potential relevance of theatrical texts to historical linguistics.

In order to contextualize Lowth's proscription of YOU WAS in 1762, as well as earlier gendered distinction in use of it in personal correspondence, we have chosen to examine the distribution of YOU WAS and YOU WERE in comic plays published by playwrights who published extensively in the decades immediately before and after it. Especially in the decade leading up to 1762, we were curious to discover any trends in playwrights' use of YOU WAS that might anticipate or explain its declining use by men in their letters and its imminent proscription by grammarians. Our scholarly predecessors have identified the 1750s as transitional, a period bridging informal and stigmatized usage. Our research fills a gap: letters, fiction and grammar books have been studied, but not plays. We have focused on the language of comic plays, drawing on Hegele (2008): tragic and historical plays have their own stylized registers. We acknowledge that literary language does not correspond to that of 'real life'. Nevertheless, as Fitzmaurice has demonstrated for negation (2012: 308, 312–314), comedies might reflect colloquial language, and sometimes associated certain linguistic tokens with social stereotypes, which when marked with humour could create or intensify stigmatization. Moreover, for the relevant fields we have special and complementary expertise: Percy has published extensively on eighteenth-century normative linguistics, while Hyett holds an Honours BA in Drama and Linguistics and an MA in Classical Acting from the London Academy of Music & Dramatic Art (LAMDA). In combination, we can analyze both the sociolinguistic and theatrical significance of YOU WAS in its immediate dramatic context.

We acknowledge that for some playwrights before 1762 variation between YOU WAS and YOU WERE is sometimes difficult to interpret. Nevertheless, in our quantitative section we find trends in the use of YOU WAS in the 1750s that support our theory that YOU WAS did indeed index some social status or identity. In our qualitative analytical section, we test the hypothesis that several playwrights used YOU WAS very much like older THOU (i.e., in its various forms), considering

its associations with low status and with such attitudes as anger, condescension, and ease (see Walker 2007: 291). In drama and in life, the choice is 'complex' and 'unstable': as Adamson explains of Shakespeare's usage, when THOU could signal attitude as well as status, it might "turn an addressee into an inferior or an intimate, thus signaling either contempt or affection" (2001: 228). Nevertheless, even when the 'meaning' of a variant like YOU WAS is somewhat obscure, we can nevertheless demonstrate that the stylized opposition of YOU WAS and YOU WERE in characters' and especially lovers' dialogue both confirms and dates the salience of the variants, in comedy and likely thus also in 'real life'.

2. Context

Our study focuses on the eighteenth century, like many contributions to normative linguistics. Tieken-Boon van Ostade explained "the rise of YOU WAS … as a kind of bridge phenomenon" in Standard English: "the construction came into being to facilitate the transition from YOU WERE as a plural to a singular construction" as well as "the extension of YOU as a plural pronoun to include the singular as well" (2002: 97, 100). Laitinen in turn used *A Representative Corpus of Historical English Registers* (ARCHER) and the *Corpus of Early English Correspondence Extension* (CEECE) to confirm that YOU WAS began to spread in the late seventeenth century, and likely "from below the level of consciousness", since usage of the form peaked in correspondence before peaking in published material (2009: 206, 215). The variant spread quickly in private correspondence: at its peak, from 1720–1739, YOU WAS accounted for 68% of the instances in men's letters (Laitinen 2009: 206, 209; see Figure 1).

By the mid-eighteenth century, YOU WAS was explicitly stigmatized in Standard English. Grammarians began to mention it in the early 1760s: the editors of the *Dictionary of English Normative Grammar 1700–1800* (*DENG*) identify Joseph Priestley (1761) and Robert Lowth (1762) as the first of many (Sundby et al. 1991: 156). Priestley (1761: 20n) expressed his opinions in a question based on analogy: "Many writers of no small reputation say *you was*, when speaking of a single person: but as the word *you* is confessedly *plural*, ought not the *verb*, agreeable to the analogy of all languages, to be plural too?" Lowth explicitly and strongly condemned YOU WAS as an "enormous Solecism" into which "[a]uthors of the first Rank ha[d] inadvertently fallen" (1762: 48n). The frequency of YOU WAS declined subsequently and steeply in fiction in the 1760s, after a rise through the 1750s (Tieken-Boon van Ostade 2002: 95; see Figure 2). Tieken-Boon van Ostade acknowledges that the strength of Lowth's condemnation and the "enormous popularity" of his grammar may have contributed to the subsequent "relegation" of YOU WAS "to non-standard registers" of English (2011: 119, 226).
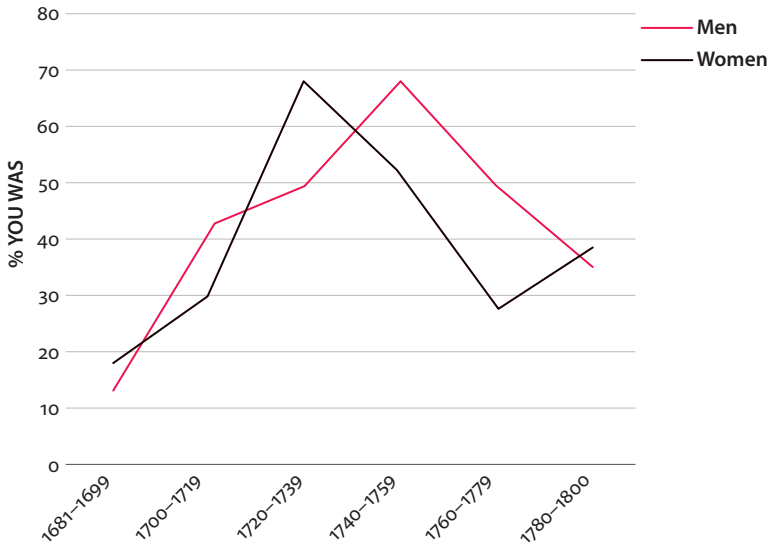
Figure 1. Proportion of YOU WAS in men's and women's personal letters (adapted from Laitinen 2009)
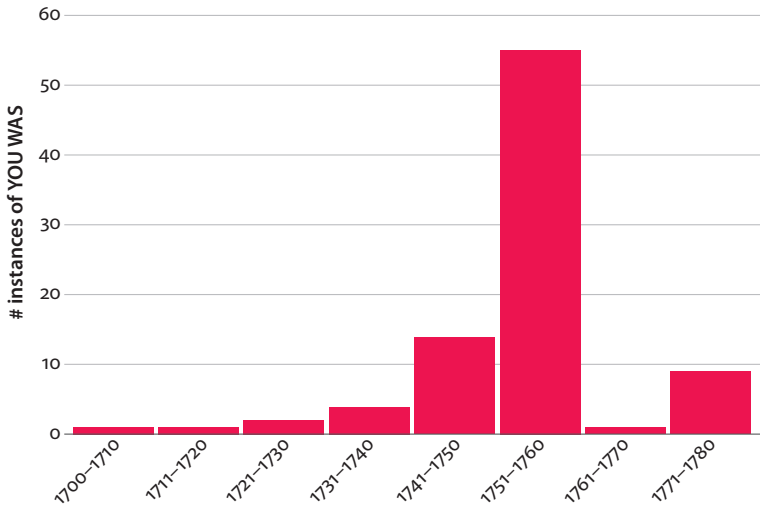
Figure 2. YOU WAS in eighteenth-century novels (adapted from Tieken-Boon van Ostade 2002)

Is there evidence of negative attitudes in the 1750s? As Figure 1 shows, Laitinen found that in men's letters YOU WAS peaked in 1720–1739, declining thereafter. Laitinen also surveyed the letters of certain better-educated 'language profession-als' in the CEECE. He reports that both before and after 1762, YOU WAS occurred less often than YOU WERE in letters written by select educated men. Even before 1762, for instance, YOU WAS appears only 29% of the time; indeed, the lexicographer Samuel Johnson used only YOU WERE from 1738–1784 in the letters selected for the CEECE. In contrast, some other authors including the poet Thomas Gray and the actor-manager David Garrick used only YOU WAS in the 1730s but by the 1750s were using YOU WERE, Gray exclusively. Johnson's and Lowth's publisher used both YOU WAS and YOU WERE in the 1750s (Laitinen 2009: 211–215; Tieken-Boon van Ostade 2002: 90–91). Garrick's variable usage in the 1750s provides some additional conti-nuity between Laitinen's study and ours. Lowth himself used YOU WAS only with his wife, using YOU WERE with all his other correspondents, and sometimes with her (Tieken-Boon van Ostade 2002: 90). Did Lowth's use of YOU WAS with his wife signal the form's potential intimacy? And as YOU WAS became less common in educated men's letters, was it marked as low and/or feminine in contemporary comedies?

3.   Methods

We have taken our texts from *Literature Online* (LION). This corpus is relatively comprehensive and easy to filter by genre (play), date of publication, authors' names, and easily searchable phrases. We initially selected all of the plays containing YOU WAS/WAS YOU and/or YOU WERE/WERE YOU that were first published in the decade before (1752–1762) and the decade after (1763–1773) the publication of Lowth's (1762) grammar. We included both broad genres of comedy and tragedy but given that there were only two instances of YOU WAS in the 28 tragedies published in that time frame, our analysis mostly focuses on comedies. We included subjunc-tive, conditional, and indeed all possible examples, like Tieken-Boon van Ostade (2002: 93–94) and Laitinen (2009: 204). An ongoing challenge involves tracking down other relevant examples (e.g., YOU NEVER WAS/WERE, YOU YOURSELF WAS/WERE). LION has many disadvantages: it does not include all plays, let alone all edi-tions, and does not support complex searches. Our further manual work included checking LION's text against the printed edition (ECCO) and adding the dates of first performance.

In order to be able to identify both influences on and the impact of grammari-ans' proscriptions, we identified all the dramatists who used both variants and who wrote extensively before and after the early 1760s. Because Laitinen's work identi-fied sex as a salient factor in variation, we also considered all the female playwrights

active in our timeframe. These decisions helped to reduce the number of texts for close reading, while building the foundation for several investigations – for the present paper, focusing on three male playwrights up to 1762; and for proposed future work, incorporating the data from the period after 1762, and incorporating women's as well as men's usage. Having determined for the present to focus on three male playwrights that wrote in the decade before 1762, we then extended these 'focused' investigations to include all the plays published during those three authors' lifetimes, beyond our initial timeframe of 1752–1773, and also included any plays that did not contain YOU WAS or YOU WERE.

4.   1752–1773: A broadly quantitative overview

YOU WAS appears almost exclusively in comic genres (Figure 3) in the decades before and after 1762, a fact which is unsurprising given Laitinen's association of the variant with more informal text-types. There are two exceptions: firstly, John
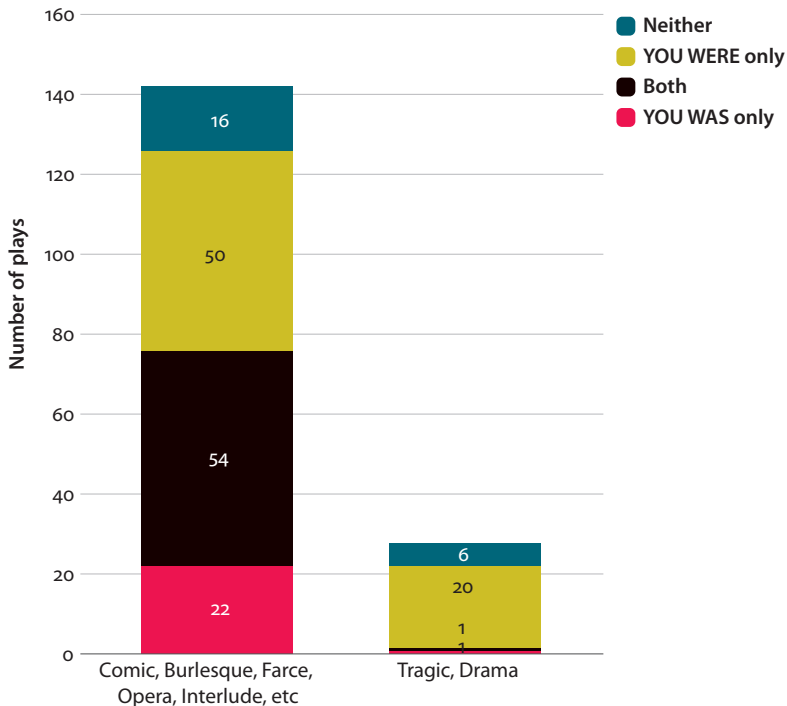
Figure 3. Distribution of YOU WAS and YOU WERE by genre

Hawkesworth's 1759 adaptation of Southerne's 1696 play *Oroonoko*, based on Behn's earlier novel. This edition of the play contains one instance of YOU WAS (in this case in the interrogative, *was you*), and six instances of YOU WERE. Hawkesworth's adaptation sought to "render *Oroonoko* a regular Tragedy" (page v), by removing the comic scenes and adding additional scenes to explain some plot holes the editor identified in the original. While all six instances of YOU WERE present in the 1759 version also appear in the 1696 original, the scene in which YOU WAS occurs (IV.i) is original writing by Hawkesworth. YOU WAS appears elsewhere in Hawkesworth's work, both in his original comedy *Edgar and Emmeline* (1761), and his adaptation of Dryden's 1690 *Amphitryon* (1756). The latter play is a similar case to that of *Oroonoko*, with the 11 instances of YOU WERE produced by the original author, and Hawkesworth adding the one instance of YOU WAS, spoken by the god Mercury. It would appear that, in his original dramatic writing at least, John Hawkesworth favoured the variant YOU WAS over YOU WERE. He only writes one instance of the latter in the three works mentioned above (in *Edgar and Emmeline*, where YOU WAS occurs four times), and so his anomalous use of that form in his adaptation of a tragedy cannot be easily incorporated into an analysis of how these forms are used in different genres and by different characters.

The second exception is *Almida: A Tragedy* (1771), the only play published (Breen 2004) by Dorothea (Mallet) Celesia (1738–1790): this play contains a single instance of YOU WAS, none of YOU WERE, and forty-nine of THOU and its forms. Post-1762 texts lie outside the scope of our detailed analysis, but it is potentially significant that the single YOU WAS coincides with the heroine's wild emotions just before her inevitable death: emotive and possibly transformative (Hodson 2014: 179). THOU sometimes occurred at the climax of contemporary tragedies (Nonomiya 2014: 226). Among more prolific playwrights, YOU WAS seems characteristic of comedy. For instance, among David Garrick's productions, YOU WAS occurs in comedies but not in tragedies or adaptations of Shakespeare (Figure 4), while in 1763 George Colman used only YOU WAS in the farce *The Deuce is in Him* and only YOU WERE in the tragedy *Philaster*.

What was the status of YOU WAS in comedies published in the decade before 1762? It would seem many playwrights who were first published in the 1750s used it. In his 1757 anthology of five plays Phanuel Bacon (1700–1783) uses YOU WAS twenty-six times, and YOU WERE once. John Hawkesworth (bap.1720, d. 1773) uses both variants in comedies published in 1756 and 1761, as mentioned above. Joseph Reed (1723–1787) uses one WAS in 1758, and one WERE in 1761. William Kenrick (1729/30–1779) used both WERE (twice) and WAS (thrice) in 1752, only WAS (five times) in 1760, but in 1767 used mostly WERE (eight times of nine). As he did later in the 1760s, George Colman (bap. 1732, d. 1794) uses primarily YOU WAS, four times in 1760 and once in 1762; in 1761 the only example of YOU WERE (alongside
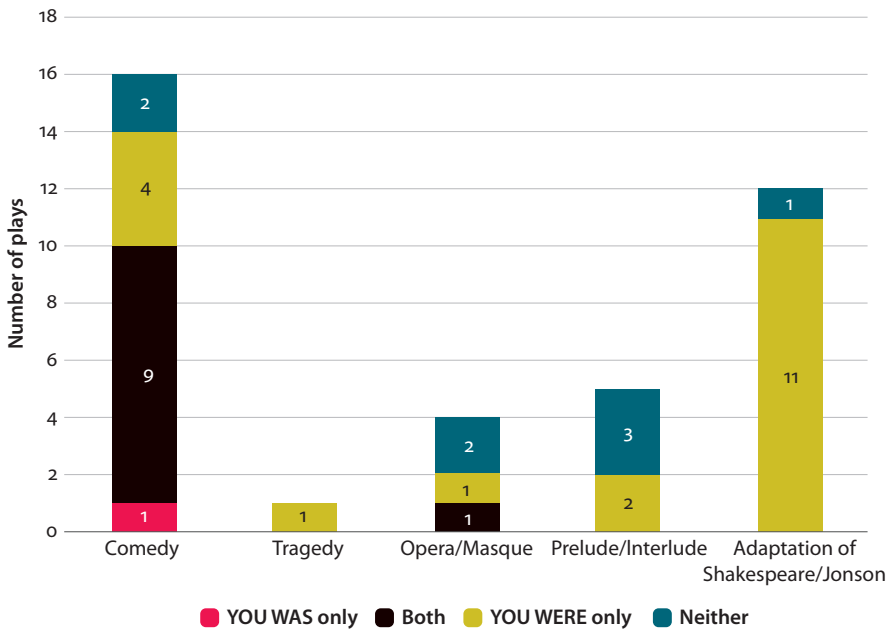
Figure 4. Distribution of YOU WAS and YOU WERE in works by David Garrick

ten with WAS) is counter-factual: "I wish *you were* married to one another with all my Heart." In summary, both variants were certainly used in the years leading up to Lowth's pronouncement.

Three prolific male playwrights wrote for a relatively long period before 1762 – and after it. For the present paper, we have thus narrowed our focus and broadened our search to include every play that was published before 1762 (inclusive) by David Garrick (1717–1779), Samuel Foote (1720–1777), and Arthur Murphy (1727–1805) and that appears in LION: 19, 7, and 8 plays, respectively, for a total of 24 plays. For purposes of comparison it is convenient that the playwrights were relatively close in age, although we note that Murphy was born in Ireland, unlike Garrick and Foote. We can make some quantitative generalizations surveying the distribution of each variant in their plays to 1762. Garrick used both variants in two farces in the 1740s, but in the early 1750s he, like Foote, used only YOU WERE. After 1756, when Murphy's first publication appears, all three use both variants. In 1762, Foote uses only WAS with YOU in *The Orators*. YOU WAS appears to have become more salient in the later 1750s, after a period of no use in the first half of that decade. Aggregating the total numbers of instances of each variant in Garrick, Foote, and Murphy's plays in this period before 1762, and looking at them in five-year chunks, this trend is clear (Figure 5).
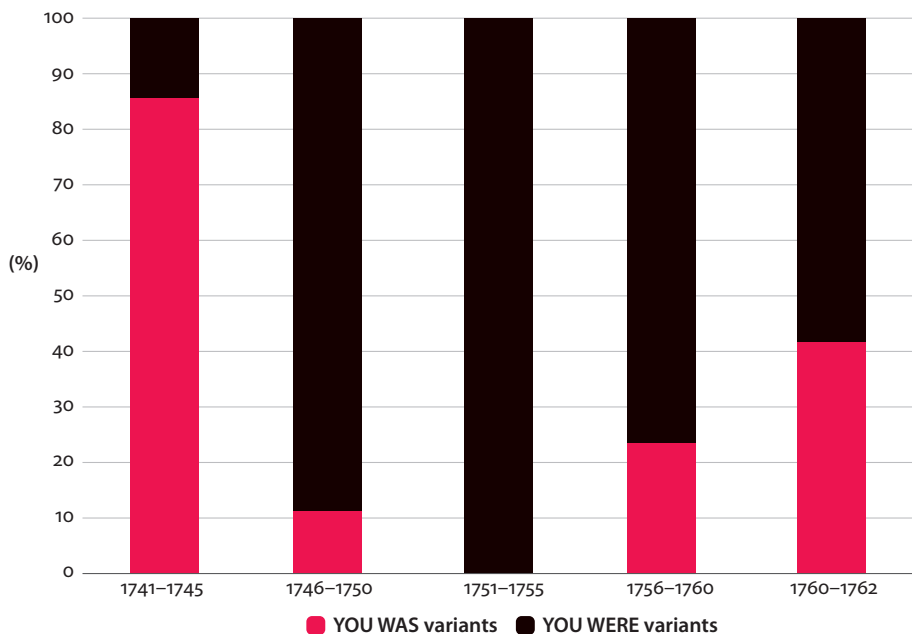
Figure 5. Proportion of YOU WAS/WERE variants in Garrick, Foote, and Murphy 1741–1762

Obviously, these three playwrights contribute to this trend in different ways, given that Murphy only published his first play in 1756, while Garrick is represented through the entire span of the data, but this trend is in line with other studies. The drop-off of the use of YOU WAS before the start of the 1750s runs parallel to Laitinen's data on men's use of the variant in their personal letters, which began to decline in this two-decade period (2009: 206). However, the apparent re-emergence of YOU WAS in the latter half of the 1750s goes against the trend established in previous studies. It could be that the period of declining use (approximately 1746–1755 in Figure 5) is the end of the subconscious use of YOU WAS, which reemerged in the late 1750s, being used consciously to index some kind of identity or status. This trend further corroborates Laitinen's assertion that YOU WAS emerged from below the level of consciousness, reaching a peak threshold of usage in everyday speech and private correspondence before rising above the level of consciousness, becoming explicitly marked, and further propagated in published texts. Examining the instances of YOU WAS in these men's plays in their dramatic context will provide a clearer picture of what this variant could have indexed.

5.   Before 1762: Qualitative analysis of YOU WAS in context

We both read each pre-1762 play, and analyzed every passage containing YOU WAS, from literary and linguistic perspectives. Overall, despite nuances in playwrights' usage, we test the hypothesis that in general and in combination these plays represented YOU WAS as equivalent to earlier THOU – signaling combinations of low rank, condescension, anger, and ease (see Walker 2007: 291). From a theatrical point of view, perhaps most interesting is a possible comic trope involving the use of YOU WAS in situations of intimacy and emotional vulnerability by elite female leads. Most significant (and least ambiguous) from a historical-linguistic point of view are several interpersonal interactions that show YOU WAS and YOU WERE in explicit contrast, as early as 1747.

5.1   Social status and YOU WAS

YOU WAS likely arose "from below," according to Laitinen, appearing earlier in seventeenth-century correspondence than in printed texts (2009: 214–215). How was it represented theatrically by dramatists who were active in the 1750s and 1760s? David Garrick (b. 1717) first uses YOU WAS in his second play, *The Lying Valet* (1741). It is used only by the servants of the main plot's love interests: once by Melissa's servant Kitty, and five times by her opponent Sharp, the titular *Lying Valet*. In this play, YOU WERE occurs only once – and is used also by Sharp, perhaps style-shifting in his ongoing attempt to persuade Melissa that his master is financially reliable.

(1)   Garrick, *The Lying Valet* (1741), I.ii
      Kitty.   'Tis great Pity such great Preparations as Mr. Sharp has made should be thrown away.
      Sharp.   So it is, as you say, Mrs. Kitty. But I can immediately run back and unbespeak what I have order'd; 'tis soon done.
      Melissa. But then what Excuse can I send to your Master? He'll be very uneasy at my not coming.
      Sharp.   Oh terribly so! – but I have it – I'll tell him you are very much out of Order – that **you were** suddenly taken with the Vapours or Qualms; or what you please, Madam.   (p. 13)

Of course, since Sharp is the title character, we must also consider that his frequent use of YOU WAS reflects the character's high number of lines rather than his low social status – or the play's low register as a farce; Garrick "built his career" on "low entertainments" (Anderson 2014: 356). Further, consider that Garrick himself used only YOU WAS in his letters in the 1730s, according to Laitinen (2009: 213), and indeed that YOU WAS had seen its highest usage in the period 1720–1739, at 63%

of the possible cases (2009: 206). Later, in Garrick's *Miss in Her Teens* (1747), the servant's wife Tag uses YOU WERE (p. 12). Now, this ambiguous pattern of usage could simply reflect the trend noted above in Figure 5. In that analysis, we would locate *The Lying Valet* in the first column, at the height of unmarked use of YOU WAS, not having been raised above the level of consciousness in Garrick's internal grammar, meaning he is not making conscious use of this linguistic variant in this early play, as he is in later texts.

