DE GRUYTER
MOUTON

# STUDIES IN THE HISTORY OF THE ENGLISH LANGUAGE VIII

## BOUNDARIES AND BOUNDARY-CROSSINGS IN THE HISTORY OF ENGLISH

*Edited by Peter J. Grund and Megan E. Hartman*

TOPICS IN ENGLISH LINGUISTICS

Peter J. Grund, Megan E. Hartman (Eds.)

**Studies in the History of the English Language VIII**

# Topics in English Linguistics

**Editors**
Susan M. Fitzmaurice
Bernd Kortmann
Elizabeth Closs Traugott

## Volume 108

# Studies in the History of the English Language VIII

Boundaries and Boundary-Crossings
in the History of English

Edited by
Peter J. Grund, Megan E. Hartman

DE GRUYTER
MOUTON

ISBN 978-3-11-062725-1
e-ISBN (PDF) 978-3-11-064328-2
e-ISBN (EPUB) 978-3-11-063985-8

**Library of Congress Control Number: 2020941980**

**Bibliographic information published by the Deutsche Nationalbibliothek**
The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available on the Internet at http://dnb.dnb.de.


Typesetting: Integra Software Services Pvt. Ltd.
Printing and binding: CPI books GmbH, Leck

www.degruyter.com

# Acknowledgments

The tenth-anniversary meeting of Studies in the History of the English Language (SHEL 10) took place at the University of Kansas (Lawrence, KS) in June 2017. Although "boundaries" was not a stated theme of the conference, this book emerged organically from the papers presented, and we supplemented the selected papers with additional, invited contributions on the same topic to produce this edited collection. While the book would obviously not have happened without the well-considered and thought-provoking chapters by the authors in the volume, many others have contributed substantially to the successful completion of the project. SHEL 10 was generously sponsored by the Office of Chancellor Bernadette Gray-Little, the College of Liberal Arts and Sciences, the Hall Center for the Humanities, and the English Department at the University of Kansas, as well as the American Dialect Society. The editing process was assisted by the gracious help of a number of colleagues in the field: our anonymous peer reviewers, who provided specific and detailed comments on the individual chapters; R. D. Fulk, who, in producing his Coda for this book, painstakingly read through the entire volume and offered very helpful editorial feedback; and Amanda Sladek and Molly Zahn, who read the Introduction and made a number of valuable suggestions. We also gratefully acknowledge Pam LeRow's help with the index of this volume.

We are also grateful to Susan Fitzmaurice, co-editor of the TiEL series at Mouton, for her guidance and penetrating questions during the editing process, and Barbara Karlson at Mouton, who smoothly guided the production process.

Finally, we humbly acknowledge the support and patience of friends and family, who have heard us talk about this volume and its ideas for many years. Having patient and caring listeners matters a great deal in a project like this! Thank you, All!

May, 2020
Peter J. Grund and Megan E. Hartman

https://doi.org/10.1515/9783110643282-202

# Contents

**Acknowledgments —— V**

**Contributor addresses —— IX**

Peter J. Grund and Megan E. Hartman
**Introduction: Boundaries and boundary-crossings in the history of English —— 1**

## Section 1: **Conceptual and methodological boundaries**

Michael Adams
**1 Scale and mode in histories of English —— 23**

Amy J. Devitt
**2 The blurred boundaries of genres-in-use: Principles and implications from rhetorical genre studies for English historical linguistics —— 45**

Edgar W. Schneider
**3 Meanderings from early English to World Englishes: A complex systems perspective on morphosyntactic changes in *wh*-pronouns —— 73**

## Section 2: **Linguistic boundaries**

Donka Minkova
**4 First or best, last not least: Domain edges in the history of English —— 109**

Edgar W. Schneider and Sarah Buschfeld
**5 Expanding boundaries of a function word: Uses of *one* in Early Modern and Modern English —— 135**

Erik Smitterberg
**6 Non-correlative commas between subjects and verbs in Early and Late Modern English sermons and scientific texts —— 167**

Raquel Vea Escarza
**7 Old English verbs of envy: Class membership and grammatical behaviour —— 187**

Section 3: **Language and language variety boundaries**

Anatoly Liberman
**8 Germanic /r/ as an isogloss, rhotacism, and the West Germanic gemination —— 211**

Angela Hoffman and Merja Kytö
**9 Migration, localities, and discourse: Shifting linguistic boundaries in Swedish-American cookbooks —— 225**

Don Chapman
**10 Specimen texts and boundaries in the history of the English language —— 249**

R. D. Fulk
**Coda: HEL-bound —— 271**

**Index —— 279**

# Contributor addresses

**Michael Adams**
Department of English
Indiana University
440 Ballantine Hall
1020 E. Kirkwood Avenue
Bloomington, IN 47405-7103
USA
adamsmp@indiana.edu

**Sarah Buschfeld**
TU Dortmund University
Emil-Figge-Straße 50
44227 Dortmund
Germany
sarah.buschfeld@tu-dortmund.de

**Don Chapman**
Department of Linguistics
Brigham Young University
4064 JFSB
Provo, UT 84602
USA
don_chapman@byu.edu

**Amy Devitt**
Department of English
University of Kansas
3001 Wescoe Hall
1445 Jayhawk Blvd.
Lawrence, Kansas 66045
USA
devitt@ku.edu

**R. D. Fulk**
Department of English
Indiana University
440 Ballantine Hall
1020 E Kirkwood Ave
Bloomington, IN 47405
USA
fulk@indiana.edu

**Peter J. Grund**
Department of English
University of Kansas
Wescoe Hall 3001
1445 Jayhawk Blvd.
Lawrence, KS 66045
USA
pjgrund@ku.edu

**Megan E. Hartman**
Department of English
University of Nebraska at Kearney
Thomas Hall 202
2404 11th Ave
Kearney, NE 68849
USA
hartmanme@unk.edu

**Angela Hoffman**
Department of English
Uppsala University
Thunbergsvägen 3 L
752 38 Uppsala
Sweden
angela.hoffman@engelska.uu.se

**Merja Kyt**ö
Department of English
Uppsala University
Thunbergsvägen 3 L
752 38 Uppsala
Sweden
merja.kyto@engelska.uu.se

**Anatoly Liberman**
Department of German, Nordic, Slavic, and Dutch
University of Minnesota
320 Folwell Hall
9 Pleasant St SE
Minneapolis, MN 55455
USA
aliber@umn.edu

https://doi.org/10.1515/9783110643282-204

**Donka Minkova**
Department of English
UCLA
Renée & David Kaplan Hall, Box 951530
405 Hilgard Ave.
Los Angeles, CA 90095-1530
USA
minkova@humnet.ucla.edu

**Edgar W. Schneider**
Department of English and American Studies
University of Regensburg
Universitaetsstrasse 31
93053 Regensburg
Germany
edgar.schneider@ur.de

**Erik Smitterberg**
Department of English
Uppsala University
Thunbergsvägen 3 L
752 38 Uppsala
Sweden
erik.smitterberg@engelska.uu.se

**Raquel Vea Escarza**
Departmento de Filologías Modernas
University of La Rioja
C/ San José de Calasanz, s/n 26004 Logroño
Spain
raquel.vea@unirioja.es

Peter J. Grund and Megan E. Hartman

# Introduction: Boundaries and boundary-crossings in the history of English

## 1 Introduction

Boundaries abound in language and linguistic research. Language users and scholars alike create, maintain, and cross boundaries in their use or study of language use (cf. Shapiro 1993). Not surprisingly considering their prominence, the concept of boundaries has received attention from a variety of perspectives in previous research on present-day English, illustrating the diverse set of boundaries existing in language. There are formal and theoretical framings (e.g., Fernández-Soriano et al. 2017), approaches focusing on particular sectors of the linguistic system (such as "prosodic boundaries" in phonetics/phonology; e.g., Fougeron and Keating 1997), and research that is concerned with more empirical, physical, and perceptual boundaries (e.g., contributions in Preston 1999; Juuso and Kretzschmar 2016; see section 2). Scholars of the history of the English language have also been interested in boundaries, notably from a conceptual or methodological angle in connection with the periodization of the history of English (e.g., Curzan 2012). At the same time, while boundaries have figured explicitly and, more often, implicitly in previous research on historical use, variation, and change in English, there appears to be little sustained attention to what it means to study historical boundaries in language, how users negotiate boundaries historically, and how scholars in English historical linguistics conceptualize boundaries (see section 3).

In this volume, we hope to begin to give such attention to boundaries and boundary-crossings. We bring together scholars who put boundaries front and center in their research, by investigating boundary phenomena, by viewing features through a boundary lens, and/or by interrogating boundaries established in the field. The contributions do not take one particular theoretical or methodological approach to boundaries and boundary-crossings. Instead, they explore through research on materials from the history of English how thinking about various types of boundaries – linguistic, conceptual, analytical, generic, material – helps us study, illuminate, and account for historical English. In their exploration

**Peter J. Grund,** University of Kansas, Lawrence, KS, USA,
e-mail: pjgrund@ku.edu
**Megan E. Hartman,** University of Nebraska at Kearney, Kearney, NE, USA,
e-mail: hartmanme@unk.edu

https://doi.org/10.1515/9783110643282-001

of a rich set of topics in the history of English, the contributions raise a range of questions: what does it mean to set up boundaries between different time periods? When do two language varieties have distinct boundaries and when do they overlap? Where do language users draw up clausal, constructional, semantic, phonetic/phonological boundaries and how are they maintained or crossed? The chapters explore not only how boundaries help us analyze synchronic and diachronic features in the history of the English language but also what happens if we ignore or question perceived or actual boundaries as we conduct historical linguistic research, or indeed if we consider conscious or subconscious boundary-crossings. While individual chapters bring their own take on these ideas, they open up larger, fruitful cross-chapter conversations about methodology and approach in English historical linguistics more broadly (see Fulk, this volume).

In this introductory chapter, we give a broad context for understanding boundaries and boundary-crossings in language and linguistic research by surveying treatments in a variety of fields of (English) linguistics (section 2). We then focus on the history of English and English historical linguistics (section 3). Section 4 outlines the topics, connections, and contributions of the chapters of the volume.

## 2 Boundaries and boundary-crossings in English linguistics (and beyond)

As we hinted above, much of language appears to revolve around boundaries and, sometimes, the crossing of those boundaries: indeed, as Fulk (this volume, p. 275) suggests, "constructing and deconstructing boundaries is an ineluctable facet of human cognition, and thus of language". Even such a "simple" example as word classes – nouns, verbs, conjunctions, etc. – involves boundaries between different sets of words, maintained for some communicative purposes but also crossed in individual cases for other communicative goals, which can of course lead to language change (e.g., a noun such as *Google* being used as a verb *google*; see section 3). But this deceptively simple example also highlights a fundamental split in boundary conception between boundaries that are language internal, created by the users of language, and boundaries that are drawn up by researchers of the language. Naturally, these types of boundaries often overlap, as researchers try to describe and account for linguistic use, but they can also be distinct or differ in nature and conceptualization. For example, boundaries discovered by researchers may not be consciously recognized by language users,

and/or the naming of those boundaries may be different for the two groups (cf. Devitt, this volume). As in the case of the word classes above, language users may maintain distinctions between lexical items of different functional groups so as to accomplish communication, but they may not recognize the terms – adjective, noun, etc. – applied to them by researchers studying the language. Even more poignantly, users may not recognize the often arbitrary boundaries of usage drawn up by prescriptivists and the prescriptive tradition if they do not reflect boundaries inherent in the language. The long-standing "rule" that *impact* is a noun is a case in point. Despite frequent and historically earlier usage of *impact* as a verb (https://www.merriam-webster.com/words-at-play/yes-impact-is-a-verb), many prescriptivists maintain that it is exclusively a noun. The distinction can be characterized broadly as one between *emic* and *etic* boundaries, to adopt and adapt a concept from anthropology and some branches of linguistics (see, e.g., Hahn 2006 for a discussion of the history and development of these terms/concepts). Here the etic perspective should be seen as the view from outside the user community, mostly from a researcher perspective (but also from the prescriptivist perspective noted earlier). The emic perspective, on the other hand, represents the boundary as conceived of and seen by the community of language users. Naturally, the distinction should be considered more of a sliding scale than a dichotomy: after all, linguists interested in boundaries (should) try to capture the emic system, but in doing so may apply etic conceptualizations or theories that may go beyond community understandings (as we shall see) (cf. Hahn 2006).

Boundaries also take on varying guises and connotations in different fields or contexts. For example, while *boundary* as a term may seem fairly neutral, the creation, maintenance, or crossing of boundaries is rarely a neutral act (even if it is subconscious or unconscious). For example, crossing usage boundaries by using a lexical item as a verb that was previously only employed in nominal contexts may be welcomed by some as creative but castigated by others as a corruption of the language (see, e.g., http://blog.writeathome.com/index.php/2013/09/verbing-weirds-language-when-nouns-become-verbs/). Boundary creation can also serve a number of inclusionary or exclusionary purposes, and hence linguistic boundaries can act as barriers (see below; also Fulk, this volume). Choices by scholars regarding how to describe or study a language phenomenon is another potential boundary-creating activity. While such boundaries are of course a normal part of the delimitation and professional practices of a field, they do affect the way we understand the nature of a feature and language in general (Adams, this volume; Fulk, this volume). We are thus dealing with a multifaceted, multilayered concept, which, as we shall see, has been addressed explicitly but more often implicitly in a variety of ways. Obviously, we cannot cover all approaches and topics related to boundaries here, but surveying some broad perspectives gives us a

sense of the richness and complexity of the concept of boundaries in language and language research.

One area where boundaries have been and continue to be central is dialectology. Surveys of especially British and American English lexis, phonology, and grammar have pointed to various regional patterns of use, establishing large-scale as well as specific boundaries (in the form of isoglosses) between regional dialects (e.g., Orton, Sanderson, and Widdowson 1978; Carver 1987; Labov, Ash, and Boberg 2006). At the same time, the nature of those boundaries and how they are constructed has been questioned from a variety of perspectives, raising issues about how regional boundaries are separated from other social categorizations and how isoglosses are drawn and generalized (e.g., Juuso and Kretzschmar 2016; Boberg, Nerbonne, and Watt 2018: 2–5). Perceptual dialectology approaches (Preston 1989) have added even more complexity to how regional boundaries should be understood. It has been shown over and over again in different locations that regional boundaries based on use may not coincide with users' perception of dialect demarcations and/or that boundaries are not neutral constructs (e.g., Evans 2011; Montgomery 2018; Stell 2018). Clearly, boundaries are essential for how we conceive of regional dialects, but much more is implied by where and how we draw those boundaries than purely objective and clearly delineated dialect difference.

Similarly complex boundary issues have emerged in the allied field of sociolinguistics. Indeed, from the inception of modern sociolinguistics, language has been shown to reflect boundaries among various social categories, be they gender, social class/socioeconomic status, or ethnicity, among others (e.g., Labov 1966; Rickford 1985). Of course, few studies have insisted that linguistic features are either-or markers, reflective of strict and immutable social categories, but second-wave and third-wave sociolinguistic studies especially (Eckert 2012) have emphasized the complex use that speakers (and writers) make of linguistic variation to negotiate identity-related boundaries (consciously and unconsciously). That is, language can certainly be used for creating inclusionary or exclusionary boundaries for communities and users, but the dynamics are often complex, involving various semiotic resources (e.g., clothing, body language, etc.) and situated "performance" of language (e.g., Eckert 2000; Barrett 2017). We also find sociolinguistic research focused on various kinds of boundary-crossings. A notable strand of such research is the negotiation of languages and dialects (or codes) in multiethnic communities whose speakers have several codes at their disposal. Studies have shown how speakers transcend the boundaries of one language to mix, mesh, or switch been varieties of language and by doing so cross national, regional, and ethnic boundaries conventionally associated with one language or the other (see, e.g., Rampton 1995; Young et al. 2014). Of course, such crossings

often entail new boundary making as users forge new identities with the help of their mixing and switching strategies. Rothman and Rell (2005: 529), for example, argue that Spanglish (the mixture and meshing of Spanish and English) in the US "is the linguistic embodiment of the juxtaposition of two very different cultures, which meet, intertwine, amalgamate and finally emerge as a unique identity for a particular cohort of people" (see also Hoffman and Kytö, this volume).

In the case of phonetics and phonology, the topic of boundaries has received particular attention in terms of boundary marking, that is, how various boundaries are signaled at the level of the word, the phrase, and the utterance (see also Minkova, this volume). A number of studies have explored the ways in which these borders are marked through articulatory strengthening, specifically lengthening (e.g., Oller 1973; Cooper 1991; Byrd 1994; Fougeron and Keating 1997), and additional studies have built on those more general analyses to identify just where the lengthening takes place (e.g., Wightman 1992) or how the lengthening differs with different pairs of sounds (e.g., Keating et al. 1994). Similarly to the way sociolinguists extend their analysis to other semiotic resources, phonologists/phoneticians extend their analysis to gestures, examining how changes in the actual physical actions of the oral articulators (Byrd and Saltzman 1998) as well as movements of the head or hands (Krivokapić 2014) can likewise reinforce the boundaries. It is furthermore clear that audience perception plays a role and that there are different ways to disambiguate ambiguous word or phrase boundaries. For example, Cho, McQueen, and Cox (2007) found that domain-initial strengthening is most effective in facilitating word recognition in ambiguous phonetic sequences, while Scott (1982) focused on syntactic grouping and showed that stress rhythms appear to be the most effective cue.

It is clear from sociolinguistics and dialectology, then, that language has external boundary-marking properties, and, from phonological/phonetic descriptions, that users mark unit boundaries in various acoustic ways. A great deal of attention has also been paid to the "internal", cognitive status of boundaries in language. As we mentioned in section 2, word classes constitute a "simple" example of how boundaries in language are maintained and crossed. As we did there, we use "simple" here within quotation marks since the status of word class – or part of speech – boundaries has been much debated and theorized (for an overview, see Aarts 2007: 10–33). Some scholars question the cognitive reality of word classes altogether (see Croft 2007, and below). Others recognize the usefulness of the categorization but doubt the validity of the sharp boundaries between them that researchers and others sometimes draw up (a kind of etic and emic distinction, in our conceptualization of boundaries; Quirk et al. 1985: 90). Among those who accept the existence of word classes, the number of word classes, their boundaries, and their constituents may vary considerably (see, e.g.,

Huddleston and Pullum [2002], where pronouns and nouns are treated as one class, and what has traditionally been seen as conjunction uses of *after*, *since*, etc. are classified as prepositions). For Aarts (2007), there are strict word class boundaries, but there is also gradience, and lexical items appear on a scale closer to or further away from "contiguous" categories depending on their grammatical properties; in Aarts's (2007: 224) terminology, a word belonging to one category can "converge" on another, but still be a full-fledged member of the first category (see also Denison [2007, 2013] in section 4).

More broadly, various formalist approaches to language have argued that boundaries are inherent in language structure and linguistic processing. Concepts such as *barrier*, *bounding node*, and *phase* have been important for theorizing properties of syntax, including so-called *movements*, *binding* and *government* (e.g., Chomsky 1986; Castroviejo, Férnandez-Soriano, and Pérez-Jiménez 2017). In the recent framework of Phase Theory, for example, phases "constitute minimal (syntactically and semantically) independent units which have the property of being structurally opaque" (Castroviejo, Fernandez-Soriano, and Perez-Jimenez 2017: 2). This means that, in this theory of language, phases are seen as bounded phenomena with important implications for processing: they "are [. . .] the locus for syntactic/semantic phenomena", such as agreement and case relations, but also phonological features such as stress and intonation (Castroviejo, Fernandez-Soriano, and Perez-Jimenez 2017: 2; see also Citko 2014).

Proponents of Construction Grammar or constructionist approaches provide a different conceptualization of the properties and cognitive storage of language, but the framework similarly relies on different bounded units, with concomitant boundaries. At the heart of the framework are *constructions*, "conventionalized form-meaning pairings" (Hoffmann and Trousdale 2013: 1). Constructions can come at various levels of abstraction, from a morpheme or word to larger syntactic templates, such as concessive adverbial clauses (Kerz and Wiechmann 2015). At the same time, these constructions are embedded in a larger network of constructions. Related constructions are connected through *inheritance links*, which entail, for example, that one construction is a more specific instantiation of a more abstract schema (Boas 2013: 244–245). While constructions are bounded units, then, there is also the sense that they are interconnected rather than isolated; there are boundaries, but not discreteness. At the same time, the boundaries of individual constructions can be fuzzy. Croft (2007: 428, n. 8), discussing Radical Construction Grammar, a subfield of Construction Grammar (see also Croft 2013), suggests that construction category boundaries "may not be sharp if acceptability judgements are not categorical, or if they vary across speakers or occasions".

Genre studies, which we here conceptualize as including both linguistic approaches as well as approaches in the allied fields of composition and rhetoric, is another area centrally concerned with boundaries, including both cognitive and social implications of such boundaries. A considerable portion of the research relates to charting genre boundaries. Interest in establishing such boundaries abounds at both the macro and the micro levels. On the macro level, criteria range from the social action a text is used to accomplish (e.g., Miller 1984) to the form of a text as understood through a range of linguistic features, often outlined by means of corpora (e.g., Biber 1988). Additional studies have sought to combine these criteria, suggesting that genre can only be established by simultaneously considering a text's function, features, and audience (e.g., Görlach 2004). On a micro level, many scholars have considered individual genres, such as tax documents (Devitt 1991), illness narratives (Frank 1995), and architect sketchbooks (Medway 2002), to investigate what characterizes these genres and distinguishes them from others. Scholarship in composition and rhetoric in particular has also been concerned with revising and breaking down genre boundaries, and linguistic research has shown the substantial changes that can take place in genres over time, demonstrating that genre boundaries are far from solid and immutable. We will return to this question in section 3.

## 3 Boundaries in the history of English and English historical linguistics

Since boundaries are so prevalent in language and language study, as we have seen in previous sections, it is little surprise to see that issues of boundaries and boundary-crossings have received attention in English historical linguistics as well. As in research on present-day language, the role that boundaries play is not always explicitly articulated; rather, they feature implicitly in many narratives of linguistic variation, change, and historical usage. The approaches and areas where boundaries have been highlighted or where boundaries play an implicit part also mirror what we see in studies of present-day language, but there are differences, as we shall see below. Of course, geolinguistic boundaries are at the center of historical atlas projects, such as the *Linguistic atlas of late mediaeval English* (*LALME*) and the *Linguistic atlas of early Middle English* (*LAEME*). *LALME* and *LAEME* have demonstrated the complex nature of Middle English dialects and their fuzzy demarcations. Actual regional boundaries signaled through language were no doubt also associated with perceptual boundaries, although we have less access to such perceptions for most of the history of English. Some

indication comes from the somewhat scattered commentary that we find in literary texts, grammars, and other writings. There we sometimes see regional boundaries that differ from boundaries of actual usage, and the boundaries are often connected to a set of social evaluations, just like in present-day contexts (see, e.g., Görlach 1999: 474–486; cf. Chapman, this volume).

Boundaries play a role not only among regional or social dialects, but also on broader levels of language and nation. In whatever location and at whatever time period it has been used, English has shared the scene with various other languages. This is reflected in, for example, the extensive borrowing taking place throughout the history of English. However, it is also becoming increasingly evident that, in these communities of multilingualism, the boundaries of different language codes (and hence of concomitant national and social identities) were not as strictly separated as has been previously thought. Rather, as shown by recent research (e.g., Schendl and Wright 2011), code-switching and code-mixing are common and well integrated in various situations and genres. This permeability of national and sociolinguistic code boundaries is also at the heart of recent theories about the creation of new World Englishes. Schneider's (2003) Dynamic Model, for instance, posits the gradual integration of "indigenous" and "settler" strands of linguistic usage, leading to the formation of new varieties of English.

Studies of diachronic change are frequently concerned with changing boundaries of usage, whether those boundaries be constructional, semantic, social, etc. In other words, over time, we see many boundaries of language being crossed (even if there are of course also boundaries that are maintained through stability in language; e.g., Kytö, Rydèn, and Smitterberg 2006: 9–10). Usage may start out with solid boundaries, which over time are increasingly blurred and eventually shifted, often with the end result of new boundaries, fuzzy or not, being established (see, e.g., van de Pol and Hoffmann [2016] for absolute constructions with *with*). Why such crossings happen is difficult to address, representing the elusive *actuation* stage of linguistic change (Weinreich, Labov, and Herzog 1968), but some of the mechanisms involved, which are many and varied, are known and at the heart of what much historical scholarship tries to pin down. For example, linguistic context may encourage a re-interpretation of lexico-semantics, with invited inference leading to the change in semantic boundaries, as in the case of certain contexts of temporal *since* giving rise to causal *since* (Traugott and Dasher 2002: 80). To come back to our word class example, functional shifts across word class boundaries are frequent and diverse, and often gradual, as in the increasing use of adjectival *key* and *fun* from original nominal uses (e.g., Denison 2007, 2013; see also Schneider and Buschfeld, this volume). Such category changes of course call into question hard and fast boundaries between word classes. A different boundary-crossing dynamic is seen in colloqualization, where features associated with

speech (or more informal genres) are adopted in more formal, written discourse, such as the increasing use of contractions in more formal genres over the twentieth century (Mair 2006: 189–190). In some ways, the process of colloqualization may be seen as one of potential erasing or increased blurring of previously reinforced boundaries between written and spoken or formal and informal usage.

Arguably, as in present-day contexts (see the sociolinguistic discussion in section 2), historical boundary-crossing can be conscious, a linguistic move that serves a social purpose, but may not be part of an overall process of linguistic change. Walker and Grund (2020) argue that trial witnesses or scribes of their depositions in the early modern period crossed boundaries of speech representation categories, exploiting the mixing of direct and indirect speech reporting for pragmatic purposes, such as clarity of exposition and negotiation of reported voices.

One area where boundaries and boundary-crossings have been a topic of intense and continuing debate is in the periodization of the history of English. This clearly represents primarily the etic side of boundary making (section 2), as periods would be of little consequence to communities of language users themselves; indeed, Curzan (2012: 1234) appropriately notes that "periodization is inherently an artificial, interpretive device imposed upon history [. . .]". In other words, it artificially splits up what is essentially "a continuous entity" (Curzan 2012: 1235), whose complexity may thereby be reduced to deceptive simplicity (cf. Schneider, this volume). Writers of history of English textbooks and authors of research articles and monographs alike rely on periodization and labels such as *Old English*, *Middle English*, etc., but there is no consensus on how such periods are delimited and exactly where the boundaries lie. Some advocate purely linguistic criteria for the boundaries between periods; others prefer extralinguistic criteria (though usually in conjunction with some linguistic features) (see Curzan 2012 for a review of different approaches; see also Chapman, this volume).

More specific period boundaries are also of importance in corpus-based work (which is beginning to dominate research in English historical linguistics). Most corpora come with a pre-set stratification, which may be based on specific (linguistic or extralinguistic) criteria or may simply represent an even "chunking" of the covered time. The utility of such divisions has been questioned, and some researchers have instead argued for a data-driven, statistical approach to periodization (e.g., Hilpert 2012: 149–153; Shao, Cai, and Trousdale 2019: 156–158). However one conceives of these period boundaries (broad or specific), the importance they have had and continue to have as heuristic devices and even constraints on our linguistic histories cannot be denied, which again stresses the centrality of boundaries in the discipline.

Classificatory boundaries similarly lie on the more etic side of the spectrum. Often, our conceptual tools and terms require sharper divisions than the data may allow. In an exploration of loanwords and borrowing, Durkin (2004: 89–90) stresses that descriptions of how new words and meanings have entered into English may not be possible to capture with neatly divided categories. Such issues are of course plentiful in English historical linguistics (see, e.g., Moore [2007] for the classification of BE *supposed to*, or Collins [2009] for modal auxiliaries), and the solution to such boundary issues is not straightforward. Perhaps our categorizations are simply not sophisticated enough to represent the emic reality of what language users do with language, or the phenomenon is not possible to pin down with sharply defined boundaries, instead representing a continuum or fluid concept (cf. Devitt, this volume; Adams, this volume; Vea Escarza, this volume).

This is evident in the study of genres. As genre scholarship works to establish boundaries (as we saw in section 2), it also troubles them. Genres are by default unstable and often show a degree of "drift" that allows characteristic features to shift over time (see Biber and Finegan 1989; Görlach 2004). Furthermore, genre boundaries have expanded in recent years due to new technologies (Biber and Egbert 2016; Miller and Kelley 2017). New media have allowed for greater discourse interactivity in genres such as newspapers, patient narratives, and games, forcing scholars to expand their concepts of what these genres entail (Miller and Kelley 2017, and contributions therein). While some scholars focus on fluidity, others, especially scholars of older literary works, have questioned the validity of certain generic boundaries, as in the case of the so-called elegy in Old English. Although the elegy was considered a central genre to the Old English poetic corpus for some time, many now argue that the category is not only etic but also anachronistic and that Old English could be better analyzed without the influence of imposed generic boundaries (Shippey 1994; Fulk and Cain 2013: 256–273). Thus, the focus on generic boundaries and how they should be perceived can vary widely depending on the researcher's perspective and the texts in question.

Naturally, as in the case of our overview in section 2, it is not possible to describe fully here the many and varied ways that boundaries manifest themselves in the history of English or how they are conceptualized and studied in English historical linguistics. But the real takeaway is that boundaries are crucial for how language has operated historically, and they are thoroughly implicated in narratives of linguistic change, even if they are not always explicitly acknowledged or theorized.

## 4 Overview of contributions

The discussion in the previous sections may raise some questions: if boundaries are so ubiquitous in language and linguistic research, is there anything new to say? What do we gain by further discussing boundaries or boundary-crossings? By focusing specifically on this topic, this volume offers a number of new perspectives on boundaries and boundary-crossings. While boundaries are part of many linguistic narratives, as we have shown above, they often feature implicitly rather than explicitly, and we suggest that there is payoff in making boundaries more central and in bringing together various boundary narratives in one place. This volume shows the diverse ways boundaries can be approached, discussed, and framed. Some of our chapters question conceptual, theoretical, and professional boundaries (Adams; Devitt; Schneider; Vea Escarza; Chapman); some investigate boundary phenomena (Minkova; Smitterberg); some consider how linguistic features or strategies create or allow us to cross (perceived) boundaries (Liberman; Hoffman and Kytö; Chapman); and some explore the dynamics and implications of linguistic features moving across various grammatical or usage categories (Schneider and Buschfeld; Liberman). Together, then, by providing these different angles on and narratives of boundaries and boundary-crossings, the chapters enrich the concept of boundaries, testifying to the complexity and broad potential of the concept for research on the history of English. We also hope that the concept of emic and etic boundaries can be helpful in separating different boundary concepts, in revealing how complicated our boundary construction can be and in provoking new questions about how the field of English historical linguistics operates.

The first section, "Conceptual and methodological boundaries", features chapters that grapple with issues and intersections of etic and emic boundaries. They suggest ways in which our methods and concepts do not fully capture the complex reality of the history of English, and they advocate various adjustments to our current conceptions. These studies are conceptually and methodologically different from the other studies in the volume in that they focus on broader, programmatic questions. While the authors provide illustrations and concrete explanation of their proposals, their main aim is to lay out frameworks that question, transgress, or shift boundaries in our current methodologies and conceptions: full-fledged implementations of the frameworks and suggestions await future studies to see how these proposals play out in various materials and time periods.

The volume begins with a striking proposal by Michael Adams that we, as scholars of the history of the English language, should revisit how we think about the boundaries imposed by and on the discipline. He imagines studies of English at both the macro and micro levels, considering what a study of the history of the English language over a year, over a single individual's lifetime, or

over multiple millennia might look like. To undertake such studies, scholars would have to rethink the conventions that define our discipline, but Adams argues that crossing that boundary and using other disciplinary conventions as models could in fact be beneficial to the field and perhaps even expand the audience and perceived relevance of the history of the English language. Adams's chapter demonstrates how intricately intertwined etic and emic boundaries are: on the one hand, he provides a view of the emic, professional boundaries as construed by the community of English historical linguistics. On the other hand, he also questions how such boundaries, which would often be further towards the etic side of the continuum when viewed from the perspective of language users, fit the emic reality of the history of English.

Amy Devitt, on the other hand, addresses the central concept of genre, which plays a significant role in descriptions of variation and change in English. Drawing on scholarship from the related fields of composition and rhetoric, she argues for shifting the focus from genres as static categories to categories with fuzzy edges and permeable boundaries (and thus bringing etic description closer to the complexity of the emic reality). By simultaneously analyzing features such as audience, purpose, author, setting, and medium under the umbrella of a genre's social action, Devitt argues, scholars can establish a fuller understanding of individual genres and their inherent instabilities. She provides an illustration of what such a conceptualization and implementation of genre would mean for a study of language variation and change by revisiting her earlier research on sixteenth- and seventeenth-century Scots. By showing the questions that may be asked and the tools that may be used for this particular material, she points to the procedures that can be adopted in future research to shine new light on dynamics of variation and change through a genre-as-social-action approach.

Edgar Schneider's contribution rethinks the way in which English is often categorized in terms of differences across time and space (a type of etic perspective), arguing that it is possible to see English as "a single overarching Complex Dynamic System". He demonstrates this idea by adopting a Complex Systems approach to *wh*-pronouns in the history of various Englishes. Focusing on typical features of Complex Systems such as systematicity, complexity, and perpetual dynamics, he argues that the system of *wh*-pronouns is constantly in motion and grows increasingly complex while still maintaining a clear systematic relationship between the different forms. In this way, analysis of the pronouns both demonstrates their own participation in a Complex Dynamic System and illustrates the validity of viewing the boundaries of time and variety in English less as boundaries and more as manifestations of the larger complex system. In other words, he demonstrates that, without a more holistic perspective, research may

fail to realize that changes attributed to individual varieties are the result of larger, intertwined forces and changes that have been operating across varieties for centuries as part of English as a Complex Dynamic System.

The second section, "Linguistic boundaries", contains articles that consider particular linguistic (syntactic, grammatical, lexical, semantic, phonetic/phonological) features in light of the concept of boundaries. While these articles are of course concerned with describing various emic boundaries inherent in language and language use, they also (to a greater or lesser extent) engage with issues of etic description of those boundaries. Donka Minkova's chapter begins the section by illustrating the payoff of putting the effect of boundaries at the center of the study. Focusing on Old English, she surveys the effect boundary edges have on phonological lengthening. Pairing these phonological observations with metrical theory, she uses the principle of lengthening to explain features such as cluster alliteration and vowel alliteration on the left edge and, notably, uses vowel lengthening at the right edge to provide a phonological justification for the long-debated metrical principle known as Kaluza's law. In her exploration, Minkova also engages in a boundary move of a different kind. By crossing the boundary between phonological and metrical studies, she is able to make enlightening observations about the former while simultaneously explaining some cruxes in the latter (cf. Adams's chapter).

Edgar Schneider and Sarah Buschfeld investigate the expanding usage boundaries of *one*, looking at how various features and factors contribute to the way the term broadens to cover a larger number of semantic categories from Early Modern English into Modern English. Using the ARCHER corpus, they examine both word class (e.g., numeral, pronoun, determiner, etc.) and syntactic usage (e.g., subject, object, prepositional complement, etc.) to map out the various usages in both British and American English. Their findings reveal that *one* has developed along a variety of trajectories, crossing various usage boundaries, occasionally diverging between British and American English, but that the word generally seems to be undergoing a process of grammaticalization that underscores its use as a pronoun without actually eliminating the other uses.

Like Minkova, Erik Smitterberg is concerned with a phenomenon at a boundary: he explores the use of "non-correlative" commas in Early Modern and Late Modern English science texts and sermons drawn from the ARCHER corpus. Non-correlative commas are found between the subject and verb, as in "*many actions externally mean and lowly*, *may* [. . .]". While non-standard today, such commas are used in a variety of contexts in earlier periods, as Smitterberg shows, although they decline in both genres. In these contexts, they highlight grammatical and/or prosodic boundaries and can thus assist in parsing information (science) and be used for rhetorical effect (sermons). His chapter underscores that users

can employ various linguistic means (including punctuation) for boundary-marking purposes in the service of larger communicative goals.

Raquel Vea Escarza considers usage boundaries from semantic as well as syntactic perspectives. In addition to seeing how users of Old English drew up usage boundaries (the emic side), she is also concerned with how we as researchers construct boundaries around groups of features (the etic side). She investigates the semantic and syntactic properties of a set of four Old English verbs expressing 'envy': *æfestian*, *andian*, *niðan*, and *ofunnan*. Systematically evaluating seven semantic and syntactic criteria, she demonstrates that, while similar on a general semantic level, the verbs behave in several overlapping but in the end different ways. Her chapter highlights how, depending on perspective and the criteria involved, researchers can delimit and construct boundaries in different ways, or indeed how strict boundaries constructed in one way are permeable when viewed from a different angle.

The third section, "Language and language variety boundaries", concerns boundaries (perceived or real) as well as boundary-crossing between varieties of English or between English and other languages. Drawing on a range of evidence, Anatoly Liberman argues that various historical developments involving the phoneme /r/ and its realization have created spatio-linguistic boundaries among Germanic languages. Particularly important for his discussion is the argument that /r/ in English developed from a process of rhoticism rather than from Indo-European /r/, as in other related languages. That distinction accounts for what has been seen as an idiosyncrasy in English compared to other Germanic languages. What we see here then is that individual features can act as boundary markers, separating languages or language varieties.

Unlike Liberman's chapter, Hoffman and Kytö's chapter is more centrally concerned with boundary-crossing than boundary maintenance, although the latter topic is also covered in their chapter. Surveying the mixing and switching of English and (Heritage) Swedish in recipe books from a number of Swedish heritage communities in Illinois, Kansas, and Minnesota, they demonstrate the permeability of the language boundaries and the shifting use of language codes in space and time. The boundary-crossing clearly takes place in particular functional spaces (such as recipe headings and naming of dishes), and, paradoxically, such acts work to define the boundaries of a particular identity as a Heritage Swede. We thus see how creating and crossing boundaries can be simultaneous and interrelated.

A different kind of boundary maintenance (and boundary-crossing) between language varieties is found in Don Chapman's chapter. In a study of the primary texts that fifty-one textbooks on the history of English draw on for exemplification, Chapman reveals a canon of sorts that reinforces certain boundaries in the

narrative of the history of English. Such boundaries include those between different periods of English, between non-standard and the standard dialect, and between literary and non-literary registers. While the boundary work that textbooks do appears to have changed in many respects, other narratives continue, whether authors are consciously aware of them or not. However, we witness boundary-crossing as well, as textbooks break with traditions and increase coverage of varieties and World Englishes. Like Adams, then, Chapman is explicitly and implicitly concerned with shifting boundaries in the profession and what it means to contribute to the story of English.

Finally, in a coda, R. D. Fulk muses on the nature of the volume, its theme, and the implications for the study of the history of English. Pulling out a number of threads and connections among the chapters' boundary discussions, Fulk suggests that the chapters

> ask us to rethink the limits on the study of English language history and to consider alternative interpretations, aims, definitions, research questions, sources of data, levels of analysis, kinds of evidence, ways of grouping, methods, scales, and modes. That is to say, they expand the boundaries of the field, inviting new kinds of research and new ways of approaching issues of long standing. They suggest where HEL is ideally bound: unboundedness (p. 276).

## References

Aarts, Bas. 2007. *Syntactic gradience: The nature of grammatical indeterminacy*. Oxford: Oxford University Press.

Barrett, Rusty. 2017. *From drag queens to leathermen: Language, gender, and gay male subcultures*. Oxford: Oxford University Press.

Biber, Douglas. 1988. *Variation across speech and writing*. New York: Cambridge University Press.

Biber, Douglas & Jesse Egbert. 2016. Register variation on the searchable web: A multi-dimensional analysis. *Journal of English Linguistics* 44(2). 95–137.

Biber, Douglas & Edward Finegan. 1989. Drift and the evolution of English style: A history of three genres. *Language* 65(3). 487–517.

Boas, Hans C. 2013. Cognitive Construction Grammar. In Thomas Hoffmann & Graeme Trousdale (eds.), *The Oxford handbook of Construction Grammar*, 233–252. Oxford: Oxford University Press.

Boberg, Charles, John Nerbonne & Dominic Watt. 2018. Introduction. In Charles Boberg, John Nerbonne & Dominic Watt (eds.), *Handbook of dialectology*, 1–15. Oxford: Wiley-Blackwell.

Byrd, Dani. 1994. *Articulatory timing in English consonant sequences*. Los Angeles, CA: University of California Los Angeles dissertation.

Byrd, Dani & Elliot Saltzman. 1998. Intragestural dynamics of multiple prosodic boundaries. *Journal of Phonetics* 26(2). 173–199.
Carver, Craig M. 1987. *American regional dialects: A word geography*. Ann Arbor, MI: University of Michigan Press.
Castroviejo, Elena, Olga Fernández-Soriano & Isabel Pérez-Jiménez. 2017. Introduction: Boundaries, phases and interfaces. In Olga Fernández-Soriano, Elena Castroviejo & Isabel Pérez-Jiménez (eds.), *Boundaries, phases and interfaces*, 1–23. Amsterdam: John Benjamins.
Cho, Taehong, James M. McQueen & Ethan A. Cox. 2007. Prosodically driven phonetic detail in speech processing: The case of domain-initial strengthening in English. *Journal of Phonetics* 35(2). 210–243.
Chomsky, Noam. 1986. *Barriers*. Cambridge, MA: MIT Press.
Citko, Barbara. 2014. *Phase theory: An introduction*. Cambridge: Cambridge University Press.
Collins, Peter. 2009. *Modals and quasi-modals in English*. Amsterdam: Rodopi.
Cooper, A. 1991. Laryngeal and oral gestures in English /p,t,k/. *Proceedings of the XIIth international congress of phonetic sciences 2*, 50–53. Aix-en-Provence: University of Provence.
Croft, William. 2007. Beyond Aristotle and gradience: A reply to Aarts. *Studies in Language* 31(2). 409–430.
Croft, William. 2013. Radical Construction Grammar. In Thomas Hoffmann & Graeme Trousdale (eds.), *The Oxford handbook of Construction Grammar*, 211–232. Oxford: Oxford University Press.
Curzan, Anne. 2012. Ch. 79: Interdisciplinarity and historiography: Periodization in the history of the English language. In Alexander Bergs & Laurel Brinton (eds.), *English historical linguistics: An international handbook*, vol. 2, 1233–1256. Berlin: Mouton de Gruyter.
Denison, David. 2007. Playing tag with category boundaries. In Anneli Meurman-Solin & Arja Nurmi (eds.), *Annotating variation and change*. Helsinki: Research Unit for Variation, Contacts and Change in English (VARIENG), University of Helsinki. http://www.helsinki.fi/varieng/series/volumes/01/denison/. np. (As accessed Sept., 2019).
Denison, David. 2013. Parts of speech: Solid citizens or slippery customers? *Journal of the British Academy* 1. 151–185.
Devitt, Amy J. 1991. Intertextuality in tax accounting: Generic, referential, and functional. In Charles Bazerman & James Paradis (eds.), *Textual dynamics of the professions: Historical and contemporary studies of writing in professional communities*, 336–357. Madison, WI: University of Wisconsin Press.
Durkin, Philip. 2004. Loanword etymologies in the third edition of the *OED*: Some questions of classification. In Christian J. Kay, Carole Hough & Irene Wotherspoon (eds.), *New perspectives on English historical linguistics*, vol. 2, *Lexis and transmission*, 79–90. Amsterdam: John Benjamins.
Eckert, Penelope. 2000. *Linguistic variation as social practice: The linguistic construction of identity in Belten High*. Oxford: Blackwell.
Eckert, Penelope. 2012. Three waves of variation study: The emergence of meaning in the study of sociolinguistic variation. *Annual Review of Anthropology* 41. 87–100.
Evans, Betsy E. 2011. “Seattletonian” to “faux hick”: Perceptions of English in Washington state. *American Speech* 86(4). 383–414.
Fernández-Soriano, Olga, Elena Castroviejo & Isabel Pérez-Jiménez (eds.). 2017. *Boundaries, phases and interfaces*. Amsterdam: John Benjamins.