The association of YOU WAS with socially low speakers is also suggested by later plays and other playwrights. In Arthur Murphy's 1756 play *The Spouter*, the stage-struck title character uses YOU WAS as he agrees that his father *Muckworm* "never was a fine Gentleman":

(2) *Murphy, The Spouter (1756), II.i*
    Muckworm. Twelve times Twelve is a Hundred and Forty-four. So, I've made
              a right Use of your Elegy written in a Country Church-Yard.
              Death and Thunder, you Booby, look at your Legs. You be a fine
              Gentleman! I never was a fine Gentleman.
    Slender.   Nobody says **you was**, Sir.   (p. 31)

Yet in *The Spouter*, only WAS occurs with YOU, as also in Murphy's *Apprentice* (1756), again in the speech of another stage-struck 'spouter': "how **was you** receiv'd?" (p. 23). Spouters are stereotyped in the prologue of *The Apprentice* (n.p.) as a "Woollen Draper," "young Tobacconist," or "Haberdasher …!" – lower in social rank, but higher in cultural aspiration. Two servants' contrasting use confirms the 'lower' status of YOU WAS in Murphy's *The Way to Keep Him* (1760): Muslin uses WAS YOU (p. 11) while the more refined Mignionet uses YOU WERE (p. 41). Garrick's *The Guardian* (1759) has similarly suggestive stylistic contrast: the servant Lucy uses the only example of YOU WAS (p. 12) as she describes the pretentiousness of Mr. Clackit, the play's only user of YOU WERE (p. 30). Moreover, when Samuel Foote mocks Thomas Sheridan's elocutionary movement in *The Orators* (1762), two of the orator's pupils use WAS, one combining it with pretentious legal jargon and bad grammar (e.g., "this here", condemned in 1757 by Ralph Griffiths, writing anonymously in the *Monthly Review* volume 17, p. 233):

(3) *Foote, The Orators (1762), Act II*
    Counsellor. May it please your worship – hem – I am counsel in this cause
              for the ghost – hem – and before I can permit her to plead, I have
              an objection to make, that is – hem – … all criminals should be
              try'd per pares, by their equals – hem – that is – hem – by a jury
              of equal rank with themselves. Now, if this be the case, as the case
              it is; I – hem – I should be glad to know, how my client can be
              try'd in this here manner.

…

Counsellor.   Pray, Mr. Paragraph, where **was you** born?

Paragraph.   Sir, I am a native of Ireland, and born and bred in the city of Dublin.

(p. 47; see also p. 57)

Nobody in *The Orators* uses YOU WERE, but we can infer that YOU WAS was marked for Foote, since in earlier comedies (1752–1756) he used only YOU WERE. But is there less circumstantial evidence for status and/or attitudes associated with YOU WAS?

5.2   YOU WAS and YOU WERE in explicit contrast

As early as 1747, explicit contrasts in single passages between YOU WAS and YOU WERE confirm their markedness as variants. After *The Lying Valet*, the next Garrick character to use YOU WAS is again a title character who style shifts. Most characters in *Miss in Her Teens* (1747) use YOU WERE, including Miss Biddy herself, in conversation with the foppish suitor Fribble; Miss Biddy's aunt, who had her engaged to the elderly Sir Simon; Sir Simon himself; and his son Captain Loveit, who has returned from real battle to engage in another one over the *Miss in Her Teens*. Example (4) shows explicit variation between the boisterous Biddy's YOU WAS and her preferred suitor's YOU WERE:

(4)   *Garrick, Miss in Her Teens (1747), II.i*

Biddy.   But pray, Mr. Fox, how did you get out of your Hole? I thought **you was** lock'd in?

Capt.   I shot the Bolt back when I heard a Noise; and thinking **you were** in danger, I broke my Confinement without any other Consideration than your Safety. [*Kisses her Hand*].   (p. 32)

The lovers' contrasting verbs show that the variants were salient, though the two uses of course show different degrees of factuality: Captain Loveit was locked in, but Biddy was not in danger, and that verb was thus subjunctive. Such stylized opposition is the engine of romantic comedy: in his survey of *English Stage Comedy*, Alexander Leggatt has described the action that blocks the ending as the deferral of marriage "by the lovers themselves" (1998: 84). In this scene, however, the conflict is resolving, and "comedy stands aside for [low] farce" (Bevis 1980: 42): Biddy has already learned that her aunt won't oppose her own choice of husband, and the Captain has just dispatched the other suitors. In the play's context, Biddy's YOU WAS differs from the five occurrences of YOU WERE, including her own (p. 22). Does YOU WAS index anything other than the difference between reality and imagination or Biddy and her suitor? Perhaps Biddy's feminine sex: according to Laitinen, by 1740–1759 women were slightly more likely than men to use YOU WAS in informal

letters. For our survey, we will hypothesize that here it marks female ease, a combination of informality and intimacy.

The explicit variants YOU WAS and YOU WERE index difference between eventual lovers in a more protracted passage of another play. In Arthur Murphy's *All in the Wrong* (1761), style-shifting indexes Beverley's suspiciousness of his lover Belinda, and her attempts to connect with him – reality and imagination. After she says "As if you were," he replies "that you was." And after she echoes "you said you was", he says "where you were":

(5) *Murphy, All in the Wrong (1761), Act I*
    Belinda.   … as if **you were** …   (p. 9)
    …
    Beverley.   … that **you was** …   (p. 10)
    …
    Beverley.   … **you was** so last night …   (p. 10)
    …
    Belinda.   … you said **you was** …   (p. 10)
    …
    Beverley.   … where **you were** concerned … the spirits **you were** in … (p.11)
    …
    Belinda.   … if **you were** to … if **you were** to …   (p. 16)

Finally, Belinda once again echoes Beverley, following two of his "you were"s with two of hers, "if you were." Although the variants YOU WAS and YOU WERE are not particularly common in comedies writ large, this sequence of linguistic divergence and accommodation subtly underscores the misunderstandings between the lovers (Hodson 2014: 178).

Indeed, the second-person singular has a more distinctive and dramatic variant: in Murphy's *The Apprentice* (1756) we see the comic conflict between lovers epitomized by YOU versus THOU. Dick's use of the archaism THOU indexes not only his devotion to Charlotte Gargle, the apothecary's daughter, but also his distraction by drama from their union and from his real-life duties as apprentice to her apothecary father:

(6) *Murphy, The Apprentice (1756), II.iii*
    Charlotte.   Who's there? my *Romeo*?
    Dick.   The same my Love, if it not **thee** Displease. –
    Charlotte.   Hush! not so loud, **you**'ll waken my Father. –
    Dick.   Alas! there's more peril in **thy** Eye.
    Charlotte.   Nay, but pri**thee** now – I tell **you you**'ll spoil all – what made **you**
               stay so long?   (pp. 29–30)

In our argument, this passage's contrast between THOU and YOU indexes the relationship between the language of comedy and of real life – Dick's artificial use of THOU, which he has adopted from the theatre, is contrasted with Charlotte's YOU, meant to indicate the 'real life' concern of Dick putting them in danger by showing off. Of course, Murphy has written both characters into a play that was intended for performance on stage, so neither is actually 'real life', but Charlotte is written in a way that would be recognizable to Murphy's original audience as representative of contemporary English usage, contrasted with Dick's archaic and dramatic speech. We now must consider what we can learn about eighteenth-century usage from eighteenth-century comedy. Although YOU WAS clearly contrasted with YOU WERE in the 1740s, can we infer what this variation indexed – in comedy, at least?

5.3 YOU WERE in formal contexts

In our focused corpus of comedies, some contexts suggest that YOU WERE was appropriate for formal use, if not an unambiguous index of it. In Garrick's *Miss in Her Teens* (1747), the Captain's YOU WERE accompanies a courtly gesture of hand-kissing that contrasts with Biddy's boisterousness. YOU WERE appears in Arthur Murphy's letter to David Garrick that was prefaced to Murphy's *The Upholsterer* (1758: i). Public print is just one of the relevant contexts: we know that Murphy had an increasingly difficult relationship with Garrick in his capacity as manager of the Drury Lane playhouse, and they had reconciled only at the end of 1757. Although in most real-life contexts YOU WERE was probably the unmarked variant, it might have indexed excessive formality for some comic characters. In Garrick's *The Guardian* (1759), the only user of YOU WERE (p. 30) is described as vain and aspirational, albeit by the female servant who in Example (7) is the only user of YOU WAS:

(7) *Garrick, The Guardian (1759), I.i*
    Lucy.   Indeed, Miss Harriet, you are very particular; **you was** tired of the Boarding School, and yet seem to have no Inclination to be married … – That smirking old Gentleman is Uncle to Mr. Clackit; and, my Life for it, he has made some Proposals to your Guardian.
    Miss Harriet.   Prithee don't plague me about Mr. Clackit.
    Lucy.   But why not, Miss? Tho' he is a little fantastical, <u>loves to hear himself talk</u>, and is somewhat self-sufficient; you must consider he is young, has been abroad, and keeps good Company.
                                             (p. 12)

Garrick had added the character of Lucy to his adaptation of Fagan's *La Pupille* as a (sometimes) reliable guide for the audience (Garrick 1980: I, 402).

Elite men's use of YOU WERE was likely unmarked, but playwrights could still exploit the discrepancy between a formal form and a ludicrous situation when well-bred young men used YOU WERE in misguided declarations of love. For instance, Garrick's *Male-Coquette* (1757) George Daffodil uses WERE YOU when mistakenly wooing a man disguised as a woman (p. 43). In *The Old Maid* (1761), so does the "young – handsome – rich" Mr Clerimont when obliviously declaring his passion for a married woman (p. 40). Finally, it could convey distance of social equals: *The Old Maid* herself used WERE when rejecting her faithful suitor, Captain Cape (p. 30).

5.4   YOU WAS from condescension to emotion

Because theatrical characters tend to appear in identifiable social contexts with clear perspectives, we can identify some attitudes appropriate for expression with YOU WAS. Sharing functions with earlier THOU (Walker 2007: 229; Nonomiya 2014: 219), YOU WAS was used not only by servants, but by some higher-class characters expressing anger – to servants and to each other. YOU WAS is evidently appropriate for a superior to use when berating a servant. Garrick's *Male-Coquette* (1757) uses WAS as he interrogates his servant Ruffle (p. 13) – after declaring that "**Thou** art a most incomprehensible Blockhead" (p. 12), THOU here combining the negative emotions and social condescension typical of its remaining uses (Walker 2007: 229, 275). In Murphy's *The Way to Keep Him* (1760), Lovemore's wife no longer bothers to dress for his pleasure, and he is wooing the Widow Bellmour under the assumed name of Lord Etheridge. In this passage, Example (8), his schemes are unravelling, and he is berating the family servant, Muslin:

(8)   *Murphy, The Way to Keep Him (1760), Act III*
   Muslin.   Madam, Madam, – here's your Letter, – I wou'd not for all the World that my Master –
   Lovemore.   What, is she mad too? What's the matter, Woman?
   Muslin.   Nothing, Sir, – nothing, – I wanted a word with my Lady, – that's all, Sir.
   Lovemore.   You wou'd not for the World that your Master, – What **was you** going to say? – What Paper's that?   (p. 68)

Just as it can be used to convey anger from a high-status speaker to a low, YOU WAS can likewise convey condescension that is in some sense friendly, again like THOU (cf. Walker 2007: 226). In Samuel Foote's *The Minor* (1760), Sir William uses YOU WAS when interviewing the shifty character Shift, whom he is enlisting in a scheme to assay his son, *The Minor*, who is already bankrupt but who Sir William hopes will prove his worth. In Example (9), the knight's casual use of WAS YOU while

attempting to "have a closer connection" with this potential employee is conde-scending, but perhaps friendly:

(9)   *Foote, The Minor (1760), Act I*
      Sir William. No apology, I beg. But as we are to have a closer connection, it may
                         not be amiss, by way of introduction, to understand one another
                         a little. Pray sir, where **was you** born?
      Shift.   At my father's.
      Sir William. Hum. – And what was he?
      Shift.   A gentleman.
      Sir William. What **was you** bred?
      Shift.   A gentleman.
      Sir William. How do you live?
      Shift.   Like a gentleman.
      Sir William. Cou'd nothing induce you to unbosom yourself.
      Shift.   Loke'e Sir William, there is a kind of something in your counte-
                         nance, a certain openness and generosity, a je ne sçai quoi in your
                         manner, that I will unlock: You shall see me all.   (pp. 19–20)

In the context of Foote's writing, YOU WAS seems marked: from 1752 through 1756 Foote uses only YOU WERE in his comedies. In this passage, it is a testament to Sir William's flexibility of character that he can style-shift so easily and cross class lines to achieve his goals. In Murphy's *The Upholsterer* (1758), the variants contrast two reunited friends in a friendly exchange, the roguish Rovewell and the more educated Bellmour:

(10) *Murphy, The Upholsterer (1758), I.i*
      Bellmour.   Poh, that is not mentioned, – you know my father took me against
                         my will from the university, and consigned me over to the aca-
                         demic discipline of a man of war; so that to prevent a dejection
                         of spirits, I was oblig'd to run into the opposite extreme, – as **you
                         yourself were** wont to do.
      Rovewell.   Why, yes, I had my moments of reflection, and was glad to dis-
                         sipate them – you know I always told you there was something
                         extraordinary in my story; and so there is still … I must see the
                         town a little this evening, and have my frolic first. But to the point
                         Bellmour, **you was** going from the tavern you say.   (p. 12)

The two friends' contrasting variants signal the social distance between them and perhaps also the honesty of their individual self-presentation. The juxtaposition of the two variants confirms their salience, and their potential sociolinguistic difference.

Between well-born men, YOU WAS is used in exchanges that ambiguously bridge distance and intimacy. In Murphy's *All in the Wrong* (1761), Bellmont's explicit request for "No ceremony" might suggest that his father's response is informal:

(11)   <u>*Murphy, All in the Wrong (1761), V.i*</u>
      Bellmont.   No ceremony, Sir.
      Sir William Bellmont. I suppose **you was** coming hither, George, to wait upon
                             your mistress, when I met you but now –   (p. 82)

But this is not a warm familial interaction: Sir William is blocking the plot by forcing his son to marry the wrong girl. Elsewhere in this play, suspicious husband Sir John Restless uses YOU WAS as he (privately) curses a footman (p. 29), while using YOU WERE both in an angry exchange with the well-born Beverley (p. 52), and when berating himself (p. 5). YOU WERE is clearly appropriate for expressing exasperation with fellow men of high status, so Sir William's "unceremonious" use of YOU WAS indicates his severity towards his son. As with THOU, what might be a signal of informality is more likely a performance of superiority from Bellmont's father.

Textual clues can confirm the negative attitudes sometimes conveyed by YOU WAS among genteel characters. In Garrick's *The Male-Coquette* (1757), YOU WAS conveys not only apparent condescension ("Familiarity") but is something said while rudely "picking [your] teeth", "humming a tune", and "turning away", as in Example (12) below. In this passage from early in the play, Sophia is disguised as an Italian nobleman – ultimately to discover the truth about Daffodil's rumoured inconstancy, and here tests the effectiveness of her disguise by seeing whether her patient lover Tukely can discover her. This is a multiply charged scene: Sophia has overheard Tukely complain about her treatment of him, and several times mentions his jealousy. Men's jealousy of women (and Sophia's of Daffodil) expresses a kind of control over them, in this period and this play. Like Garrick's Miss Biddy, Sophia is an active, madcap heroine. Here, she capitalizes on the variable indexing of YOU WAS, conveying a certain intimacy but also irreverent offence – like THOU in apparent social condescension, and likely also emotion.

(12)   <u>*Garrick, The Male-Coquette (1757), Act I*</u>
      Sophia.   I say, Sir, wear your Cloak as long as you please, the Hoof will peep
                  out, take my Word for it.
      Tukely.   Upon my Word, Sir, you are pleas'd to honour me with a <u>Familiarity</u>
                  which I neither expected, or indeed desired, upon so slight an
                  Acquaintance.
      Sophia.   I dare swear you did not. [*Turns off, and hums a Tune*].
      Tukely.   I don't understand this!
      Arabella.   This is beyond Expectation – [*Aside*].
      Sophia.   I presume, Sir, **you never was** out of England – [*Picking her Teeth*].
                                                       (p. 8)

In this play, characters use WERE with YOU only for counterfactual *were you*, both times somewhat heightened – Daffodil's passion for a man in disguise, and Arabella's final humiliated accusation of Daffodil. Arabella then uses YOU WAS to another woman betrayed by Daffodil, in shared humiliation and anger: she reassures Lady Pewitt that "Your Ladyship is better than **you was**" (p. 47). While we remember that YOU WAS might simply have marked female more than male speech by the end of the 1750s (Laitinen 2009: 209), we also interpret this particular exchange as showing Arabella's ease with a sympathetic equal as the play's complications unravel.

A few years later, Murphy uses YOU WAS in a squabble between relatively genteel sisters-in-law: the *Old Maid* (1761) Miss Harlow lives with her brother and his young wife Mrs. Harlow, who resents her presence:

(13)   *Murphy, The Old Maid (1761), Act I*

Mrs. Harlow.	Well! I vow passion becomes you inordinately – It blends a few roses with the lillies of your cheek, and –
Miss Harlow.	And tho' you are married to my brother, ma'am, I would have you to know, ma'am … my fortune is in my own hands, ma'am, and ma'am –
Mrs. Harlow.	Well! do you know now when your blood circulates a little, that I think you look mighty well? – But **you was** in the wrong not to marry at my age – sweet three and twenty! – you can't conceive what a deal of good it would have done your temper and your spirits, if you had married early – (p. 4)

Here YOU WAS conveys Mrs. Harlow's contempt for her sister-in-law's appearance, "temper" and "spirits." But YOU WAS does not unambiguously index contempt, as seen in our examples above involving reunited friends (Example 10) and a desire for "a closer connection" (Example 9).

5.5   Unaffected YOU WAS used by lead female characters

We have established some associations between YOU WAS and the low status or the condescending attitude of its users. We have also noted its use by several lead female characters with their lovers. Sometimes it conveys their negative emotions: romantic comedies after all are sustained by lovers' misunderstandings. As we saw in Example (12), in Garrick's *Male-Coquette* (1757), Sophia disguises herself as an Italian nobleman and uses WAS with her jealous suitor Tukely, in seeming condescension and genuine provocation. Foote's Arabella uses YOU WAS after she misinterprets and is affronted by the behaviour of her suitor in *The Author* (p. 29[b]), shortly after having used YOU WERE with her pretentious sister-in-law (p. 29[a]).

Angry young women's use of YOU WAS is unsurprising in the context of comedies of the 1750s, since in contemporary real-life letters women were more likely than men to use YOU WAS (Laitinen 2009: 209).

We have likewise seen women's friendlier use of YOU WAS – for instance among the genteel women as their shared humiliation yields to the plot's denouement in Garrick's *Male-Coquette* (p. 47). Below, we also observe several female leads use YOU WAS with their lovers, in situations of subverted intimacy, where they have the upper hand over their ignorant lover. In Murphy's *The Way to Keep Him* (1760), YOU WAS is used by the Widow Bellmour to her suitor. They have had a long conversation about beauty and love, but their seeming closeness is undermined by several facts. First, the Widow Bellmour's lover does not know that she realizes he is not "Lord Etheridge" but Mr. Lovemore, a bored married man. Second, Mr. Lovemore also doesn't know that the conversation is being observed by Mrs. Lovemore – invited by the Widow Bellmour to learn how to keep a man. The Widow Bellmour is in a position of superiority and confidence, like the disguised Sophia in *The Male-Coquette*:

(14)   *Murphy, The Way to Keep Him (1760), II.i*

Lovemore.	Why Ma'am, in general, one does not see the Talents of a Wife, dedicated to the Happiness of the Husband. – I have known Ladies, who on the Eve of their Wedding appear'd like the very Graces, in a few Weeks after the Ceremony become very Slatterns, both in their Persons and Understandings: no Solicitude on their side to appear amiable. – … and so good-night to all real and solid Happiness. – But with one accomplish'd as you are, Ma'am –
Mrs. Bellmour.	To be sure, with me no-body cou'd be otherwise than happy; – was not that what **you was** going to say? – I know it was.
	(p. 49)

We also argue that YOU WAS marks the deceiving speaker's genuine lack of affectation. The Widow Bellmour is an unusually and explicitly sympathetic "other woman", with no hint of lowness: we've already been told that she has "kept the best Company" (p. 20); even her servant Mignionet uses YOU WERE (p. 41); and the cultivation and good sense implied by her "Harpsicord," "Book-case," and reading of Pope (p. 29) is confirmed by her apparent agreement that wives should obey their husbands (Nussbaum 2010: 245–246). The betrayed Mrs. Lovemore is witnessing the familiarity of this (albeit staged) conversation – but soon afterwards disconcerts her husband by enlivening her behaviour, using YOU WERE as she announces she will "pursue her own Plan of Diversion" (p. 64–6). In this play, neither woman uses both variants – though the Widow Bellmour's lesson to Mrs. Lovemore is that women's nature must integrate art (p. 37).

Some style-shifting by other lead females more explicitly suggests how YOU WAS might convey a combination of superiority and comfort as they engineer the fulfilment of their desires and hope to subordinate art to nature. Garrick's *Miss in Her Teens* (1747) used YOU WAS with her favourite suitor as her actions helped to resolve the plot, but YOU WERE with the unsuitable Fribble (p. 22). In Murphy's *Old Maid* (1761), Miss Harlow rejects her faithful suitor Captain Cape using YOU WERE (p. 30), while (in Example 15) using YOU WAS to express her confident passion for Clerimont, who has been courting her by correspondence, mistaking her for her sister-in-law. In a series of asides we can contrast this "young – handsome – rich" man's (p. 8) contempt for this ill-natured woman with her enthusiastic use of YOU WAS as she mistakenly contemplates her "power over him" (p. 20):

(15)   *Murphy, The Old Maid (1761), Act I*

Clerimont.	She rather looks like Miss Harlow's mother than her sister-in-law – [*Aside*].
Miss Harlow.	He seems abashed – his respect is the cause – [*Aside*].
	– My sister told me, Sir, that **you was** here – I beg pardon for making you wait so long –
Clerimont.	Oh, ma'am *(bows)*

the gloom disappears from her face, but the lines of ill-nature remain – [*Aside*.

Miss Harlow.	I see he loves me by his confusion; – I'll cheer him with affability –
[*Aside*].	
	– Sir, the letter **you was** pleased to send, my sister has seen – and – (p. 18)

Easy expression of feeling makes a genteel woman vulnerable – especially if she doesn't know it. The low associations of YOU WAS underscore the difference between the ageing spinster and the cultivated youth. For Murphy, women's confidently unaffected YOU WAS was recognizable enough to be burlesqued.

6.   Conclusion

In the analysis of our corpus of comedies published before 1762 by three prolific playwrights, we can corroborate Laitinen's conclusion from a survey of correspondence that YOU WAS likely emerged "from below" (2009: 206). Our Figure 5 shows an early peak in the subconscious usage of YOU WAS followed by a sharp drop and then an increase in its use as a marked feature in comedies of the late 1750s. Several other observations can be made about the use and distribution of YOU WAS and YOU

WERE. It is clear that deliberate juxtaposition of the variants by these playwrights indicates that there was a clear difference between their usages at the time of writings, for instance as a means of contrasting not only the difference between reality and hypotheticality, but also the different social standings of two reunited friends in Murphy's *Upholsterer* (1758), the characters of the boisterous Miss and the courtly Captain in Garrick's *Miss in Her Teens* (1747), or the mistrust between lovers, such as Beverley and Belinda in *All in the Wrong* (1761). Garrick's *Lying Valet*'s (1741) uncharacteristic use of YOU WERE when manipulating a character of higher social status, and conversely Sir William's employment of WAS YOU to style-shift closer to Shift in Foote's *Minor* (1760) may be further cogent examples of how these playwrights capitalized on these linguistic variants to convey social dynamics. The associations of excessive formality with the less-marked variant YOU WERE, and the colloquial quality of YOU WAS for at least some speakers indicates that both variants had sociolinguistic salience.

We can also see that YOU WAS came to be used in dramatic contexts where the then-deprecating second person pronoun THOU and its forms had once been used: for elite men berating servants (as in Garrick's *Male-Coquette* (1757) and Murphy's *The Way to Keep Him* (1760)), and similarly for a stern father asserting his preference of daughter-in-law in *All in the Wrong* (1761). Also, following Laitinen's analysis that YOU WAS had come to be used more by women than men by 1740–59 (2009: 209), there is particular use of YOU WAS in these comedies by female leads. These characters use this variant in a wide range of emotional contexts, from solidarity with another woman scorned by Garrick's *Male-Coquette* (1757), to disrespect of a spinster sister-in-law in Murphy's *Old Maid* (1761), to anger at a lover in Foote's *Author* (1757). There are likewise several scenes across these plays of lead female characters using YOU WAS in moments of unaffected confidence and self-assured ease, such as when deceiving their lover with a disguise in *The Male-Coquette* (1757), ensnaring their suitor under the gaze of his wife in *The Way to Keep Him* (1760), and even when their confidence and ease is entirely misplaced, as with Murphy's *Old Maid* (1761).

Surveying the work of three prolific playwrights up to 1762, we have suggested that as THOU receded even from theatrical use YOU WAS temporarily reinforced its marking of status and attitude. We have also identified potential comic tropes perhaps intuitive to actors – for instance, elite females' use of YOU WAS in moments of temporary superiority. While acknowledging the difference between literature and life, we have corroborated historical sociolinguistic surveys that identify YOU WAS as informal. We have further established that YOU WAS was associated with social lowness well before it was proscribed in Lowth's 1762 grammar, and identified that the variants were salient as early as 1747, contrasted in an exchange between

Biddy and her suitor. Indeed, for historical sociolinguists we have identified one useful property of comedy and especially romantic comedy: the stylized opposition between characters (particularly lovers) can sometimes date the salience of sociolinguistic variation.

Acknowledgements

References

Adamson, Sylvia. 2001. Understanding Shakespeare's grammar: Studies in small words. In Sylvia Adamson, Lynette Hunter, Lynne Magnusson, Ann Thompson & Katie Wales (eds.), *Reading Shakespeare's dramatic language: A guide*, 210–236. London: Arden Shakespeare.

Anderson, Misty G. 2014. Genealogies of comedy. In Julia Swindells & David Francis Taylor (eds.), *The Oxford handbook of the Georgian theatre, 1737–1832* (Oxford Handbooks Online), 347–367. Oxford: Oxford University Press.

Bevis, Richard. 1980. *The laughing tradition: Stage comedy in Garrick's day*. Athens: The University of Georgia Press.

Breen, Jennifer. 2004, September 23. Celesia [née Mallet], Dorothea (bap. 1738, d. 1790), playwright and poet. *Oxford dictionary of national biography*. http://www.oxforddnb.com.myaccess.library.utoronto.ca/view/10.1093/ref:odnb/9780198614128.001.0001/odnb-9780198614128-e-4986 (27 March, 2019.)

Eighteenth Century Collections Online = ECCO. 2021. Gale Cengage Learning. http://find.galegroup.com.myaccess.library.utoronto.ca/ecco/

Fitzmaurice, Susan. 2012. Social factors and language change in eighteenth-century England: The case of multiple negation. *Neuphilologische Mitteilungen* 113(3). 293–321.

Garrick, David. 1980. *The plays of David Garrick*. 2 volumes. Harry William Pedicord & Fredrick Louis Bergmann (eds.). Edwardsville: Southern Illinois University Press.

[Griffiths, Ralph.] 1757. Review of *The Epigoniad*, by William Wilkie. *Monthly Review* 17. 228–238.

Hegele, Arden. 2008. *Bad grammar in the eighteenth century: Mid-century dramatists and public opinion*. https://cpercy.artsci.utoronto.ca/prescrip/18thcComedy/home.html

Hodson, Jane. 2014. *Dialect in film and literature*. New York: Palgrave Macmillan. https://doi.org/10.1007/978-1-137-39394-4

Laitinen, Mikko. 2009. Singular YOU WAS/WERE variation and English normative grammars in the eighteenth century. In Arja Nurmi, Minna Nevala & Minna Palander-Collin (eds.), *The language of daily life in England (1400–1809)*, 199–217. Amsterdam: John Benjamins. https://doi.org/10.1075/pbns.183.12lai

Leggatt, Alexander. 1998. *English stage comedy, 1490–1990: Five centuries of a genre*. London: Routledge.

Literature Online = LION. 1996–2021. Search Texts. https://literature-proquest-com.myaccess.library.utoronto.ca/gotoSearchTexts.do

Lowth, Robert. 1762. *A short introduction to English grammar, with critical notes*. London: A. Millar, R. & J. Dodsley. ECCO.

Nonomiya, Ayumi. 2014. *Thou* and *you* in eighteenth-century English plays. *Token: A journal of English linguistics* 3. 211–234.

Nussbaum, Felicity. 2010. *Rival queens: Actresses, performance, and the eighteenth-century British theater*. Philadelphia: University of Pennsylvania Press. https://doi.org/10.9783/9780812206890

Oxford Dictionary of National Biography [ODNB]. https://www.oxforddnb.com

Priestley, Joseph. 1761. *The rudiments of English grammar, adapted to the use of schools. With observations on style*. London: R. Griffiths. ECCO.

Sundby, Bertil, Anne Kari Bjørge & Kari E. Haugland (eds.). 1991. *The dictionary of English normative grammar, 1700–1800* [= DENG]. Amsterdam: John Benjamins. https://doi.org/10.1075/sihols.63

Tieken-Boon van Ostade, Ingrid. 2002. *You was* and eighteenth-century normative grammar. In Katja Lenz & Ruth Mohlig (eds.), *Of dyuersitie & chaunge of langage*, 88–102. Heidelberg: Universitätsverlag C. Winter.

Tieken-Boon van Ostade, Ingrid. 2011. *The bishop's grammar: Robert Lowth and the rise of prescriptivism in English*. Oxford: Oxford University Press.

Walker, Terry. 2007. THOU and YOU in early modern English dialogues: Trials, depositions, and drama comedy. Amsterdam: John Benjamins. https://doi.org/10.1075/pbns.158

CHAPTER 12

Towards a companionate marriage in Late Modern England?

Two critical episodes in Mary Hamilton's courtship letters to John Dickenson

Anne-Christine Gardner
Université de Lausanne

Using a Critical Discourse Analysis approach (Fairclough 1992) which combines micro and macro level analysis, this article explores how the social relationship of a privileged couple, Mary Hamilton (1756–1816) and John Dickenson (c.1757–1842), is negotiated during their courtship in 1784–85 on the basis of Hamilton's letters in the *Mary Hamilton Papers* (HAM/2/15), held by The University of Manchester Library. The focus lies on two moments of crisis in which Hamilton criticises Dickenson's behaviour concerning letter-writing and hunting practices. The couple seems to follow the ideals of the companionate marriage which emerged in the eighteenth century: they treat each other as equals, which allows Hamilton to be outspoken about her opinions, and the partners accommodate to each other's positions.

Keywords: Late Modern English, historical pragmatics, critical discourse analysis, intertextuality, letters, diaries, social relationships, courtship

1. Introduction

The eighteenth century has often been described as the "golden age of letter-writing" (Tieken-Boon van Ostade 2009: 1). Serving a variety of functions, letters were a fundamental means to negotiate social relationships on paper (e.g., Fitzmaurice 2002; Gardner 2018). Hannan (2016: 156) observes that "[l]etters were used to maintain relationships during separation, but they also provided an arena for the deepening of intimacy between two people", and that the "process of sitting down and writing encouraged the necessary reflection and introspection to speak openly about important topics such as religion, philosophy, love or friendship". The aim of this chapter is to determine whether these principles also underlie the courtship

letters written by Mary Hamilton (1756–1816) to her fiancé John Dickenson (c.1757–1842) in 1784 and 1785. Born into an aristocratic family, Hamilton was a member of the Bluestocking Circle and one-time sub-governess at the court of King George III and Queen Charlotte.[1] On 18 June 1784 she became engaged to John Dickenson, whom she had known since 1771 and who was associated with Birch Hall, an estate historically in the county of Derbyshire (now Cheshire) which had been in the family for over twenty generations. Mary and John were married in London on 13 June 1785, and their daughter Louisa was born in 1787. The courtship letters survive in manuscript form and are held by The University of Manchester Library as part of the *Mary Hamilton Papers* (GB 133 HAM/2/15).[2] With Mary residing in London and John in Taxal near Chapel-en-le-Frith, these are letters which were written 'during separation'. This article investigates how the relationship between Mary and John is negotiated and developed on paper – and whether Mary does indeed 'speak openly'.

When studying historical correspondence and social relationships in the past, different perspectives and various interpretative layers need to be incorporated in the analysis. Fitzmaurice (2009: 233) speaks of "multiple contexts – from the cotext of the text to historical events – that are salient in aiding the interpretation of the ways in which individuals communicate, negotiate, and develop their relationships." The present study considers such "multiple contexts" in the interpretation of the courtship correspondence pertaining to Mary Hamilton and John Dickenson. On the micro level, this article investigates how the relationship between Mary and John is negotiated during the engagement period. The analysis focuses on two early moments of crisis documented in the letters and centres around two requests made by Mary to John concerning letter frequency and hunting practices. In both case studies I reconstruct John's reactions to Mary's request on the basis of intertextual references in HAM/2/15 and related manuscript sources. In the first case study, a censored passage in Mary's letter adds an additional layer of interest for the analysis of intertextuality and the relationship between the betrothed. On the macro level, this study considers the social, cultural and historical contexts in which their correspondence is situated. Particular attention is paid to letter-writing, courtship and marriage in late eighteenth-century England.

---

1.   The biographical information on Mary Hamilton and John Dickenson given in the following is derived from Anson & Anson (1925) and the *Mary Hamilton Papers* (HAM/1/3/1/2, HAM/2/10, 14–15). See also Baker (2004), Crawley (2014), and Eger (2017).

2.   The letters have not yet been edited; excerpts are contained in Anson & Anson (1925). The *Mary Hamilton Papers* are being digitised and transliterated at the University of Manchester (http://www.projects.alc.manchester.ac.uk/maryhamiltonpapers/; accessed 18 October 2021) and at the University of Zurich (see Gardner et al. 2017).

A framework which usefully combines the micro and the macro levels, and also enables the investigation of power relations between individuals, is Critical Discourse Analysis (CDA). Originally intended for the analysis of discourse in modern societies, Wood (2004, 2009) has successfully applied Fairclough's (1992) framework to historical data. The present study adopts the CDA approach, which Wood sees situated in the wider field of historical pragmatics. The following section describes this framework in more detail, while section 3 offers additional information on the manuscript sources consulted for this study. The macro-level contexts and social practices in which the courtship is situated are outlined in section 4. The analysis of the correspondence between Mary and John on the basis of two case studies follows in section 5. Section 6 concludes with a summary of main findings and an outlook on further avenues for research.

2. Framework

Fairclough considers language to be "a form of social practice" (1992: 63) and in consequence to have three primary functions, i.e., "identity", "relational" and "ideational":

> The identity function relates to the ways in which social identities are set up in discourse, the relational function to how social relationships between discourse participants are enacted and negotiated, the ideational function to ways in which texts signify the world and its processes, entities and relations.
>
> (Fairclough 1992: 64)

The three functions of language become manifest in discourse in the sense that discourse contributes to the construction of social identities, social relationships, as well as "systems of knowledge and belief" (Fairclough 1992: 64).

In Fairclough's three-dimensional conception of discourse (see Figure 1), the micro and the macro levels are presented as interdependent areas of investigation. The dimension 'text' lies at the core of the model and represents the micro level. For the analysis of texts, Fairclough proposes 'vocabulary', 'grammar', 'cohesion' and 'text structure' as fundamental topics of interest. The dimension 'text' is positioned within the higher-level dimension 'discursive practice', which is mainly concerned with the production, distribution and consumption of text. In the context of the present study this relates to the mechanics of letter-writing, in particular the ink used, as well as the time and place of writing, when letters were sent, and who read them. Together with 'text', 'discursive practice' is encompassed by the macro-level dimension of 'social practice'. Social practice can be investigated through a range of domains, including social status, gender and religion, as relevant to the interactants

```
┌──────────────────────────────────────────────────────────────────┐
│  Social practice (social status, gender, religion)                 │
│  ┌──────────────────────────────────────────────────────────────┐ │
│  │  Discursive practice (production, distribution, consumption)   │ │
│  │  ┌───────────────────────────────────────────────────────┐    │ │
│  │  │  Text (vocabulary, grammar, cohesion, text structure)  │    │ │
│  │  └───────────────────────────────────────────────────────┘    │ │
│  └──────────────────────────────────────────────────────────────┘ │
└──────────────────────────────────────────────────────────────────┘
```
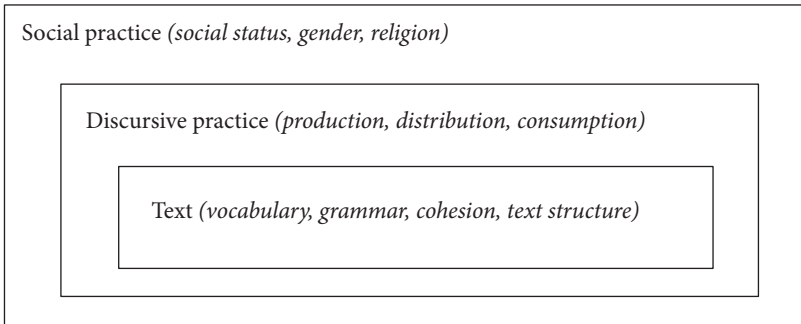
Figure 1. Three-dimensional conception of discourse, adapted from Fairclough (1992: 73) and Wood (2004: 233)

and object of study. 'Discursive practice' plays a mediating role between the two other dimensions: "it is the nature of the social practice that determines the macro-processes of discursive practice, and it is the micro-processes that shape the text" (Fairclough 1992: 86). Formal properties of texts also include 'force' (of utterances, i.e., the nature of speech acts), 'coherence' and 'intertextuality'. Crucially, however, these three categories also aid the analysis of 'discursive practice' in that they additionally provide insights into the production and interpretation of a text. Therefore, these categories in particular illustrate the interdependence of the three dimensions of discourse.

When adopting a CDA approach, it is possible to take either the dimension of social practice as a starting point and progress via discursive practice to the textual dimension (e.g., Wood 2004), or to begin at the level of the text and then move through the higher-level dimensions (e.g., Wood 2009). In the studies just mentioned, Wood examines how Margaret Paston projects her own identity in her letters and how her relationships with family members are negotiated in writing. Variation in salutations, closing formulae and other conventional epistolary elements reveal the powerful position she had within her family. Wood finds this linguistic variation to be strongly linked to important events in Margaret Paston's life as well as emotional states such as anger. On the level of discursive practice Wood shows, for instance, that during the production of letters Paston has influenced or even determined the wording penned by the scribe. On the macro level, linguistic variation is possible thanks to changing social practices, which saw a decline of traditional formula-heavy and tightly structured letter-writing appropriate for the feudal system in favour of an adapted, more expressive rhetorical style suitable for merchants like the Pastons.

In her analysis of letters written by Frances Leonora Macleay to her brother William in the first half of the nineteenth century, the CDA approach helps

Chiavetta "understand changes in the discursive and social practices of the sender" (2012: 101). These changes are linked to an increase in geographical distance between the siblings. The first set of letters analysed by Chiavetta date from the time when Frances still lived in London and her brother in Cambridge and later in Paris. In 1825 the family moved from London to Australia, with the exception of William, who remained in Europe and subsequently moved to Cuba. The second set of letters Frances sent to William from Australia. The siblings never saw each other again, and the writing of letters thus constituted the only means of maintaining their (close) relationship. With the increased geographical distance, the meta-discussion of discursive practices gains in urgency. Frances frequently made intertextual references to letters she received from her brother or expected from him, stressing the importance of letters in the maintenance of "communication and affective relations despite geographical distance" (Chiavetta 2012: 102). The letters are also a platform through which Frances constructs and performs a social identity as a sister. While she voices the trope of the neglected and inferior sister common at the time, practicing an ironic self-humiliating discourse, Frances is also able to express her true views and admonish her brother with the help of hedging devices.

In the qualitative analysis of Mary Hamilton's courtship letters, the present chapter proceeds from the outer to the inner dimensions, or from the macro to the micro level. Regarding social practice, section 4 explores the nature of courtship and marriage in the upper classes of late eighteenth-century England, as well as the roles available to women in (pre-)marital relationships at the time. In the discussion of 'discursive practice' and 'text' in section 5, the focus lies on requests voiced by Mary in moments of crisis and John's reactions to these requests. This is illustrated in two case studies: the first is concerned with letter frequency, which exemplifies the aspects of 'production' and 'consumption' in discursive practice (§ 5.1); the second case study, on the topic of hunting, is situated in the dimension 'text' (§ 5.2). With the focus on requests, the analysis centres on Fairclough's category of 'force'. However, with the intention of also establishing John's reaction to Mary's requests, this category is investigated in conjunction with 'intertextuality'. In the present context this term is taken to denote that letters "respond to prior texts both written and spoken" (Wood 2004: 231). How intertextuality will be investigated on the basis of the *Mary Hamilton Papers* is outlined in the following section.

3.   Material

Mary and John's courtship is documented in the letters she wrote to him from late December 1784 to May 1785 (HAM/2/15). Catalogued as diary letters in the *Mary Hamilton Papers*, they sometimes contain typical letter elements such as direct address, salutation, closing formula, subscription and superscription, but are mostly written like a diary, recording daily events, with letter elements interspersed. An ideal place to learn of John's reactions to Mary's requests would be his letters to her, written in direct response and containing intertextual references to her requests. However, while the *Mary Hamilton Papers* do preserve some of John's letters (HAM/1/2), only three survive from the engagement period, and none of these are direct replies to the requests investigated in the two case studies. The analysis in section 5 therefore relies in part on intertextual references within Mary's own letters, i.e., occasions in which she reprises a topic at a later date, after its first mention. Additional evidence can be found in his later letters to Mary, which were composed after they were married.

This study also takes into account a volume containing mainly diary extracts (but also letters and other notes) which Mary assembled (DDX 274/18), and which is held alongside her husband's diaries (DDX 274/19–25) by the Lancashire Archives. Mary started writing in this book from both ends, and on the respective inside covers wrote in pencil, "Extracts from my old journals &c" and "Extracts from my old journals wch I have destroyed".[3] It is uncertain when Mary compiled this volume, but evidence points towards a main period of activity between 1802 and 1806. Mary copied three diary extracts or letters from 1782, which were not destroyed after all and survived in the *Mary Hamilton Papers*. However, when writing the date she accidentally noted 1802 instead of 1782 each time. It is a mistake typically made when the current year, in which much dated correspondence and writing takes place, is salient in the writer's mind. The latest dated additions to the volume are newspaper extracts and an anecdote from 1806, as well as a note from 1813. This main period of activity is relevant for the discussion in section 5.1 of textual alterations in HAM/2/15.

---

3.   Hamilton uses the terms *journal* and *diary* interchangeably, even within a single entry (e.g., HAM/2/9, 3 April 1784).

4.   Social practice: Courtship and marriage in the upper classes in eighteenth-century England

The eighteenth century saw a range of social changes, such as the declining importance of the patriarchal system and the "rise of the individual" and of the "culture of sensibility" (Barclay 2011: 110). As a result, the expression of personal emotions and wishes, also in correspondence, became more widely accepted at a societal level. This also heralded a gradual, but major move away from marriages in which the family dictated choices and terms, which often foregrounded economic and social factors, to marriages based rather on personal choice and promise of future happiness (Barclay 2011). Among the upper classes, the so-called companionate marriage gained ground, in which partners were united in friendship and "put the prospects of emotional satisfaction before the ambition for increased income or status" (Stone 1977: 217). The notion of the companionate marriage also assisted in "equalizing relationships between husband and wife" (Stone 1977: 217), which opposed the traditional view of women as subservient and obedient wives.[4] Concomitantly, private letters between spouses decrease in formality over the course of the eighteenth century, writers instead employing more emotive language and more intimate terms of address (Nevala 2004: 245–248).

Nevertheless, this new outlook, which prioritised the interests of the individual over those of the family and strengthened the position of women in particular, did not entirely supplant the traditional expectations of the role women were expected to play in married life. For instance, household management still lay firmly in the hands of women, while men were expected to provide for their family. Upper-class couples typically did not marry until the man was able to set up a new household, separate from his family home. Despite many examples to the contrary (see, e.g., Stone 1977 and Barclay 2011), there were also cases where the relationship between (prospective) partners was evidently not equal. To give one example, Heffron (2010) sketches the development of the courtship between Louisa Catherine Johnson and John Quincy Adams at the end of the eighteenth century. Their correspondence reveals an ongoing struggle on Louisa's part between the social pressure to conform with tradition and her desire to assert her individuality: "John Quincy's precedence was tacitly acknowledged", whereas Louisa "vacillated between defiance and submission" (Heffron 2010: 216). This struggle neatly illustrates Barclay's observation on the role of courtship: "Just as courtship was a time to negotiate the economic terms of marriage, it was also the time to negotiate power in married life" (2011: 95).

---

4.   For a critical in-depth discussion of the companionate marriage and its implications for the role of women in society see Barclay (2011).

The principles of how relationships are negotiated and intimacy is deepened in letters, as described by Hannan (2016) and quoted in section 1, seem to underpin the ideals of the companionate marriage. Equality and friendship could be argued to form the foundation which allows partners to exchange their views freely and openly, and consequently to negotiate and develop their relationship across the divides of time and space. A couple struggling in these negotiations was Mary Pierrepoint and Edward Wortley, who did not 'speak openly'. Fitzmaurice (2000, 2009) extensively reconstructs the explicit and implicit pragmatic meanings conveyed in their courtship correspondence, and how failure to communicate in a clear and more straightforward manner resulted in a lover's spat. Like Louisa Catherine Johnson, Edward Wortley employs the rhetorical device of self-depreciation, however to "convey indirectness and tentativeness of approach" (Fitzmaurice 2009: 225). Eventually the couple eloped to get married, but the path to this conclusion was filled with challenges, such as refusals to fulfil requests, (perceived) misrepresentations of intentions and subsequent misunderstandings. It is important to note, however, that this couple lived in the early eighteenth century, when the ideal of the companionate marriage was not yet widespread. The case studies in the following section investigate how, towards the end of the same century, Mary Hamilton and John Dickenson negotiated their roles and power relations during their courtship. Do they see each other as equals, do they 'speak openly'?

5.   Two case studies

On the very first page of her first surviving courtship letter to John Dickenson (Figure 2), Mary Hamilton offers what could in modern terms be described as a mission statement underlying her correspondence with him. Mary's fiancé had been with her in London and was travelling back home to Derbyshire. Missing John, she describes her sadness and how she tries to control her emotions. Perhaps wondering whether she should really convey these struggles to John, Mary makes it clear in (1) that she intends to be "transparent" in revealing her "heart" to him, and that she does not want to keep anything from him.