Fougeron, Cécile & Patricia A. Keating. 1997. Articulatory strengthening at edges of prosodic domains. *The Journal of the Acoustical Society of America* 101(6). 3728–3740.
Frank, Arthur W. 1995. *The wounded storyteller: Body, illness and ethics*. Chicago: University of Chicago Press.
Fulk, R. D. & Christopher M. Cain. 2013. *A history of Old English literature*. 2nd edn. Malden, MA: Blackwell.
Görlach, Manfred. 1999. Regional and social variation. In Roger Lass (ed.), *Cambridge History of the English Language*, vol. 3, *1476–1776*, 459–538. Cambridge: Cambridge University Press.
Görlach, Manfred. 2004. *Text types and the history of English*. Berlin: Mouton de Gruyter.
Hahn, Christina. 2006. Clear-cut concepts vs. methodological ritual: *Etic* and *emic* revisited. *Beiträge zur Geschichte der Sprachwissenschaft* 16. 245–262.
Hilpert, Martin. 2012. Diachronic collostructional analysis: How to use it and how to deal with confounding factors. In Kathryn Allan & Justyna A. Robinson (eds.), *Current methods in historical semantics*, 133–160. Berlin: Mouton de Gruyter.
Hoffmann, Thomas & Graeme Trousdale. 2013. Construction Grammar: Introduction. In Thomas Hoffmann & Graeme Trousdale (eds.), *The Oxford handbook of Construction Grammar*, 1–12. Oxford: Oxford University Press.
Huddleston, Rodney & Geoffrey K. Pullum. 2002. *The Cambridge grammar of the English language*. Cambridge: Cambridge University Press.
Juuso, Ilkka & William A. Kretzschmar, Jr. 2016. Creation of regions for dialect features using a cellular automaton. *Journal of English Linguistics* 44(1). 4–33.
Keating, Patricia, Björn Lindblom, James Lubker & Jody Kreiman. 1994. Variability in jaw height for segments in English and Swedish VCVs. *Journal of Phonetics* 22. 407–422.
Kerz, Elma & Daniel Wiechmann. 2015. Register-contingent entrenchment of constructional patterns: Causal and concessive adverbial clauses in academic and newspaper writing. *Journal of English Linguistics* 43(1). 61–85.
Krivokapić, Jelena. 2014. Gestural coordination at prosodic boundaries and its role for prosodic structure and speech planning processes. *Philosophical Transactions of the Royal Society: Biological Sciences* 369. 1–10.
Kytö, Merja, Mats Rydén & Erik Smitterberg. 2006. Introduction: Exploring nineteenth-century English – past and present perspectives. In Merja Kytö, Mats Rydén & Erik Smitterberg (eds.), *Nineteenth-century English: Stability and change*, 1–16. Cambridge: Cambridge University Press.
Labov, William. 1966. *The social stratification of English in New York City*. Washington, DC: Center for Applied Linguistics.
Labov, William, Sharon Ash & Charles Boberg. 2006 *The atlas of North American English: Phonetics, phonology, and sound change*. Berlin: Mouton de Gruyter.
*LAEME* = *A Linguistic Atlas of Early Middle English*, 1150–1325, compiled by Margaret Laing [http://www.lel.ed.ac.uk/ihd/laeme2/laeme2.html]. Edinburgh: Version 3.2, 2013, © The University of Edinburgh.
*LALME* = Angus McIntosh, M. L. Samuels & Michael Benskin. 1986. *A linguistic atlas of late mediaeval English*. 3 vols. Aberdeen: Aberdeen University Press.
Mair, Christian. 2006. *Twentieth-century English: History, variation, and standardization*. Cambridge: Cambridge University Press.
Medway, Peter. 2002. Fuzzy genres and community identities: The case of architecture students' sketchbooks. In Richard M. Coe, Lorelei Lingard & Tatiana Teslenko (eds.),

*The rhetoric and ideology of genre: Strategies for stability and change*, 123–153. Cresskill NJ: Hampton Press.
Miller, Carolyn R. 1984. Genre as social action. *Quarterly Journal of Speech* 70. 151–167.
Miller, Carolyn R. & Ashley R. Kelley (eds.). 2017. *Emerging genres in new media environments*. London: Palgrave Macmillan.
Montgomery, Chris. 2018. The perceptual dialectology of England. In Natalie Braber & Sandra Jansen (eds.), *Sociolinguistics in England*, 127–164. Basingstoke: Palgrave.
Moore, Colette. 2007. The spread of grammaticalized forms: The case of *be+supposed to*. *Journal of English Linguistics* 35(2). 117–131.
Oller, D. Kimbrough. 1973. The effect of position in utterance on speech segment duration in English. *Journal of the Acoustical Society of America* 54. 1235–1247.
Orton, Harold, Stewart Sanderson & John Widdowson. 1978. *The linguistic atlas of England*. Atlantic Highlands, NJ: Humanities Press.
Preston, Dennis. 1989. *Perceptual dialectology: Nonlinguists' view of areal linguistics*. Berlin: Mouton de Gruyter.
Preston, Dennis (ed.). 1999. *Handbook of perceptual dialectology*, vol. 1. Amsterdam: John Benjamins.
Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech & Jan Svartvik. 1985. *A comprehensive grammar of the English language*. London: Longman.
Rampton, Ben. 1995. *Crossing: language and ethnicity among adolescents*. London: Longman.
Rickford, John R. 1985. Ethnicity as a sociolinguistic boundary. *American Speech* 60(2). 99–125.
Rothman, Jason & Amy Beth Rell. 2005. A linguistic analysis of Spanglish: Relating language to identity. *Linguistics and the Human Sciences* 1(3). 515–536.
Schendl, Herbert & Laura Wright (eds.). 2011. *Code-switching in early English*. Berlin: Mouton de Gruyter.
Schneider, Edgar. 2003. The dynamics of New Englishes: From identity construction to dialect birth. *Language* 79(2). 233–281.
Scott, Donia R. 1982. Duration as a cue to the perception of a phrase boundary. *Journal of the Acoustical Society of America* 71. 996–1007.
Shao, Bin, Yingying Cai & Graeme Trousdale. 2019. A multivariate analysis of diachronic variation in *A bunch of* NOUN: A Construction Grammar account. *Journal of English Linguistics* 47(2). 150–174.
Shapiro, Michael. 1993. Presidential address: The boundary question. *The American Journal of Semiotics* 10(3/4). 5–25.
Shippey, T. A. 1994. *The Wanderer* and *The Seafarer* as wisdom poetry. In Henk Aertson & Rolf H. Bremmer, Jr. (eds.), *Companion to Old English poetry*, 145–158. Amsterdam: VU University Press.
Stell, Gerald. 2018. Representing variation in creole continua: A folk linguistic view of language variation in Trinidad. *Journal of English Linguistics* 46(2). 119–139.
Traugott, Elizabeth Closs & Richard B. Dasher. 2002. *Regularity in semantic change*. Cambridge: Cambridge University Press.
Van de Pol, Nikki & Thomas Hoffmann. 2016. With or without *with*: The constructionalization of the *with*-augmented absolute in English. *Journal of English Linguistics* 44(4). 318–345.
Walker, Terry & Peter J. Grund. 2020. Free indirect speech, slipping, or a system in flux? Exploring the continuum between direct and indirect speech in Early Modern English. In

Peter J. Grund & Terry Walker (eds.), *Speech representation in the history of English: Topics and approaches*, 156–182. Oxford: Oxford University Press.

Weinreich, Uriel, William Labov & Marvin I. Herzog. 1968. Empirical foundations for a theory of language change. In W. P. Lehmann & Yakov Malkiel (eds.), *Directions for historical linguistics*, 97–188. Austin: University of Texas Press.

Wightman, Colin W. 1992. Segmental durations in the vicinity of prosodic phrase boundaries. *Journal of the Acoustical Society of America* 91. 1707–1717.

Young, Vershawn Ashanti, Rusty Barrett, Y'Shanda Young-Rivera & Kim Brian Lovejoy. 2014. *Other people's English: Code-meshing, code-switching, and African American literacy*. New York: Teachers College Press.

# Section 1: **Conceptual and methodological boundaries**

Michael Adams

# 1 Scale and mode in histories of English

## 1 Introduction

What do we mean when we talk about language "history"? Our dominant model privileges change in linguistic features and accounts for it over fairly long chronological arcs. For instance, one might examine the development of free adjuncts in English and Dutch for three centuries until they diverge in the nineteenth century (Fonteyn and Cuyckens 2015) or attempt to account for the phenomenon we call The Great Vowel Shift (see Giancarlo 2001). We focus on dates and periods even when we distrust such categories (see, e.g., Curzan 2012). We are less attentive, however, to change at either small scales or grand scales, and we rarely conceive of history on terms other than change, though there are alternatives. When we use the term *history*, we often seem to be talking about facts of the past and the knowledge to which they lead but not about modes of telling about the past – there's very little historiography of language history. We might productively reconsider scales and modes of "history" in histories of the English language, that is, examine the generic boundaries of our historiography critically and then redraw the terms on which we write the history of English.

Social, political, and cultural history offer models of methods and narrative mostly absent from language history – their boundaries not only differ from those of English historical linguistics but reach beyond them in several directions, both of scope and variety. While a historian might write about England in 1603 CE (Lee 2003) – putatively everything worth knowing about England in that single year – we do not write histories of English in 1776 or 1901. While a historian might write about Mediterranean culture over millennia (Braudel [1966] 1972) – its economies and other mechanisms of exchange – we do not write about the linguistic markets of what are now England and the Netherlands over the millennia. At a radically different scale, we rarely write the history of yesterday's English or the history of one person's English, although, of course, sometimes we do (e.g., Tieken-Boon van Ostade 2011: 137–253, and 2014).

There is nothing wrong with the currently favored modes and scales – language history should be done in that fashion but not only nor even mostly in that fashion. There are several reasons to reconsider historical modes and scales.

**Michael Adams,** Department of English, Indiana University, Bloomington, IN, USA,
e-mail: adamsmp@indiana.edu

https://doi.org/10.1515/9783110643282-002

For instance, the current modes and scales appeal to professionals, but narrative engages a much wider audience, and we might introduce language history into the intellectual lives of educated people by means of it.[1] It would be worth accounting for the English of a lifetime, at a human scale, in other words – to bring the language into the context of living and thus make what linguists know about it newly relevant to the inheritors of English and its history. Yet, for the professionals, history told in different modes and at different scales facilitates unexpected connections among language phenomena, as well as connections between language-internal history and cultural or "outer" history (see Millward 1988), the effective synthesis of which should figure significantly in a language historian's work (and, for essentially the same reasons, in teaching of the history of English, for which, see Adams 2012).

## 2 Conventions

All writing depends on conventions of argument, structure, and style, which together constitute generic boundaries in professional discourse. Within an academic discipline, writing conventions tend to be formalized and specific. (Try to convince a linguist to capitalize words in article titles in a Reference section and you will elicit evidence of both qualities.) There is nothing wrong with disciplinary and even sub-disciplinary writing conventions. Indeed, they are to be expected, not just to promote efficiency in both production and reception – everyone knows what to expect – but because conventions define the social group of researchers in a particular field, that is, knowing and following the conventions are in-group activities that promote group identity and solidarity. Conventionally, you might

**1** *Narrative* is a complex term, and I use it here in its broadest possible implications, so as not to exclude modes of explanation that might align at least as well with *narrative* as with *analysis*, though those two categories are not mutually exclusive. Its value to the present discussion is rooted in the polysemy of Greek *ἱστορία* and reflected in French *histoire*, both of which link concepts of history and story (see, e.g., Zgusta 2006: 1–5) to the extent that each depends on event structure (Herman 2002: 43–45; Zgusta 2006: 5), which is the key to narrative and narrativity. Different modes of history evince different degrees of narrativity, as does each essay in each mode – narrativity is a variable property (Herman 2002: 100–105). Some would argue that many elements of stories – for instance, settings and characters – lie outside of narrative, but such elements figure in historical modes described here, and when I refer to a mode or text within that mode as "fully narrative", I mean that it operates with the full repertoire of storial elements.

expect me to cite work by John Swales (e.g., 1990) at this point, but anyone reading this chapter, I believe, will know from experience the truth of what I just wrote. Of course, we bend or break conventions regularly – traverse discourse boundaries – for all kinds of rhetorical reasons. As Devitt (2015: 51) has argued, "genres are grounded in shared communicative purposes and realized in particular linguistic utterances", and, I would add, we need to promote "improvisation" (Devitt's term) that challenges generic and stylistic conventions and expands the domains of disciplinary discourses.

Consider, for the sake of example, just one subdisciplinary template, that of historical sociolinguistics. The typical – and, yes, it could be said, stereotypical – article begins with a research question, followed by a literature review, perhaps, or a section on method or experimental design, another that presents data, one that analyzes the data, sometimes, when the disciplinary approach tips towards sociolinguistics, a too-short section titled "Discussion", and a conclusion that points helpfully towards further research. Chapters in books often follow the same template. We employ it so often, not because it is habitual and mindless – although, of course, it can be either or both – but because it works well for us. It has certain virtues. For instance, it tends towards comparability of method and data such that research participates in a developing conversation on agreed intellectual and methodological terms, and it expects that researchers account for their data.[2] Nevertheless, it also tends to homogenize historical scale and mode, even given Devitt's (2015) point about the variety and individuality of performances with in a genre (see also Devitt, this volume).

Templates that generally govern the writing of historical linguistics reflect a disciplinary emphasis on *linguistics* and a tendency to interpret *historical* in its weakest sense (see Considine 2015). Here, I suggest that we should shift emphasis more often, so that historical linguists sometimes write more like historians than

**2** Philosophically, the case for analysis over narrative, represented by the template, is still more profound. With Dray (1985: 130), "let us ask whether there is some kind of incompatibility, or at any rate serious tension between the ideas of narrating and of adequately explaining a sequence of events. The concern is presumably that narrative, being necessarily linear, is explanatorily thin; analysis, relieved of the obligation to tell a story, can be as thick as you please". One may resist narrative history from this reasonable principle. In what follows, however, I suggest that no single mode or scale by itself adequately explains events and that narrating is compatible with explanatory adequacy because it provides information essential to adequacy, even if it is not self-sufficiently adequate. Further, analysis and narrative are thus complementary explanatory modes, far from incompatible between themselves in history, and, indeed, compatible within a multimodal historical text.

like linguists. In doing so, we would find ourselves not only following different conventions – history written by historians who aren't also linguists has conventions, too – but in negotiating the respective shares of history and linguistics in any given case developing interesting idiosyncratic forms of historical linguistic writing. Here, I will present several models – from both history and linguistics – of historical writing that challenge our current disciplinary conventions, the conceptual boundaries we observe and reinforce in our disciplinary practice. Some of them are outside the mainstreams of both history and linguistics, though not far from the mainstreams, but I think their adventurous stances are more likely than history-as-usual to provoke thinking about historical scale and mode in our own linguistic historiography. I am not arguing to a conclusion, really, nor am I developing a typology of modes and scales in the writing of English historical linguistics. Instead, I am performing a heuristic exercise in historiography that I hope will lead to yet more imaginative practice of our discipline.

## 3 Varieties of scale and mode

One can write multi-volume histories of long periods, or large areas, or big events, or all three together, as in the decline and fall of the Roman Empire (Gibbon 1776). Or one can write shorter histories focused on social movements within much shorter timeframes –Hill's *The world turned upside down* (1972) – or the history of one life – Hill's *God's Englishman: Oliver Cromwell and the English Revolution* (1970) – or historical episodes of quite short duration – Steil's *The Battle of Bretton Woods* (2013). Indeed, one can write still shorter histories, in articles rather than books, and even in footnotes to articles and books not otherwise historical. Political and social history is written at many scales, and so we might ask whether language history can similarly scale large and small, as well as scaled as convention leads us to scale our histories. Obviously, too, history appears in many modes, from the fully narrative to the lexicographical, though, I will suggest, historical lexicography is an unexpectedly narrative genre. We can measure every work of history along lines of scale and mode, and the two almost necessarily intersect. Here follow examples of some scales and modes less familiar in language than in other types of history. I identify examples of each in language history, too – although, on the periphery of our practice – suggesting that some of us already feel the impulse to vary scale and mode in the histories we write.

## 3.1 La longue durée

At the widest scale, we might consider a classic of the *annales* school, Braudel's *The Mediterranean and the Mediterranean world in the age of Philip II* ([1966] 1972: 23).[3] At the "beginning" of the book, Braudel writes,

> The first part of this book, as its title suggests, is concerned with geography: geography of a particular kind, with special emphasis on human factors. But it is more than this. It is also an attempt to convey a particular kind of history. Even if there had been more properly dated information available, it would have been unsatisfactory to restrict our enquiries entirely to a study of human geography between the years 1550–1600 – even one undertaken in the doubtful pursuit of a determinist explanation. Since in fact we have only incomplete accounts of the period, and these have not been systematically classified by historians – material plentiful enough it is true, but insufficient for our purpose – the only possible course, in order to bring this brief moment of Mediterranean life, between 1550 and 1600, out of the shadows, was to make full use of evidence, images, and landscapes dating from other periods, earlier and later and even from the present day. The resulting picture is one in which all the evidence combines across time and space, to give us a history in slow motion from which permanent values can be detected. Geography in this context is no longer an end in itself but a means to an end. It helps us to rediscover the slow unfolding of structural realities, to see things in the perspective of the very long term.

I quote at such length because, as far as I know, no contemporary historian of English language approaches her subject from this angle. I am not suggesting it as a model, exactly, but as a prompt to experiment with scale and mode. What would it mean to take on a similarly big subject in English historical linguistics?

We write very little about English in geographical regions – I'm thinking here only of Inner Circle English (see Kachru 1982) – that reflect complex international political, social, or economic relationships at one time or over some span of time, though English is rarely isolated within a national boundary. Some of our colleagues have considered the relationship of Celtic languages to English in a maximized region of relevance (Filppula, Klemola, and Paulasto 2008). Townend (2002) has written about the Norse/English connection, and Bremmer (2017) has been talking and writing about relations between English and Frisian – given that last example, obviously we can traverse complexity in works of various size, not only

**3** The *annales* school of history was founded by French historians Marc Bloch and Lucien Febvre with their journal *Annales d'histoire économique et social* in 1929. Other members of the *annales* school familiar to Anglophone readers include Jacques Le Goff and Emmanuel Le Roy Ladurie. *Annales* historians disdain events, the stuff of chronicle, and prefer to investigate *mentalités*, or world-views. While *la longue durée* is associated with the *annales* school – and rightly so, since it introduced the term to characterize its preoccupations – not all *annalistes* write only long-term history, as we shall see later in the case of Le Roy Ladurie.

in books. But I can think of nothing that sweeps across time and space in a way that resembles Braudel but takes in all of the North Sea. That would certainly require extended treatment, if only to arrange what we've learned from shorter feature studies into a narrative frame. Why can't "we" write a history of language focused on the North Sea, or on southeast England and the Low Countries, language of a millennium rather than part of a century? Why couldn't such a work embed analysis in narrative, as history on the *annales* model tends to do?

Recently, we have seen some work in the history of English on the scale of *la longue durée*. For instance, D'Arcy's *Discourse–Pragmatic Variation in Context: Eight Hundred Years of LIKE* (2017a) claims that scale in its title. And while the book's mode is not uniformly narrative, it includes narrative passages. Of the many different pragmatic uses of *like* over centuries, D'Arcy (2017a: 47) writes, "[t]hey represent neither the rogue liberties of youth nor contemporary folly, but the transmission of linguistic features from one generation to the next", a conclusion one can reach validly only after accounting for use of *like* over centuries, attending to variously scaled diachronic and synchronic wheels within wheels of history (D'Arcy 2017b).

Even if you don't find Braudel's example compelling, the theory of history he articulates in the preface to his big work bears consideration. Later, he would elaborate the underlying principles: "History takes place at various levels, I would willingly say three levels, but that's just a manner of speaking, and it oversimplifies things. There are ten, one hundred levels to consider, ten, a hundred spans of time. On the surface, a history of events stands for the short term: it is microhistory" (Braudel 1969: 112).[4] One notices immediately how well Braudel's (1969) view aligns with Curzan's (2012: 1233): "The history of the English language, in theory, could be broken up into as many periods as the imagination allows. Defined historical periods could span a year, a decade, a generation, a century, or multiple centuries. Or, again in theory, the history of English could be told as one continuous narrative, with no defined historical periods at all". But, for Braudel (1969), these levels are not discrete but mutually implicated: "In fact, the spans of time we distinguish are in solidarity, one with the others: we do not so much create the broad span of time in our minds as we do the segments

4 My translation of Braudel (1969: 112): "L'histoire se situe á des paliers différents, je dirais volontiers trois paliers, mais c'est façon de parler, en simplifiant beaucoup. C'est dix, cent paliers qui'il faudrait mettre en cause, dix, cent durées diverses. En surface, une histoire évenémentielle s'inscrit dans le temps court: c'est une micro-histoire".

of it" (Braudel 1969: 117).[5] Intellectually, one cannot deny the relevance of any level, but that does not mean one can write a history of all levels all at once. Braudel (1969) attempts to weave history out of many levels, but not all of them.

One may very well reject the totalizing assumptions of the *annales* method and perhaps even its plausibility. Can we really, as Braudel (1969: 114) puts it, "saisir l'ensemble, la *totalité* du social" (the emphasis is original), that is, 'to grasp all of it, the totality of social life'? And while he acknowledges that history can be subdivided into all sorts of scales, he is absolutely against attending to short-term history, which militates against a number of historical modes I will suggest momentarily: "To transcend the event is to transcend the short span of time in which it occurs that which belongs to chronicles or to journalism – those rapidly passing moments of awareness, from day to day, whose traces render such a vivid sense of events and lives of the past" (Braudel 1969: 103).[6] Today, many cultural historians would question such a claim if not dismiss it outright, and so would many historians of English. But, with all his talk of transcendence, Braudel (1969: 75) admits that "[i]ndeed, the historian can never avoid the issue of time in history: time sticks to his thinking like dirt to the gardener's spade".[7] He excuses himself with a challenge to his discipline that might also apply to our own, overlapping one: "I also hope that no one will reproach me for being overly ambitious in my desire and need to take the long view. History should not be condemned to the study of well-walled gardens" (Braudel 1969: 13).[8] And so, we historians of English may also need to take our spades out into open fields – beyond disciplinary fences – to turn new soil, whether we follow the *annales* method or take directions Braudel would not approve.

**5** My translation of Braudel (1969: 117): "En fait, les durées que nous distinguons sont solidaires les unes des autres: ce n'est pas la durée qui tellement creation de notre esprit, mais les morcellements de cette durée".

**6** My translation of Braudel (1969: 103): "Dépasser l'événement, c'était dépasser le temps court qui le contient, celui de la chronique, ou du journalism – ces prises de conscience des contemporains, rapides, au jour le jour, dont les traces nous rendent, si vive, la chaleur des événements et des existences passes. Autant se demander si, au delà des événements, il n'y a pas une histoire inconsciente cette fois, ou mieux, plus au moins consciente, qui, en grande partie, échappe à la lucidité des acteurs, les responsables ou les victimes: ils font l'histoire, mais l'histoire les emportes".

**7** My translation of Braudel (1969: 75): "En fait, l'historien ne sort jamais du temps de l'histoire: le temps colle à sa pensée comme la terre à la bêche du jardinier".

**8** My translation of Braudel (1969: 13): "J'espère aussi que l'on ne me reprochera pas mes trop larges ambitions, mon désir et mon besoin de voir grand. L'histoire n'est peut-être pas condamnée à n'étudier que des jardins bien clos de murs".

## 3.2 The English of a lifetime

I have long been attracted to what Braudel (1969) calls microhistory, the sort of history that addressing *la longue durée* is supposed to transcend. I don't put much stock in transcendence, and I think it is all too easy to overlook the importance of small things in the grand design of history. Mostly, our personal histories, yours and mine, are small and intersect with histories on other scales, but bigger is not necessarily better, and historical writing should not always attempt to paint the big picture. Many historians of English already write on a microhistorical scale. Feature studies are often micro, and another robust tradition of microhistory in philology is the scholarly edition, focused on just one text focused on one person, as in Henry Machyn's diary (Bailey, Miller, and Moore 2006), or on a single event, like the Salem witch trials, as viewed through trial records (Rosenthal et al. 2009).[9] Historians who are not historians of English, however, write in some other microhistorical modes worth our consideration; on occasion, historical linguists have adopted these modes, yet their example has not been taken up with any regularity and remains peripheral to convention.

Historical linguists extract much data from English letters, but historians who are not linguists construct narratives out of letters. So, Ozment wrote *Magdalena & Balthasar: An intimate portrait of life in 16$^{th}$-century Europe revealed in the letters of a Nuremburg husband and wife* (1989), which he introduces in this way: "Intrigued historically by the editor's complaint, I plunged into the letters, spurred on by the rare prospect of hearing two ordinary, but literate, people speak their minds intimately and at length over a sixteen-year period. Thus began my long acquaintance with the Nuremberg Balthasar Paumgartner and his wife, Magdalena Behaim" (Ozment 1989: 12). Conceivably, one could write the history of a relationship between two ordinary people in which language took a much more prominent role than it usually does in biography; instead of worrying about business details, one might focus on how language participates in constructing the relationship and how language contextualizes

**9** Grund et al. (2009) account for historical linguistic aspects of the Salem witch trial records in an introductory essay, a conventional explanatory genre within scholarly editions. Another often overlooked genre of historical writing, the footnote, is also frequent in philology, but many view footnotes as dispensable rather than as significant historical writing. A robust defense of the footnote is possible (Grafton 1997), and historians of English should consider them as sites for self-contained stories of the language. Also, should a historian of English work in a cognate field – literature or linguistics, for instance – works on those subjects might occasionally bear supplement by footnoted microhistories of English.

both the relationship and its members, as in Fitzmaurice's (2011) investigation into the sociopragmatics of a long-ago lover's spat.