> (1)   I will not erase what I have written – shall I ever hide a thought from You? …
> No No You shall see my heart as through a transparent Mirror.
>
> (HAM/2/15, 27 December 1784)

So Mary does indeed seem to 'speak openly'. To what extent this is also the case later-on, and what form her openness takes, will be explored in the following case studies. They investigate two early moments of crisis documented in the courtship letters, the first relating to letter frequency (§ 5.1), the second to the topic of hunting (§ 5.2).

Decbr 27th 1784. It is not necessary for me to tell You all I felt & thought when we parted last night — could I have wept I should have been relieved but it was not till I sat alone in my breakfast room this morn that the friendly aid of tears eas'd [in some degree] the oppression of my heart — I check'd every melancholy presage whilst I saw You, but when I knew You were hastening on Your road from me all thoughts of future happiness seemed to abandon me. I have been taking myself severely to task — but I will not erase what I have written — shall I ever hide a thought from You? — shall I ever be guilty of making You believe me more worthy of Your approbation than I really am? No No You shall see my heart as through a transparent Mirror.   The dear Sister was obliged to leave me for ye whole day She went to ye Wifes house to see the Widow'd Sister — I denied myself to every one & sat alone till dinner time, I read, I wrote, I work'd & made it a duty to [empty myself &] keep up my spirit The perverse one & I din'd tête à tête, who left me immediately after dinner — a letter from the good [Duchess of Portland] Dowager cheer'd my spirit in some degree, I will tell You what She said, at least so far as will interest You = The Duchess & Wm Delany call'd Mr Dickenson [always ye King of Poland]
"We hope ye King of Poland will stay till 12th day & yn Ye will
"be Queen, You have feasted them about so much yt he is not
"condescend to a vulgar Beef stake but for a particular relish
"he knows he shall meet with. Would you believe it my Dear
"— the D.D.R. ordered a poor Rook to be shot to death & a whole
"Colony frightened out of their wits who were all starving, for
"picking up a few grains of Corn fr ye Hares, whilst ye screaming
"Peacocks, squalling Guineas & cackling Bantams gobbled it
"all up wth impunity? we don't doubt Ye had a merry Xmass &
"return Yr kind wishes wth great warmth — the Friend [Mr Delany] says that

Figure 2. First page of Mary Hamilton's diary letter to John Dickenson, 27 December 1784 (Mary Hamilton Papers, HAM/2/15. Copyright: the University of Manchester)

5.1   Discursive practice: Letter frequency

In the eighteenth and nineteenth centuries, letter-writing was "fundamentally important in fostering, gauging, and testing a romantic bond" (Holloway 2019: 47), so much so that couples also engaged in corresponding with each other even if they did not live far apart. Holloway therefore argues that letter-writing represented "a distinct stage of courtship" (2019: 47) at the time. Both men and women frequently expressed anxiety and impatience when awaiting missives from their partners. According to Holloway (2019: 48), such behaviour revealed "a writer's commitment to a correspondence – and by extension a relationship – through reinforcing the high value of the letters they received".

   Similarly, the topic of waiting for letters also recurs in Mary Hamilton's correspondence with John Dickenson. On 28 December 1784, one day after her fiancé left her to return home, Mary expects to hear from John and expresses in (2) her disappointment when a promised message does not arrive. Using an imperative construction, she commands John to pay attention to "these kind of requests of mine". By this Mary presumably alludes to a prior conversation in which she asked John to write to her, to let her know of his journey's progress. For Mary the sending and receiving of letters clearly play an essential role in maintaining their relationship, and letters represent tokens of love. Failure on John's part in recognising this could result in a diminishment of her feelings towards him. The following day, however, relieves the initial tension; John's letter is delivered. Mary immediately apologises for her earlier reproach, also changing the way she addresses John from "my friend" (see (2)) to the more positively polite "my Dearest friend" in (3).[5]

(2)   I enquired if there was not a letter – there was not & I was disappointed – You had promised to send me a line … I tell You once for all my friend that You must either be very attentive to these kind of requests of mine, or I must learn to grow more indifferent about You.   (HAM/2/15, 28 December 1784)

(3)   I return'd to my writing & was agreeably interrupted by a Letter from You, forgive me my Dearest friend for what I said in yᵉ begin[n]ing of this page
                                        (HAM/2/15, 29 December 1784)[6]

---

5.   On address terms in historical correspondence, see Raumolin-Brunberg (1996) and Nevala (2004).

6.   The following potentially unfamiliar abbreviations appear in the manuscript passages cited in this chapter: "abᵗ" (*about*), "coudᵈ" (*could*), "havᵍ" (*having*), "Lʸ" (*Lady*), "recᵈ" (*received*), "Revᵈ" (*Reverend*), "shᵈ" (*should*), "wᶜʰ" (*which*), "wᵈ" (*would*), "wᵗʰ" (*with*), "yᵉ" (*the*), "yʳ" (*your*), and "yᵗ" (*that*).

In his next letter, which does not survive, John seems to have explained that on his travels he had not had enough time to write sooner. For Mary this reason is unacceptable. She begins by writing in (4) that she may be "wrong" and John may think her "unreasonable". With the following "Yet surely" Mary at first seems to have intended to make a suggestion as to how John should have behaved. Changing her approach, she proceeds to outline her own principles when it comes to letter-writing. Mary would rather sacrifice some sleep for writing than neglect "claims" of "friendship & affection". She goes on to state that she would have been "satisfied" if John had merely spent "five Minutes" in dropping her a line, less time than even the busiest statesman would manage "to devote to friendship". Mary feels disappointed by the fact that he was oblivious to her "suffering" and did not honour their relationship by dedicating this small amount of time to her.

(4)   perhaps I am in y^e wrong – & You may think I am unreasonable ~~Yet surely~~ well I wont say what You might have done – I will only say for myself – that I never in my life even when I had not y^e command of a Moment $^{\text{in y=e= day}}$ but that I found when friendship & affection had claims upon $^{\text{me}}$ that I could <u>steal</u> either an <u>early</u> or <u>late</u> hour from my repose to satisfy them – … even a $^{\text{first}}$ Minister of State will find one ½ hour in y^e 24 to devote to friendship – I should have been satisfied with <u>five Minutes</u> [deleted words, illegible] had You bestowed them upon me <u>last Wednesday</u>. I believe that I poſseſs Your affection & that You love me entirely – how comes it then that Your heart should be at ease & Your Mind at rest when You must know that I am suffering the torture of suspense, put Yourself in my place for a moment & then judge how anxious I was till I heard of Your safety.   (HAM/2/15, 4 January 1785)[7]

The topic is not reprised until after John's next visit to Mary, which lasted from mid-February until 19 March 1785. As in December, Mary starts writing to John on the day he leaves, and the question soon arises again when she will hear from him. However, this time there is no sign of her earlier impatience. Mary does not speak about John writing until two days after his departure and even turns the wait into a game. In (5) she tells John how she and her house mate Anna Maria Clarke agreed to a bet regarding the arrival of his letter, with Mary putting her stakes on a later delivery date. John, in turn, fulfils Mary's original request and ensures she receives a letter the day after his departure. Mary records her delight and "pleasure" in (6), referring to his "dear letter" as "proof of Your attention".

---

7.   Superscript letters above the line are marked by equal signs, as in "y=e=" for "y^e".

(5)   we laid a wager whether or no I sh^d receive a letter from You to day – I laid the reading 3 Acts of any Play of Shakespear AM [Anna Maria Clarke] chose, that I <u>should not</u>, she the playing any 3 leſsons [on the harpsichord] I liked best ^{y=t=} I should   (HAM/2/15, 21 March 1785)

(6)   I fill'd up my time wth writing to You & reading Your dear letter w^ch ~~I rec^d~~ came by this days Post, did You know y^e pleasure You afforded ^me by sending me this proof of Your attention & writing so much from Your heart

(HAM/2/15, 21 March 1785)

It is conceivable that, when seeing each other in early 1785, the couple talked about the issue of the late letter after John's first visit in December 1784. Both Mary and John now seem to accommodate to each other. John honours Mary's wish to hear from him as quickly as possible after his departure and to find proof of his commitment to their relationship in his writing, while Mary awaits his letter less anxiously. The earlier crisis appears to have been resolved, and the betrothed respect each other's positions, in line with the ideals of a companionate marriage (see § 4).

Intriguingly, however, there is an element of censorship in Mary's letter from 4 January following her reproach in (4). Mary originally apologises to John for this reproach in (7), and explains that her feelings for him are so strong that she cannot but speak openly and from her heart (cp. also the wording in (1)). This passage is almost entirely inked out (see Figure 3), but can be deciphered with the help of a magnifying glass and an LED-light, which makes it easier to distinguish the two different layers of ink. The original apology was written in a lighter shade of brown, whereas the censoring was done with a different type of ink which is almost black, as well as with a thicker pen stroke. Who censored the passage in (7), when, and why?

(7)   forgive me this little reproach – I must speak to You always from my heart – & I must cease to exist when I become indifferent ab^t You.

(HAM/2/15, 4 January 1785)
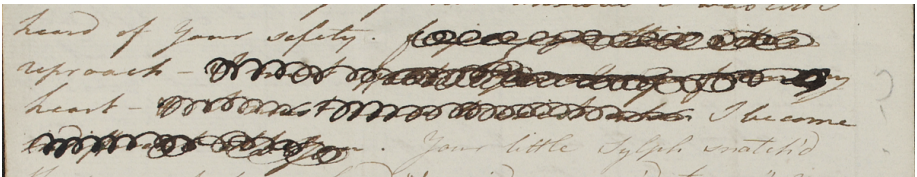
Figure 3. From Mary Hamilton's diary letter to John Dickenson, 4 January 1785 (Mary Hamilton Papers, HAM/2/15. Copyright: University of Manchester)

Mary's letters to John in HAM/2/15 contain numerous censored passages, as well as additions above the line or in the margins in the same darker ink. All of these additions are in Mary's handwriting. Some were made in 1806, so 21 years after the engagement period, as can be gathered from the reference to this year in (8) and (9), as well as the use of present tense *is* and the temporal adverb *now*, respectively. In these two additions Mary provides additional information on persons mentioned in the text. Elsewhere, she reveals the identity of those who originally were only mentioned with code-names. These include, amongst many others, "the Friend" and "Nestor", which Mary glosses as "Mrs Delany" (see Figure 2) and "Mr Glover" in additions to the entries from 27 December 1784 and 14 January 1785, respectively. Using such ciphers rather than real names was a common strategy to protect identities at a time when postal service was unreliable and letters could easily fall into strangers' hands (Nevala 2004: 54; Whyman 2009: 69). It is conceivable that Mary felt safe in lifting the identities behind code-names during her marriage since the potential audience had changed.[8] During the courtship, with her letters travelling from London to Derbyshire, Mary protected the identity of members of her London network from unknown third parties, or "eavesdroppers". During married life, Mary's letters would remain at home. However, also then there were potential parties other than the addressee, John, who could gain access to these letters. These readers would primarily be family members, so known "overhearers", for instance daughter Louisa, as well as John's father and his sister Sarah, who at first lived in the same house and later close by. This shift in third-party readers could explain a second type of censorship found in HAM/2/15. Mary heavily inked out passages which critically reflect on family members. In (10), for example, she writes negatively about her uncle Frederick Hamilton, whom she does not like as much as an old family friend, Lord Dartrey. By obliterating such passages Mary may have wanted to preserve an image of peaceful family relations, also for the next generation(s). These letters in fact remained in family possession until they were acquired by The University of Manchester Library in 2007.

(8)   NB [Mr Stanhope's] wife L^y Catherine is still alive & hearty <u>1806</u>
                    (HAM/2/15, addition in entry from 12 January 1785)

(9)   Rev^d M^r [Fisher] – Now Bishop of Exeter 1806
                    (HAM/2/15, addition in entry from 14 January 1785)

---

8.   For an application of Bell's audience design framework (1984, 2001) to historical correspondence, see Nevala (2004).

(10)   I have wounded him by not having thought it necesſary to consult him about my Affairs … . I must conceal from him that I prefer'd Lord D[artrey] for had he not [sic] loved me he w^d not have used such <u>bitter words</u>

(HAM/2/15, 9 May 1785)

Having established the author, potential date and reasons of censored as well as added passages in HAM/2/15, the question remains why Hamilton deleted her apology to John in (7). Evidence from her copy volume (DDX 274/18) provides a clue. As outlined in section 3, Mary copied into this book letter and diary sections which she wanted to preserve for herself and potentially for posterity. Here we find an almost verbatim record of Mary's reproach to John from (6). Mary introduces this passage by stating in (11) that this is her response to "a friend" who cited "want of time" as the reason why he did not answer her letter. Incidentally, this confirms the assumption made earlier regarding the content of John's lost apology which prompted Mary's reproach. Mary worked on this volume in 1806, the same year she made additions in HAM/2/15. It is conceivable that in this year Mary both deleted her apology in HAM/2/15 and copied her reproach into DDX 274/18. Also, when Mary inked out her reproach, she left four words standing, "reproach", "heart" and "I become". Had she censored this passage in 1785 and sent it to John, this would not have made sense to him. It is possible that, even though the situation seemed resolved in 1785, the issue was still rankling with Mary years later. Upon re-reading her letters to John, she may have felt justified in her original position and consequently deleted her apology.

(11)   Extract from a letter I sent to a friend who excused Himself for not hav^g answered a letter for want of time.   (DDX 274/18, page 31)

5.2   'Text' dimension: Hunting

The second case study focuses on another early moment of crisis in the courtship letters, concerning the topic of hunting. The episode starts innocently with Mary thanking John for a gift he sent her, "a very fine Hare", in (12). She likens this poetically to "triumphs in ye Sylvan Chase" and regards his gift as a physical manifestation of his efforts of maintaining and strengthening their social relationship across the distance of space. As Holloway (2019: 69) notes, "the exchange and use of romantic gifts provided a key way for courting couples to negotiate the path to matrimony". Although Holloway focuses on material objects of more permanent quality such as portraits, gloves or rings, which often carried symbolic meaning as well, Mary's reaction in (12) makes it clear that she sees John's (perishable) gift as a love token.

(12)   I rec^d one of Your triumphs in ye Sylvan Chase wch presented itself in ye form of a very fine Hare – I thank ^{You} for this as well as every other mark of Your recollection & attention.   (HAM/2/15, 18 January 1785)

Yet John seems to have miscalculated his choice of gift. In his reply to Mary's thanks he must have outlined his technique for ensuring the meat is tender. Upon learning that the animal had suffered for five hours, Mary condemns his hunting practices in no uncertain terms in (13). She starts her reproach with an imperative, a bald on-record request ("send me no more of Your Hares"), and vows that she would "rather live upon Bread & Water all my life" than benefit from "a diversion only fit for Savages". Hamilton effectively accuses her fiancé of ungentlemanlike behaviour. Driving the point home and using underlining for emphasis, she speaks of animals "in <u>an agony</u>" and "<u>tormented</u>", as well as the "<u>barbarous</u> sounds" made by Dickenson's hounds during the hunt.

(13)   send me no more of Your Hares – I cannot eat animals that I know have been ~~tormented~~ $^{in\ \underline{an\ agony}}$ for 5 hours – if food is to be made <u>tender</u> by that means – for You took ^{care} to inform me y^e hare You sent cou^d not be <u>tough</u> – I had rather live upon Bread & Water all my life. say what You please Hunting is a diversion only fit for Savages   (HAM/2/15, 21 January 1785)

To support her argument against hunting, Mary then launches into quoting a poem by William Somerville, *The Chase*, which describes the suffering of a hunted hare. Having spent four pages citing parts of this poem, Mary explains in (14) why she feels justified in voicing her critical opinion "in so free a manner", without apologizing: on an earlier occasion John alluded to hunting as a waste of time. At this stage we can observe a shift from a negative (13) to a more positive description and evaluation of her fiancé and his activities, beginning with Mary calling John "my Dearest friend" (14). She also uses this form of address when apologizing to John in (3), and elsewhere in the diary letters to index closeness.

(14)   I should my Dearest friend think it necefsary to make You an apology for having given my sentiments of Your favorite diversion in so free a manner, & do not think I should have done so, had <u>You</u> not said – "I begin to think y^t it is murdering time to bestow so great a part upon Dogs and Horses."
                    (HAM/2/15, 21 January 1785)

Afterwards Mary appears to qualify her earlier criticism. She acknowledges John's need to "exercise", but suggests that "other objects than a poor humbling Hare" might be more suitable for a hunt. Mary ends on a conciliatory note, focusing on positive qualities of her fiancé's character. In (15), she emphasises her desire for John's happiness (while removing her initial reference to "merriment", at odds

with her views on hunting) and her belief in his good judgement. She affirms her universal trust in his "principles" and "understanding", which convince her that it would be appropriate for her, or potentially even her duty, to approve of all his actions. Even more, his "satisfaction" and freedom to do what he likes is her ultimate goal. Also on an emotional level he is deserving of Mary's praise, as becomes clear from her use of positively connotated adjectives in (16) to describe his "heart", i.e., "generous", "sympathizing" and "Benevolent".

(15)   do not however imagine y^t I shall ever wish You to be debarr'd of what <u>may be</u> a perfectly innocent amusement, for I shall always be most happy when You do what $^{\text{You like, \& what}}$ is most likely to procure You satisfaction <s>& merriment</s> $^{\text{& afford}}$ $^{\text{recreation}}$. – I have too high an opinion of Your principles, too good an opinion of Your understanding not to be certain that <u>I ought to approve</u>, <u>every thing You do</u>.   (HAM/2/15, 21 January 1785)

(16)   I do not mean to leave out Your <u>heart</u> – for I know <u>that</u> is blefs'd wth $^{\text{all y=e= finer}}$ feelings, & that it is equally generous, sympathizing, & [deleted word, illegible] Benevolent   (HAM/2/15, 21 January 1785)

In a complete reversal from the beginning, Hamilton concludes by portraying herself as the one who may be lacking in quality in the eyes of the partner (17), relinquishing the moral high ground to him. Self-depreciation is a rhetorical strategy also found in other letters of the period in the context of relationship work and identity construction, for instance in the collections studied by Fitzmaurice (2009) and Chiavetta (2012) mentioned in section 2. To address her fiancé, Mary chooses the positively polite form of address "my Dear $^{\text{M=r=}}$ Dickenson" here, with "M^r" added perhaps as an additional show of respect.

(17)   the only uneasy apprehension I feel, my Dear $^{\text{M=r=}}$ Dickenson, is, that if providence should grant our wishes of being United that <u>I</u> shall fall short of y^e opinion You have form'd of me   (HAM/2/15, 21 January 1785)

The hunting episode is thus particularly revealing concerning the negotiation of power in this relationship, with a marked shift in Hamilton's linguistic choices (13 vs. 14–17). At the beginning, Hamilton expresses her displeasure towards John very directly in an unmitigated way. In the second part of the episode, in contrast, Hamilton is concerned with what might be called repair work. After all, she does not want to risk their relationship over the hunted hare, on the contrary, she still very much wishes to marry Dickenson.

John's immediate response to Mary's views is not extant, and the topic is not mentioned again in Mary's courtship letters. She may simply have dropped the subject, but it seems more likely that John at least stopped sending her game during

their engagement. One would expect Mary to acknowledge receipt of his gifts in her letters to honour convention, even if this did not reflect her sentiments. The topic of hunting is reprised, however, in John's later letters to Mary. When writing to his wife, John often sent a detailed list of his hunting exploits, as in (18). It is evident that John did not quit this pastime, and it would appear that Mary at least tolerated it to the extent that she accepted to read about his hunting narratives in his letters.

(18)   I went a shooting on Saturday, killed a Land rail & that was all – … Monday – Went a Shooting killed a Leveret a Hawk & a brace of Partridges
(HAM/1/2/51, letter from John Dickenson to Mary Hamilton, 1813)

The situations prompting Hamilton's two requests which were studied in section 5 can be interpreted as moments of crisis. Holloway (2019) and Lystra (1989) found that in courtship letters women in the eighteenth and nineteenth centuries often tested their fiancé's commitment to their relationship. To this end emotional crises were "frequently orchestrated", and the satisfactory conclusion of such a crisis could be regarded as a "rite of passage" (Holloway 2019: 63). In Hamilton's case, the emotional crisis surrounding Dickenson's neglect in correspondence is certainly not fabricated. As could be observed in her copy diary volume, this topic still held significance for her about two decades later. As regards Hamilton's request that Dickenson stop hunting, this may have been a slightly more artificial moment of crisis, in the sense that Hamilton must have been aware that her chances of convincing her partner were slim indeed. Hunting was one of the most common gentlemanly diversions at the time, and also a way of providing food for the family on country estates.

6.   Concluding remarks

The two case studies have shown that Mary Hamilton's courtship letters to John Dickenson, including her textual alterations, reflect contemporary discursive and social practices. In negotiating her social relationship with John and her own identity on paper, she portrays herself as an independent woman with her own interests, as well as a critical mind. Mary voices her opinions and wishes very openly, uttering her requests in direct speech acts with imperatives. She explicitly criticises John for his actions – even likening him to a savage in the hunting episode. At the same time, though, Mary expresses her affection for him during potentially volatile phases of negotiation. She is quick to emphasize the closeness of their relationship, through positively polite forms of address, and stressing her high opinion of John, which becomes particularly apparent through her choice of adjectives in the second case

study. When her reproaches could appear too strong, Mary even shows a degree of deference towards her fiancé, relinquishing the moral high ground in his favour.

Neither of the two points of criticism Mary levels at John appear to have been detrimental to their relationship. Overall, the case studies reveal that Hamilton and Dickenson show respect for each other as equals, reflecting the ideals of the companionate marriage. They act upon each other's wishes, yet without compromising their identities as individuals. Regarding Mary's request that John write to her more quickly on his return home from visits to her, the two partners accommodate to each other. After his next visit, Mary does not await his letter with such urgency, while John posts a letter within the expected time frame. Concerning Mary's second request, John does not give up hunting for her, but appears to at least stop sending her game during their engagement, and Mary does not reprise the topic in her letters to him.

In a further step it would be interesting to investigate how Mary responds to requests made by John. One recurring topic in the courtship letters, for instance, concerns her bodily and mental well-being, and coupled with this his conviction that going to bed early would be beneficial to Mary. Since the courtship letters span a period of about half a year, this also offers the opportunity to investigate how the relationship develops and deepens over time. Further avenues of research include changes in address terms and closing formulae, as well as in topics covered in the courtship letters (also in John's three surviving letters to Mary from this period), and how the couple deals with trying circumstances in terms of linguistic self-discipline (Sairio 2013, 2017).

Moreover, it would also be useful to study the views expressed by members of Mary's and John's networks in the surviving correspondence and diaries, both on courtship and marriage in general, as well as on the relationship between Mary and John in particular. The Dowager Duchess of Portland (1715–1785) and Mary Delany (1700–1788), for example, supported their prospective union. When Hamilton was staying with the two older women at the Duchess's home as a house guest in autumn 1784, her fiancé was invited to join her at Bulstrode when he came back to the London area for his first visit since their engagement. Mary Delany herself, Bluestocking and a close friend of Mary Hamilton's despite the generational difference, was convinced that marriage should be based on rational friendship, affection and equality (Brown 1982). Yet not everybody shared Delany's and Hamilton's more progressive views. One such person seems to have been Frances Burney (1752–1840), who was only four years older than Hamilton. Burney clearly did not approve of the marital behaviour of the Dickensons and the higher status allowed to the female partner which contrasted more traditional relationship settings. In her diary entry for 26 January 1789, Burney recorded her

criticism, voiced together with Stephen Digby (who eventually married Hamilton's friend Charlotte Gunning): "[Mrs Dickenson's] conjugal conduct, in displaying her superior power over her Husband, was our particular theme, &, of course, our mutual censure" (Sill 2016: 74). This statement needs to be considered with caution. Burney had a history of perceiving Hamilton's actions and words differently than they were intended: where Hamilton offered affection and friendship, Burney saw the affectation of an acquaintance (Gardner 2018). In a similar way, Burney could have (mis-)perceived the way the Dickensons embodied the ideal of the companionate marriage, believing Mary to flaunt her superiority and behaving in a way which was inappropriate for a married woman.