And the speakers don't even have to be ordinary. Newsome (2017) has shown how the pragmatics of Margaret Tudor's letters – holograph and not – manage familial relationships but also participate in diplomacy and political alliance. The queen's "voice", which we can examine linguistically, of course, speaks at historical scales outside of linguistics, the regnal and dynastic and cultural periods to which political and social historians respond, but with which language histories rarely intersect. Newsome's (2017) story could be told at a micro scale and engage various other histories while never abandoning linguistics, and it can be told, like Ozment's (1989) story, in a narrative mode.

Sometimes, we find recognizable historical figures who illustrate language or language attitudes of their times (e.g., see Tieken-Boon van Ostade 2011, 2014). Writing about them can focus linguistic issues but also contextualize them beyond what we usually consider the language system. Mugglestone (2011: 88) opens an essay as follows:

> [Richard W.] Bailey's own work has, of course, consummately served to reorient the linguistic gaze, revealing the nineteenth century as a period of "unprecedented linguistic self-consciousness" with significant – and distinctive – shifts in language as social and cultural practice [. . .]. This essay will explore one particular manifestation of this "unprecedented linguistic self-consciousness" – here by means of the (equally neglected) historical intersection of Michael Faraday, the largely self-educated son of a blacksmith who was to become one of the foremost scientists of the nineteenth century, and Benjamin Smart, one of its most prominent elocutionists.

If we wrote more microhistories of this kind, the history of English would become, however briefly, the subject of many a biography. Intertextually, then, we would be writing more history – the history we write would be re-written into different historical scales by other historians. Most important, though, examples like these remind us that one can resist the template and *la longue durée* by writing about language within the span of a human lifetime, a perfectly legitimate scale of inquiry that naturally appeals to both professional and amateur historians of English.

While we are on the subject of biography, it is worth noting that even an autobiographical history of English is not inconceivable. Consider how Sonja Lanehart announces what drove her to write *Sista, speak!: Black women kinfolk talk about language and literacy* (2002: 1): "Maya, Grace, Reia, Deidra, Sonja: all African American women in one family – my family – whose stories have spoken to me for as long as I can remember. They were purposely chosen for this study because I want to share their stories with others, who I know have mothers, grandmothers, sisters, and daughters just like mine [ . . . .] I present these stories

as narratives. Their narratives are important because they are stories to be told and not forgotten". Lanehart's synchronic experience of her language depends on its role in a diachronic experience measured in generations of a family, understood on personal terms.

Lanehart's (2002) book is narrative sociolinguistic inquiry into African American English and language attitudes of a few generations of its speakers, generations that can be known in a lifetime – a topical-generic boundary-breaker if there ever was one. But we should also consider autobiographies of a more traditional mode, like Crystal's splendidly written *Just a phrase I'm going through: My life in language* (2009), both their role in language history and their historiographical example. Some historians of English resist the notion that the history of historians deserves a place in the history of the language, but I insist that it does, that the language attitudes of which we should be most aware – and of which others should be most aware – are our own.

Also, accounts of what we knew and when is far from irrelevant to the history of English; we need to assess methodologies, our modes of writing history, just from such autobiographical witness. Tagliamonte's astonishingly innovative history of variationist sociolinguistics, *Making waves* (2016) – which weaves together narrative history and frequent quotations from interviews with prominent variationists – shows just how compellingly the testimony of sociolinguists of a certain stripe can be brought to bear on language matters. A similar case, of different significance, could be made from the testimony of historical linguists. At any point, an autobiographer might write about the usual material of histories of English: the author's own production of features, perceptions of others' production, developments of variation, change noted, if not systematically. In one dimension, however, autobiography exceeds all other modes at all other scales – it explains beyond refutation how English matters to the autobiographical subject, whose English readers easily compare and contrast with their own.

## 3.3 Annals and other microhistories

If personal history seems too far from "objective" history for comfort, we can reduce scale, not by talking about all things linguistic in a person's life, but by talking about all things linguistic in a single year, a historical mode exemplified by Lee's *1603* (2003: 1): "This is the story of the year 1603. It is therefore the tale of one of the great step changes in the history of these islands. Noted historians, for example G. R. Elton, thought 1603 unimportant. Yet the events of any twelve-month period are rarely unimportant [. . .] Our view of 1603 starts in its late winter when in the bedraggled early hours of 24 March, Elizabeth the first turned her face

to the wall". Literary historians also have had recourse to this scale, for instance, James S. Shapiro, a serial perpetrator of annual history in *1599: A year in the life of William Shakespeare* (2005) and *1606: William Shakespeare and the year of Lear* (2015). If you wanted to write the history of a year's English, you could avoid all the methodological problems of periodization, mystification, and Whiggism by choosing a year – 1605, perhaps – whose linguistic events were far from unimportant, but on which no other argument depended.

If one questions the validity of the annal as a historical linguistic scale, perhaps because it seems too arbitrary, that is, because language does not change according to calendars, one might instead consider the language made public in an event, on what we might call the event scale. Carlo Ginzburg, an avowed micro-historian and so anti-*annales*, provides an example in his popular *The cheese and the worms* (1980), published originally in Italian in 1976. The book is about the strange, subversive, and heretical cosmology of a sixteenth-century Italian miller named Menocchio, not the *mentalité* of the *annalistes* but an idiosyncratic worldview. While it limns the social background, it foregrounds Menocchio's trial, such that Ginzburg will shift focus in a transition like this one: "Let's put aside Menocchio's conception of the universe for the time being and follow instead the progress of the trial" (Ginzburg 1980: 7) – the trial keeps the account on the event scale. The event scale is, in fact, so beguiling that some *annalistes* have failed to resist its charms: Emmanuel Le Roy Ladurie followed his eminently *annaliste Peasants of Languedoc* (1977), originally published in French in 1966, with the event-oriented *Carnival in Romans* (1979), an account of the 1580 Mardi Gras massacre in the city of Romans.

At the micro-est level, one can write the microhistories of words, a scale of historical lexicography at which I myself have long been engaged, as a Contributing Editor to the *Barnhart dictionary companion*, effectively the Review Editor of "Among the New Words" while I was editor of *American Speech*, and as the author of *Slayer slang: A* Buffy the Vampire Slayer *lexicon* (2003). All these works track usage of new or newish words from month to month, rather than at the conventional, "even unit" scale of a decade usual in historical dictionaries identified by Fitzmaurice (2017) or yet broader spacings of quotations, as Sheidlower (2011: 204–205) describes the *Oxford English Dictionary*'s (*OED*) editorial practice. The *OED*, of course, represents *la longue durée*, but it cannot represent all English at all scales of use, so we rely on dictionaries of smaller scope to provide the missing historical material. Opportunities to engage English lexis on this scale are practically infinite, and, surely, we would welcome historical treatment of subcultural words, valuable to sociolinguistics, too, where it intersects with the history of English (for the rationale and method underlying this practice, see Adams 2014).

Studying the micro-histories of words is in fact essential, because lexical *actuation* – a crucial event in any word's history – must be approached diachronically, and I would add, at the chronologically finest scale (for examples, see Adams 2018: 94–98, 102–103; Taylor and Christensen 2018; Zimmer and Carson 2018: 69–72). The terms of actuation are obscured in large-scale historical dictionaries. We invest so much in earliest attestations of use that once we think a first date is settled we ignore the process of actuation that follows the first quotation and is complete by the time the next quotation comes up in an entry.[10] The matter is further complicated by the fact that we are often wrong about those first dates (Shapiro 2018), and only micro-historical research can lead to antedatings. It seems unwise to consign the record thus recovered by such research to the dustbin, as though the first date of use were the only interesting or important thing about the origin and early history of a word, or, indeed, any linguistic feature.

In other words, one need not always do historical lexicography over *la longue durée* in multiple volumes. Indeed, one should not, because lexicography at that scale may – no, will – overlook lexical history at a minimized scale. If we don't write histories of English at lots of scales, then all the histories we do write will be impoverished. In a fully articulated history of English, D'Arcy's (2017b) synchronic wheels would spin within diachronic wheels like clockwork. Of course, all history is partial, but that seems a poor excuse for overlooking knowledge about language captured at different historical scales and in different historical modes.

## 4 The narrative mode

Certain themes run through my descriptions of historical scales, modes, and styles less familiar in writing the history of English than they should be. The modes are all at least partly narrative – they tell stories. Braudel's ([1966] 1972; 1969) practice may privilege the largest possible geographical and temporal

**10** *Actuation* is a tricky concept, a mixture of event and process. Of course, in one sense, each time anyone uses a word (or sound, or syntactic structure, etc.), the feature is actuated. Let's call that sort of actuation "*a*". It's the sort of actuation – $a^1$ – that goes unnoticed unless there are other *a*-level actuations – $a^2$, $a^3$, $a^4$ . . . $a^n$ – of the same feature. Mere expression of the feature does not itself ensure actuation at the *A*-level, which depends on other parameters, what Metcalf (2002: 152) calls the FUDGE factors: "**F**requency of use[,] **U**nobtrusiveness[,] **D**iversity of users and situations[,] **G**eneration of other forms and meanings[, and] **E**ndurance of the concept". A word's microhistory – a selection of *a* events such as that for ADJ/N + *much* (Adams 2018: 102–103) – registers on what terms and how well the item satisfies these *A* factors.

scales, and there are people who like to think about human issues at that scale, just as there are people who like to think in geological, astronomical, or evolutionary time. Even with all the economic and social data dear to the *annaliste*, there is a story there, laden and long, but a story, nonetheless. Besides Braudel's, the modes I have illustrated tend towards microhistorical scales. We all have lives, so we are interested in and capable of appreciating Balthazar and Magdalena's lives, or Sonja Lanehart's life. Microhistorical histories of language can get *i*-umlaut in through a back door, so to speak, but they are very likely to engage language ideology, and I believe many historians of English – even those devoted to *i*-umlaut – are willing to compromise delivery of some facts if they can capture the attention of various audiences, from students in classrooms to a public always ready to hear stories about English. We take history best when we see its relation to our own experience, but also when it is contained at a humanly comprehensible scale. We all know what a year is; most of us know family from previous generations; a few of us even know what it feels like to write letters in our own hands.

I am not prone to prescription, but I admit that I wince, internally, when students say that a poem or an essay is "relatable". A good teacher breathes and asks, "What do you mean, relatable? *How* does it relate to you? What do you make of it?" I resist the vagueness, but I listen to students anyway. They want something. They need something I can give them – a story of English that speaks to them. That they prefer narrative accounts of language or anything at all comes as no surprise, because narrative is more natural – I mean this cognitively and culturally – than analysis.[11] We can locate significant knowledge about language on a scale that speakers, students, people understand. Analysis can be banked until someone cares to withdraw it, but history requires a listener or reader or viewer. We should ask ourselves, when we are doing historical English linguistics, are we writing history or linguistics about language of the past? Does that matter? It does, if we want a wide audience, and if historical linguists will be public intellectuals.

**11** Indeed, stories are, as Boyd (2009: 9) puts it, a "human universal", and it's thus no surprise, in linguistic as well as cultural terms, that we tell our children folk tales to arm them in a hostile world, rather than require that they attend after-dinner PowerPoint presentations of crime statistics. Boyd (2009: 15) actually dismisses history from his story about stories: "true stories need little special explanation", he argues, "and language allows humans to share information about the past", a rather naïve view of narrative history that avoids the question of narrative's historiographical potential. Still, he proposes the evolutionary position and value of narrative, which we might apply to narrative history as well as fiction, despite Boyd's position.

Audience raises other questions of mode and scale. Although younger colleagues may be well-advised to write only research articles for publication in leading journals (or, in some departments, books) in order to secure their careers, we no longer need assume that serious discussion of language is possible only in those modes. Blogs abound and some are amazingly good at doing linguistics professionally but publicly – some with several writers, like *Language log*, *Lingua franca*, and *Strong language*, and some singly authored, like Gretchen McCulloch's *All things linguistic*, Neal Whitman's *Literal-minded*, and others come to mind. In other words, between the book and the blog, we can write the history of English at a variety of intermediate scales variously suitable for subjects and audiences.

Blogs accommodate many rhetorical positions: polemical, narrative, explanatory, and even analytical. For the blog-reading audience, each such position requires a voice different from that one usually hears in journal articles. Pretty much everyone has read a blog post or two these days, but considering the array of possible voices, one must consider whether one writes equally well in all of them, and in any event writing well in any of them except the analytical voice we employ in the standard research mode takes experience and practice. This would seem to be the greatest obstacle to the narrative experiments I have been promoting here. There is no guarantee that a linguist of historical English reads narrative history, and, without that experience, such a linguist might not know what to practice in developing a historical style. The brief quotations from narrative history sampled here are no substitute for reading non-linguistic history deeply and widely, but they do signal the ear that some non-linguistic modes are in the air.

Even historical lexicography – whether of *la longue durée* or microhistorical scale – is a form of narrative. The event status of quotations as we present them in historical dictionaries motivates narrative interpretation – that is, the writing of history. The entries aren't narratives per se, but lexicographers think narratively when compiling evidence into an entry. Thus, an entry implies narratives anticipated by its author. Just as important, the reader will construct narratives from the entry, or, rather, co-creates the narrative with the lexicographer. The beauty of the historical dictionary genre is that it invites reading rather than mere consultation.

Herman (2002: 91) describes the narrativity I have in mind, though he does not discuss the narrativity of dictionary entries specifically:

> Sequences that have a minimal narrativity, which distinguishes them from nonnarrative sequences (of zero narrativity, by definition), are less readily configured into chronologically and causally organized wholes – less readily interpreted as stories – than are sequences with higher degrees of narrativity. Narrativity is a function of the more or less richly patterned distribution of script-activating cues in a sequence. Both too many and too few script-activating cues diminish narrativity.

Dates and the chronologies constructed with them, not exclusively but significantly, serve as script-activating cues in historical narrative. Arguably, historical dictionary entries are sequences of minimal narrativity, because they include too many script-activating cues – indeed, they may be composed of nothing but script-activating cues. But when readers narrate, they spread the cues within the developing story they inspire, which slows their occurrence and enables a higher degree of narrativity.

Narrative essays, books, blog posts, and footnotes about the history of English, sometimes biographical or autobiographical but more often in third-person historical voices, as well as historical dictionary entries of whatever scale, all draw the linguistics we do closer to the practice of history.

## 5 Multimodal histories

One might misinterpret my argument as suggesting that there is a templated way of researching and writing about the history of English, on one hand, and a flamboyantly narrative way, on the other, and that one must choose between the two. But examples presented here suggest several possible modes and scales of narrative history, which might productively be combined in a historical text about English. Indeed, one could present systematically collected and analyzed data in a narrative frame – typical of the *annales* approach – or one could alternate sections/ chapters between the two modes. Or, one could be yet more adventurous formally, sometimes even creating modes while mode-shifting, thus looking at historical problems from alternative perspectives within the same text.

Consider again Tagliamonte's history of variationist sociolinguistics, *Making waves* (2016). Tagliamonte begins by outlining the movement's origins in annals: William Labov's Martha's Vineyard study (1960) (Tagliamonte 2016: 2), other events in the intervening years, "William Bright's Conference – 1964" (Tagliamonte 2016: 6–7), "The LSA Summer Institute – 1964" (Tagliamonte 2016: 7–8), some more essential events, "NWAV 1 – 1972" (Tagliamonte 2016: 21–22), etc. The annals are narrative, as you'd expect. Here's a slice of the one about NWAV (Tagliamonte 2016: 21): "[i]n the early 1970s the new Sociolinguistics program was getting going at Georgetown University with Roger [Shuy] at the helm. The new faculty were hanging out and enjoying their lives as sociolinguists. They all had nice offices and ideas were flying. They came up with a plan to bring like-minded people together", that is, in NWAV.

Later, though, in a chapter titled "A Crescendo of Research", the narrative pace and perspective sometimes change interestingly. Here is how Tagliamonte (2016: 57) begins her section on Anthony Kroch's Constant Rate Effect:

> Tony Kroch is working at Penn alongside Bill [Labov]. He's wondering about how to combine what he knows about generative grammar with VSLX [variationist sociolinguistics] methods [. . .] He is also intrigued by a big problem: the paradox between different schools of thought in Linguistics and how they handle language change. For generative grammar, gradual diachronic development is untenable. For VSLX, it is an empirical fact.

Then, VSLX's version is depicted in a graph. So, in order to reconcile the two positions, "Tony thinks to himself, 'What I need to find is a diachronic development of something syntactic", so "Tony runs a regression on the different contexts of *do* – questions, declaratives, negatives, and so on" (Tagliamonte 2016: 58). But while shifting from past to present narrative, Tagliamonte sticks with small-scale accounts of historical events, rather than transferring the template onto the history of a discipline – possible, but much less manageable, elegant, provocative.

Among these narratives, Tagliamonte (2016) intersperses testimony from linguists culled from her interviews with them, and they at once advance the book's argument but also illustrate language attitudes worth taking into a historical account of sociolinguistics, as well as English in our time. Here is the sixth quotation taken from Tagliamonte's (2016: 124) interview with Joan Beal:

> So, I think there's a personal thing and I think there often is. I think you know if you look at people who are in Sociolinguistics they're very often people who, you know, have come from a background where they don't speak with a standard variety and they have had personal experience – discrimination would perhaps be too harsh a word. But they've had experience of the kind of attitudes you can get. And I think that it can very well be a very personal thing.

In many instances, Tagliamonte (2016) quotes serially from interviews with two or more linguists, without providing any intervening, interpretive narrative. The approach is as rare in history generally as it is in histories of English.

Peter Gilliver mixes modes in his recent biography of the *OED*, also published in 2016, apparently a good year for historiographical innovation. The book chronicles the *OED* from its origins to the present, not in short-takes but in stretches, such 1861–1875 and 1957–1972. The stretches are at a human, generational scale, but of course the story of the *OED*'s first edition is not, let alone the continuing project. The overarching mode is narrative, yet two interludes – "The work of Furnivall's sub-editors" and "Method: from quotation slip to published entry" – switch into an explanatory mode. In addition, in order to explain lexicographical preoccupations more vividly, accounts of editorial work on specific words, distinguished from the rest of the text in shaded boxes, accompany the narrative history,

and vary the reader's perspective on and points of contact with that history. It is an ingenious way to write history, and one can imagine transferring it easily to other works on the history of English, some of them discussed here and some of them as yet unrealized in the imaginations of historical linguists and philologists.

## 6 Conclusion

Can we write the history of English at a personal scale in a narrative mode? Is that plausible? I think it is. But doing so – indeed, writing in any number of unusual modes at a potentially infinite number of scales – requires us to think and practice outside of disciplinary boundaries. As illustrated here, those boundaries are fluid enough to accommodate unusual history, dotted rather than solid lines on the disciplinary map. So, one might ask, why challenge them? Aren't they essential to the political geography and topography that distinguish disciplines across the sciences from one another? By implication, aren't they also important to disciplinary identity? Must historical linguists practice history as well as linguistics? The answer: English historical linguistics would be a richer and more significant discipline if more but not all historians identified as historians first and linguists second, rather than the other way around. Further, we should not assume that historians always color within the lines, and they may very well redimension history, moving generic boundaries we thought were fixed. We need not deny the value of boundaries while also challenging them and operating across or beyond them.

One boundary on which we've depended too restrictively is that between inner and outer histories of English, between language-internal change and language in and of a culture. A linguist writing the history of English tends to favor the former, and we need more historians to develop the latter, yet what we most need are linguistic historians and historian linguists who nimbly traverse that boundary. Consider, for instance, Donka Minkova's contribution to this volume. Therein, she identifies a paradoxical space in which boundary signals – a linguistic matter – are difficult to distinguish from supposedly extra-linguistic metricality judgments, which are potentially motivated by all sorts of cultural experience and expectation, yet to be described. In such cases, there is an articulated relationship between something inner and outer and thus, in principle, the linguistic problems at issue can be approached in an outer history mode – the story is told from outside the box without ever being disconnected from the linguistics inside the box. Arguably, more often, we might think and write outside the box.

For a long time, from a series of contingencies, the history of English has resided in English departments, but it is fair to say that, nowadays, many English professors do not see the study of English language as central to English studies. Some historians of English – English historical linguists but also philologists – have wondered whether the history of English should move from English to Linguistics. I doubt that English historical linguistics would be much more comfortable in Linguistics departments – English is almost literally the last thing linguists study. Perhaps we need to note that there is no "linguistics" in "The History of English". Why wouldn't we move from English to History? Why don't we historians of English language write more like historians of things other than language? If the answer is easy, I suspect we aren't taking the question as seriously as we should.

## References

Adams, Michael. 2003. *Slayer slang: A* Buffy the Vampire Slayer *lexicon*. New York: Oxford University Press.

Adams, Michael. 2012. Resources: Teaching perspectives. In Alexander Bergs & Laurel J. Brinton (eds.), *English historical linguistics: An international handbook*, vol. 2, 1163–1177. Berlin: Mouton de Gruyter.

Adams, Michael. 2014. Vocabulary analysis in sociolinguistic research. In Janet Holmes & Kirk Hazen (eds.), *Research methods in sociolinguistics: A practical guide*, 163–176. Malden, MA: Wiley-Blackwell.

Adams, Michael. 2018. Periodization in historical lexicography revisited. *Dictionaries* 39(1). 75–104.

Bailey, Richard W., Marilyn S. Miller & Colette V. Moore (eds.). 2006. *A London provisioner's chronicle, 1550–1563, by Henry Machyn: Manuscript, transcription, and modernization*. Ann Arbor, MI: University of Michigan Press and the Scholarly Publication Office, the University of Michigan Library.

Barnhart, Clarence L. & David K. Barnhart (eds.). 1982–present. *The Barnhart dictionary companion*. Springfield, MA: Merriam Webster (1982–2001); Hyde Park, NY: Lexik House (2001–present).

Boyd, Brian. 2009. *On the origin of stories, evolution, cognition, and fiction*. Cambridge, MA: Belknap.

Braudel, Fernand. 1969. *Écrits sur L'Histoire*. Paris: Flammarion.

Braudel, Fernand. 1972 [1966]. *The Mediterranean and the Mediterranean world in the age of Philip II*. 2 vols. Trans. Siân Reynolds. New York: Harper & Row.

Bremmer, Rolf H. 2017. Latin loans in Old English and Old Frisian: Parallels and divergences. Paper presented at Historical English Lexicography and Lexicology (HEL-LEX) 5, Zürich, 16–18 February.

Considine, John. 2015. Historical dictionaries: History and development; current issues. In Philip Durkin (ed.), *The Oxford handbook of lexicography*, 163–175. Oxford: Oxford University Press.

Crystal, David. 2009. *Just a phrase I'm going through: My life in language*. Abingdon, UK: Routledge.

Curzan, Anne. 2012. Interdisciplinarity and historiography: Periodization in the history of the English language. In Alexander Bergs & Laurel J. Brinton (eds.), *English historical Linguistics: An international handbook*, vol. 2, 1233–1256. Berlin: Mouton de Gruyter.

D'Arcy, Alexandra. 2017a. *Discourse–pragmatic variation in context: Eight hundred years of LIKE*. Amsterdam: John Benjamins.

D'Arcy, Alexandra. 2017b. Variation, change, and the *longue durée*. Paper presented at Studies in the History of the English Language (SHEL) 10, Lawrence, Kansas, June 2–4.

Devitt, Amy J. 2015. Genre performances: John Swales' genre analysis and rhetorical-linguistic genre studies. *Journal of English for Academic Purposes* 19. 44–51.

Dray, W. H. 1985. Narrative versus analysis in history. *Philosophy of the Social Sciences* 15(2). 125–145.

Filppula, Markku, Heli Paulasto & Juhani Klemola. 2008. *English and Celtic in contact*. Abingdon, UK: Routledge.