However, the *Mary Hamilton Papers*, as well as John Dickenson's letters to Hamilton (HAM/1/2) and his personal diaries (DDX 274/19–25), document the spouses' continued mutual respect, love and happiness throughout their married life. Stone sees here "the epitome of the new companionate marriage among the upper classes of the late eighteenth century" (1977: 240). Indeed, Dickenson professes his continued attachment to his wife in his letters. For instance, he closes one of his letters with "I am y^r <u>happy</u> Husband" (HAM/1/2/58, undated) and in (19) elaborates on these sentiments more extensively on the occasion of their eighteenth wedding anniversary:

(19)   Adieu my dear dear Wife – may you enjoy many Years of Health & Happineſs & live to see y^r Louisa united to a Man, who will be as grateful for the matrimonial Bleſsings he may enjoy, as is at this moment Y^r faithful & Aff[ectionate] Husband JD   (HAM/1/2/42, letter from John Dickenson to Mary Hamilton, 13 June 1803)

Acknowledgements

Sources

DDX 274. *Dickenson family of Birch Hall*. Lancashire Archives.

Mary Hamilton Papers. The University of Manchester Library (GB 133 HAM). http://luna.manchester.ac.uk/luna/servlet/Manchester~11~11 (24 April, 2019)

Sill, Geoffrey M. (ed.). 2016. *The court journals and letters of Frances Burney*. Volume 5: 1789. Oxford: Clarendon Press.

References

Anson, Elizabeth & Florence Anson. 1925. *Mary Hamilton, afterwards Mrs. John Dickenson, at court and at home: From letters and diaries, 1756 to 1816*. London: Murray.

Baker, Anne Pimlot. 2004. Hamilton, Mary (1756–1816). *Oxford Dictionary of National Biography*. Oxford: Oxford University Press. https://doi.org/10.1093/ref:odnb/48934 (24 April, 2019)

Barclay, Katie. 2011. *Love, intimacy and power: Marriage and patriarchy in Scotland, 1650–1850*. Manchester: Manchester University Press. https://doi.org/10.7228/manchester/9780719084904.001.0001

Bell, Alan. 1984. Language style as audience design. *Language in Society* 13. 145–204. https://doi.org/10.1017/S004740450001037X

Bell, Alan. 2001. Back in style: Reworking audience design. In Miriam Meyerhoff & Erik Schleef (eds.), *The Routledge sociolinguistics reader*, 32–52. London: Routledge.

Brown, Irene Q. 1982. Domesticity, feminism, and friendship: Female aristocratic culture and marriage in England, 1660–1760. *Journal of Family History* 7(4). 406–424. https://doi.org/10.1177/036319908200700408

Chiavetta, Eleonora. 2012. A critical discourse analysis approach to the letters of Frances Leonora Macleay. In Marina Dossena & Gabriella Del Lungo Camiciotti (eds.), *Letter writing in late modern Europe*, 89–104. Amsterdam: John Benjamins. https://doi.org/10.1075/pbns.218.06chi

Crawley, Lisa. 2014. A life recovered: Mary Hamilton 1756–1816. *Bulletin of the John Rylands Library* 90(2). 27–46. https://doi.org/10.7227/BJRL.90.2.3

Eger, Elizabeth. 2017. Bluestocking circle [bluestockings]. *Oxford Dictionary of National Biography*. Oxford: Oxford University Press. https://doi.org/10.1093/ref:odnb/63013 (24 April, 2019)

Fairclough, Norman. 1992. *Discourse and social change*. Cambridge: Polity Press.

Fitzmaurice, Susan. 2000. Like talking on paper? The pragmatics of courtship and the eighteenth-century familiar letter. *Language Sciences* 22(3). 359–383. https://doi.org/10.1016/S0388-0001(00)00011-5

Fitzmaurice, Susan. 2002. *The familiar letter in early modern English: A pragmatic approach*. Amsterdam: John Benjamins. https://doi.org/10.1075/pbns.95

Fitzmaurice, Susan. 2009. The sociopragmatics of a lovers' spat: The case of the eighteenth-century courtship letters of Mary Pierrepont and Edward Wortley. *Journal of Historical Pragmatics* 10(2). 215–237. https://doi.org/10.1075/jhp.10.2.04fit

Gardner, Anne-Christine. 2018. Downward social mobility in eighteenth-century English: A micro-level analysis of the correspondence of Queen Charlotte, Mary Hamilton and Frances Burney. In Anita Auer & Mikko Laitinen (eds.), special issue *Mobility, variability and changing literacies in modern times, Neuphilologische Mitteilungen* 119(1). 71–100.

Gardner, Anne-Christine, Marianne Hundt & Moira Kindlimann. 2017. Digitization of the Mary Hamilton Papers. *ICAME Journal* 41. 83–110. https://doi.org/10.1515/icame-2017-0004

Hannan, Leonie. 2016. *Women of letters: Gender, writing and the life of the mind in early modern England*. Manchester: Manchester University Press.

Heffron, Margery M. 2010. "A fine romance": The courtship correspondence between Louisa Catherine Johnson and John Quincy Adams. *The New England Quarterly* 83(2). 200–218. https://doi.org/10.1162/tneq.2010.83.2.200

Holloway, Sally. 2019. *The game of love in Georgian England: Courtship, emotions, and material culture*. Oxford: Oxford University Press. https://doi.org/10.1093/oso/9780198823070.001.0001

Lystra, Karen. 1989. *Searching the heart: Women, men, and romantic love in nineteenth-century America*. Oxford: Oxford University Press.

Nevala, Minna. 2004. *Address in early English correspondence: Its forms and socio-pragmatic functions*. Helsinki: Société Néophilologique.

Raumolin-Brunberg, Helena. 1996. Forms of address in early English correspondence. In Terttu Nevalainen & Helena Raumolin-Brunberg (eds.), *Sociolinguistics and language history: Studies based on the corpus of early English correspondence*, 167–181. Amsterdam: Rodopi.

Sairio, Anni. 2013. Cordials and sharp satyrs: Stance and self-fashioning in eighteenth-century letters. In Marijke van der Wal & Gijsbert Rutten (eds.), *Touching the past: Studies in the historical sociolinguistics of ego-documents*, 183–200. Amsterdam: John Benjamins. https://doi.org/10.1075/ahs.1.10sai

Sairio, Anni. 2017. Grey goose quills and scripts of self-discipline: The linguistic biography of Elizabeth Montagu (1718–1800). Paper presented at the symposium "The eighteenth century in ego documents: The individual in society", Zurich, March 9–10.

Stone, Lawrence. 1977. *The family, sex and marriage in England 1500–1800*. London: Weidenfeld and Nicolson.

Tieken-Boon van Ostade, Ingrid. 2009. *An introduction to Late Modern English*. Edinburgh: Edinburgh University Press.

Whyman, Susan. 2009. *The pen and the people: English letter writers 1660–1800*. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199532445.001.0001

Wood, Johanna. L. 2004. Text in context: A critical discourse analysis approach to Margaret Paston. *Journal of Historical Pragmatics* 5(2). 229–254. https://doi.org/10.1075/jhp.5.2.05woo

Wood, Johanna L. 2009. Structures and expectations: A systematic analysis of Margaret Paston's formulaic and expressive language. *Journal of Historical Pragmatics* 10(2). 187–214. https://doi.org/10.1075/jhp.10.2.03woo

CHAPTER 13

On the development of OE *swā* to ModE *so* and related changes in an atypical group of demonstratives

Ekkehard König and Letizia Vezzosi

FU Berlin / Università di Firenze

Building on previous comparative studies and comparative reconstructions (König 2012, 2015), we trace the syntactic and semantic development of OE *swā*, of its adverbial counterpart *swylc* and of *þus* to Modern English *so, such and thus* on the basis of relevant text corpora for OE and ME. In contrast to previous studies of *swā* in OE (cf. Schleburg 2002) and of *so* in ME (Nummenmaa 1973), it is shown that *swā, swelc, thus* and their counterparts in ModE are not isolated particles or adverbs, but are more adequately analysed as demonstratives of manner, quality and degree. Starting from a basic exophoric (gestural) use and its typical extensions to anaphoric and cataphoric uses, these expressions develop into a wide variety of grammatical markers in ModE.

Keywords: demonstratives (of manner, quality, degree), exophoric and endophoric uses, 'similarity', grammaticalisation, text corpora, Germanic languages, Romance languages

1.   Introduction

The uses of OE *swā* and ME *so* have been the object of several studies since the first decades of the last century (Johnson 1913; Ericson 1932; Fettig 1935; Peltola 1959), up to the more recent and detailed studies of Schleburg (2002) for Old English (henceforth OE) and Nummenmaa[1] (1973) for Middle English (henceforth ME). Different as their approaches might be, they all are synchronic and primarily focused on the syntactic functions of this expression and on the semantic classification of its various uses in OE or ME. Not too different is the analysis of these expressions in historical grammars (cf. Mustanoja 1960; Mitchell 1985), which generally

---

1.   We are indebted to Terttu Nevalainen for drawing our attention to Nummenmaa's book.

identify their uses in terms of their syntactic role as adverbs or conjunctions and in terms of their meaning as denoting manner and degree. In none of these grammars is *swā/so* considered in relation to functionally similar expressions such as *þus/thus* or *swylc/such* or to their counterparts in other Germanic languages, where, more clearly than in English, they manifest all the hallmarks of demonstratives.

Demonstratives are a typologically well-established, basic and possibly universal grammatical category. They are a subclass of deictic expressions, i.e., expressions whose reference can only be determined relative to a centre of orientation, as proximal, medial, distal, etc. This centre of orientation is typically provided by the coordinates of the speech situation, i.e., by the place, the time and the participants involved in an utterance. In their basic, deictic (exophoric) use, demonstratives are typically accompanied by a pointing or mimicking gesture. Demonstratives are highly flexible in their syntactic behaviour and grammars usually assign them to several syntactic categories, but as far as their semantics is concerned, they have a very simple structure, expressing typically values in only two dimensions of meaning: in addition to the deictic dimension described above, they also have a content dimension, which characterises an entity in terms of basic ontological categories such as 'human being' (*him, her*), 'inanimate entity' (*this, that*), location (*here, there, yonder*), direction (*hither, hence, thither, thence*), etc. The preceding characterisation shows that we are using the term 'demonstrative' not in the traditional sense of deictics categorised as pronouns, but more in the sense found in cross-linguistic and typological studies, such as Diessel (1999) or Dixon (2003), i.e., as a subclass of deictics characterised by the following properties: (i) a simple semantic structure (deictic + ontological components), (ii) specific use types (exophoric/gestural vs. anaphoric and cataphoric), (iii) their association with pointing or mimicking gestures, (iv) their role in the evolution and acquisition of language, (v) their syntactic flexibility, (vi) the relatedness in their sub-morphemic components (*this, there, thither, then, thus*) and to those of interrogatives (*which, where, whither, when*), and (vi) their important role as sources of grammaticalisation. Thus our use of the term demonstratives covers a similar set of expressions identified by the term deictics in other terminological traditions.

Our paper is concerned with the historical development of an atypical and somewhat neglected subclass of demonstratives from Old English to Modern English (ModE), namely the demonstratives denoting the content dimensions of 'manner' (*swā > so; þus > thus*), of 'quality' (*swylc > such*) and of 'degree' (*swā > so; swylc > such*) (cf. König 2012, 2015; König & Umbach 2018). Given the frequency and prominence of OE and ME *swā* in the historical development of English and the important roles of *so* in the structure of Modern English, it is not surprising that our study can be based on earlier work, in particular the rich and nearly exhaustive collections of data compiled by Schleburg (2002) for *swā* in OE and by Nummenmaa

(1973) for *so* in ME. We use these data to a large extent, but apart from that our study differs from and goes beyond these two in crucial and important respects:

i.   Our study is a diachronic one and outlines the development of OE *swa* and related expressions (*þys/þus, swylc*) from their origins to their use in Modern English on the basis of relevant text corpora (see below). This diachronic approach also requires a detailed analysis of the relevant phenomena in Modern English, not found in the studies mentioned above, but partly provided by dictionaries, in synchronic studies of specific use types of *so* in PDE (Johnson 2002; Tagliamonte & Roberts 2005; Bolden 2008, 2009; Raymond 2004), and in handbooks, such as Huddleston & Pullum (2002) or Aarts & McMahon (2006).

ii. Our study goes beyond the analysis of a single expression and examines the development of the whole subclass of atypical demonstratives, whose members are related in their meaning and partly also in their form. *Such* (*swylc/swelc*) derives from the adjectival form of *swā* (*swā + līc > swylc/swelc*) and *þus* is related to the adnominal demonstrative *þis*. An analogously derived adjectival form (*þuslic/þyllic*) is also attested and can still be found in dialects (*thusly*).

iii. Instead of analysing *swā/so* as an isolated particle (Schleburg 2002), adverb (Nummenmaa 1973) or discourse marker (Boldon 2008, 2009) we embed its analysis and that of related forms into the larger context of the historical development of demonstratives, to which they clearly belong, since the predominant endophoric (anaphoric and cataphoric) use of these expressions clearly points to their deictic (exophoric) roots.

iv. Our analysis is largely based on insights gained within 'grammaticalisation theory' for the development of grammatical items from deictics and relate the developments reconstructed for English to those found in a variety of other European languages (cf. Diessel 1999, 2013; Heine & Song 2011; König 2012, 2015).

2. Data and methodology

Given the high frequency of the words under analysis and the relevance of their co-text of occurrence, it was necessary to gather data from different sources and through different methods. As already mentioned, we are not dealing with totally uncharted territory. Our analysis can partly be based on earlier studies (cf. the list given above) and traditional historical grammars (Mustanoja 1960; Mitchell 1985), which mainly focused on classifying the diversity of uses of OE *swā* and, to a much lesser extent, of ME *swa/so*. The most valuable starting points for our inquiry, however, were the rich collections of data presented in Schleburg (2002) and Nummenmaa (1973). Nevertheless, every example taken over from these collections was checked against the original editions.

The material examined in the present study also consists of data extracted from large corpora: besides the Helsinki Corpus online (HC), for Old English, our data derive from the *Dictionary of Old English Corpus – on the World Wide Web* (DOEC), supported by the *York-Toronto-Helsinki Parsed Corpus of Old English Prose* (YCOE) and the *York-Helsinki Parsed Corpus of Old English Poetry* (York Poetry Corpus); as for ME, we used the *Corpus of Middle English Prose and Verse* (CMEPV), the *Middle English Grammar Corpus* (MEG-C) and the *Penn-Helsinki Parsed Corpus of Middle English* (PPCME2). Furthermore, we included information provided by dictionaries, mainly the OED, the *Dictionary of Old English* (DOE) and the *Middle English Dictionary* (MED), as well as the MED online (McSparran et al. 2000–2018).

Our syntactic analyses are based on the usual distributional tests, but are also inspired by recent syntactic theorising and by typological work, so that they lead to much more fine-grained distinctions than previous classifications in the same domain. Detailed comparisons between data from Old English and Middle English with their counterparts in Modern English are an essential part of our semantic analysis, which is additionally based on information provided by dictionaries and grammar handbooks (such as Huddleston & Pullum 2002) and our own introspective judgements as well as on available information on insights into typical developments of demonstratives provided by typological and comparative studies (cf. Diessel 1999; König 2015). Overall the methods used in this study are qualitative rather than quantitative ones and so our occasional remarks on frequency are more or less impressions based on the lists and classifications found in Schleburg (2002) and Nummenmaa (1973). As already mentioned, the hypotheses generally associated with 'grammaticalisation theory' (cf. Hopper & Traugott 2003; Heine & Kuteva 2007; Lehmann 2017) also play a role in our investigation, since it is assumed that changes on the formal side, i.e., de-categorisation and re-categorisation, and those on the semantic side (e.g., loss of exophoric meaning, contextual enrichment) occur in tandem and are unidirectional (exophoric > endophoric > discourse-structuring).

3.   Demonstratives of manner, quality and degree in Old English

As far as their distribution and syntactic behavior are concerned, demonstratives are highly versatile, so their traditional categorisation as determiners or articles (*this/that thing*), as pronouns (*this, that*) and as adverbs (*here, there*) excludes a wide variety of expressions clearly classifiable as demonstratives on the basis of the criteria mentioned above. It does not come as a surprise, therefore, that the expressions under study (*swā, swylc, þys/þus*) are not assigned to the category of demonstratives (deictics) in the two historical studies mentioned above, but are

categorised as particles (Schleburg 2002) or as adverbs, conjunctions and relative markers (Nummenmaa 1973). So, how can we justify our categorisation? These expressions are parallel in their semantic structure to other demonstratives/deictics (Table 1) in expressing two dimensions of meaning: a content dimension relating to 'manner' (OE *swā, þus*; arch. *thus*; ModE *like this*), to 'quality' (arch. *such*; ModE *like that*) and to 'degree' (*so, this/that*), and a deictic dimension, differentiated in many other languages (e.g., Serbian, Finnish, Japanese, Armenian), though not in English, in terms of two-term or three-term systems (proximal, distal, medial). So, despite appearances to the contrary, English also had and still has demonstratives of manner, quality and degree ('MQD-demonstratives' henceforth), even if the forms expressing these meanings have changed considerably in the course of its history. What has also changed are the possible uses: the three forms directly derived from their roots in Old English (*thus, such, so*) have almost completely lost their exophoric (deictic) use and are primarily used anaphorically (retrospective use) and cataphorically (prospective use). Due to this loss, each of the two semantic components, the deictic/endophoric one and the ontological one, is often either spelt out separately in ModE (*do it so > do it like so / do it like this; such* N > *suchlike* N) or the demonstrative in question is totally replaced by a basic local one (*so* Adj > *this/that* Adj). The sources and endpoints of our historical analysis can roughly be described as follows:[2]

Table 1. From OE to ModE: The 'global' picture

	Manner	Quality	Degree
Old English (exophoric)	þys > þus; swā	swelc/swylc/(þuslic)	swā; swylc
ModE (exophoric/deictic)	like so; like this	like this/that	so/this/that
ModE (endophoric/anaphoric)	(thus), so	such, suchlike	(so) this, that

Just like *swā*, the demonstratives *þys* and *þus*[3] can express the content dimension of manner. The content dimension of quality is expressed by *swylc* in Old English and by its cognate forms *such, suchlike* in Modern English. In the context of adjectives, *swā* may also express the dimension of degree in addition to that of manner. The following examples illustrate the meanings and uses of the OE demonstratives listed in Table 1, as far as their content dimensions are concerned:

---

2. We note that the inventory of MQD-demonstratives in Old English is very similar to that found in two closely related West Germanic languages, namely German (*so, solch, so*) and Dutch (*zo/dus/zus, zulk, zo*), where the relevant expressions have however preserved their exophoric use.

3. *Þys* is the instrumental form of the demonstrative pronoun *þis* (masc/neuter) and it is quite plausible to assume that *þus* and *thus*, its counterpart in Modern English, derived from this instrumental form as a result of unrounding the vowel.

(1) *Swā*[4] *ða   drihtguman drēamum lifdon*   [Beow 99][5] MANNER
so   the retainers   in.delights lived
"In this way the king's entourage lived happily"

(2) *Mid þæm bryne hio wæs* **swā** *swīþe forhiened þæt hīo nǣfre siþþan*
with that fire   she was so   much destroyed that she never afterwards
swelc næs   [Or 6 1.133.15] DEGREE
such not.was
"With that fire it was so much destroyed that afterwards it was never such [as
it had been] since"

(3) a. *Ealle hīe   sceolan þonne arīsan … on* **swylcum** *hēowe swā hie   ǣr*
all   they should then   rise   in such   form as   they before
hīe   sylfe ȝefrætwodan
them self adorned
"They all should then rise on such forms as they had adorned themselves
before"   [HomS 26 (BlHom 7) 226] QUALITY (attributive)

b. *Bēo þīn wīf* **swylc swā** *Uenus, þīn fūle gyden,   wæs, and bēo ðu*
be thy wife such as   Venus thy foul goddess was and be you
swylc swā *Iouis þīn sceandlica god, wæs …*
such as   Jove thy shameful god was
"Be thy wife such as Venus, your foul goddess, and be you such as Jove,
your shameful god"   [ÆLS (Agatha) 65] QUALITY (predicative)

Whether some of the relevant Old English forms also differed in their deictic di-
mension is not clear. In ModE coordination of two demonstratives denoting the
same content dimension (*here and there, this and that, hither and thither*) provide a
relevant test by losing their deictic meaning and expressing variation and quantifi-
cation in their referents without precise identification. In our Old English data there
are almost no instances of such constructions, but even when they start occurring,
during the Middle English period, clear evidence for the interpretation and deictic
differentiation of *so, thus* and *such* is not available, since the relevant meaning can
also be expressed by coordinating two instances of one and the same demonstrative
(*þus and þus; swelc and swelc*):

---

4.   [Wider context for (1)] … *se þe his wordes geweald wide hæfde·/ he beot ne aleh· beagas dælde
/ sinc æt symle … þæt he dogora gehwam dream gehyrde / hludne in healle· þær wæs hearpan sweg
/ swutol sang scopes·*[Beow 76–90] " … he whose words weight had everywhere; / he did not lie
when he boasted; rings he dealt out, / riches at his feasts. … he that every day heard noise of
revelry / loud in the hall; there was the harmony of the harp, / the sweet song of the poet".

5.   All the examples are quoted either from the DOE and the MED, whenever possible, or from
DOE and MEC, if not indicated otherwise.

(4)   *Somme seyden **thus**; somme seyde, 'it shal be **so**.'*[6]

              [Chaucer CT. Kn. (Manly-Rickert) A.2516]

    "Some said so; others said "it shall be like that""

(5)   *No deede of oure gouernaunce may be a moral vertu, howe euer hiȝe or lowȝe or in meene, here or þere, **þus or so**, myche or litil þilk deede falle.*

              [Pecock Rule (Mrg M 519) 294 (Greet 1927)]

    "No deed of our governance can be a moral virtue, however high or low or common, here or there, thus or so, much or little the same deed happens to occur"

(6)   *Penance I giue þe nan, For scrift* [Frf: *crist*] *sal mak þi saul scirre, Wit penance thol in clenssing fire; Bo, if þat þou mai liue, in chance, For **þus** and **þus**, þou do penance*   [Cursor (Richard Morris) (Vsp A.3) 26203]

    "I give you no penance, for Christ shall make your soul bright, with penance suffered in cleansing fire; both, if you may live, by chance, for thus and thus you do penance"

(7)   *þe blinde ȝaff he … to sen … dumbe men. To spekenn. All **þuss & tuss** he dide god Amang Judisskenn lede*   [Orm (Jun 1) 15520]

    "he gave the blind to see, the dumb to speak. Al thus and thus he did good things among the Jewish people"

(8)   *For that **so or so or so** (and in noon other wise) it is writun in storie or cronicle.*

              [(c1449) Pecock Repr. (Cmb Kk.4.26) 350]

    "For so or so or so (and in no other way) it is written in histories and chronicles"

Thus, while providing evidence for the categorisation of *swā/so*, *swelc* and *þus* as demonstratives, these examples do not justify the assumption that there were deictic differentiations of meaning (e.g., proximal vs. distal) in the relevant domain.

    Table 1 provides some basic orientation in the analysis and development of MQD-demonstratives in English, but the straightforwardness of this overview is partly due to some simplification. We have to add at this point that the semantic specialisation and the division of labour between our three demonstratives, not only in Old English, but also in Middle English (see below) and Early Modern English, is less clear-cut than is suggested by Table 1. *Thus* and *such* can be combined (cf. (9)), both seeming to express quality, and in the environment of *much* in particular, *thus* and *so* seem to be partly interchangeable in their use as demonstratives of degree.