Fitzmaurice, Susan. 2011. The sociopragmatics of a lovers' spat: The case of the eighteenth-century courtship letters of Mary Pierrepont and Edward Wortley. In Jonathan Culpeper (ed.), *Historical sociopragmatics*, 37–59. Amsterdam: John Benjamins.

Fitzmaurice, Susan. 2017. "War doesn't change": Modelling change in the social media reception of a concept. Paper presented at Studies in the History of the English Language (SHEL) 10, Lawrence, Kansas, 2–4 June 2017.

Fonteyn, Lauren & Hubert Cuyckens. 2015. The development of free adjuncts in English and Dutch. In Michael Adams, Laurel J. Brinton & R. D. Fulk (eds.), *Studies in the History of the English Language VI: Evidence and method in histories of English*, 15–49. Berlin: Mouton de Gruyter.

Giancarlo, Matthew. 2001. The rise and fall of the Great Vowel Shift: The changing ideological intersections of philology, historical linguistics, and literary history. *Representations* 76(1). 27–60.

Gibbon, Edward. 1776–1789. *The history of the decline and fall of the Roman empire*. 6 vols. London: Strahan and Cadell.

Gilliver, Peter. 2016. *The making of the* Oxford English Dictionary. Oxford: Oxford University Press.

Ginzburg, Carlo. 1980. *The cheese and the worms: The cosmos of a sixteenth-century miller*. Trans. John and Anne Tedeschi. Harmondsworth, UK: Penguin.

Grafton, Anthony. 1997. *The footnote: A curious history*. Cambridge, MA: Harvard University Press.

Grund, Peter, Risto Hiltunen, Leena Kahlas-Tarkka, Merja Kytö, Matti Peikola & Matti Rissanen. 2009. Linguistic introduction. In Bernard Rosenthal, Gretchen A. Adams, Margo Burns, Peter Grund, Risto Hiltunen, Leena Kahlas-Tarkka, Merja Kytö, Matti Peikola, Benjamin C. Ray, Matti Rissanen, Marilynne K. Roach & Richard B. Trask (eds.), *Records of the Salem witch-hunt*, 64–86. Cambridge: Cambridge University Press.

Herman, David. 2002. *Story logic: Problems and possibilities of narrative*. Lincoln, NE: University of Nebraska Press.

Hill, Christopher. 1970. *God's Englishman: Oliver Cromwell and the English Revolution*. London: Weidenfeld & Nicholson.
Hill, Christopher. 1972. *The world turned upside down: Radical ideas during the English Revolution*. London: Temple Smith.
Kachru, Braj. 1982. Models of non-native Englishes. In Braj Kachru (ed.), *The other tongue: English across cultures*. Urbana, IL: University of Illinois Press.
Lanehart, Sonja L. 2002. *Sista, speak!: Black women kinfolk talk about language and literacy*. Austin, TX: University of Texas Press.
*Language log*. http://languagelog.ldc.upenn.edu/nll/
Lee, Christopher. 2003. *1603: A turning point in British History*. London: Hodder Headline.
Le Roy Ladurie, Emmanuel. 1977. *The peasants of Languedoc*. Trans. John Day. Urbana, IL: University of Illinois Press.
Le Roy Ladurie, Emmanuel. 1979. *Carnival in Romans*. Trans. Mary Feeney. New York: George Braziller.
*Lingua franca*. https://www.chronicle.com/blogs/linguafranca
McCulloch, Gretchen. *All Things Linguistic*. https://allthingslinguistic.com/
Metcalf, Allan. 2002. *Predicting new words: The secrets of their success*. Boston: Houghton Mifflin.
Millward, Celia M. 1988. *A biography of the English language*. New York: Holt, Rinehart & Winston.
Mugglestone, Lynda. 2011. Benjamin Smart and Michael Faraday: The principles and practices of "talking proper" in nineteenth-century England. In Michael Adams & Anne Curzan (eds.), *Contours of English and English language studies*, 87–107. Ann Arbor, MI: University of Michigan Press.
Newsome, Helen. 2017. "Vryten vyth my hand": The pragmatic value of Margaret Tudor's holograph letters. Paper presented at Studies in the History of the English Language (SHEL) 10, Lawrence, Kansas, June 2–4.
Ozment, Steven. 1989. *Magdalena and Balthasar: An intimate portrait of life in 16th-century Europe revealed in the letters of a Nuremberg husband and wife*. New Haven, CT: Yale University Press.
Rosenthal, Bernard, Gretchen A. Adams, Margo Burns, Peter Grund, Risto Hiltunen, Leena Kahlas-Tarkka, Merja Kytö, Matti Peikola, Benjamin C. Ray, Matti Rissanen, Marilynne K. Roach & Richard B. Trask (eds.). 2009. *Records of the Salem witch-hunt*. Cambridge: Cambridge Univ. Press.
Shapiro, Fred R. 2018. Confessions of the antedater. *Dictionaries* 39(1). 23–42.
Shapiro, James S. 2005. *1599: A year in the life of William Shakespeare*. London: Faber & Faber.
Shapiro, James S. 2015. *1606: William Shakespeare and the year of Lear*. New York: Simon & Schuster.
Sheidlower, Jesse. 2011. How quotation paragraphs in historical dictionaries work: The *Oxford English Dictionary*. In Michael Adams & Anne Curzan (eds.), *Contours of English and English language studies*, 191–212. Ann Arbor, MI: University of Michigan Press.
Steil, Benn. 2013. *The Battle of Bretton Woods: John Maynard Keynes, Harry Dexter White, and the making of a new world order*. Princeton, NJ: Princeton University Press.
*Strong language*. https://stronglang.wordpress.com/
Swales, John M. 1990. *Genre analysis: English in academic and research settings*. Cambridge: Cambridge University Press.

Tagliamonte, Sali. 2016. *Making waves: The story of variationist sociolinguistics*. Malden, MA: Wiley-Blackwell.
Taylor, A. Arwen & Kjerste Christensen. 2018. Dating and chronology in the lexicology of Mormon English. *Dictionaries* 39(1). 43–59.
Tieken-Boon van Ostade, Ingrid. 2011. *The bishop's grammar: Robert Lowth and the rise of prescriptivism*. Oxford: Oxford University Press.
Tieken-Boon van Ostade, Ingrid. 2014. *In search of Jane Austen: The language of the letters*. Oxford: Oxford University Press.
Townend, Matthew. 2002. *Language and history in Viking Age England: Linguistic relations between speakers of Old Norse and Old English*. Turnhout: Brepols.
Whitman, Neal. *Literal-minded*. https://literalminded.wordpress.com/
Zgusta, Ladislav. 2006. *Lexicography then and now*. [Frederic S. F. Dolezal & Thomas B. I. Creamer (eds).] Tübingen: Max Niemeyer.
Zimmer, Benjamin & Charles E. Carson. 2018. "Among the New Words": The prospects and challenges of short-term historical lexicography. *Dictionaries* 39(1). 59–74.

Amy J. Devitt

# 2 The blurred boundaries of genres-in-use: Principles and implications from rhetorical genre studies for English historical linguistics

## 1 Introduction

Genres are categories that everyday language users know well but that scholars in multiple disciplines have struggled to define. In linguistics, scholars have differentiated what they variously refer to as text types, genres, registers, and styles (e.g., Halliday 1976; Biber 1988; Tsiplakou and Floros 2013). These textual categories and concepts have also received substantial attention in English historical linguistics (for overviews, see Claridge 2012; Taavitsainen 2016). While scholars in this field have always shown an interest in these concepts, "genre" moved front and center with the advent of historical corpora, whose compilers made genre differentiation a cornerstone of their corpus collection, and it has continued to be so in recent projects (e.g., Rissanen 1996; Schmied and Claridge 1997; Taavitsainen et al. 2005). This readily available textual classification has produced a large body of scholarship on the linguistic characteristics of and developments across genres (see Taavitsainen 2016). Research within historical sociolinguistics has further theorized genre as a lens through which to view the correlation between social variation and linguistic variation in historical periods (e.g., Devitt 1989a; Nevalainen and Raumolin-Brunberg 2012; Romaine 2016: 30–32). At the same time, focused research on the nature and theory of genre (and related concepts and terms) in the history of English has revealed that questions remain about how to understand this complex topic (e.g., Kohnen 2001; Moessner 2001; Görlach 2008).[1]

The aim of this chapter is to show what insights from the field of rhetorical genre studies (see 2.2) can bring to this discussion. Görlach (2008: 13) defines

**1** I owe enormous thanks to this volume's editors, Peter Grund and Megan Hartman, for their productive feedback on my chapter and generous sharing of their extensive knowledge and references in English historical linguistics. Any errors, of course, remain my own.

**Amy J. Devitt,** Department of English, University of Kansas, Lawrence, Kansas, USA, e-mail: devitt@ku.edu

https://doi.org/10.1515/9783110643282-003

rhetoric as a linguistic discipline that contributes to text type linguistics, and I argue that this perspective has much to add to current conceptions in English historical linguistics. Like everything else that matters to our study of language, genres are meaningful but complex. To study genre as something more than an artificial construct and to understand its effects on language requires studying genres as they are employed by language users and meaningful to them, even though the resulting concept can at times seem messy. A rhetorical conception of genre grounds its approach in the forms and meanings these categories have for language users and their communities, giving the genre linguistic and rhetorical validity. In her foundational articulation, Miller (1984) defines genres as typified actions that fuse linguistic form and shared context, making contextual variables more reliable and tangible (see 2.2). Rhetorical genre studies also reveal genre's inherent fluidity and instability. Genre-in-use, like all language-in-use, is dynamic, flexible, unpredictable, social, cultural, and imprecise – with blurred boundaries and fuzzy edges.

For this chapter, I will synthesize and highlight five of the field's key discoveries:

1. Genres are recurring types of social actions, blurring the distinction between form and meaning, linguistic and extralinguistic components.
2. Genres are constructed, defined, and named by communities of users, whose boundaries are themselves ever-changing.
3. Genres interact within genre sets and uptakes, blurring the separation of one genre from another.
4. Genres change with every instantiation and are only ever temporarily stabilized, leaving their boundaries forever blurry.
5. Genres are normalizing and ideological, in ways that can blur historical change.

I argue that these key points enable language scholars to investigate form and context as a coherent, interactive whole; to account for relevant context without specifying hidden, unknowable, or static components; to recognize genres through their relationships with other genres and their interactions; to acknowledge the simultaneity of language change and stability; and to attend to genres' influence on users' language critically as well as descriptively.

I devote substantial attention to laying out these five principles and some of the foundational research leading to the discoveries in sections 2–6. These sections facilitate access to this scholarship, demonstrate its foundations, and support easy application of the principles in the future. I also show that some of the principles have received attention in English historical linguistics, particularly in Görlach (2008), who considers some rhetorical research. However, to

my knowledge, the key discoveries have not received attention as a systematic approach to genre in English historical linguistics. In section 7, I illustrate what it means to consider the full range of principles in a study of genre and the history of English. I do so by looking at and reconsidering my approach to and interpretation of my own historical linguistic research (Devitt 1989b). I do not attempt here a full-fledged re-evaluation of my research results, which is not within the scope of this chapter, but I hope to demonstrate the practical potential of systematically considering these key principles, which could potentially raise new questions, invigorate old topics of study, and offer different interpretations of the relationships between genre and language. What this exploration highlights is that genre boundaries are inherently blurry because language-in-use entails such blurriness, and acknowledging that simple fact can open up new avenues of research.

# 2 Defining genres

## 2.1 Genre as linguistic, extralinguistic, and both

Previous linguistic scholarship has worked to define boundaries of genre by locating a number of identifying features and observing ways in which those features are grouped together, though there has been some disagreement as to how much those features should be primarily linguistic or extralinguistic. There is further disagreement about what term to use for these text categories and how the terms correlate with a focus on linguistic or extralinguistic characteristics. I have so far used the term *genre*, but this term is by no means undisputed or uniformly accepted in linguistic scholarship. Indeed, genre, as well as the related text type, register, and style, has been much debated (see, e.g., Kohnen 2001; Lee 2001; Moessner 2001; Görlach 2008; Biber and Conrad 2009). Lee (2001: 41) points out that linguists tend to use text type and genre interchangeably, a point with which Whitt (2018: 1) concurs, and certainly some like Douglas Biber do, as Lee (2001: 45) describes, at times throw their hands up in frustration at trying to distinguish the terms. I use the term genre in part because I am exploring ideas from rhetorical genre studies (see further 2.2). Within this field the term represents a distinctly different concept from the others, a fusion of linguistic form and extralinguistic variables through its capacity as social action.

In linguistic research, English historical linguistics included, then, we find the use of both text type and genre. Although often used without distinction, some who prefer the term text type (e.g., Biber 1988, usually; and Görlach 2008)

tend to emphasize linguistic patterns, while some who favor the term genre (e.g., Yunick 1997; Lee 2001; Moessner 2001) tend to emphasize extralinguistic criteria. However, the separation of the two terms and underlying concepts is often blurred.

Biber (1988) and work adopting the approach of multidimensional analysis (for recent work, see, e.g., Kruger, Van Rooy, and Smith 2019) represent perhaps the most explicitly formal approach. In his ground-breaking study, Biber (1988) charted the co-occurrence patterns of 67 pre-selected linguistic features, using the statistical method of factor analysis. The patterns are interpreted in functional terms along a number of dimensions (e.g., involved vs. informational production), and texts that cluster together in occurrence patterns are interpreted as text types; these text types often cut across genre categories (Biber 1988: 70, 170). At the same time, Biber's continued extensive research into language variation by text type demonstrates repeatedly the significance to language patterns of such extralinguistic and rhetorical concepts as audience and situation.

In English historical linguistics, work by Görlach (esp. 2008) is central, as it has focused on the theorization and diachrony of what he calls text types. Görlach (2008: 105) defines a text type as "a specific linguistic pattern in which formal/structural characteristics have been conventionalized in a specific culture for certain well defined and standardized uses of language". This definition clearly incorporates uses as well as forms of language but at its core conceptualizes text type as linguistic pattern. His analysis of the history of particular text types in the same volume includes dedications, cooking recipes, various types of advertising, church hymns, lexical entries, and jokes, as well as sets of text types in the history of Scots and Indian English. Throughout, he itemizes their textual qualities, both linguistic and "constituent", but he also argues for the significance of the users' knowledge, which goes far beyond form, and interprets the texts in light of their cultural contexts as well as conventional uses.

Other English historical linguists define genre as solely extralinguistic. Rissanen (1996: 231), for example, suggests that "texts should be assigned their genre labels by extralinguistic criteria, such as subject matter, purpose and discourse situation – not by a mixture of extralinguistic and linguistic features". Much of the research that takes this approach appears to follow Biber's (1988: 170) division of genre as defined by extralinguistic and text type as determined by linguistic features (see also Biber and Conrad 2009: 16; Taavitsainen 2016: 274).

Most consistent with rhetorical genre studies' definition of genre (in spite of his use of the term text type) is Kohnen's (2001: 198) perspective (cf. 2.2): "I suggest a definition of text type which includes both functional-situational and formal aspects. Text types may be thought of as dynamic patterns of communication combining aspects of function, context and form". In fact, Kohnen's

(2001) definition echoes one by early rhetorical genre scholars Campbell and Jamieson (1978: 17), who define genre as a "constellation of substantive, stylistic, and situational characteristics". Within rhetorical genre studies, as I will explain in the next section, genre has been defined in a way that fuses both linguistic and extralinguistic components by recognizing that genres are neither form nor context alone but rather social action.

## 2.2 Genre as social action

*Genres are recurring types of social actions, blurring the distinction between form and meaning, linguistic and extralinguistic components.*

In her (1984) foundational article, rhetorician Miller elaborated a semiotic theory of genre as the fusion of form and substance (to use Miller's terms taken from semiotics), redefining genre as "typified rhetorical actions based in recurrent situations" (Miller 1984: 159), a pragmatic rather than solely syntactic or semantic definition. Rhetorical situations encompass exigence – a social motive for discourse – as well as the speaker/writer, audience/reader, and rhetorical, social, and cultural constraints. As situations recur, so too do their associated actions – genres. As Miller (1984: 163) concludes, "[g]enre refers to a conventional category of discourse based in large-scale typification of rhetorical action; as action, it acquires meaning from situation and from the social context in which that situation arose". Genres are not the sets of conventionalized features but the actions they perform, whether eulogizing, sermonizing, joking, recommending, recipe-ing, making a scholarly argument, or performing multiple actions within one genre.

Drawing similarly on semiotics and speech act theory, Freadman (1994) explains the pragmatic nature of genres as actions through the metaphor of tennis. Genre is a game, like tennis. Although it has conventionalized rules for playing the game, genre is constituted not by the rules but by players taking shots. Players must not only take shots but do so in particular locations with another player across the net to return the shot. Someone describing the rules of tennis or bouncing a tennis ball against a garage wall with a racket isn't "playing tennis", any more than someone who utters the words "George Bailey was a good man" is giving a eulogy. Like speech acts and utterances, genres are language made meaningful by users acting in situational contexts. Genre exists not as a set of linguistic features tied to sociocultural variables but as performances enacting those features in socially recognized ways.

Communicative purpose is a privileged component of situation for John Swales, a foundational genre scholar and applied linguist. While similarly

conceiving of genre as "comparable rhetorical action" (1990: 58), Swales emphasizes the purposes of genres within "discourse communities": "[a] genre comprises a class of communicative events, the members of which share some set of communicative purposes. These purposes are recognized by the expert members of the parent discourse community, and thereby constitute the rationale for the genre. This rationale shapes the schematic structure of the discourse and influences and constrains choice of content and style" (Swales 1990: 58).

These definitions of genre as social action support and extend linguistic definitions that recognize the simultaneity of language form and context. Such a definition treats genres as both language features and contextual meaning, Campbell and Jamieson's (1978: 17) "constellation of substantive, stylistic, and situational characteristics". Although "[g]enre is distinct from form", in Miller's (1984: 163) semiotic hierarchy, it "is a form at one particular level that is a fusion of lower-level forms and characteristic substance". Kohnen's (2001) definition noted above most obviously argues for that fusion, but Görlach (2008: 105), too, as we saw in 2.1, defines a genre not only as "a specific linguistic pattern" but one that has "become conventionalized in a specific culture for certain well defined and standardized uses of language". Those "uses" of genres need to be broadened to include all the recurring, typified actions of language users that language users know, from eulogies to legal briefs, recipes to sermons, diaries to book dedications, and even scholarly articles. As Miller (1984: 163) writes, "the number of genres current in any society is indeterminate and depends upon the complexity and diversity of the society".

Although defining genre as primarily contextual, Lee (2001) captures the essence of why genre is such a significant linguistic variable for historical linguists – and all linguists – to consider: "when we study differences among genres, we are actually studying the way the language varies because of social and situational characteristics and other genre constraints (register variation), not the way texts vary because of their categorisation" (Lee 2001: 47–48). Rather than determining and itemizing isolated characteristics or constraints, scholars studying genres as social actions are studying all of the relevant characteristics and constraints at once, combined and made visible through the fusion that is a genre. No scholar or user of a genre, for that matter, could possibly specify or control all of the contextual variables that might influence a particular type of language use. Freedman (1993) argues persuasively that in fact no one can articulate accurately or fully the textual and linguistic conventions of any genre. These typified actions based in recurrent situations within particular locations encompass multiple contextual variables at once, including ones researchers cannot specify precisely, control tightly, or, especially in historical work, understand fully. Scholars can, however, identify a genre, and in identifying a genre they are accounting for those contexts

and constraints that are relevant enough to have shaped the genre's action. Defining genres as actions focuses attention on the balance of linguistic form enacting contextual variables while allowing the constituents and the boundaries between them to remain fuzzy and blurred.

# 3 Genres within communities

*Genres are constructed, defined, and named by communities of users, whose boundaries are themselves ever-changing.*

## 3.1 Community construction of genres

With genres conceived as social actions, studying language within generic contexts makes language features interpretable in light of what they do, not just how they appear. The actions genres perform are social actions because they develop within communities of users and are constructed and defined by their cultural communities. Thus, what genres do is a matter for communities and cultures to define, not critics or researchers (Devitt 2004: 9). In Freadman's (1994) game metaphor, genres must be played only in particular locations and with other players. In Miller's (1984) terms, typification and recurrence, even exigence, develop not from some pre-existing reality but from a rhetorical community's shared understanding. Even situations themselves are "social constructs that are the result, not of 'perception,' but of 'definition'" within the culture (Miller 1984: 156). Since no two situations are identical, the existence of a genre depends on a group's shared understanding of what's comparable and what matters.

Within rhetorical genre studies, the concept of community has been defined in different ways, with different emphases depending on the object and purpose of the study. The two most common terms used are *discourse communities* (emphasizing texts; see Swales 1990) and *rhetorical communities* (emphasizing situation; see Miller 1994). It is these communities that define their typified social actions, their genres. For Swales (1990: 58), communicative purposes are established within a "parent discourse community",[2] a community

**2** The concept of discourse community has been challenged and refined over the years (e.g., Harris 1989), and Swales himself has continued to refine his use of the concept (most notably, Swales 1998). Others have shifted the focus to the larger institutional systems that regulate and overlap those communities (see Russell 1997 on activity systems).

defined by its shared discourse and tasks, one in which more active or expert members of the community recognize and shape its genres, influencing the choices of less active members. In disciplinary communities, for a simplified example, editors and reviewers of scholarly journals define the nature and action of that discipline's scholarly article.

None of these conceptions of community excludes the others, and none draws distinct boundaries or claims anything other than overlapping and blurred boundaries. In their take on the "situated cognition" of genres, Berkenkotter and Huckin (1995) drew from sociologist Anthony Giddens' structuration theory to explain the paradoxical nature of genres as at once constructing and constructed by the community. The community defines the genres; the genres define the community. For all rhetorical genre scholars, however, whatever their conceptions of community, those genre and community definitions are not universal or precisely delineated. The most active and expert members of a community influence genres more substantially than novices or more peripheral participants, according to Swales (1990: 77), but even the most powerful members cannot control a genre, and community membership is always changing. The genres thus constructed remain fluid and without firm boundaries.

The connection between genre and community has been noticed in English historical linguistics as well. In his description of the compilation of the Helsinki Corpus, Rissanen (1996: 229) notes that "the study of texts and their grouping into types or genres means an attempt to define and describe man's aims and purposes of communication and to relate them to the social, cultural and political conditions of the speech community". Going beyond such broad formulations of community, researchers have drawn on the concepts of discourse community and community of practice to show the linguistic and discursive practices that develop in such communities, including as represented by and in different genres (e.g., Watts 1995, 2008; Fitzmaurice 2010; Gotti 2013). These studies underscore the potential of exploring different communities' needs and negotiations, especially in combination with the other principles I lay out in this article.

## 3.2 Community identification of genres

Although rhetorical genre scholars may offer different terms for and conceptions of communities, they agree on the centrality of groups of users with degrees of shared purposes, values, and norms – so much so that scholars prioritize those community's names and recognition in identifying genres. Swales (1990: 55) explains that the expert and active members of a community "give genre names to classes of communicative events that they recognize as providing recurring

rhetorical action". In my own study of discourse within present-day tax accounting firms (Devitt 1991), the partners across firms consistently offered samples of the same genres with the same names, or they recognized and named those genres when asked. Once the partners identified a tax accounting genre, its linguistic and rhetorical patterns could be both discovered and interpreted in light of the genre's action in that context. The names given genres by members of a community provide the most valid information about what genres exist for that community and what therefore might significantly influence the genre users' actions, including their language.

When there is evidence of a named genre within a community, then, that name provides substantial evidence of a genre action. Applying that insight is not always simple, however, and sometimes tugs against the desire for systematicity. Importantly, the information from the community does not always agree with an initial linguistic analysis. The tax accounting experts, for example, insisted on identifying two separate genres out of letters to clients that had apparently similar forms. The community members understood a critical distinction of the genre action that wasn't visible in the text or available to the analyst: one type of letter to a client constituted a legally binding opinion; another apparently similar type of letter to clients constituted merely informal information (Devitt 1991). Once those two sets of texts were separated into distinct genres, previously unnoticed yet quantifiable formal patterns became visible and interpretable. Such insight into language use is made possible through privileging the community's expertise about the distinct genres, allowing the users to blur boundaries that might otherwise seem clear to analysts.

Finding the users' names is not always easy or even possible. Sometimes names are inconsistent, or there might be no written evidence of what seem to be visible generic actions. As Görlach (2008: 123) claims for the cooking recipe, "some well-defined types have no names". Some recurring social actions might be recognized within a community but not named (Swales 1990: 57), and at some point new genres might be emerging but not yet named (Miller and Kelly 2017). Swales (1990: 56–57) describes some genre names that might be generated by community members with no typified action behind them aspirational rather than enacted. Some established genre names might be institutionalized or inherited, not necessarily reliably representative of the rhetorical action occurring (Swales 1990: 55–56).

Even flimsier as social actions would be genre names created by analysts rather than community members, especially ones based on textual patterns without confirmation of the genre's meaningful action within its communities. Such genre identification is especially problematic for historical research where

contemporary language users cannot be interviewed. In Old English scholarship, for example, scholars have tried to label the patterns they uncovered based on modern or classical ones, but doing so lacks authenticity. Tolkien (1936) famously pointed this out in relationship to *Beowulf*. He argues that most of the negative criticism of *Beowulf* is the result of evaluating the poem according to the standards of something it is not, most notably an epic. Since then, other scholars have called into question the genre of a group of poems often categorized as elegies, arguing that some of these poems could be recategorized (see Shippey 1994) or that the entire genre is artificially constructed and skews our understanding of the poems (see Fulk and Cain 2013: 256–273). Despite the difficulties, when genre names do appear within a community, they provide potentially meaningful insight and a starting point for further investigation, even though the identification of genres might remain blurry.

Even after identifying a community's genre name, formal patterns may remain invisible to analysts. Görlach (2008: 116), in his study of dedications, encountered a genre well-established and frequently named but for which he could not identify quantifiable recurring linguistic features. This is a result similarly found by Medway (2002) for early twenty-first-century architecture students' sketchbooks. The community explicitly named and identified this genre but the textual instances did not show recurring structure, content, format, or language forms. Some genres indeed have recurring features, but ones within genres unseen by the users and potentially the analyst without additional efforts to bypass their typical audiences. These "occluded" genres (Swales 1996) include confidential genres such as letters of recommendation and reviewers' reports or ones that prohibit observation or public record, such as jury deliberations (Devitt 2016). Since the shared naming of a genre indicates a meaningful social action, those genre names always indicate an influence on language, one that requires investigation and may lead to new insights about language use, even when the results might be difficult to find or interpret.

Genre names thus provide the most valid evidence of which categories might influence language behavior, though they cannot always be accepted without caution. Görlach (2008: 8) again shares a perspective with rhetorical genre scholars: "[t]he assumption is that although not all conventionalized uses of a language have a term relating to them, those that have can probably be correlated with specific functions and recurring linguistic features as well as writers' intentions and readers' expectations". While genres have not been explored extensively from this perspective in English historical linguistics, some studies have acknowledged the importance of such labels for understanding text categories. Schmied and Claridge (1997), for example, include a text's own label as one of the text classifications in the Lampeter Corpus (see also

Grund and Walker 2011; Grund 2012; cf. also Hübler and Busse 2012: 6). Those labels then signify Görlach's features, functions, writers' intentions, and readers' expectations, all fused into genre actions that are constructed by the communities in which those components act.

## 4 Genres with other genres

*Genres interact within genre sets and uptakes, blurring the separation of one genre from another.*

One way of confirming as well as understanding the influence of genres, whether named or unnamed, is through their relationships with other genres, both those that surround them in genre sets and those that take up their genre actions. Within rhetorical genre studies, the discussion about *genre sets* (Devitt 1991), *systems* (Bazerman 1994), *repertoires* (Orlikowski and Yates 1994), and *uptakes* (Freadman 2002) is extensive and involves multiple theoretical debates. For the purposes of this chapter, most important are two ideas: that genres act not in isolation but within a community's set of genres; and that genres respond to other genres, the actions of both in part defined in that uptake.

My study of tax accounting (see 3.2) exemplifies the interaction of genres within their communities, as well as introducing the concept of genre sets (Devitt 1991). The genres these tax accountants recognized and used constituted their work. If all their genres were gathered, oral and written, that set of discourse would be the ways tax accountants act in their world. In fact, one of the senior partners had all of the week's texts bound for him so that he could review the firm's work in a single volume, creating a material representation of the firm's genre set. The highly intertextual nature of those genres also provides evidence of that interaction – a memorandum for the files references a client phone consultation, the opinion letter to a client cites the IRS Rules and Regulations. Each of these genres represents its particular action, but its action depends on the actions of other genres in that community, blurring the boundaries between them. Language users in those communities know of that interrelatedness and act out of that knowledge.

No genre operates in isolation, then, whatever the action and whatever the community. Book dedications exist within other volumes; death notices, obituaries, and eulogies share occasions; witness testimonies connect to accusations, charges, depositions, opening and closing statements, jury deliberations, and verdicts. These genre sets reveal themselves in multiple ways. Community members might reference multiple genres, establishing a user-recognized interconnection.

One study revealed that users represent their genre knowledge in part in contrast to other genres, in terms of what genres they are not (Reiff and Bawarshi 2011). Intertextual references conjoin genres, whether within a community or more blurrily crossing the boundaries between activity systems (for example, from the corporate tax accounting system to the IRS institutional system). Sets of documents might also be positioned together materially, whether bound or archived (though of course an archival gathering might represent the judgment or even convenience of the archivist rather than the actions of the originating community). The action of a genre must always be interpreted in the context not only of its fused form and situation, not only of the community that constructs it, but also of the surrounding genres whose actions partially define it.

One particularly important relationship is what Freadman (2002) calls, drawing from speech act theory, uptake. Genre actions often exist as reactions to other genres, in the ways they take up another genre's invitation to action. To illustrate, Freadman (2002) elaborates the case of a trial that resulted in the defendant being executed. She tracks the uptakes required from sentence to verdict through psychiatric assessments, church petitions, prisoner's record of conduct, among others, before reaching a death warrant. The notion of uptake is not simply a clean call-and-response relationship, a direct determination of one genre from another. Just as multiple genres might act within the same genre set, fulfilling different and overlapping needs of the community or even across activity systems, multiple genres might serve as uptakes of multiple other genres. All genres act in part through interaction, existing not in their textual instantiations and generic boundaries but in what happens in the space between them and other genres.

The histories of genre sets and uptakes have been studied through those interactive relationships. Some genres have been tracked as antecedents of others: the king's speech, for one notable example, is taken up in the newly formed United States by both the presidential state of the union address and the congress's response to it (Jamieson 1975). One particularly well-detailed historical study of genre sets examined business genres. Yates (1989) investigated how changing ideas about efficiency within business communities produced and changed the genres in use, including business memoranda, company newsletters, and many other genres – all acting together toward developing that community of efficient business practices (see also Tachino 2012; Rounsaville 2017).

In English historical linguistics, Görlach (2008: 22) counts as an area still to be explored genres' "intertextual histories – how did types influence each other or how did they become differentiated?". While rhetorical genre studies has contributed several studies to that effort, much more remains to be done to

describe the intertextual relationships among genres through uptakes and genre sets as well as genre users' knowledge and use of those interactions.

## 5 Genre stability and change

*Genres change with every instantiation and are only ever temporarily stabilized, leaving their boundaries forever blurry.*

One of the most intriguing qualities of genres is their tension between stability and change. For communities of users to recognize and define genres, the social action must have become stabilized enough to be defined as recurrent and typified. At the same time, because genres are defined by the communities they enact and communities are always changing, genres are constantly changing as well. Because genres develop through users' actions within contextual situations, every instance of a genre contributes to constructing that genre. Genres thus are always changing, yet always recurring; boundaries always being defined, but those boundaries always blurring. Schryer (1993: 204) captured this tension by describing genres as only ever "stabilized-for-now" or "stabilized enough".

Of course, the degree of change will vary with the genre and its rhetorical and cultural circumstances. Focusing especially on changes in linguistic characteristics and possible social reasons for such changes, research on the history of English has treated such genre shifts extensively in various domains, such as religion, science, and the legal system (for an overview, see Taavitsainen 2016: 281–284) Particularly interesting examples of stability and change are found in Görlach (2008): compare the cooking recipe, in which he found regularly recurring formal features and social actions across texts and throughout time, with the utter lack of stability he found in the dedication, with nothing shared across textual instances but its position in the book and its dedication "to" someone. He notes that more stability might have been seen if he'd narrowed the time period to, say, five years, during which linguistic features would have changed less. But to do so, Görlach (2008: 119) says, would be "to admit defeat – to concede that the concept and/or its expression exhibits little stability through time except for the fact that a book is dedicated by an author to a person, and the sole linguistic feature of a text is that it is placed at the beginning of a book and begins with the word 'to'".

As Görlach's (2008) research illustrates, stability exists throughout time only in the "fact" of its social action (an author dedicating a book to a person), its fusion of purpose, setting (placed at the beginning of a book), rhetorical situation, and linguistic conventions (beginning with the word "to"), not in the

linguistic conventions alone. And that social action changes over time and in different writers/speakers. Note that in the cooking recipe, too, Görlach (2008) found so much individual variation that even the most obvious features might be dominant but never universal. Genres will always have blurry boundaries when defined as action rather than form because genres are constructed by and therefore change with each instantiation. Genres reflect the simultaneous synchronic and diachronic nature of language in use.

The inherent instability of genres over time comes in part through cultural change (cf. Taavitsainen 2016: 281–282). Scholars like Lee (2001: 47) who emphasize extralinguistic, contextual components of genres see change as inevitable: "[g]enres can come and go, or change, being cultural constructs which vary with the times, with fashion, and with ideological movements within society". For Lee (2001: 46), changes in genre conventions derive directly, though sometimes imperceptibly, from those cultural changes. Tsiplakou and Floros (2013: 120) also recognize that "both stability and flux are inherent properties of genre", but they attribute its flux solely to contextual changes and its stability solely to "some kind of linguistic, textual or even 'functional' stability of generic schemes" required for social interaction. To cite a specific example from the history of English, Fachinetti (2015: 2) notes how news writing has "continuously adapted to new environments", suggesting the influence of context and media consumption as promoting change in news "as a textual type".

The difference for rhetorical genre studies is that this dynamic and inherently blurry nature of genres stems from that fusion of the linguistic and extralinguistic, not from cultural changes alone. Every text and utterance requires variation as well as standardization in ways that carry out its social action at that unique moment for that unique situation (Devitt 1997). Not only recognizable patterns of change over time – linguistic or extralinguistic patterns – but rather every enactment of a genre simultaneously constructs and changes that genre, so that all genres are always stabilized only enough to permit momentary typification for the purposes of a community's needed actions. That tension between stability and change within every genre is part of what gives genres such power.

## 6 Genres as normalizing and ideological

*Genres are normalizing and ideological, in ways that can blur historical change.*

As genres and genre sets do the work of their communities, they reinforce that community's values and norms. While that normalizing force provides some necessary stability, it also inhibits change in those generically typified actions.

Yet here, too, that normalizing effect is not complete but rather blurred around the boundaries since genre change does still happen and even individuals do still resist those norms.

Perhaps the most telling study of genres as ideological and normalizing is Paré's (2002) study of Inuit social workers in Canada. Working within established recordkeeping genres, these social workers felt a tension between their own community's values and those of the dominant southern social worker community. The existing records genre enforced a distancing and disembodied professional role for the social workers, a role visible in some linguistic patterns, that contradicted their personal relationships with their clients. The records' requirement for detailed specificity about the clients seemed unnecessarily invasive. As one Inuit social worker said, "white people are greedy for other people's problems" (Paré 2002: 62). The two communities and their interactions created for these language users an "ideological conflict", Paré (2002: 65) observed, "between their dual and contradictory role as advocate for the client and agent of the state". While these workers did try to resist the normalizing force of the records genre, recognizing a genre's ideological norms is difficult. As Paré (2002: 59) writes, "[t]he automatic, ritual unfolding of genres makes them appear normal, even inevitable; they are simply the way things are done". This "illusion of normalcy" (Paré 2002: 61) becomes more easily recognized and challenged through what Paré (2002) calls "chinks", created when situations, institutions, and participants change enough to create a mismatch between the prior norm and the currently needed action. Such chinks create openings for historical linguists to recognize and observe the ideological and normalizing influences of genres on language change.

In the conclusion to her (1984) article, Miller summarizes the defining influence of genres: "what we learn when we learn a genre is not just a pattern of forms or even a method of achieving our own ends. We learn, more importantly, what ends we may have" (Miller 1984: 165). For linguists in general, awareness of this normalizing effect makes genre an even more important variable in language use. Linguistic features may pattern in a particular way not only for gender, class, date, region, identification, face-saving, and a myriad of other variables but also for genre (Devitt 1989a), including for the normalizing influence of meeting others' expectations. Language users are using particular language and discourse patterns, too, not only because they are the convention but because those conventions represent participation in a community, acceptable ways of behaving, expected roles and actions. Such unconscious habits and unrecognized motives deserve to be studied and acknowledged for what they are – powerful factors influencing language use and language change.

# 7 Implications for English historical linguistics – An illustration

To illustrate some of the benefits and complications I see from this rhetorical-linguistic view of genre-in-use, I will consider in the rest of this chapter some of the potential methods and implications were this view to be applied to my own historical study of the shift from Scots to English standards in the sixteenth and seventeenth centuries (Devitt 1989b).[3] This is not a detailed re-evaluation of my results and conclusions. Indeed, my conclusions, theories about linguistic change, and contribution to the understanding of Scottish English have been confirmed in other studies. What I attempt here is to offer an approach, set of questions to ask, and some of the possible implications of doing so for a dataset and context such as mine that can be extended to other materials and data.

Briefly, in my 1989 study I counted in 121 texts written by Scots from 1520 to 1659 occurrences of five linguistic features: the relative pronoun spelled *quh-* or *wh-*; the preterite as *-it* or *-ed*; the indefinite article as *ane* or *a* and *an*; the negative particle as *na*/*nocht* or *no*/*not*; and the present participle as *-and* or *-ing*. In each of the five features, the first variants are associated with a Scots standard written variety and the second with an English standard written variety. My object of study was the process of standardization, especially the shift or diffusion of one standard to another, as happened with the reunification of Scotland and England and the subsequent anglicization of Scots-written texts. I controlled for three extralinguistic variables: medium, audience, and genre. When I ran statistical tests, the results revealed that only one variable in addition to time was significant in the use of Scots or English features – genre. The process of standardization differed significantly across different Scottish genres, diffusing in a clear S-curve, making genre as significant a variable as 140 years of historical change.

To interpret these surprising results, I reconsidered the meaning of genre, which I had conceived as simply type of text. The linguistic features I had studied were not conventionalized features of specific genres. For example, religious treatises shifted early to English standards, but spelling the relative pronoun with *wh-* was not a marker of religious treatises. National public records, which were slowest to anglicize, weren't known for their use of *quh-* spellings.

**3** These results were confirmed in a second study of early American English texts and led to my argument that rhetorical genre must be considered as a significant variable in sociolinguistic research (Devitt 1989a).

What then would make genre have such a significant correlation with the small linguistic details in this process of standardization? Since I am a specialist in rhetoric and composition as well as English language study, I sought a rhetorical explanation.

## 7.1 Treating genre as social action

I interpreted my results by offering a new insight: genres weren't just a single extralinguistic variable; genres encompass multiple extralinguistic variables in one – the whole of the rhetorical situation, including not only audience but also purpose, writer, setting, medium, and other cultural and contextual constraints. For example, religious treatises were often printed, written for unknown, usually English readers, meant to be persuasive, including in their subject a vernacular English Bible, and more. National public records were usually not printed, set in a highly formal situation, written to general, unknown, and probably future readers, with a purpose of recording government acts, about legal subject matters, and more. Other factors would be more important in other genres, to the extent that, as Görlach (2008: 220) captures so well, "the progress of anglicization (or, the continuance of Scots) has to be looked at genre by genre, and register by register".

Although my overall explanation, that genre encompasses multiple aspects of context, still makes sense, the ad hoc selection of contextual features – formality, audience, literariness, cultural meaning – can become more precise and consistent with genre redefined as social action. Specifying the social actions, communities, genre sets, uptakes, and ideologies of a genre makes it more likely that scholars would specify more relevant contextual features, systematically investigating, for example, what actions religious treatises perform, within what communities, in relation to what other genres and uptakes, and how that helps account for their early adoption of English textual features. In different and unique ways, the actions of the national public records – continuing a long-standing record of all actions taken by the Privy Council, for example – within a governmental community are actions and communities that help explain the maintenance of older language forms as well as the ideological cultural identity associated with them.[4] From this historical distance, scholars are not able to

4 Of course, research on change and genres in Scots-English has continued and made further discoveries about variations and contextual factors at the time (see, e.g., Meurman-Solin 2001; Görlach 2008).

recognize all the influences on those Scots writers, any more than they could have recognized or articulated the influences at the time.[5] Nevertheless, treating genre as social action provides not only an interpretation but also a method for encompassing multiple contextual factors, thereby making visible those which are less obvious though no less telling.

Rather than having to select some relatively obvious components of context as explaining the observed language features, the components of the generic action more precisely identify relevant contextual factors. Rather than standing apart from the linguistic features, the interpretation through genre recognizes the necessarily mutual construction of language and its meanings. Defining genres by what they do within and for their communities offers a more focused framework for interpreting linguistic facts in light of both linguistic and contextual variables.

## 7.2 Privileging community genre definition

For genre to account for contextual as well as textual variables in this way, the genres included must represent genres meaningful to the language users at the time. This definition of genre also reveals a flaw in my research: I determined the genres I would include from my own relatively broad categorizations rather than seeking the generic actions identified by contemporary users. This kind of broad categorization of course has been and continues to be a common strategy in, for example, corpus compilation (e.g., Rissanen 1996; Meurman-Solin 2001; Taavitsainen and Pahta 2010).

My broad types – religious treatises, official correspondence, private records, personal correspondence, and national public records – were not, as Görlach (2008: 205) notes, "watertight categories, nor are the types in any way comprehensive", even though they did prove to be "diagnostic for the anglicization process happening in the period and provide enough contrasts to show that standardization is not a monolithic process, but that its speed depends on social and stylistic/text type variables". As confirmed later by both Görlach (2008) and Meurman-Solin (2001), some of these categories were likely meaningful to writers at the time, and my study did tease out some significant factors. What else, though, might have appeared had I selected genres based on

5 I found no evidence of the features I studied being discussed at the time, as markers of Scots language or anything else.

evidence of contemporaneous recognition? In fact, one of the "genres" I used, national public records, included[6] many texts from the Register of the Privy Council, from Records of conventions, from Acts and proceedings, and from trials – all clearly named at the time and hence distinct genres. My other broad categories might have masked other meaningful distinctions. For example, personal correspondence showed a different use of Scots features in my corpus than in the texts studied by Görlach (2008). While his study had a different purpose, and he was looking for texts that were explicitly written in Scots, my own study might have had different results had I subdivided the broad category into types of correspondence (as Meurman-Solin 2001 discusses), but particularly into correspondence genres recognized by writers or recipients at the time. As I showed in section 3.2, in a modern-day context, tax accountants informed me that what seemed like one type of letter was actually two, performing importantly different actions (Devitt 1991). Research into sixteenth- and seventeenth-century Scots may or may not have been able to find evidence of more distinct contemporaneous genres – through names, evidence of distinct social actions, or inferences from genre sets and uptakes – but I wish I had known to try. By sweeping all correspondence into two categories – official or personal – and by creating a convenient category of private records rather than distinguishing what people called "diaries" from "journals" and other records, my study was less likely to reveal even more distinct processes of anglicization due to different social actions.[7] Indeed, the distinction between official and personal letters has been shown to be important for historical linguistic studies (e.g., Nevalainen 2013), and Biber and Gray (2013) have shown the significant linguistic differences that may occur depending on whether we look at what they call "registers" (e.g., newspaper reportage) or "sub-registers" (e.g., magazine articles from *TIME* and articles from the *New York Times*).

If genres are social actions constructed by users and if those actions define their recurrence within communities, then genres must be defined by users and their communities, not by researchers inventing categories or pre-determining which linguistic or extralinguistic variables matter. The process is messier, since communities' genres have fuzzy edges, blurred boundaries across systems, and a constellation of linguistic and extralinguistic features, only some of which can be known. The point is not for the analyst to refine even narrower subcategories

**6** I selected the texts to fill cells in my corpus using numbers generated randomly.

**7** Since the data were collected by hand with no digital corpus, unfortunately I cannot simply rerun the analysis using these refined and more valid genres.

of genre but to identify which categories the genre users themselves recognize. Beginning with genres named or apparently recognized by users has the advantage of ensuring an action exists that has meaning for that community and therefore likely shapes linguistic features.

## 7.3 Incorporating evidence from genre sets and uptakes

My study of Scots texts did not include genre sets or uptakes, though doing so might have enriched my data and offered deeper interpretations not just of the process of change in one genre but change within a community over time. Perhaps the clearest example is religious treatises, which take up sacred texts and interact with them in religious genre sets. Sermons, scripture, commentaries, and more perform different actions within religious communities; examined as a set, the language changes in these genres might have revealed with more complexity the cultural changes within the Scots religious community over that century-and-a-half. Public records similarly act as part of larger genre sets that do the work of the Scots government, and studying language change across that larger set might diversify the understanding of how Scots officials and their language use influenced and were influenced by reunification. The influence of language use over time in such institutions as schools or legal systems, too, might be seen more clearly by investigating that language use of multiple genres that together do the work of that institution. Identifying genre sets could also help to identify communities whose boundaries or histories are not well established, through indicating the shared discursive practices that might have constructed a discourse community. While enlarging an investigation considerably, the purposes of some studies would be well served by including, yet still differentiating, multiple genres from a community and their interactions.

Another approach that can prove revealing is studying genres' uptakes. In the case of my study of Scots anglicization, the religious treatises I studied surely were affected in some way by the language of the Bibles they took up. As Görlach (2008: 208) has pointed out, it may have mattered if a Scots Bible had been created and used regularly, rather than English Bibles. Just as tax accountants took up their Bibles of IRS rules and regulations and referred to them differently in different genres (Devitt 1991), writers of treatises likely echoed the language of their Bibles. Similarly for other genres, correspondents may have reflected the language of the correspondence to which they were replying, and national public records interacted with the language of trials, acts, and proceedings in ways that might prove evident even without access to the originals

themselves. In these days of computerized databases and searchable, quantifiable language features, such linguistic echoes are more easily noticed, confirmed, or disproven.