---

6. This example can also tell us something different: *thus* is already the preferred option with *verba dicendi*, while *so* is restricted to other, non-cataphoric uses (degree, manner). There do not seem to be any instances of so+*say in the Middle English corpora we have consulted.

There seems to be, however, a clear preference for *thus* in the cataphoric and in the anaphoric uses of these combinations in Chaucer[7] (cf. (10)).

(9) *… for þus suche tormentes þou schalt somtyme se me wyth sayntes in blis*
                                    [St. Kath. (3) (Richardson 44) 45]
     "For such tortures you will see me one day with saints in beatitude"

(10) a. *Ther been ful fewe whiche that I wolde profre To shewen hem **thus much** of by science*   [Chaucer CT.CY (Manly-Rickert) G.1124]
          "There are but few to whom I would proffer to make my science clear and evident [*lit.* so much of my science for them]"
      b. ***Thus muchel** wol I seyn: that whan thou prayest …*
                                    [Chaucer CT.Pars. (Manly-Rickert) I 1043]
          "So much I will say: that when you pray … "
      c. *Al this seith seynt Bernard, and **thus myche** of these vertues at this tyme*
                                    [Love *Mirror* (Brsn e.9) 49]
          "Saint Bernard says all this, and so much of these virtues at this time"

Comparative reconstruction makes it plausible to assume that the English demonstratives of manner, quality and degree (MQD-demonstratives) largely lost their exophoric and gestural use in the course of their historical development (cf. König 2015), with a few exceptions like the degree use of *swā > so* (*The fish was so long* + gesture). Moreover, even that use tends to be replaced by the basic demonstratives *this/that* (*Was it really this long*?), just as for the manner and quality uses. The question we need to ask at this point, however, is whether there is any direct evidence for an exophoric use in the early historical stages of English. It is generally assumed on the basis of evolutionary and ontogenetic arguments that exophoric (gestural) uses are basic for demonstratives,[8] but such uses are difficult to find and to clearly identify in the available, predominantly narrative, texts of Old English. The following Examples (11a)–(c), all of them direct speech in the present tense, seem to be cases in point, however: in (11a) a question is uttered by a donkey in a co-text explaining the situation when he was beaten *swyðe* (hard) by Balaam, that is, *swa* is exophorically referring to the manner of the beating. Similarly in (11b), the sentence is a direct question addressed to Satan, plausibly with a gesture pointing at a light in front of their current location (*her* "here"), and (11c) is an unquestionable case of directly quoted command.

---

7.   A valuable help for the analysis of the uses of *thus* in Chaucer was Oizumi (1991–2012).

8.   The loss of an exophoric use of the original manner demonstratives is also attested for Romance languages (Italian *si*, French *si*). In those cases the relevant forms were not replaced by local demonstratives, however, but were reinforced by presentative demonstratives in certain uses (Italian *ecco + si > così*; French *accum si > ensi > ainsi*).

(11) a. *Hē cwæð: "Hwī bēatst ðu mē **swā** swȳðe?"* [Num 22.28]
"He said: "Why do you beat me so strongly?""

b. *Hwæt is þiss liht þæt hēr **swā** fǣrlice scīnð?* [Nic (C) 193]
"What is this light that all of a sudden shines here so?"

c. *Ic sylf am nū   ān mon on anwealde iset, and ic habbe under mē*
I self am now a   man on power   set and I have under me
monigæ cnīhtæs on fare,   and ic cwæðe to ðissum, Fār ðu, and
many   knights on journey and I say   to this.one go you and
he fǣrð, alswā eft   to ōðrum,   Cum þu, and he cymæð sōnæ,
he goes also again to other.one come you and he comes at.once,
*and to mīne <ðēowe>, Dō **þus**, and hē dēþ*
and to my   servant   do thus and he does
"I myself am now a man set on power and I have under me many knights
on expedition and I say to this one "You go!" and he goes, so as to the other
"You come" and he comes at once, and to my servant "Do (like) this" and
he does" [ÆHomM 6 (Irv 1) 112]

d. ***So, thus**, keep her **thus**!* [Sturmy, Mariner's Mag. i.ii.18][9]

Instances of such exophoric use are still found in ME and EModE texts, as witnessed by cases such as (11d): the text is a sort of 'diary' where the author reports on some of his journeys to a hypothetical addressee 'you' to whom the sailing techniques are described as if he were on that ship. In this instance, the author is describing how the addressee should keep the rope. In the surrounding text there are no clues about the manner in which the rope should be kept.

4. OE *swā* as a source of grammaticalisation: Some frequent patterns

Before we examine the historical changes in detail, let us take a look at the targets of these changes. A detailed look at the uses of *so* and related forms in Modern English (*as, also, yes*, etc.), where almost 30 different uses are generally distinguished for *so* in major dictionaries,[10] reveals patterns of change that are typical instances of grammaticalisation: they involve extensions in the possible contexts of use and decategorialisation, i.e., a loss of their basic categorial properties. They are unidirectional and involve concomitant changes in both the syntactic and the semantic domain.

---

9. Taken from *THE Mariners Magazine; OR, STURMY's Mathematical and Practical ARTS. The Practick Part of Navigation, in working of a Ship in all Weathers at Sea* available at <https://quod.lib.umich.edu/e/eebo2/A61915.0001.001/1:19?rgn=div1;view=fulltext>.

10. 34 in the OED, 21 in Samuel Johnson's *Dictionary of the English Language*, 26 in the *Collins Cobuild Advanced Dictionary*.

Moreover, many of the relevant changes are found in a wide variety of languages inside and outside of Europe, thus revealing general tendencies of language change (cf. König 2012, 2015; König & Nishina 2015). It is, in particular, the semantic rather than the formal changes that exhibit a remarkable degree of generality or overlap, since they can be shown to lead to the same overall semantic result, even if the forms may be somewhat different. The development of approximative markers from manner demonstratives provides a simple example. Latin *qua-si*, which has also been borrowed by many European languages (e.g., French), is composed of the ablative form of *quae* "which, as" and the older Latin form *si* "so", which was later reinforced to *sīc*.[11] In English and German the modern counterparts of the old demonstratives are used for this meaning, though with different syntactic properties:

(12)   a.   ENGLISH There are fifty students or **so** in this room.
        b.   GERMAN *In diesem Raum sind* **so** *(ungefähr) fünfzig Studenten.*
        c.   FRENCH *Les raisins sont qua**s**i mûrs.* "The grapes are almost/practically ripe."

Figure 1 presents an overview of the changes frequently attested in European languages, using primarily examples from English.[12]
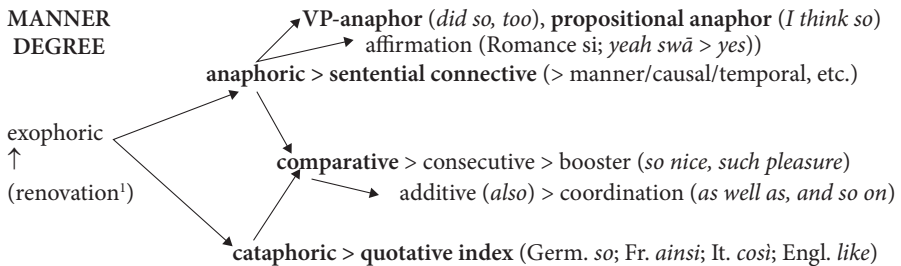
Figure 1. A-typical demonstratives as sources of grammaticalization:
The case of *so* (König 2012: 23)

---

11. The Online Etymological Dictionary (https://www.etymonline.com/). The assumption sometimes found in etymological dictionaries of Latin that "if" is the older use of *si* in Latin (*Si veniat, me videat.* "If he were to come, he would see me") is highly implausible. There is no path of grammaticalisation leading from 'conditionality' ("if") to 'manner' and 'degree'. Only the opposite direction provides an explanation for the development of *si* to a marker of affirmation (Ital. *si*), to a marker of degree (Fr. *si*), to a marker of conditionality (Fr. *si*, Span. *si*, Ital. *se*) and to a marker of manner (Ital. *così*). Moreover, in Germanic languages (German, English), too, demonstratives of manner (*so*) are used as conditional connectives.

12. By "renovation" in Figure 1 we mean a new construction, such as *like this/that*, replacing *so* in its exophoric use.

This overview starts from the plausible assumption that the basic use of demonstratives is an exophoric one, a use in which they are accompanied by gestures. In this basic use demonstratives can be assumed to provide a transition between systems of communication largely based on gestures and elementary uses of verbalisation (cf. Arbib 2012). Brugmann (1904) even regarded these basic lexical elements as the primary source in the development of nearly all grammatical categories.

It was presumably in the course of developing textual and discourse functions that the language extended the use of demonstratives from the exophoric to endophoric uses, i.e., to anaphoric (retrospective) uses relating to preceding stretches of discourse and to cataphoric (prospective) uses relating to subsequent stretches of discourse. Note that MQD-demonstratives are atypical members of this class, since, unlike *that* or *there*, they do not directly identify the entity pointed at, but an entity or subset of entities similar to the one pointed at, e.g., approximations to the length of a fish in *The fish that I caught yesterday was so long* (+ gesture) and approximations to a perfect manner of dancing in the German example *Den Zumba tanzt man so* (+ enacting gesture) "This is how you do the Zumba" (cf. Umbach & Gust 2014). The typical gestures accompanying them are indeed mimicking (enacting) rather than pointing ones. Most of the additional changes and processes of grammaticalisation are based on the anaphoric and the cataphoric uses.

Below we take a closer look at some of the branches of the chains of grammaticalisation summarized in Figure 1, focusing on *swā/so*, and discussing more briefly *þus/thus* and *swylc/such*. The three MQD-demonstratives *swā/þus/swylc* overlap in their functions and meanings in Old English texts – they can all be used as intensifiers, as in (13) – but get specialised over the course of time.

(13) *Ða   cwæð Petrus: 'Sege mē, becēapode gē   **ðus** micel lands?' Hēo*
      then said Peter say me sell   you thus much of.land she
      *andwyrde: 'Gēa, lēof, **swā** micel'*   [ÆCHom I, 22. 358.97]
      answered yes sir so   much
      "Then Peter said: "Tell me, did you sell so much land?" She answered "Yeah, sir, so/this much""

For reasons of space, we cannot cover all the changes identified in Figure 1, but are forced to select what we think are the most relevant developments. Our selection of changes starts from the initial developments of the exophoric uses to the anaphoric and cataphoric ones and proceeds to their additional differentiations. It includes processes not found in other Germanic languages (the use of *so* as propositional anaphor or VP-anaphor: § 4), but also processes that are widely attested in European languages (the development of equative comparatives: § 5). We introduce as well two other significant developments, the change of *swā/so* to quotative markers and to adverbial connectives, but we leave aside the additive and affirmative uses.

Given a certain imbalance in the quantity of our data,[13] we will primarily focus on the development of *swā*, without completely neglecting the related expressions *þys/þus* and *swylc*.

5.   From exophoric to anaphoric

5.1   VP-anaphora *do so*

Anaphoric uses can be found for all demonstratives. What kind of stretches of preceding discourse and what kind of entities are picked out by the anaphor, however, depends on the demonstrative in question. Anaphoric *there* picks out a location; *this* and *that* relate to objects (*I did not want that*), but also to propositions (*I don't understand that*). Modern English *so* may relate to a verb phrase (cf. (14a)), but also to a complete proposition, expressed by an embedded sentence. The following examples show that the former use can also be found in Old and Middle English, but differs from its modern use in a few interesting aspects: in Modern English only *do* is admissible as an auxiliary verb preceding *swā/so*, whereas in Old and Middle English both modal verbs (cf. (15)) and the verb *have* (cf. (16)) may occur in these constructions. As far as the antecedents are concerned, Modern English requires action predications. This is the most frequent case in Old English and Middle English too. Nevertheless, in Middle English stative predications can often be found as antecedents (cf. (17)). It is interesting to note that in Old English *þus* also appears to have an analogous anaphoric function, although limited in frequency and restricted to the verb *do* and action predicates, as will be shown below. In this configuration *swā/so* and *þus* continue to alternate freely throughout the entire Middle English period.

(14) a.   John writes papers in the library and Fred does **so** at home.
      b.   *Se cing hēt   hī   feohtan agien Pihtas, and hī   **swā** dydon.*
            the king ordered them fight   against Picts and they so   did
            "The king told them to fight against the Picts. And they did so"
            [ASC 12 (449)] (Schleburg 97ff.). **ACTION VERB AS ANTECEDENT**

(15) *wolde   frēadrihtnes feorh ealgian mǣres þēodnes ðǣr   hie   meahton **swā***
      wanted lord's   life defend famous prince there they could   so
      "they wanted to defend the lord's life, the famous prince's, where they could do so"   [Beow 796–797] **MODAL VERB IN VP-ANAPHOR**

---

13. This imbalance is partly due to the absolute frequency of the different items, as a result of a different range of meanings and functions. There are about 15,000 occurrences of *thus* and of *such* and about 145,000 occurrences of *swa/so* in the Corpus of Middle English prose and verse.

(16)   *Gē   secgað foroft,   þæt Petrus, se   apostol, hæfde wīf   and cīld;   and*
      you say   very.often that Peter   the apostle had   wife and child and
      *wē ēac   secgað, þæt hē   **swā** hæfde*
      we also say   that he so   had
      "you say that Petrus had a wife and a child, and we also say that he did"
                            [ÆLet 2 (Wulfstan 1) 84] HAVE-PREDICATION IN ANAPHOR

(17)   *Alle solde crystende bee, And thay that wolde noghte do **swaa**, He wolde thame*
      … slaa.
        [Isumb.(Thorton) 693–695] (≠ ModE) BE-PREDICATION AS ANTECEDENT
      "Everyone should be Christian. And he would kill those who did not want to
      do so"

Overall, we can say that what started out as a productive replacement of a preceding predication (VP) by a combination of an auxiliary verb combined with the anaphoric use of *swā* (> so) became more restricted in its syntactic properties and more tightly organised as a construction.

5.2   *So* as a propositional anaphor

In contrast to the use as a VP-anaphor, the use of *so* as 'propositional anaphora' is not found until ME (cf. (19)). In Modern English this use is found after verbs expressing propositional attitudes, such as *think, suppose, imagine, guess, hope, expect, say, believe*, etc. In combination with a few verbs, the anaphor is in contrast with the equally acceptable anaphor *it* (cf. (18)):

(18)   a.   She said it. – She said so.
      b.   I believe it. – I believe so.
      c.   I expect it. – I expect so.

The basic meaning of *swā* and *so* as expressing similarity in manner, quality or degree is also visible in such minimal pairs. Whereas the first sentence in (18a) relates to a *verbatim* report, the second one reports on the basic content of an utterance. The use of demonstratives of manner as VP-anaphors or propositional anaphors is not exclusively manifested by *swā-so*, but is frequently also found for *þus*[14] (cf. (19f)). This is even more significant if one keeps in mind that the use of manner demonstratives as propositional anaphors is a rare phenomenon in European languages and is not even found in closely related Germanic languages.

---

14.   *Þus* as sentence anaphor tends to collocate with *thinken* in Middle English almost as if they constituted one formula.

(19) a. A. The meeting has been postponed? B. I suppose/assume/think/guess/
      imagine … so.   (CONTEXT: verbs of propositional attitude)

  b. *Andswarast þu   swā ðām   bisceope?*
      answer   you so   to.the bishop
      "Do you answer like that to the bishop?"
                                    [Jn (WSCp) 18.22] (BRIDGING CONTEXT?)

  c. "*A! goode sir*," quod I, "*say not soo!* [Chaucer BD (Benson-Robinson) 714]
      ""A, good sir", I said "do not say that / do not speak like that""

  d. *I tolde hym so, & euer he seyde nay.*   [Hoccl. RP (Hrl 4866) 717]
      "I told him so/that and he always said "no""

  e. *Offreþ þe lombes of Innocensye, ffor he comaundet so.*
                                          [Psalt.Mariae (1) (Vrn) 223]
      "[He] offers the lambs of innocence, for he ordered so"

  f. *Now hoope ye not, hynde fader, ne in hert thinke, / That I carpe **thus** for
      cowardys … Or for the sake of my selfe in sauyng alone*
                                          [*Destruction of Troy* 2292][15]
      "Now don't think, tormented father, nor believe in your heart, that I say
      so/that for cowardice or for the sake of saving myself only"

German sentences like (20a)–(b) look similar to English, but the manner demonstrative *so* has not really acquired the grammar and meaning of an object, which is found elsewhere in these sentences. Its syntactic position and interpretation is more like that of a manner adverb:

(20) a. *Die Koalition ist nicht stabil. So hat sich der Minister jedenfalls ausgedrückt.*
      b. *Die Koalition ist nicht stabil. Der Minister hat sich jedenfalls so ausgedrückt.*
          "The coalition is not a stable one. The minister thought so anyway."

Such a comparison between English and other Germanic languages supports our hypothesis according to which the use of *swā* as propositional anaphor is a further development of the anaphoric use of that expression as manner adverb: that is, sentences like (19b) can be regarded as bridging contexts between the use of *swā* as manner adverb and as propositional anaphor. The same holds good for the occurrences of *þus* as propositional anaphora as in (19f), where *þus* could be both the object of the proposition and an adverb expressing the way the action is performed. Unlike *swā, þus* never seems to develop any further. The German example could be analysed as manifesting an intermediate step in the direction of the structures possible in English.

---

15. For the complete text, see <https://quod.lib.umich.edu/c/cme/APE7380.0001.001?rgn=full+text>.

5.3   *So* as a sentence connective

The use of *so* as sentence connective in PDE,[16] as in (21), is based on its anaphoric use, in that *so* relates to a preceding part of a sentence or discourse. From a vague connective function with a possible interpretation of 'manner', still visible in uses of type (21b)–(c), the demonstrative has developed further uses expressing a variety of rhetorical relations. In (21d), this basic meaning is enriched to a causal one, a process of pragmatic or semantic enrichment as common as the one from temporal sequence to causality on the basis of the fallacy 'post hoc ergo propter hoc' or from simultaneity to concessivity (cf. König 1988). Manner and degree demonstratives used as sentence connectives are frequently attested in OE texts, but only indicate clausal connection in a non-specific, vague way. Additional semantic differentiations (Boulonnais 2006), such as those evident in (21e)–(g), are not well established until the ModE period.

(21)   a.   And so Christmas passed. VAGUE INDICATION OF A CONNECTION
       b.   He went into lower gear, so (that) his car would slow down/slowed down. PURPOSIVE/RESULTATIVE
       c.   He is very sick. Even so he goes to work. CONCESSIVE
       d.   I did not like it. So I wrote to him. CAUSAL
       e.   So you are a linguist, eh? INFERENTIAL (cf. Blakemore 1988)
       f.   So, how did they vote in your constituency? (INTERACTION-INITIATING USE, INCIPIENT MARKER OF DELAYED MOVES TO BUSINESS (cf. Bolden 2008)
       g.   So what? ASKING FOR THE RELEVANCE OF A PRECEDING UTTERANCE

The current polysemy of the connective *so* starts to emerge in Middle English texts (cf. (22)), first the temporal and causal meaning (already quite frequent at the beginning of the 13th century, as shown in Laȝamon's *Brut* examples in (22b)–(c), and later the conditional and purposive meaning, attested from Gower or Chaucer onwards, as in (23a)–(d).

---

16. We labelled as PDE those phenomena which are attested since the 1960s, following Smitterberg (2005: 8), according to whom "Present-Day English refers to the period from 1961 onwards, the year in which texts that make up the Brown and LOB corpora were published". Consequently, ModE refers to the period beginning in 1700.

(22) a. *Þa   agann se   arcebiscop Landfranc atywian mid openum*
          then began the archbishop Landfranc show   with open
          gesceade. þæt he mid rihte crafede   þas þa   he crafede   ⁊
          distinction that he with right requested that which he requested and
          mid   strangan cwydan þæt ylce   gefæstnode toforan þam papan
          with strong   speech that same confirmed before the   pope
          Alexandre. ⁊   toforan eallan þam concilium þe   þar   gegadered was.
          Alexander and before all   the   council   that there assembled was
          ⁊   swa ham   foran.
          and so   home went
          "Then began the Archbishop Landfranc to show with clear distinction,
          that what he craved he craved by right; and with strong arguments he
          confirmed the same before the Pope Alexander, and before all the council
          that was collected there; and then/so they went home."
                              [ChronA (Bately) 1070.24] VAGUE ADVERBIAL MEANING

      b. *Heo letten lude   clepian … þat Brutus þe sele   to þare sæ wolde;*
          they let   loudly proclaim that Brutus the good to the sea wanted
          Swa heo ferden to heora scipa.
          so   they went   to their ships
          "They caused it to be loudly proclaimed that Brutus the good would go to
          the sea. So they proceeded to their ships"
                                          [Lay. Brut (Clg A.9) 888] TEMPORAL/CAUSAL

      c. *Utlaȝen hefden i-ræued þat lond … swa hit wes al   west.*
          outlaws had   ravaged the land   so   it   was all wasted
          "Outlaws had ravaged the land … and so was it all wasted"
                                                  [Lay. Brut (Clg A.9) 564] CAUSAL

(23) a. *He had y-swore to hire … Thar so she saved hym hys lyfe, He wolde haue*
          take hir to hys wife.   [Chaucer H. Fame i. 423] CONDITIONAL
          "he had sworn her then that if she had saved his life, he would have taken
          her as his wife"
      b. *That can I do wel, Be so my lif therto wol laste.*
                                              [Gower CA (Frf 3) 1.187] CONDITIONAL
          "I can do that well, provided that my life would last thereto"
      c. It is no matter how dirty a bag it is conveyed to him in, … **so** the money
          is good.   [OED, 1750; s.v. *so*] CONDITIONAL, NECESSARY CONDITION
      d. *Thanne seide I to my-self, so Patience it herde*
                                              [Lang. P.Pl. B. xii. 64] PURPOSIVE
          "then I said to myself, in order that Patience heard it"

As these examples show, *swa* and *so* provide the source of a variety of adverbial meanings enriching a basic and vague meaning 'manner' ("in this/the way"). The temporal use is simply the consequence of the linear order of events (cf. (22b)–(c)) since, as is well known, there is a close semantic connection between 'manner', 'temporal sequence' and 'causality'. The development of a conditional use of manner demonstratives is a well-known phenomenon in Romance languages (e.g., French/ Spanish *si*, Italian *se*) and in German (*so*). In English this development is more restricted than in the other languages mentioned. Conditional *so* in ME and later centuries typically and perhaps exclusively expressed necessary conditions, i.e., conditions signaled by "provided that" in Modern English (cf. (23c)). The example in (23a) could be one remarkable exception. The inferential use, which is not attested in earlier stages of English, is based on a reversal in the order of cause and consequent (q, because p $\implies$ p, so q). The concessive meaning of the adverbial *even so* is the compositional result of the 'manner' use of *so* combined with a scalar use of *even* that developed in Modern English from a use still attested in Shakespearean plays (roughly paraphrasable by "precisely", "exactly", "indeed"; König 1991: 135). This use and its interpretative enrichment to concessivity can still be found in the combination *even as* in constructions like the following and more generally in all constructions expressing remarkable simultaneity or concomitance. Concomitance and, more specifically, simultaneity of two situations is always remarkable when the two situations co-occur that do not normally go together (cf. König 1988: 157ff.).

(24) a. What you will have it named, even that it is.

[Sh. *Taming of the Shrew*, III.iv]

b. Even as it admits of a serious pollution problem, the country is substituting cheap brown coal for imported oil.

Additional manifestations of a connective use of *so* are frequently discussed in the framework of conversation analysis and relevance theory as markers of specific interactional moves (cf. Johnson 2002; Raymond 2004). In Bolden (2008), the initial use of *so* in examples like (21f) is analysed as a marker of 'delayed moves to business' and in Bolden (2009) *so* is characterised as a marker of 'emergence of incipiency'. Such 'pragmatic uses', as they are often called, are far removed from the earlier stages of English discussed in this paper and it is by no means clear how many such uses need to be distinguished and what the relevant generalisations are.