## 7.4 Accounting for genre instability and change

One of the most problematic aspects of genre for any historical study is its stability and instability, the recognition that genres are at best only stabilized-for-now or stabilized enough (Schryer 1993). On the one hand, historical studies are particularly well suited to acknowledge this instability due to constant change, as they have done, noted in section 5. For English historical linguistics, the ever-changing diachronic nature of language is its frame of reference and key assumption. On the other hand, studying language change, like studying genres, requires a temporary stability, a momentary fixedness in synchronic slices that enables the object to be studied.

In my own study of the diachronic process of anglicization (diffusion, in particular), I took texts from within 20-year historical time periods and traced the process of change over 140 years: 1520–1539, 1540–1559, 1560–1579, and so on up to 1640–1659. Already, of course, this early study acknowledged the arbitrariness of dividing one time period from another. A text from 1539 would fall into a different time cell than a text just a year later, from 1540. The question of periodization has been a topic of debate lately in English historical linguistics, several scholars arguing for more dynamic, statistically-driven approaches to periods (e.g., Hilpert 2012: 149–153; Shao, Cai, and Trousdale 2019: 156–158; see also Grund and Hartman, this volume). At the same time, that the method I used of separating time periods did not invalidate the study is confirmed by the evidence that the process of change was less than random (that is, the change over time was statistically significant, even forming an S-curve of diffusion).

A different sort of curve develops from the instability of genre, one that linguists have recognized (e.g., Biber and Finegan 1997) and that blurs rather than sharpens the picture of ongoing change: while the language use is changing, the genre is changing, too. Over 140 years, the genres I studied likely were changing not only in their anglicization but in their actions. Records of the Privy Council, for example, likely performed different actions for their communities in 1659 than they had in 1520. I could have at least partly taken that change into account if, when seeking initial evidence of which genres were meaningful to the communities, I had sought evidence of those genres' meaning at different slices of time. How did treatises function for religious communities at the beginning of

my study? The middle? The end? Were they ever named differently? How were Records of the Privy Council defined or referred to at different points in that time period? Even more troublesome to the need for stability is the understanding that the genres studied never do hold still. Every use of the genre changes it, just as every use of language changes language. I might examine each genre at different times throughout my study, but the community's construction of the genre might change with the very next utterance/text. As is true for all language research, whether through a genre lens or not, all identified patterns are variable and not universal, a limitation to acknowledge. But as was true for the process of linguistic anglicization over 140 years, the fact that genre proved statistically significant as a variable confirms that the method of distinguishing genres over time matters, even without being able to define sharp boundaries to genres or genre change.

## 7.5 Acknowledging genre's normalizing effects on change

Genres have another way of affecting changing over time: their normalizing nature can inhibit change. To the degree that they are typified and stabilized-for-now, genres are potentially hegemonic within communities, not just defining their actions but prescribing them. Within my study, the genre that would seem most obviously affected by this aspect of genre is public records. Those records and registers likely had established conventions of the genres, ways that things were done, and those internal traditions may have affected whether and how the texts within that genre anglicized. Such genre norms can outlive their originating functions for the community, lagging behind social and cultural change. It is possible, then, that the public records I examined were slowest to shift from Scots to English standards not only because of the actions they performed for the Scots government or the cultural meaning those forms enacted but also because those actions and meanings had become the usual way things were done.

The same might be true for all genres, even ones less obvious. Not just public records but all genres have such normalizing effects. That the balance has tipped toward stability might be indicated by a certain predictability of textual form. If I had seen, for example, that public records had a relatively rigid structure, used formulaic phrases, or predictable syntactic patterns, such inflexibility in form would indicate inflexibility in meaning and hence action. Since all genres must have some degree of stability and typification, genre's normalizing and hence inhibiting effects should be something any linguistic study would examine.

Though I did not frame it this way at the time, my particular study illustrates the opportunity created when those norms are changing or communities are clashing – the "chinks" that make ideology more visible (Paré 2002). As part of the process of anglicization, the particular linguistic features I selected for my study already carried some known ideological import: they were conventions of Scots that gave way over time to conventions of English, carrying all the freight of those national identities. My study, then, was already alert to the ideological effects visible through this chink, and my discovery of genre's significance led to my considering the potential ideological meaning of each genre. What I did not consider enough was the normalizing effects of every genre, no matter the ideological meaning of the particular linguistic features studied. That normalizing effect would not be the same for every genre, would not inhibit change to the same degree. Some communities are more narrowly controlled than others, perhaps including the religious professions as well as the keepers of the public records. The actions of some genres are more occluded, as we saw in 3.2, with a less visible textual tradition to follow: personal records like diaries and journals are less public, creating fewer formal expectations and internal traditions, perhaps – though still well-known are their social actions of selecting and shaping the details of daily living. Had I begun with genres defined by their users rather than my own broad categories, their normalizing traditions might have been more evident. As it is, my selection of small linguistic features to study may have avoided the more obvious genre norms, but then again the fact that genre was significant for these small features makes the effects of genre even more remarkable and unpredictable.

# 8 Implications for English historical linguistics – Scholarly abstractions and complications

As the illustration from my own study indicates, I see great potential for this conception of genre in English historical linguistics. Research in linguistics as well as rhetoric has found genre to be a significant variable in language change, but scholars have struggled to define it in a way that acknowledges both its linguistic and extralinguistic components, and "the notion of genre itself remains elusive" (Whitt 2018: 10). The conception I describe in this chapter enables genre to fuse both text and context. By recognizing genre as social action, it accounts for the textual patterns of genre while expanding the contextual elements it includes. Importantly, it recognizes that the impact of genre on

language comes from the meaningfulness of genre to its users. Without that meaning for the community of users, an analyst's text category may or may not capture the factors that affected language and language change, whatever they might be.

While sketching these five foundational conceptions of genre from rhetorical studies, this chapter of course simplifies, ignores some debates within the field,[8] and has to stop at some point. Still, the complexity of genre's blurred boundaries reveals itself in several aspects. Genres can have multiple names or no names, different names in different communities serving similar functions, or the same name at different times in spite of significant textual changes. The genre action itself interacts with and defines itself in terms of other genres and is never more than momentarily stabilized. The relevant contextual factors can never be fully articulated, even by expert users of the genre, and they are always changing as the community and its culture change. Genres can inhibit change, but in ways often invisible and unpredictable, like all carriers of ideology. In short, genres-in-use (unlike what might be termed genres-in-the-lab) can be difficult to pin down precisely and hard to hold still for long. Language is like that, and genres are language-in-use.

That complexity of genre, though, makes it an even richer concept and should not deter English historical linguists from including genre as an important factor in language change. As Görlach (2008: 100) writes so eloquently:

> It is evident that historical linguists attempting to describe socio-stylistic conditions and how they correlate with linguistic variation are up against a great number of problems. But these should not hinder us from grasping the nettle. I am convinced that many of the important developments of English (and of the other European languages, especially in their standard forms) took place in the field of text syntax and the emergence of text types; thus any kind of historical linguistics that fails to take account of this field cannot explain why English developed the way it did.

I won't pretend to know the many directions English historical linguistics might take in continuing to study genre, though I see potential and promise. I do know, though, that in order to "explain why English developed the way it did", the field must treat genres as meaningful actions for language users, as genres-in-use. To do so requires acknowledging and working with genres' inherently blurry boundaries.

**8** In addition to Devitt (2004), interested readers might see Bawarshi and Reiff's (2010) overview of genre issues from different disciplinary perspectives or Coe et al's (2002) collection on genre and ideology, stability and change.

# References


Bawarshi, Anis S. & Mary Jo Reiff. 2010. *Genre: An introduction to history, theory, research, and pedagogy*. West Lafayette, IN: Parlor Press; Fort Collins, CO: WAC Clearinghouse.

Bazerman, Charles. 1994. Systems of genres and the enactment of social intentions. In Aviva Freedman & Peter Medway (eds.), *Genre and the new rhetoric*, 79–101. London: Taylor & Francis.

Berkenkotter, Carol & Thomas N. Huckin. 1995. *Genre knowledge in disciplinary communication: Cognition/culture/power*. Hillsdale, NJ: Lawrence Erlbaum.

Biber, Douglas. 1988. *Variation across speech and writing*. Cambridge: Cambridge University Press.

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

Biber, Douglas & Edward Finegan. 1997. Diachronic relations among speech-based and written registers in English. In Terttu Nevalainen & Leena Kahlas-Tarkka (eds.), *To explain the present: Studies in the changing English language in honour of Matti Rissanen*, 253–275. Helsinki: Société Néophilologique.

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

Campbell, Karlyn Kohrs & Kathleen Hall Jamieson. 1978. Form and genre in rhetorical criticism: An introduction. In Karlyn Kohrs Campbell & Kathleen Hall Jamieson (eds.), *Form and genre: Shaping rhetorical action*, 9–32. Falls Church, VA: Speech Communication Association.

Claridge, Claudia. 2012. Styles, registers, genres, text types. In Alexander Bergs & Laurel Brinton (eds.), *Historical linguistics of English: An international handbook*, vol. 1, 237–254. Berlin: Mouton de Gruyter.

Coe, Richard M., Lorelei Lingard & Tatiana Teslenko (eds.). 2002. *The rhetoric and ideology of genre: Strategies for stability and change*. Cresskill, N.J: Hampton Press.

Devitt, Amy J. 1989a. Genre as textual variable: Some historical evidence from Scots and American English. *American Speech* 64(4). 291–303.

Devitt, Amy J. 1989b. *Standardizing written English: Diffusion in the case of Scotland, 1520–1659*. Cambridge: Cambridge University Press.

Devitt, Amy J. 1991. Intertextuality in tax accounting: Generic, referential, and functional. In Charles Bazerman & James Paradis (eds.), *Textual dynamics of the professions: Historical and contemporary studies of writing in professional communities*, 335–357. Madison: University of Wisconsin Press.

Devitt, Amy J. 1997. Genre as language standard. In Wendy Bishop & Hans Ostrom (eds.), *Genre and writing: Issues, arguments, alternatives*, 45–55. Portsmouth NH: Boynton Cook/Heinemann.

Devitt, Amy J. 2004. *Writing genres*. Carbondale IL: Southern Illinois University Press.

Devitt, Amy J. 2016. Uncovering occluded publics: Untangling public, personal, and technical spheres in jury deliberations. In Mary Jo Reiff & Anis Bawarshi (eds.), *Genre and the performance of publics*, 139–156. Boulder, CO: Utah State University Press.

Facchinetti, Roberta. 2015. Introduction. In Roberta Facchinetti, Nicholas Brownlees, Birte Bös & Udo Fries (eds.), *News as changing texts: Corpora, methodologies and analysis*, 1–4. Newcastle: Cambridge Scholars.

Fitzmaurice, Susan M. 2010. Mr Spectator, identity and social roles in an early eighteenth-century community of practice and the periodical discourse community. In Päivi Pahta, Minna Nevala, Arja Nurmi & Minna Palander-Collin (eds.), *Social roles and language practices in Late Modern English*, 29–53. Amsterdam: John Benjamins.

Freadman, Anne. 1994. Anyone for tennis? In Aviva Freedman & Peter Medway (eds.), *Genre and the new rhetoric*, 43–66. London: Taylor & Francis.

Freadman, Anne. 2002. Uptake. In Richard M Coe, Lorelei Lingard & Tatiana Teslenko (eds.), *The rhetoric and ideology of genre: Strategies for stability and change*, 39–53. Cresskill NJ: Hampton Press.

Freedman, Aviva. 1993. Show and tell? The role of explicit teaching in learning new genres. *Research in the Teaching of English* 27(3). 222–251.

Fulk, R. D. & Christopher M. Cain. 2013. *A history of Old English literature*. 2nd edn. Malden, MA: Blackwell.

Görlach, Manfred. 2008. *Text types and the history of English*. Berlin: Mouton de Gruyter.

Gotti, Maurizio. 2013. The formation of the Royal Society as a community of practice and discourse. In Joanna Kopaczyk & Andreas H. Jucker (eds.), *Communities of practice in the history of English*, 269–285. Amsterdam: John Benjamins.

Grund, Peter J. 2012. Textual history as language history? Text categories, corpora, editions, and the witness depositions from the Salem witch trials. *Studia Neophilologica* 84(1). 40–54.

Grund, Peter J. & Terry Walker. 2011. Chapter 2: Genre characteristics. In Merja Kytö, Peter J. Grund & Terry Walker, *Testifying to language and life in early modern England*, 15–56. Amsterdam: John Benjamins.

Halliday, M. A. K. 1976. *Language as social semiotic: The social interpretation of language and meaning*. London: Edward Arnold.

Harris, Joseph. 1989. The idea of community in the study of writing. *College Composition and Communication* 40(1). 11–22.

Hilpert, Martin. 2012. Diachronic collostructional analysis: How to use it and how to deal with confounding factors. In Kathryn Allan & Justyna A. Robinson (eds.), *Current methods in historical semantics*, 133–160. Berlin: Mouton de Gruyter.

Hübler, Axel & Ulrich Busse. 2012. Introduction. In Ulrich Busse & Axel Hübler (eds.), *Investigations into the meta-communicative lexicon of English*, 2–16. Amsterdam: John Benjamins.

Jamieson, Kathleen. 1975. Antecedent genre as rhetorical constraint. *Quarterly Journal of Speech* 61. 406–415.

Kohnen, Thomas. 2001. On defining text types within historical linguistics: The case of petitions/statutes. *European Journal of English Studies* 5(2). 197–203.

Kruger, Haidee, Bertus van Rooy & Adam Smith. 2019. Register change in the British and Australian Hansard (1901–2015). *Journal of English Linguistics* 47(3). 183–220.

Lee, David Y. W. 2001. Genres, registers, text types, domains, and styles: Clarifying the concepts. *Language Learning and Technology* 5(3). 37–72.

Medway, Peter. 2002. Fuzzy genres and community identities: The case of architecture students' sketchbooks. In Richard M. Coe, Lorelei Lingard & Tatiana Teslenko (eds.), *The rhetoric and ideaology of genre: Strategies for stability and change*, 123–153. Cresskill NJ: Hampton Press.

Meurman-Solin, Anneli. 2001. Genre as a variable in sociohistorical linguistics. *European Journal of English Studies* 5(2). 241–256.

Miller, Carolyn R. 1984. Genre as social action. *Quarterly Journal of Speech* 70(2). 151–167.
Miller, Carolyn R. 1994. Rhetorical community: The cultural basis of genre. In Aviva Freedman & Peter Medway (eds.), *Genre and the new rhetoric*, 67–78. London: Taylor & Francis.
Miller, Carolyn R. & Ashley R. Kelly (eds.). 2017. *Emerging genres in new media environments*. London: Palgrave Macmillan.
Moessner, Lilo. 2001. Genre, text type, style, register: A terminological maze? *European Journal of English Studies* 5(2). 131–138.
Nevalainen, Terttu 2013. Words of kings and counsellors: Register variation and language change in early English courtly correspondence. In Esther-Miriam Wagner, Bettina Beinhoff & Ben Outhwaite (eds.), *Scribes as agents of language change*, 99–119. Berlin: Mouton de Gruyter.
Nevalainen, Terttu & Helena Raumolin-Brunberg. 2012. Historical sociolinguistics: Origins, motivations, and paradigms. In Juan Manuel Hernández-Campoy & Juan Camilo Conde-Silvestre (eds.), *The handbook of historical sociolinguistics*, 22–40. Malden: Wiley-Blackwell.
Orlikowski, Wanda J. & JoAnne Yates. 1994. Genre repertoire: The structuring of communicative practices in organizations. *Administrative Science Quarterly* 39(4). 541–574.
Paré, Anthony. 2002. Genre and identity: Individuals, institutions, and ideology. In Richard M. Coe, Lorelei Lingard & Tatiana Teslenko (eds.), *The rhetoric and ideology of genre: Strategies for stability and change*, 57–71. Cresskill NJ: Hampton Press.
Reiff, Mary Jo & Anis Bawarshi. 2011. Tracing discursive resources: How students use prior genre knowledge to negotiate new writing contexts in first-year composition. *Written Communication* 28(3). 312–337.
Rissanen, Matti. 1996. Genres, texts and corpora in the study of medieval English. In Jürgen Klein & Dirk Vanderbeke (eds.), *Anglistentag 1995 Greifswald, proceedings*, 229–242. Tübingen: Max Niemeyer.
Romaine, Suzanne. 2016. The variationist approach. In Merja Kytö & Päivi Pahta (eds.), *The Cambridge handbook of English historical linguistics*, 19–35. Cambridge: Cambridge University Press.
Rounsaville, Angela. 2017. Genre repertoires from below: How one writer built and moved a writing life across generations, borders, and communities. *Research in the Teaching of English* 51(3). 317–340.
Russell, David R. 1997. Rethinking genre in school and society: An activity theory analysis. *Written Communication* 14(4). 504–554.
Schmied, Josef & Claudia Claridge. 1997. Classifying text- or genre-variation in the Lampeter Corpus of Early Modern English Texts. In Raymond Hickey, Merja Kytö, Ian Lancashire & Matti Rissanen (eds.), *Tracing the trail of time. Proceedings from the second diachronic corpora workshop*, 119–135. Amsterdam: Rodopi.
Schryer, Catherine. 1993. The lab vs. the clinic: Sites of competing genres. In Aviva Freedman & Peter Medway (eds.), *Genre and the new rhetoric*, 105–124. London: Taylor & Francis.
Shao, Bin, Yingying Cai & Graeme Trousdale. 2019. A multivariate analysis of diachronic variation in *A bunch of* NOUN: A Construction Grammar account. *Journal of English Linguistics* 47(2). 150–174.
Shippey, T. A. 1994. *The Wanderer* and *The Seafarer* as wisdom poetry. In Henk Aertsen & Rolf H. Bremmer Jr (eds.), *Companion to Old English poetry*, 145–158. Amsterdam: Vrije University Press.

Swales, John M. 1990. *Genre analysis: English in academic and research settings*. Cambridge: Cambridge University Press.
Swales, John M. 1996. Occluded genres in the academy: The case of the submission letter. In Eija Ventola & Anna Mauranen (eds.), *Academic writing: Intercultural and textual issues*, 45–58. Amsterdam: John Benjamins.
Swales, John M. 1998. *Other floors, other voices: A textography of a small university building*. Mahwah NJ: Lawrence Erlbaum.
Taavitsainen, Irma. 2016. Genre dynamics in the history of English. In Merja Kytö & Päivi Pahta (eds.), *The Cambridge handbook of English historical linguistics*, 271–285. Cambridge: Cambridge University Press.
Taavitsainen, Irma & Päivi Pahta (eds.). 2010. *Early Modern English medical texts*. Amsterdam: John Benjamins.
Taavitsainen, Irma, Päivi Pahta & Martti Mäkinen. 2005. *Middle English medical texts*. Including *MEMT presenter* by Raymond Hickey. Amsterdam: John Benjamins.
Tachino, Tosh. 2012. Theorizing uptake and knowledge mobilization: A case for intermediary genre. *Written Communication* 29(4). 455–476.
Tolkien, J. R. R. 1936. *Beowulf*: The monsters and the critics. *Proceedings of the British Academy* 22. 245–295.
Tsiplakou, Stavroula & Georgios Floros. 2013. Never mind the text types, here's textual force: Towards a pragmatic reconceptualization of text type. *Journal of Pragmatics* 45(1). 119–130.
Watts, Richard J. 1995. Justifying grammars: A socio-pragmatic foray into the discourse community of early English grammarians. In Andreas H. Jucker (ed.), *Historical pragmatics: Pragmatic developments in the history of English*, 145–185. Amsterdam: John Benjamins.
Watts, Richard J. 2008. Grammar writers in eighteenth-century Britain: A community of practice or a discourse community? In Ingrid Tieken-Boon van Ostade (ed.), *Grammars, grammarians and grammar-writing in eighteenth-century England*, 37–56. Berlin: Mouton de Gruyter.
Whitt, Richard J. (ed.). 2018. *Diachronic corpora, genre, and language change*. Amsterdam: John Benjamins.
Yates, JoAnne. 1989. *Control through communication: The rise of system in American management*. Baltimore: Johns Hopkins University Press.
Yunick, Stanley. 1997. Genres, registers and sociolinguistics. *World Englishes* 16(3). 321–336.

Edgar W. Schneider

# 3 Meanderings from early English to World Englishes: A Complex Systems perspective on morphosyntactic changes in *wh*-pronouns

## 1 Introduction

Ever since Ferdinand de Saussure's classic *Cours de linguistique générale* of 1916, the conventional, established view in synchronic linguistics has been the assumption that languages are separate, well-ordered systems which consist of distinct units on several levels between which systematic syntagmatic and paradigmatic relationships hold. On the abstract, *langue* level these units are viewed as clear emic entities (phonemes, morphemes, etc.), with variants (allophones, allomorphs, etc.) in *parole*. Similarly, dialects and varieties tend to be seen as independent systems, with overlapping but also varying sets of units, such as, for example, Standard English, African American English (AAE), or Indian English (IndE), each regarded as a variety of English which is internally systemic, well-ordered, and distinct from all other varieties. This categorizing thinking appears to be grounded in principles of human cognition, a deeply rooted desire for clear-cut divisions, boundaries, and conceptualizations which allow us to cope with and successfully operate in a bewilderingly complex reality. It constitutes a legitimate heuristic strategy, but in some respects it is too clear and simplifying, too strongly reductive – reality is more complex and messy!

At the same time, while this categorial view underlies much of modern systemic linguistics, it has been recognized that linguistic behavior is much more fuzzy, and the boundaries that we draw tend to be artificial. There are borderline varieties, for instance, which defy easy categorization. Görlach (1996), in a paper entitled "And is it English?", discussed "utterances which are only marginally English" (1996: 171). In a similar vein, Mesthrie and Bhatt (2008) posit the existence not of simply "the English language" but of an "English Language Complex", distinguishing twelve different "variety types", including "Hybrid Englishes", "Jargon Englishes", and more.

**Edgar W. Schneider,** Department of English and American Studies, University of Regensburg, Regensburg, Germany, email: edgar.schneider@ur.de

https://doi.org/10.1515/9783110643282-004

Individual varieties of a language, in our case English, constitute snapshots in time and space. In some core respects they are all similar; in others they are all different. And, fundamentally, they are all constantly evolving, all more or less distantly related to each other. In this article, I argue that a Complex Dynamic System's approach to English(es) allows us a way to provide a novel and alternative conceptualization of the fundamental character of a language system, of the relationship between different varieties of a language in space and time, and of the nature of linguistic change that enables a closer mapping of some aspects of linguistic evolution than other explanations do. Applied to varieties of English, ranging from Old English to New Englishes, this line of thinking identifies long-term continuities and changes and even disruptions of linguistic evolution in a coherent fashion, thus allowing theoretically new, fruitful perspectives. The approach is also useful as it brings together both internal linguistic dynamics and extralinguistic factors into one unified system, which acknowledges that social as well as intralinguistic factors are important for how we conceptualize English(es) and linguistic variation and change. This framework has received some attention within linguistics in the past (see section 2), and individual components and concepts that are brought together under the Complex Dynamic Systems (CDS) umbrella have been discussed in isolation for English and the development of English. Here I explore the various dimensions and principles of CDS jointly to demonstrate the advantages of using this framework systematically. I do so by investigating and documenting the uses and development of *wh*-pronouns in English through a CDS lens.