6.   From exophoric to cataphoric

6.1   Comparative constructions

In their typology of equative comparative constructions, Haspelmath & Buchholz (1998) and later Haspelmath & the Leipzig Equative Constructions Team (2017) distinguish six major types on the basis of their language sample. The type that is particularly widespread in Indo-European languages uses degree demonstratives like ModE *as/so* or German *so* as grammatical degree markers and has the following basic structure:

> (25)   a.   COMPAREE – (degree marker) – parameter – standard marker – STANDARD
> b.   Bill is as old as George.
> c.   George is as old as Bill. Fred is not so/as old as Bill.
> d.   [German] *Georg ist so alt wie Willi.*

The cognate constructions in Old English are of the same basic type and use *swā* (to a lesser degree also *swylc*,[17] though never *þus*) both as degree marker and as standard marker. This pattern continues to be attested in some ME texts, although the standard marker is gradually replaced by *as* over the course of time.[18] In combination with adjectives, *swā* expresses similarity of degree both in Old English (cf. (26a)–(c)) and in Middle English (cf. (26d)). In verbal contexts, by contrast, it is similarity of manner or quality that is expressed by this demonstrative (cf. (27)). The relevant marker in Modern English (*as*) derives from a reduction of the combination *eall swā* (*eal swā > as*). In nominal contexts, *swylc* is also found and the construction resembles a relative clause, just as its ModE counterpart (cf. (28)):

> (26)   a.   *þa   wearð   Tiberius Romanus* **swā** *wrāð and* **swā** *heard* **swā** *he*
> then   became Tiberius Romanus so   angry and so   hard as   he
> *ǣr   wæs milde and īeþe.*   [Or6 2.134.20] DEGREE
> earlier was mild and kind
> "Then Tiberius Romanus became as angry and as hard as he had been mild and kind before"

---

17.   Cf. *Swylc scolde eorl wesan, æþeling ærgod swylc Æschere wæs* [Beow 1328] "so ought a man be experienced and noble as Æschere was"

18.   *As* can be used as standard marker even in such comparative constructions until the 15th century as *I hadde never more neede for to have help of my goode, as I have at this tyme* [Paston 3.241]. Used as a degree marker, *as* is not frequent in ME, except for the pattern with indeterminate quantifiers, such as *many* and *much*.

b. *Hit wæs þā   swā leoht swā se   merȝenlica steorra*
   it   was then so   light so the morning   star
   "It was as light as the morning star"   [LS 20 (AssumptMor) 12] DEGREE

c. ***Swā oft   swā** Galli wið   Romanum wunnan, swa wurdon*
   so   often so   Gauls against Romans   fought   so   became
   Rōmāne gecnysede.   [Or 3 11.77.3] DEGREE
   Romans overcome
   "As often as Gauls fought against the Romans, [so] the Romans were overcome"

d. *þe arewes come **so** þykke **so** reyn*
   the arrows come so thick so rain
   "The arrows come as densely as rain"
                          [Mannyng *Chron.Pt.1* (Lamb 131) 6827] DEGREE

(27) a. *Ðā   woldon hī   dōn, **swā swā** hī   ǣr   dydon.*
        then wanted they do   so   so they before did
        "Then they wanted to act as they had acted before."
                          [ÆCHom II, 38.283.109] SIMILARITY OF MANNER

b. *Hē him gehēt   his ǣriste **swā** hē þā   mid sōðe gefylde gelīce*
   he them promised his rising so he then with truth fulfilled just
   ***swā** hē ǣr   þā þrōwunge dyde*
   so   he earlier the passion   did
   "He gave them the promise of his resurrection, which he truly performed, as he had also done with his martyrdom"
                          [HomS 8 (BlHom 2) 34] SIMILARITY OF MANNER

c. ***Swā** hit is **swā** þu   sæȝst*   [Bo 39.126.29] SIMILARITY OF QUALITY
   so   it is so   you say
   "It is [so] as you say"

(28) *for þanne he wile ðere   gelden elch man his hwile mid **swilch** mede*
    for then   he will there pay   each man his time with such   reward
    ***swo**[19] he ernede here.   [*Trin.Hom.*(Trin-C B.14.52) 5] RELATIVE MARKER
    so   he earned here
    "For then he will there requite each man his time with such reward as he earned here"

Our view that the use of *so* in equative constructions derives primarily both from a cataphoric and an anaphoric use of *swā* is based on the fact that the constituent expressing the standard follows the demonstrative and provides new information. On the other hand, the referent of the standard must be known to the interlocutor

---

19. In correlative constructions with *such* and with a function similar to that of a relative pronoun, *as* had already replaced *swā/so* by the time of Chaucer.

in order to be relevant. A sentence like *Your son is as intelligent as Pope Gregory III* does not make communicative sense if the interlocutor does not know the personality denoted by the standard, in this case Pope Gregory III.

Special mention must be made of the example in (26c). Here we find a comparison between two situations encoded as comparee and as standard along a dimension of frequency. In such constructions, the combinations of comparative 'marker + parameter + standard marker' have developed into complex conjunctions, which in ModE may introduce a subordinate clause preceding a main clause. The relevant development is also manifested in spatial and temporal expressions like *as far as, as much as, as recently as, as late as, as soon as*, etc.[20]

(29) a. *Þā   wearð   eft   Godes wracu   on Rōme:* **swā longe swā** *sēo*
          then happened next God's vengeance on Rome so   long so   the
          ēhtnes   wæs þāra cristenra monna, swā longe him   wæs
          persecution was of.the Christian people   so   long   them was
          ungemetlic moncwealm getenge, þæt nān hūs   næs   binnan þǣre
          immense   slaughter   nearby that no house not.was inside the
          <byrig> þæt hit næfde   þǣre wrace angolden.   [Or 6 23.144.13]
          city   that it not.had   the rage paid.for
          "then God's rage was on Rome: as long as there was the persecution of the
          Christians, so long the immense manslaughter was oppressive to them,
          that there was no house within the town that hadn't paid for that rage"
      b. *Secgað me* **swā sōna swā** *gē   oncnāwaþ þæt he cucu ne byð*
          tell   me so   soon so   you understand that he alive not is
          "Tell me as soon as you understand that he is not alive"
                                                              [ÆLS (Vincent) 48.185]

For reasons of space many other developments of *swā/so* indicated in Figure 1 cannot be discussed in detail and are only briefly mentioned. The development of additive markers (*also, as well as*; German *ebenso, genau so*; Fr. *aussi*; Swed. *också*) is a widespread phenomenon in European languages, based as it is on the notion of 'similarity'. Equally widespread is the development of a scalar (intensifier, booster) use of *so* and of *such* (*It was so kind of you! It was such a treat!*), which is also manifested by the counterparts of the English expressions in other Germanic and Romance languages. The exact steps in the development of these uses are not perfectly clear. A direct path leading from the exophoric use to the identification of an unexpectedly high value or a reduced version of a resultative clause (*She was so kind that you could hardly believe it! > She was so kind!*) are the two possibilities most discussed in the literature.

---

20. Note that in Middle English the first *swā* could be replaced by *þus: Hue is asoilid **þus** sone ase hure lef lykeþ* [PPl.C (Hnt HM 137) 4.181]. "She is absolved as soon as she wishes"

6.2   The quotative use of *swā*

The use of MQD-demonstratives as introductory markers of direct and indirect speech is a widespread phenomenon in European languages (cf. Buchstaller & van Alphen 2012). The following two examples illustrate this quotative use for German *so* and French *ainsi*, which is clearly based on the cataphoric (prospective) extension in the use of demonstratives.

(30)   [German] *Angela Merkel hat das so formuliert. „Wir müssen Theresa May helfen.“* *„Wir müssen“, so Angela Merkel, „Theresa May helfen“.*

(31)   [French] *Angela Merkel s'est exprimée ainsi: 'Il faut que nous aidions Theresa May'.*
"Angela Merkel put it like this: "We have to help Theresa May.""

An analogous use of *swā* is also found in Old English:

(32)   **Swā** him   Crīst   bebēad   þæt hīe   mancynn lǣrden &   **swā** cwæð:
so   them Christ ordered that they mankind teach   and so   spoke
Euntis in universum mundum …'   [HomS 1 (ScraggVerc 5) 49]
"So Christ commanded them that they should teach mankind and [thus] said this/the following: Euntis in universum mundum … "

In OE texts it is possible, though uncommon, to come across *swā* introducing direct speech, but *þus* is clearly the preferred option to such an extent that it can be regarded the quotative marker par excellence (see below). According to the DOEC, *verba dicendi* occur with *swā* in reference to something previously said only in the pattern *swā swā X cwæð* "so as X said", where the verb of speaking is often *cweðan*, though not necessarily. If a manner demonstrative is used to introduce direct speech, *swā* manifests certain restrictions: it is limited to a particular author – for instance, Wulfstan, who alternates freely between *swā* and *þus* – or limited to a manuscript – for example, the homilies in the Vercelli Book present both *swā* and *þus* as quotative markers. Moreover, *swā* as quotative maker is restricted to the verb *cweðan*. As a matter of fact, in the DOEC there are no instances of other possible collocations. The use of *þus*, by contrast, is less restricted, being not limited to any author, nor to specific manuscripts or to predicates. The quotative use of *thus* is still found in the translation of Nietzsche's well-known book *Thus spoke Zarathustra*.

In Modern English, this quotative use of *so* or *thus* has disappeared and been replaced by two expressions which spell out the two components of the demonstratives ('manner' + 'deixis': *like this*), as is shown in the translation of (33). In recent developments of colloquial English both *like this* or *like* (+ vocal gesture) are used (see also d'Arcy 2017):

(33)   And I'm like … And he's like …

7.   The development of *thus* and *such*

7.1   *Þus > thus*

Due to the availability of the rich and specific collections of data mentioned above (Nummenmaa 1973; Schleburg 2002) the focus of the preceding discussion was on the development of *swa > so* in the history of English with only occasional glimpses of analogous or similar developments in the two other MQD-demonstratives. It is now time to present a more coherent and comprehensive analysis of the development of *þus* in this section and of *swylc/swelc* in the one following. Again, we will begin with an overview of the processes of change manifested by *thus* in the course of its historical development (Figure 2).

MANNER

(marginal)   **anaphoric > manner > causal/reason** > elaboration
    exophoric
                          cataphoric > quotative index

DEGREE   (spatial > temporal) degree (*thus far*)
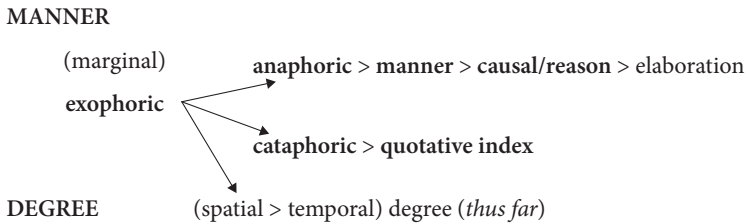
Figure 2. A-typical demonstratives as sources of grammaticalization: The case of *thus*

As pointed out above, the three demonstratives under analysis overlap to a certain extent in their distribution and meaning and acquired their specialisation and differentiation in the course of their historical development. Even in Samuel Johnson's *Dictionary of English* (1799) we still find that one expression is paraphrased by the two others. This overlap is particularly large in the case of *swa* and *þus*: both expressions can be categorised as adverbs (adverbial demonstratives) and they both can express the content dimensions of manner and of degree. Given its properties discussed so far and those of its counterpart in ModE, there can be no doubt that *þus* belongs to the system of demonstratives in OE: the adverb occurs in coordinate structures (cf. (5)– (7)),[21] it manifests the anaphoric and cataphoric uses characteristic of deictic expressions, and it must have also been used exophorically, as confirmed by such possible (although marginal) occurrences of its descendant in ModE as the expressions *thus far* or the sentences like *Fold the sheets thus* + gesture and *The investigation is thus concluded*. Just like *swā*, *þus* combines with quantifiers (*much*) and may even combine with the MQD-demonstrative *such* (cf. (9)–(10)).

---

21. In regional varieties of ModE we also find the following combinations: *thus and such, thus and so* (cf. Webster's *New World College Dictionary*, s.v. *thus*).

As far as *thus* in ModE is concerned, as few as two (Collins Cobuild) or as many as four (Merriam Webster) use types are distinguished in major dictionaries: manner, causal/reason/consequence, deictic specification of extent or degree, and rhetorical relation of elaboration. These are illustrated by the following examples.

(34) a. Stated thus, the problem seems trivial. ANAPHORIC
 b. The judge expressed it thus: "Our obligation is to discover the truth." CATAPHORIC
 c. The demand was high and thus prices increased. CONSEQUENCE
 d. Thus far our results have been quite promising. (IDIOM, up to now)
 e. Few of the nation's largest cities are state capitals; thus neither New York nor Chicago is the seat of its state's government. ELABORATION

If we distinguish between an anaphoric and cataphoric manner use we end up with five uses. On the other hand, one of these uses (*thus far*) is a frozen idiom and may not deserve a separate entry. The somewhat formal quality of some of the other sentences with *thus* are a clear indication of changes that are reducing this expression to a marginal position in the demonstrative system of English.

In older stages of English, by contrast, the position of *þus* was by no means marginal. Besides expressing manner, *þus* in Old English could anaphorically refer not only to a manner phrase in the preceding text, but also to a complete verb phrase, thus overlapping with *swā*:

(35) a. *For þām bysmrað se unrihtwisa Drihten; for ðām hē cwyð on his*
 for that mocks the unjust Lord for that he says in his
 *mōde, Ne recþ God, þēah ic **þus** dō*
 mind not cares God though I thus do
 "This is why the unjust mocks God, because he says in his mind: God does not care although I act like this" [PPs (prose) 9.33] VP-ANAPHOR
 b. *I haue brought the a childe, and I pray the that thow do it norisshe. And.yef thow do **thus**, grete gode shall come to the and thyn heyres*
 [Merlin (Cmb Ff.3.11) 91] VP-ANAPHOR
 "I have brought you a child and I pray you to nourish it. And if you do so, great wealth will come to you and your heirs"

In Old English the instrumental origin of this expression is still obvious in that it mainly occurs as a manner demonstrative denoting the way in which or the condition under which an action is performed (36a). This is still the most frequent use of the connective *thus* even in late Middle English (36b). Accordingly, in Old English texts there is no attestation of *thus* as a sentence connective. In Middle English, however, *þus* could also be used in this way (36c)–(e).

(36) a. *Wē endiað nu   þisne cwide **þus**, þanciende ðām ælmihtigan ealra*
      we end   now this speech thus thanking the almighty for.all
      his gōdnissa,   se ðe   æfre rixað on ēcnisse, AMEN
      his goodnesses he who ever reigns in eternity amen
      "Now we end this discourse in this way, thanking the Almighty for all his
      goodness, who always reigns in eternity. AMEN" [ÆJudgEp 88] MANNER

   b. *He dide do take Sidrak And bounde hise hondes And kepte him Eight dayes*
      In grevous torment. **Thus** [=in this way] *Sydrak in prisoun lay, Til it was þe*
      nynthe day   [Sidrak & B.(Lnsd 793) 501] MANNER
      "He took Sidrak and tied his hands and kept him eight days in great tor-
      ment. Thus/under these conditions Sidrak lay in prison for eight nights."

   c. *My maistre Chaucier saide þat þe fader Beqwaþe no-thing his vertue with*
      *his hous Vn-to his sone. By avncetrye **þus** may yee no-thing clayme.*
             [Scogan MB (Ashm 59) 97] MANNER / CAUSAL CONNECTIVE
      "My master Chaucer said that the father did not bestow his house as a
      legacy to his son and thus/therefore you may not claim it on the basis of
      ancestry."

   d. *Þereas pouert enpresses, Þaȝ mon pyne Þynk, / Much, maugre his mun, he*
      *mot nede suffer; / **Thus** pouerte & pacyence arn nedes playferes.*
             [Patience (Nero A.10) 43–45] MANNER / CAUSAL CONNECTIVE
      "Where poverty oppresses a man, although he may hate to admit it, he must
      necessarily suffer a lot. Therefore/Thus poverty and patience are playmates
      through necessity."

   e. *But þe comyns of engliȝschmen knowen it best in þer modir tunge, & **þus** it*
      were al oon to lette siche knowing of þe gospel …
             [Wycl.OPastor.(Ryl Eng 86) 430] MANNER / CONSECUTIVE
      "But the communities of English men know it best in their mother tongue,
      and thus/consequently it would be just as good to permit such [viz. Pater
      noster] knowledge of the Gospel"

Unlike *swā*, whose meaning can unambiguously be temporal, causal or conditional
at a clause level, *þus* continues basically to indicate manner in all its Middle English
occurrences as well as in ModE. Temporal, consecutive, casual or other meanings
are only contextually inferable, but rarely become lexicalised (36c)–(e). In Middle
English prose text, however, it is undoubtedly the case that *þus* specialises as the
narrative connector par excellence in collocation with *and* (cf. (37)).

(37) a. *Đa   syððon   þurh   heora gemelest 7   þurh   heora*
then afterwards through their negligence and through their
druncenhed on ān niht forbærnde þa cyrce. 7   eall þet
drunkenness on one night burned.down the church and all that
*þær innæ wæs. **Đus wæs se mynstre of Burch   forbærnd***
there in was thus was the monastery of Peterborough burned
*7   forhærgod. … **And þus se abbot Turolde com to Burh.***
and destroyed   and thus the abbot Turolde came to Peterborough
"Then afterwards through their negligence and through their drunken-
ness one night the church and all that was inside burned down. Thus/
Consequently the monastery of Peterborough was burnt and destroyed
… and then the abbot Turold came to Peterborough."
[Peterb.Chron. (LdMisc 636) an.1070]

b. *Tho six kyngis by hir meanys gate unto them fyve othir kyngis, **and thus** they
began to gadir hir people.*   Malory Wks. (Win-C) 25/22]
"Then six kings by their means got five other kings under their control and
then they began to gather their people"

In its cataphoric and quotative use, *þus* alternates with *swā*, but is clearly the more
frequent and typical quotative marker: as said above, in this function *þus* appears
to be fully productive with all kinds of *verba dicendi* (*cweðan* "to say" above all, but
also *specan*, and *swerian* and so on) and for both direct and indirect speech, in all
kinds of patterns. If in (38a) *þus* functions as a manner adverb, sentences such as
(38b) and (38c) can be regarded as bridging contexts through which *þus* acquired
its new function as a quotative marker as a result of its cataphoric use.

(38) a. *Nāt   ic hit be   wihte,   butan ic wēne **þus**, þæt þǣr*
not.know I it about anything except I think thus that there
screoda wære gescyred rime   siex <hun><a> searohæbbendra
chariots were destroyed number six hun[dred] of.warriors
"I do not know anything about it, except that I think this: that there a
number of six hundred armed chariots were destroyed"   [Pha 2. 4]

b. *Ic eam sōð god and sōð ic ēow secge and **þus** swerige þurh mē*
I am true god and truth I you tell and thus swear by   my
sylfne and þurh mīne halgan englas Cherubin and Seraphin þæt
self   and by my   holy angels Cherubim and Seraphim that
mīn yrre cymð ofer ēow ǣr   on lytlan fyrste
my anger comes over you before in little while
"I am the true God and I tell you the truth and swear this [thus] through
myself and through my holy angels the Cherubim and the Seraphim that
my anger will come over you in a little while"[HomU 54 (Priebsch) (58)]

c.   *… forði hit is **þus** awriten: Ne dō þu   ōðrum,   þæt þu   nelle,*
therefore it   is thus written   not do you to.another that you not.want
þæt þe   sylue gedōn sy.   [BenRW 61.125.19]
that to.you self   done be
"Therefore this is written [*lit.* it is written like this]: don't do anything to
another that you do not want to be done to yourself"

Moreover, in Middle English only *þus/thus* was still used as quotative marker for
indirect (39) and direct (40) speech, unlike *swa* that disappeared in this function:

(39) *But telle thou thy kynge **thus**, that I owghe hym,* [*none homage*] *ne none of myne*
elders   [Malory Wks. (Win-C) 55/9]
"But tell your king this: that I owe him no homage nor do any of my officers"

(40) *Then came ayen Melusyne to her two sones, and **thus** she said to them: 'Fayre*
children, thinke from hensfourthon to doo wel …'
[Melusine XX. 108 (folio 63b)]
"Then came Melusine again to her two sons and said this to them: "Good boys,
think from now on to behave well""

The use of *þus* as a degree marker is attested in both OE and ME. In ME this use
was flourishing especially in combination with quantifiers, but was not restricted
to them nor to idiomatic expressions, such as *þus fer* "thus far", *þus longe* "for so
long", *þus sone* "at once", as is shown by the examples in (41a)–(b). The loss of such
a use is another example of the somewhat marginalised position of *thus* in ModE.

(41) a. *An þareuore ich am **þus** þriste*   [Owl & Nightingale 758] DEGREE
"and therefore am I so bold"

b. ***Đus** ic wille freon þis   mynstre   þet hit ne be underþed buton*
thus I   want to.free this monastery that it   not be subjected except
Rome ane.   [Peterb.Chron.(LdMisc 636) an.656] DEGREE
Rome only
"To such an extent I want to free this monastery that it is subjected to Rome
only"

7.2   *Swylc/swelc > such*

In this subsection we elaborate on a point briefly mentioned above without de-
tailed exemplification, namely the distributional and syntactic flexibility of de-
monstratives, and the origins and the historical development of this flexibility. The
quality-denoting demonstrative *such* is a particularly suitable case for this discus-
sion, since its syntactic categorisation has been a major puzzle for all grammars of

English and has recently been reexamined from a new and cross-linguistic perspective (van der Auwera & Coussé 2016; van der Auwera & Sahoo 2016, 2020). In these papers it is noted that *such* has been assigned to as many as eight different syntactic categories in recent grammars and descriptions of English (pre-determiner, determiner, semi-determiner, pronoun, adjective, and post-determiner, part of complex preposition or subordinator) on the basis of examples like the following:

(42) a.   John bought himself a new e-bike. I would like to have such a bike, too.
      b.   I have seen several such buildings.
      c.   It was such a shock!
      d.   Such is life.
      e.   It was such an amount as you would hesitate to accept.
      f.   She manifested grief such as would have destroyed another.
      g.   She asked to see Mr. Johnson, but there was no such person there.

To make things worse, this classification does not even cover several other uses:

(43) a.   He is such a fool that you would be surprised.
      b.   Let's assume you go to such and such a shop and you ask for such and such.

Given this multitude of possible syntactic classifications, the question arises whether our categorisation of ModE *such* as a demonstrative is justified, especially since this expression lost its exophoric use. After stressing the semantic and syntactic 'ad hoc' character of *such*, van der Auwera & Sahoo (2016, 2020) oscillate between a classification as 'demonstrative similative', 'similative demonstrative' and simply 'similative', underlining the observation that this expression lost its neighbours in the correlative paradigms of similatives and demonstratives of the types still found in Sanskrit and Latin.

 We take a somewhat different view in our analysis of *such, so* and *thus*, analysing them as somewhat defective, atypical demonstratives, which (largely) lost their basic exophoric use, but otherwise bear all the hallmarks of demonstratives. Here are the relevant arguments for our analysis of *swylc*:

i.   Like most other demonstratives, OE *swylc*/*swelc* is at the origin of several processes of grammaticalisation leading to the development of various grammatical markers and function words, which cannot easily be subsumed under specific lexical or grammatical categories. Such processes of de-categorisation are well-known for the distal demonstrative *that* and its development to a relative marker, to a complementiser, to a degree adverb, etc. (cf. Diessel 2006, 2013) and were outlined in some detail for the development of OE *swā* in Figure 1. Analogous changes can also be found in the development of OE *swylc*/*swelc*.

ii.   Although *such* lost its exophoric use, it clearly manifests two uses essentially associated with demonstratives, namely the anaphoric and the cataphoric uses, as illustrated in (44) and (46).

iii. Another property shared by *such* with more or less all other demonstratives is its use in coordinate structures (*Let's suppose you go into such-and-such a shop …*). In contrast to two- or three-term demonstrative paradigms (*here and there, now and then, this and that*), the same item occurs twice in such coordinate structures, since there is no deictic contrast in the system. Note that, in the coordinate structures, demonstratives are neither used exophorically nor endophorically, but simply indicate that there is a variety of referents without identifying them.

As already pointed out, the three 'demonstratives' and their counterparts in OE or other West Germanic languages are atypical, because a pointing gesture (if possible) or an anaphoric or cataphoric reference does not pick out a referent, but a subclass of entities related to the intended referent by similarity. Subsuming such demonstratives together with adjectives like *similar* under one category of 'similatives' does not make sense, however. Not only do adjectives like *similar* lack the deictic component which the source of *such* originally had and which its counterparts in other languages (Germ. *solch*; Du. *zulk*) still have, but there are also clear semantic differences between *such* and *similar* (cf. Umbach & Gust 2014).

Let us now summarise what we know about the development of *swylc* from OE to ModE. Figure 3 provides a rough overview of the historical developments from *swylc/swelc* to *such*:
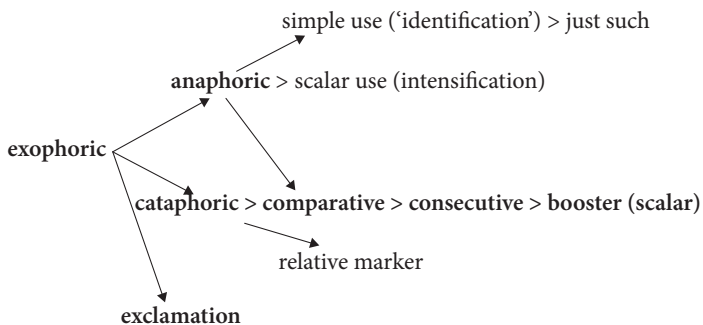
Figure 3. A-typical demonstratives as sources of grammaticalization: The case of *such*

As far as its basic syntactic properties are concerned, *swylc* occurs both in attributive (pre-nominal) and in predicative position. Some relevant examples are given in (3a)–(b) and (44a)–(b).

(44) a. *His feðeren synden ængles feðeren gelice*, *his breost &   bile.. scineð*,
          his feathers are   angels' feathers like   his breast and beak shine
          *fæȝere   &   fage;   Feawe synden **swylce**.*
          beautiful and coloured few   are   such
          "His feathers are similar to the feathers of angels, his breast and beak shine,
          beautiful and coloured, few are of that sort."

                                                              [Vsp.D.Hom.(Vsp D.14)147/9]
      b. ***Soche** was the a-vision that I saugh in my slepe.*

                                                              [Merlin (Cmb Ff.3.11) 632]
          "Such was the vision that I saw in my sleep"

On the basis of comparisons with historical and current use of counterparts in
German (*solch*) and Dutch (*zulk*), it is highly plausible that *swylc* was also used
exophorically. The most plausible candidates for such a use in OE are exclamations
like the following:

(45) a. *'Lo,' þai seyd, '**Swiche** a man!*   [Orfeo (Auch)43/505]
          ""Lo" they said, "such a man / what a man!""
      b. *Allas … to be for al our bale brouȝt to **swiche** an hende!*

                                                              [WPal.(KC 13)2333]
          "Alas, to be for all our evil-doing brought to such an end!"

Anaphoric uses, on the other hand, are clearly attested (46a)–(b):

(46) a. *Of Vðere Pendragune scal arisen **swilc** a sune.*   [Lay.Brut (Clg A.9)9422]
          "From Uther Pendragon such a son should arise"
      b. *Þe alre richeste kinge Ne dorste biginne **swch** a þing.* [Floris (Vit D.3) 202]
          "The richest king of all should not begin such a thing."

The anaphoric (retrospective) use of *such* in prenominal position is also one of the
primary uses in ModE, but is subject to several subtle restrictions relating to regis-
ter, to the nominal head and to the modal character of the clause containing them
(Bolinger 1972: 58–90). Sentences with anaphoric *such* are more acceptable if the
modified noun is gradable, so that the result is some intensification:

(47) a. Such an expression always frightens me.
      b. Such a scowl always frightens me.

(48) a. *The place is such a house!
      b. The place is such a pigsty!

In minimal pairs like (47) and (48) the first member is either totally unacceptable or at least very formal style.[22] The third restriction is illustrated by examples like (49)–(51), where *such* is only acceptable in non-episodic contexts.

(49)   a.   *Did you live with such a person?
       b.   Could you live with such a person?

(50)   a.   *Such a telescope was in the window.
       b.   Such a telescope is hard to find.

(51)   By coincidence just such a change was promoted in the Scottish Parliament on Thursday.

The relevant contexts which exclude anaphoric *such* and require *just such* instead express episodic events and occur in a context of "definiteness" (Bolinger 1972: 63).

The cataphoric use of *swylc* provides the basis for equative comparative constructions. A demonstrative denoting similarity between the intended referent and a subclass of entities identified either anaphorically or cataphorically requires a second argument, which is either found in the preceding or in the following text. Compared to its OE and ME roots, the use of *such* in equative comparatives in ModE has also been subjected to further restrictions. A clear distinction between degree markers, parameters of comparison and standard markers, as described for the analogous constructions with *swā* (> *so*), is not possible in comparative constructions with *swylc*. Both in OE and ME we find equative comparative constructions with *such* in predicative positions, i.e., without accompanying nouns, and thus without specification of the parameter of comparison. *Swylc* and its orthographic variants in OE and ME may occur alone or twice in correlative structures, with *swylc* functioning both as degree marker and parameter of comparison in the first clause and as standard marker in the second. This 'fusion' of meanings is expressed by *like* in ModE:

(52)   a.   *ðonne ic wæs mid Iudeum ic wæs **suelc** hie*   [CP 16.101.5]
             "when I was with the Jews I was so like them"
       b.   *Ðonne bið se   sacerd **suelc suelc** ðæt folc   bið ðonne he ðæt ilce*
             then   is the priest such as   that people are when he that same
             deð   ðæt hie doð   [CP 18.133.7]
             does that they do
             "The priest becomes the same as other people when he does the same that they do"

---

22. Bolinger (1972: 61ff.) uses the somewhat unfortunate term "identifying" for this use. As mentioned above, *such* never identifies a referent, but characterises the intended referent as being similar to an ad hoc subclass of the class denoted by the nominal modified.

    c.   *efne swylce mǣla* **swylce** *hira   mandryhtne þearf gesǣlde*   [Beo 1249]
        even such   times such   their liege-lord   need occurred
        "for just such times as for their liege-lord the need arose"

    d.   **Such** Capitein, **such** retenue.   [Gower CA (Frf 3)3.2421]
        "Like captain/status, like entourage."

Except for some formulaic phrases (*Like father, like son*), such correlative constructions are no longer used in ModE. Two other types of constructions with cataphoric uses of *swylc* (> *such*), by contrast, are found in all major historical periods of English: a construction with a resultative or purposive *that*-clause as second component (cf. (53)) and a construction with *swylc* (> *such*) … *as* surrounding or following a noun (cf. (54)):

(53) a. *Crist   him ȝeueð* **swilcne** *mete þet him nefre eft*
        Christ him gives such   food that him never afterwards
        ne hungreð.   [Lamb.Hom. (Lamb 487) 37]
        not.is.hungry
        "Christ gives him food of such a kind that never makes him hungry again."

    b. *He was … in his tyme* **swich** *a Conqueror that gretter was ther noon vnder*
        the Sonne.   [Chaucer CT. Kn.(Manly-Rickert) A. 863]
        "He was in his time such a conqueror that no one was greater under the sun."

(54) a. *on swelce healfe* **swelce** *hīe   þonne winnende bēon woldan, swā sūþ*
        on such   side   that   they then fighting be   wanted so   south
        swā norþ swā east swā west   [Or 3 5.59.3]
        so   north so   east so   west
        "on such a side as they then wanted to be fighting, either south or north or east or west"

    b. *He schale have theme at* **schwesche** *a pryse as ȝe kane akorde.*
                          [Acc.Howard in RC 57171]
        "He shall have them at such a price as you can agree on."

    c. *I shal warnestore myn hous with toures* **swiche** *as han castelles.*
                    [Chaucer CT.Mel.(Manly-Rickert) B.2523]
        "I will provide my house with such towers as castles have"

Both the constructions *such* N…*that* and *such* N *as/* N *such as* may have a scalar reading or a non-scalar one (cf. Bolinger 1972: 61–68). These constructions might, therefore, provide the source of the intensifying (booster) use of *such*, which is not attested until ME.

8.   Summary and conclusion

On the basis of the relevant text corpora and collections of data we have tried to reconstruct the development of a subset of demonstratives from Old English to Present-Day English with the goal of contributing both to the characterisation of demonstrative systems in different synchronic stages of English and to the analysis of widespread processes of change that start out from *demonstratives* (cf. Diessel 1999, 2013; Dixon 2003). The demonstratives selected for this study are members of a largely neglected subset of expressions which in addition to their deictic component include the content components of 'manner', 'quality' and 'degree' in their meaning and are atypical in the sense that gestures accompanying their exophoric use do not identify a referent, but an ad hoc subclass of referents which are similar to the intended referent (cf. Umbach & Gust 2014). In contrast to the systems in many other languages (cf. König 2012, 2015, 2017), the identification of the relevant demonstratives turned out to be particularly difficult, since the relevant expressions in Old English overlap in their distribution, meaning and use and largely lost their basic, exophoric, use in the course of their historical development. In spite of these difficulties it has been clearly established in our paper that *swā, swylc* and *þus* are instances of demonstratives and that even their counterparts in PDE still manifest essential properties of such deictic expressions, even if they have largely lost this basic use and developed a rich inventory of uses as function words.

The changes taking MQD-demonstratives in Old English and Middle English as their point of departure manifest the characteristic features of grammaticalisation processes: extension in the possible contexts of use, semantic change, de-categorisation, loss and renovation of properties, etc. Moreover, these changes are unidirectional, occur in tandem and were therefore analysed in the framework of 'grammaticalisation theory'. In complete harmony with Brugmann's (1904) view that demonstratives are one of the most important sources for the development of grammatical categories, it has been shown that *swā* (> *so*) and *swylc* (> *such*) give rise to a wide variety of different and new uses in the course of their historical development – 26 are listed for *so* and 11 for *such* in many dictionaries – whereas the evolution of *þus* (> *thus*) results in a loss of uses and in a more marginal position overall. The concomitant changes on the formal side are mainly processes of de-categorisation, which makes a clear assignment of these expressions to one specific part of speech or grammatical category nearly impossible, leading either to a multitude of different categorisations or to analysis in terms of an ad hoc category.

Based as it is on previous typological and comparative studies, our analysis also contributes to the cross-linguistic analysis of MQD-demonstratives, their formal and semantic properties, their characteristic differentiations of meaning and their possible paths of change leading to a wide variety of formal elements that straddle

the line between lexicon and grammar. Our analysis also confirms Diessel's general hypothesis (Diessel 1999, 2013) that demonstratives primarily develop into markers expressing transphrastic relations, i.e., connections across clauses (anaphora, cataphora, comparison, etc.) rather than strengthening the internal relations between the constituents of a simple sentence or construction.

Finally, by considering the synchronic situation in PDE as the outcome of pervasive grammaticalisation processes leading from demonstratives to a wide variety of function words, our analysis provides a more adequate perspective for a synchronic analysis of Modern English.

Acknowledgements

References

Aarts, Bas & April McMahon (eds.). 2006. *The handbook of English linguistics*. Oxford: Blackwell. https://doi.org/10.1002/9780470753002

Arbib, Michael A. 2012. *How the brain got language: The mirror system hypothesis*. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:osobl/9780199896684.001.0001

Blakemore, Diane L. 1988. *So* as a constraint on relevance. In Ruth M. Kempson (ed.), *Mental representations: The interface between language and reality*, 183–195. New York: Cambridge University Press.

Bolden, Galina B. 2008. *So what's up*? Using the discourse marker *so* to launch conversational business. *Research on Language and Social Interaction* 41(3). 302–337. https://doi.org/10.1080/08351810802237909

Bolden, Galina B. 2009. Implementing incipient actions: The discourse marker 'so' in English conversation. *Journal of Pragmatics* 41(5). 974–998. https://doi.org/10.1016/j.pragma.2008.10.004

Bolinger, Dwight. 1972. *Degree Words*. The Hague: Mouton. https://doi.org/10.1515/9783110877786

Boulonnais, Dominique. 2006. *So* et les phenomènes d'expansion contextuelle. In Geneviève Girard-Gillet (ed.), *Aux marges du texte: Texte et co-texte. Etudes réunies par Geneviève Girard* (C.I.E.R.E.C. Travaux 128), 49–69. Saint-Etienne: Publications de l'Université de Saint-Etienne.

Brugmann, Karl. 1904. *Die Demonstrativpronomina der indogermanischen Sprachen*. Leipzig: Teubner.

Buchstaller, Isabelle & Ingrid van Alphen (eds.). 2012. *Quotatives: Cross-linguistic and cross-disciplinary perspectives*. Amsterdam: John Benjamins. https://doi.org/10.1075/celcr.15

CMEPV = *The Corpus of Middle English Prose and Verse. The Humanities Text Initiative*, University of Michigan. https://quod.lib.umich.edu/c/cme/

Collins Cobuild Advanced Dictionary. 2009. Boston: Heinle.

D'Arcy, Alexandra. 2017. *Discourse-pragmatic variation in context: Eight hundred years of LIKE*. Amsterdam: John Benjamins. https://doi.org/10.1075/slcs.187

Diessel, Holger. 1999. *Demonstratives: Form, functions and grammaticalization*. Amsterdam: John Benjamins. https://doi.org/10.1075/tsl.42

Diessel, Holger. 2006. Demonstratives, joint attention, and the emergence of grammar. *Cognitive Linguistics* 17(4). 463–489. https://doi.org/10.1515/COG.2006.015

Diessel, Holger. 2013. Where does language come from? Some reflections on the role of deictic gesture and demonstratives in the evolution of language. *Cognitive Linguistics* 5(2–3). 239–249.

Dixon, R. M. W. 2003. Demonstratives: A cross-linguistic typology. *Studies in Language* 27(1). 61–112. https://doi.org/10.1075/sl.27.1.04dix

DOE = Cameron, Angus, Ashley Crandell Amos & Antonette diPaolo Healey. 2007. *The Dictionary of Old English: A to G online*. Toronto: Dictionary of Old English Project.

DOEC = Healey, Antonette diPaolo, John Price Wilkin & Xin Xiang. 2009. *The Dictionary of Old English Corpus on the World Wide Web*. Toronto: Dictionary of Old English Project.

Ericson, Eston Everett. 1932. *The use of 'swa' in Old English*. Göttingen: Vandenhoeck & Ruprecht.

Fettig, Adolf. 1935. *Die Gradadverbien im Mittelenglischen* (Anglistische Forschungen 79). Heidelberg: Carl Winter.

Greet, William C. (ed.). 1927. *The Reule of Crysten Religioun* (Early English Text Society 5). London: A. Moring.

Haspelmath, Martin & Oda Buchholz. 1998. Equative and simulative constructions in the languages of Europe. In Johan van der Auwera (ed.), *Adverbial constructions in the languages of Europe*, 277–334. Berlin: Mouton de Gruyter. https://doi.org/10.1515/9783110802610.277

Haspelmath, Martin & the Leipzig Equative Constructions Team. 2017. Equative constructions in world-wide perspective. In Yvonne Treis & Martine Vanhove (eds.), *Similative and equative constructions: A cross-linguistic perspective*, 9–32. Amsterdam: Benjamins. https://doi.org/10.1075/tsl.117.02has

HC = *Helsinki Corpus TEI XML Edition*. 2011. 1st edn. Designed by Alpo Honkapohja, Samuli Kaislaniemi, Henri Kauhanen, Matti Kilpiö, Ville Marttila, Terttu Nevalainen, Arja Nurmi, Matti Rissanen & Jukka Tyrkkö. Implemented by Henri Kauhanen & Ville Marttila. *Based on The Helsinki Corpus of English Texts* (1991). Helsinki: The Research Unit for Variation, Contacts and Change in English (VARIENG), University of Helsinki.

Heine, Bernd & Tania Kuteva. 2007. *The genesis of grammar*. Oxford: Oxford University Press.

Heine, Bernd & Kyung-An Song. 2011. On the grammaticalization of personal pronouns. *Journal of Linguistics* 47(3). 587–630. https://doi.org/10.1017/S0022226711000016

Hopper, Paul J. & Elizabeth Closs Traugott. 2003. *Grammaticalization*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9781139165525

Huddleston, Rodney & Geoffrey K. Pullum. 2002. *The Cambridge grammar of the English language*. Cambridge: Cambridge University Press. https://doi.org/10.1017/9781316423530

Johnson, Samuel. 1799/1913. *A dictionary of the English Language. A digital edition of the 1799 classic*. https://johnsonsdictionaryonline.com/so/

Johnson, Alison. 2002. So…? Pragmatic implications of *so*-prefaced questions in formal police interviews. In Janet Cotterill (ed.), *Language in the legal process*, 91–110. London: Palgrave Macmillan. https://doi.org/10.1057/9780230522770_6

König, Ekkehard. 1988. Concessive connectives and concessive sentences: Cross-linguistic regularities and pragmatic principles. In John A. Hawkins (ed.), *Explaining language universals*, 145–166. London: Blackwell.

König, Ekkehard. 1991. *The meaning of focus particles: A comparative perspective*. London: Routledge.

König, Ekkehard. 2012. Le rôle des déictiques de la manière dans le cadre d'une typologie de la deixis. *Bulletin de la Société de Linguistique de Paris* 107(1). 11–42.

König, Ekkehard. 2015. Manner deixis as a source of grammatical markers in Indo-European languages. In Carlotta Viti (ed.), *Perspectives on historical syntax*, 33–60. Amsterdam: John Benjamins. https://doi.org/10.1075/slcs.169.02kon

König, Ekkehard. 2017. The deictic identification of similarity. In Yvonne Treis & Martine Vanhove (eds.), *Similative and equative constructions: A cross-linguistic perspective*, 143–164. Amsterdam: John Benjamins. https://doi.org/10.1075/tsl.117.06kon

König, Ekkehard & Yoko Nishina. 2015. Deixis der Art und Weise, der Qualität und des Grades im Deutschen und Japanischen: Eine kontrastiv vergleichende Analyse. *Linguistische Berichte, Sonderhefte* 20. 7–32.

König, Ekkehard & Carla Umbach. 2018. Demonstratives of manner, of quality and of degree. In Marco Coniglio, Andrew Murphy, Eva Schlachter & Tonjes Veenstra (eds.), *Atypical demonstratives: Syntax, semantics and typology*, 285–328. Berlin: Mouton de Gruyter. https://doi.org/10.1515/9783110560299-010

Lehmann, Christian. 2017. *Thoughts on grammaticalization*. Berlin: Language Science Press.

McSparran, Frances, Paul Schaffner, John Latta, Alan Pagliere, Christina Powell & Matt Stoeffler. 2000–2018. *Online edition in Middle English Compendium*. Ann Arbor: University of Michigan Library. http://quod.lib.umich.edu/m/middle-english-dictionary/

MED = Hans Kurath, Sherman Kuhn & Robert E. Lewis. 1952–2001. *Middle English Dictionary*. Ann Arbor: University of Michigan Press.

MEG-C = Stenroos, Merja Riitta, Martti Mäkinen, Simon Horobin & Jeremy Smith. 2011. *The Middle English Grammar Corpus, version 2011.1*. University of Stavanger. http://www.uis.no/research/culture/the_middle_english_grammar_project/

Mitchell, Bruce. 1985. *Old English syntax I*. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198119357.001.0001

Mustanoja, Tauno F. 1960. *A Middle English syntax. Parts of speech*. Helsinki: Société Néophilologique de Helsinki.

Nummenmaa, Lusa. 1973. *The uses of so, al so and as in Early Middle English*. Helsinki: Société Néophilologique de Helsinki.

Oizumi, Akio (ed.). 1991–2012. *A complete concordance to the works of Geoffrey Chaucer*, volumes 1–16. Hildesheim: Olms-Weidmann.

Peltola, Niilo. 1959. On the 'identifying' swa (swa) phrase in Old English. *Neuphilologische Mitteilungen* 60. 156–173.

PPCME2 = Kroch, Anthony & Ann Taylor. 2000. *The Penn-Helsinki Parsed Corpus of Middle English (PPCME2)*. Department of Linguistics, University of Pennsylvania. https://www.ling.upenn.edu/hist-corpora/PPCME2-RELEASE-4/index.html

Raymond, Geoffrey. 2004. Prompting action: The stand-alone "so" in ordinary conversation. *Research on Language and Social Interaction* 37(2). 185–218. https://doi.org/10.1207/s15327973rlsi3702_4

Schleburg, Florian. 2002. *Altenglisch ,swa'. Syntax und Semantik einer polyfunktionalen Partikel*. Heidelberg: Carl Winter.

Smitterberg, Erik. 2005. *The progressive in 19th-century English: A process of integration*. Amsterdam: Rodopi. https://doi.org/10.1163/9789004333086

Tagliamonte, Sali A. & Chris L. Roberts. 2005. So weird; so cool; so innovative: The use of intensifiers in the television series *Friends*. *American Speech* 80(3). 280–300. https://doi.org/10.1215/00031283-80-3-280

Umbach, Carla & Helmar Gust. 2014. Similarity demonstratives. *Lingua* 149. 74–93. https://doi.org/10.1016/j.lingua.2014.05.003

van der Auwera, Johan & Evie Coussé. 2016. *Such* and *sådan* – the same but different. *Nordic Journal of English Studies* 15. 15–32. https://doi.org/10.35360/njes.374

van der Auwera, Johan & Kalyanamalini Sahoo. 2016. On comparative concepts and descriptive categories *such* as they are. *Acta Linguistica Hafniensia* 47. 136–173. https://doi.org/10.1080/03740463.2015.1115636

van der Auwera, Johan & Kalyanamalini Sahoo. 2020. *Such* similatives: A cross-linguistic reconnaissance. *Language Sciences* 81. 101217. https://doi.org/10.1016/j.langsci.2018.12.002

Webster's New World College Dictionary. 2014. Boston: Houghton Mifflin Harcourt.

YCOE = Taylor, Ann, Anthony Warner, Susan Pintzuk & Frank Beths. 2003. *The York-Toronto-Helsinki Parsed Corpus of Old English Prose* (YCOE). Department of Linguistics, University of York. http://www-users.york.ac.uk/~lang22/YcoeHome1.htm

York Poetry Corpus = Pintzuk, Susan & Leendert Plug. 2002. *The York-Helsinki Parsed Corpus of Old English Poetry*. Department of Linguistics, University of York. http://www-users.york.ac.uk/~lang18/pcorpus.html

Index

This volume contains a set of articles based on papers selected from those delivered at the 20th International Conference on English Historical Linguistics (ICEHL, Edinburgh 2018). It focuses on cutting-edge research in the history of English, while reflecting the diversity that exists in the current landscape of English historical linguistics. Chapters showcase traditional as well as novel methodologies in historical linguistics (the latter made possible by the increasing quality and accessibility of digital tools), work on linguistic interfaces (between segmental phonology and prosody, and syntax and information structure) and work on mechanisms of language change (such as Yang's Tolerance Principle, on the threshold for the productivity of linguistic rules in language acquisition). The volume will be of interest to those working on the historical phonology, morphology, syntax and pragmatics of English, language change, corpus linguistics, computational historical linguistics, and related sub-disciplines.

JOHN BENJAMINS PUBLISHING COMPANY