I will start by introducing the notion of CDS and some of their distinctive properties in sections 2 and 3. Since this framework has not been employed much in the study of the history of English, I lay out these principles in some detail. The main part will outline the central evolutionary processes which have shaped the system of *wh*-pronouns in English, based on the literature, earlier research of mine, and some new data in section 4. Three core processes involving *wh*-pronouns will be outlined and interpreted in CDS terms: the reallocation of interrogative pronouns to relativizer functions and the changing set of relativizer forms (4.1); the loss of case marking in *whom*, being increasingly replaced by *who* (4.2); and the changing conditions of choice in the genitive between *whose* and *of which* (4.3). In section 5, I give a summary overview of the developments in CDS terms highlighted in the individual sections. I return to the broader issue of the utility of the CDS framework and "boundaries" in section 6.

## 2 The theory of Complex Dynamic Systems (CDS)

Since roughly the 1940s Complex Systems theory has been established and growing vigorously in the natural and social sciences, based on foundations in mathematics. There has been successful applications in a range of disciplines, including biology, medicine, social organization, natural phenomena, technology, and business. The theory captures basic properties of many domains in life which are determined by complex systemic relationships, are interconnected, and perpetually developing (cf. Johnson 2009; Holland 2014; Mobus and Kalton 2015). The framework comes under varying labels (Systems Science, Complexity Science, Theory of Complex Adaptive Systems). It is closely related to Chaos Theory (e.g., Gleick 1987), a precursor theory based on nonlinear equations in mathematics.[1] CDS theory argues against reductionism and categorial, deterministic thinking. It is a "holistic", "cross-disciplinary" approach, though at this stage not yet a fully-fledged theory with a "canonical" form but rather a "meta-science," according to Mobus and Kalton (2015: 3). Characteristic properties of CDS will be discussed in section 3.

While many application domains are discussed in the literature on CDS, language hardly ever is. The only exception I am aware of is Holland (2014), who regards language as a complex system, but does so in passing. In linguistics, some claims of the applicability of CDS theory to language have been voiced in the recent past, but mostly in terms of individual CDS properties (see section 3), and often in passing and in more abstract terms (see, e.g., Lindblom et al. 1984; Hopper 1987; Lightfoot 1991; Larsen-Freeman 1997; Schneider 1997; Mufwene 2001; Ellis 2008, 2011; Bybee 2010). More full-fledged discussions of CDS in linguistics are found in Larsen-Freeman and Cameron (2008), Ellis and Larsen-Freeman (2009), Beckner et al. (2009), Kretzschmar (2015), and Schmid (2020). For example, in a broad and comprehensive analysis inspired by usage-based linguistics, Schmid (2020) shows how communicative usage, social conventionalization, and cognitive entrenchment operate as a continuously interacting feedback system to forge collective and individual grammars, modeled also as a "dynamic complex-adaptive system" (2020: 295).

Despite these creative approaches, CDS thinking in linguistics is not yet a mainstream approach. Its relevance has been intuitively sensed, projected, and suggested (most recently, see Massip-Bonet et al. 2019), but it needs to be filled with life, so to speak. The present paper wishes to contribute to this process.

**1** The precise relationship between the two frameworks is close but disputed. Not all complex systems need to be "chaotic" in the mathematical sense (marked by nonlinearity), but almost all books on CDS discuss chaotic behavior as well: "[n]on-linear behaviour is one of the cornerstones of complexity" (Bossomaier and Green 2000: 7).

The aim is to exemplify and substantiate the claim that all (varieties of) English are manifestations of an overarching CDS of English(es), the branches of which keep "meandering" through space and time and displaying features found in other CDS as well. By way of specific illustration, I want to show how the CDS approach can offer further insight into the dynamics of long-term changes and reorganization processes of *wh*-pronouns in Englishes.

## 3 Properties of CDS in English(es)

The basic idea underlying the notion of Englishes as CDS is that forms and elements of English(es) have been passed on continuously through space and time for millennia and across continents, in an unbroken chain of transmission, but continuously modified in this process by internal developments and changes and by external adjustments and contact effects. And this applies to all varieties, standard and nonstandard, regional, social, ethnic or national, and to all periods of time, including Postcolonial Englishes or World Englishes (Schneider 2007, 2011a). In other words, the East Anglian dialect, Southern US English, IndE, and Nigerian Pidgin all constitute and should be seen as perpetuations and offspring of what once were dialects of Old English (OE), with varying degrees of modifications and contact effects integrated in the course of time.

Language evolution viewed as CDS implies a constant interaction between individual human activity and collective system development, largely along the lines described by Schmid (2020), and this ongoing multiple agency with feedback makes the system complex. Humans act as agents in CDS via their individual utterances, i.e., usage, and these spontaneous and momentary utterances contribute to the ongoing entrenchment of linguistic habits (i.e., emerging structures) shared in a speech community. Continuous feedback loops thus built into usage-based linguistic behavior produce and steer systemic developmental changes.

In this section, I outline some core properties of CDS to provide a backdrop for the discussion of *wh*-pronouns. These principles include systemness, complexity, perpetual dynamics, network relationships, the alternation between order and chaos, the so-called "butterfly effect", emergentism, and a self-organizing capacity of CDS (cf. Mobus & Kalton 2015: 17–30). I leave out nonlinearity and fractals (for the time being).[2]

**2** For more information and application examples, see Kretzschmar (2018) and Schneider (2020).

## 3.1 Systemness

In CDS, "a system is a whole of some sort made up of interacting or interdependent elements or components integrally related among themselves" (Mobus and Kalton 2015: 73). This is self-evident and axiomatic for linguistics as well, considering, for example, Saussure's structuralism, which views language as a "system of signs". Units on several levels (phonemes, morphemes, lexemes; phrases, clauses; constructions) share mutual syntagmatic and paradigmatic relations. Any construction could be used to exemplify systematic relations of several kinds: e.g., words built from sounds, constituents built from words, and functionally equivalent units excluding each other.

## 3.2 Complexity

In CDS, *complexity* is defined as the interaction of large numbers and different types of objects and agents; and this interaction results in many new hierarchical levels of organization. This seems equally self-evident and uncontroversial with reference to languages and for Englishes as well.[3] Complexity can be identified from two perspectives. In language-internal relations, complexity manifests itself in hierarchies of units and utterances which are mutually constitutive (when sounds build morphemes, which compose lexemes, which are elements of phrases and clauses, etc.). Language-externally, interactions between many speakers as agents equally involve highly complex relations, as, for example, in the evolution of new varieties as posited in the Dynamic Model of the growth of Postcolonial Englishes (Schneider 2007), which involves changing relations between settlers and locals.

**3** Complexity in language has been a fashionable topic of linguistics recently (see the useful and comprehensive survey of pertinent publications by Kortmann and Schröter 2020). However, in most of these publications, the notion is applied in a wider, non-technical sense, implying a system that is 'very complicated', with many units, formal choices, and hierarchy levels involved (as opposed to a system that is 'structurally simple', with few units, choices and levels, and relatively short utterances). In CDS theory, the implication is that a system has the other properties discussed in 3.1–3.8 as well.

## 3.3 (Perpetual) Dynamics

CDS are always in flux:[4] the process of continuous change, adjustments, and modifications of all system components in interaction over time never stops. This is of course a core property of English(es) as well. Transmission and change in English(es) keep perpetually rolling; there is an unbroken chain of transmission not only from one generation to another but actually with every single utterance – which manifests and realizes the system but at the same time has the potential to modify it (cf. Devitt, this volume). Hence there is both a basic degree of continuity, with sounds, words, and patterns passed on and shared across varieties, and the potential for continuous change, including internal innovation and modification as well as a possible impact of contact effects, which can bring in novel forms and habits from other languages. Thus, on the one hand, all varieties comprise and share elements of continuity (e.g., a sound /p/, a lexeme *hand*, and a Det-Adj-N NP pattern), but, on the other hand, they are partially distinguished by some discontinuity picked up at some point in time through change, innovation, or borrowing, such as subject omission in *Can!* in Singlish, the lexeme *dhoti* in IndE, or a second person plural pronoun *una* in Nigeria (Schneider 2011a: 164, 21–22, and 146, respectively).

Interestingly enough, the mechanism of innovations spreading in a population has been described by the model of an S-curve in population biology and in CDS theory, a "logistic function [. . .] generated when processes are characterized first by an exponential rise followed by an exponential deceleration to level off at a maximum" (Mobus and Kalton 2015: 215). Similarly (though originally independently), in linguistics, the S-curve model has been well established and found to be effective in modeling the spread of innovations in English (e.g., Ellegård 1953; Kroch 1989; Nevalainen 2015).

## 3.4 Networks

In CDS, "systems are [. . .] networks of relations", where "the components are connected in various relations", often indirectly, in a chain-like fashion (Mobus and Kalton 2015: 137). This property manifests itself clearly in English(es) (as well

4 In some contexts, I use "dynamism" instead of "dynamics". The difference, as I see it, is slight, but there is one. Dynamics, which is the term preferred generally in systems science, is the property of systems, a kind of holistic view; dynamism is the behavior, force, and energy of individual factors to change or to cause change.

as in language in general). From a language-external perspective, social networks and interactions between speaker peer groups have been found to be a decisive effect shaping one's speech, as explored by Jim and Lesley Milroy in Belfast (e.g., Milroy and Milroy 1985). Similarly, in the emergence of Postcolonial Englishes, Schneider (2007) has posited a network dependency relation among external (political) conditions, the social relationships and identity projections of speakers, and internal structural properties. In a language-internal perspective, the interrelationships between items, constructions, and language levels which shape language evolution can be understood as network-like relationships. An example is the complete reshuffling of the English system of determiners and demonstratives: a highly complex set of forms in OE disintegrated completely and just a small fraction of these forms (or their descendants) resurfaced to re-construct a new system with new mutual relationships and ordering parameters, namely *the*, *this*/*these*, and *that*/*those*, which expressed definiteness, singular/plural, and proximity/distance (McColl Millar 2000; cf. Brunner 1960/1962, II: 130–135; Lass 1992: 112–114; see also 3.6).

## 3.5 Order and chaos

So-called Chaotic systems, "whose behavior appears to follow a regular pattern but not entirely" (Mobus and Kalton 2015:202), marked by reiterations of nonlinear equations and thus ultimately unpredictable feedback loops (Mobus and Kalton 2015: 251–252), constitute one major manifestation of CDS. Such systems are characterized by the co-existence of "order and chaos" (in different sub-systems and at different times): there are "pockets of order" inside "chaos" (Gleick 1987), "a complicated mix of ordered and disordered behavior" (Johnson 2009: 15), with processes known as "turbulence" leading from order and simplicity to chaos and complexity. In language, and in Englishes specifically, at any point in time, there are sub-systems which are relatively orderly (i.e., consisting of a relatively small number of units, with clear functional assignments and mutual delimitation) and sub-systems which are disorderly (with large numbers of units, fuzziness, and unclear functional assignments). And this situation is subject to change: ordered subsystems may be disrupted and break down in the course of time, while systematicity may emerge in formerly chaotic sub-systems.

An example from the history of English, extending back across millennia, is the dissolution of the English strong verb system. Indo-European is assumed to have had a perfectly regular and predictable system: verbs had four functionally distributed principal forms which had different degrees of stress assignment and, correspondingly, vowel variants. (At the same time, based on the uniformitarian

principle,[5] it may be questioned whether a reconstructed Proto-Indo-European system that is largely free of variation ever was a reality). By OE, turbulence moved the system towards increasing complexity: various interfering conditioned sound changes, triggered by different following sound environments, produced a rather complex but still somewhat predictable set of strong verb classes and forms, with different modified stem vowels. In the transition to Middle English (ME) the reduction of four principal verb forms to three substantially increased the gradual collapse of whatever order may have been left. Variability as to which form survived produced the rather chaotic set of irregular verb forms which we find in Modern English (ModE). Even more turbulence effects, irregularity, and chaos[6] can be identified in some regional and social dialects, where forms such as *bring – brung*, *fight – fit*, *sit – sot*, *fetch – fotch*, *catch – cotch*, *take – tuck/tooken*, and others appear (Schneider 1989: 90–114). At the same time, English(es) have also shown and developed counter-tendencies towards maintaining or regaining some degree of order. One is regularization (starting with the establishment of weak verbs with dental suffixes for the "past" in Germanic); this is a process that aligns with the cognitive principles of regularity, compositionality, and isomorphism, a consistent association of form (dental suffix – *ed*) and meaning ('past'). Some verbs became re-classified accordingly (e.g., OE *helpan–healp–hulpon–holpen* became ModE *help-helped-helped*), and, again, even more so in regional and social dialects, where we find forms like *knowed*, *growed*, *gived*, etc. Secondly, some new regularities of past tense formation have emerged. For example, Cheshire (1994) showed how the "ideaphone" /ʌ/, originally found in past tense forms like *flung*, *wrung*, or *stung*, became more widely associated with the notion of "past-ness", contributing to the formation of further, especially dialectal past tense forms such as *run*, *struck*, *done*, *sunk*, *drug*, *tuck*, or *brung* (see Schneider 1989: 90–114), and to the ongoing spread of *snuck* for *sneaked* in Canada (Chambers 2006–2007).

**5** This principle, introduced in linguistics in Labov (1972: 275; see also Bergs 2012) basically suggests that principles and processes observed in the present can be assumed to have operated in the past as well.

**6** Admittedly, the notion of chaos is used in a rather non-technical, metaphorical sense here, where *chaotic* should be taken to mean something like 'disorderly, unstructured, close to anything goes'. How this relates to the mathematical notion of chaos (caused by nonlinearity) is left to future discussion and research.

## 3.6 Butterfly effect

The so-called "butterfly effect", popularly known as the flapping of the wings of a butterfly causing a tornado somewhere else much later, and technically defined as sensitivity to initial conditions, constitutes a well-known characteristic of chaotic systems and some CDS; it results from nonlinear equations. In such systems very slight differences of initial states lead to unpredictably large divergences down the road. Because amplification in feedback loops may lead to a qualitative leap in the system, small-scale distinctions have far-reaching, qualitative effects.

A strong example of such a process from the history of English is its fundamental change from a basically synthetic language type up to OE to a largely analytic language today. The initial trigger of this was a relatively inconspicuous phonotactic change some 2000 years ago; in the end, it fundamentally transformed the typological character of English (Fennell 2001: 35). The Germanic Main Stress Rule caused word stress to be fixed on the first syllable in Germanic languages. Consequently, with much of the articulatory air pressure expended on the first syllable, the phonation stream was weakened in later and final syllables. In turn, vowels in these syllables were reduced, commonly first to schwa, then often lost altogether, so inflectional endings got lost (Brunner 1960/1962, I: 68–69, 344–347).[7] With grammatical suffixes lost and the functions which they carried no longer expressed, syntactic functions came to be expressed by alternative grammatical means, like the fixing of the SVO word order and the strengthening of function words (with *to* marking a former dative role, or modal verbs and adverbs expressing the functions of subjunctives, for example). It can thus be argued that the radical change of the character of English is the product of a chain of effects, from a stress detail to phonology, to morphology, and ultimately affecting syntax – a minute cause having a far-reaching effect over time. And the trend continues further in some World Englishes, e.g., in China English where, due to a Sinitic substrate, the *-s* on verbs or nouns is often not realized, yielding an even more analytic language variety (Schneider 2011b: 151–154).

**7** This process is assumed to have been accelerated substantially by language contact with Old Norse; Germanic languages without such a contact history move along the same trajectory but have retained inflectional endings more strongly.

### 3.7 Emergentism

Emergentism is a core property of CDS: smaller, less complicated entities interact, form new linkages, and jointly build gradually more complex units with new functions. It may be interpreted as a reflection of a fundamentally cooperative, complexity-building principle in life which can be taken to counter an emphasis on competition (e.g., the survival of the fittest) and fragmentation (as in structuralism) as basic organizing forces in life.

Again, this property is straightforwardly applicable to languages and English(es), where, for example, in line with the principle of compositionality, small units constantly build larger constituents, for instance in the emergence of compounds, word formation, collocations, idioms, etc. Beckner et al. (2009: 18) explicitly state: "[l]inguistic patterns are not preordained by God, genes, school curriculum, or other human policy. Instead, they are emergent" (cf. Hopper 1987).

### 3.8 Auto-organization

Auto-organization is a "twin process" to emergentism. This property involves the self-organizing capacity of systems evolving towards higher-order, more complex sub-systems and functional organization levels. Mobus and Kalton (2015: 476) describe auto-organization by stating that "systems can self-organize (i.e., become more complex)", although this development does not imply any planned directionality: "evolution is understood as a systematically produced trajectory of increasing complexity that need not be teleologically headed anywhere".

Obviously, growth of systemic relations and a functional increase of complexity are also properties of English(es) and languages. Construction Grammar approaches, for example, posit a change from rather simple and concrete form-meaning pairings (words) to increasingly abstract schematic constructions (e.g., Hoffmann and Trousdale 2013). Related notions are Sapir's (1921) concept of "drift" in languages or Keller's (1994) idea of an "invisible hand" steering developments.

## 4 Meanderings of *wh*-pronouns

In the following sections, I discuss changes in the system and morphosyntax of *wh*-pronouns in the light of CDS principles.

## 4.1 Theoretical background and methodology: "Meanderings" and sources

I use the notion of "meanderings" as a metaphor for how changes proceed in Englishes as a CDS, partly in a random, non-directional fashion but partly also motivated by the impact of fundamental structuring principles, imposed by, e.g., cognition, functional needs, and social parameters. Again, many of these factors have of course been discussed in the literature before, but usually in isolation. What I emphasize here is the interaction of all of these forces and factors as part of an overarching system. The notion implies boundaries being overcome and edges being redefined. In the following sections it will be shown that my approach of pointing out "meanderings" involves looking at:

- changes in the set of available forms, with some innovative ones appearing and some older ones disappearing in the course of time;
- changes in distributional principles, as to which forms are used in which contexts, and why;
- changes in frequency distributions and preferences of individual forms and patterns in specific contexts, a continuous waxing and waning; and
- changes in the formal expression of functions, e.g., the morphosyntactic marking of meaningful grammatical categories and relations.

These processes may be partly motivated (e.g., by well-known semantic categorizations or syntactic principles), and partly they appear to display random frequency fluctuations, but in any case we witness perpetually ongoing dynamics, in line with the properties of CDS.

My reconceptualization of the history of *wh*-forms through the lens of CDS thinking builds upon standard work on the history of English, especially on historical morphosyntax and on the subject specifically (e.g., Brunner 1960/1962, II; Graband 1965; Romaine 1980, 1982; Lass 1992; Aarts 1994; Ball 1996; de Haan 2002; Grund 2011), and integrates findings from some earlier work of mine (Schneider 1992a, 1992b, 1993, 1996), which will be outlined only briefly. As to word class, *wh*-forms can function as interrogative and as relativizers. My first case study (section 4.2) looks into the expansion of *wh*-interrogatives into the set of relativizer functions (and hence the variation among the set of relativizer choices). The second one (4.3), loss of case marking, applies to both word classes, and the third one (4.4), variation between genitive forms, applies to relative pronouns only. These survey descriptions of evolutionary processes focus upon essentials, thus disregarding some potentially relevant details, such as dialectal differences in earlier stages of English (on modern dialects, see Ball 1996; Tagliamonte 2002; Levey 2006). Original sample data

stem from a range of electronic corpora, all of which are well-known in corpus linguistics: the Helsinki Corpus (covering OE, ME, and Early Modern English [EModE] in about 1.6 million words); Shakespeare's drama texts; ARCHER (A Representative Corpus of Historical English Registers, version 3.2), some 3.3 million words from EModE and late Modern English (LModE) texts between 1600 and 1999; Brown and LOB (representing printed American English [AmE] and British English [BrE] from the year 1961, in one million words each); Frown and FLOB (similarly but from 1991); and sub-corpora of the International Corpus of English (www.ice-corpora.uzh.ch/en.html) from India and The Philippines (PhilE) with one million words roughly from the 1990s, and ca. 60% of them representing transcribed speech.

## 4.2 Meanderings of *wh*-pronouns: Set of relativizer forms

The set of relativizers expanded considerably between OE and EModE, with forms and choices partly motivated by structural or extralinguistic conditions which contribute to emergentism and auto-organization processes. A central process in this context was the move of *wh*-forms, which were originally interrogatives, into the relativizer paradigm.

Across the history of English, we see an expansion of *wh*-forms and relativizers in general, a manifestation of the complexity increase found in CDSs. OE had a single relative pronoun form, *þe*, which was invariant, i.e., not marked for case, gender, or number. Hence, it was frequently used in conjunction with a demonstrative pronoun, which then expressed the grammatical categories and relations needed, e.g., *se þe*, and demonstratives themselves could also be used for relative clause linkage (Romaine 1982: 56–57). Relativization with a zero form was also available, also as the subject of a relative clause, unlike in standard ModE.

From OE through ME into EModE a series of complex changes expanded the range of forms available (cf. Brunner 1960/1962, II: 146–156; Nevalainen and Raumolin-Brunberg 2002), with three particularly important changes. Interrogatives became relativizers, a process that created a new system-internal network relation. Between the thirteenth and the fifteenth centuries formerly interrogative pronouns (*wh*-forms) intruded into relative functions. Case-marked forms (*whom* and *whose*) appeared earlier than the base form *who*. The late appearance of subject *who* is systemically "unexpected", according to Nevalainen and Raumolin-Brunberg (2002: 109), since it violates the well-known Noun Phrase Accessibility Hierarchy, which predicts an innovation to spread first to subjects (cf. Beal and Corrigan 2002: 125–126). Two causes have been posited, which most likely interacted: the impact of the high-status models of Latin and/

or French (*wh-* forms being derived from forms like *qui*; Romaine 1980: 223, 1982: 61–63), and the structural need to express more complex syntactic relationships between the relativizer in its clause and its antecedent (implied by the primacy of the case-marked forms, which offer precisely such possessive or object-like connections).

Secondly, in the thirteenth century *that* begins to appear as a relativizer. It may be surmised that, while there is no continuity of form functionally and stylistically, this use may represent a continuation of OE *þe*, since *that* is also uninflected and unmarked for case, number, and gender (cf. Romaine 1982: 59–60). As such, in CDS terms, it represents a small pocket of order retained in the system.

Finally, the use of zero relativizers (often called "contact clauses") also expands (Romaine 1982: 74–78), but then, by ModE, it is reduced to non-subject functions (which were rare in OE). The omission of subject relative pronouns is no longer acceptable according to prescriptive rules in standard ModE, although the structure is possible in nonstandard dialects, as in (1).

(1) *Its de devil ⊘ makes folkses do bad.* (Earlier AAE; from Schneider 1989: 216)

It is noteworthy that stylistic differences are persistent: to the present day *that* and zero are more informal than *wh*-relativizers (Romaine 1980).

It is interesting and telling to consider the principles behind choices of *wh*-pronoun forms. In ME and EModE, the rules of pronoun choice were recategorized drastically, based on animacy of the referents. Initially, this affected personal pronouns: the loss of inflectional endings in nouns in ME and the correlated breakdown of the article and demonstrative system resulted in the radical disappearance of grammatical gender, which consequently was replaced by natural gender in pronominalization: *he*, *she*, and *it* came to be used for males, females and nonhumans, respectively (Lass 1992: 106–108; Stenroos 2008; cf. Siemund 2008: 135–143).[8] *Wh*-pronouns, both interrogative and relative, followed suit and adopted this pattern in a slightly simplified version, conflating males and females in an animate category. Originally the choice of relativizers was not semantically determined. For example, in the King James Bible of 1611 we read "Our father which art in heaven", with *which* referring to a human (or divine) antecedent. But in EModE a systematic distinction between *who* and *which* by semantic class of referent gradually spread in the community and

**8** This process has been characterized by a large amount of variation, though – much more than earlier research had assumed (Grund 2011).

became established: *who* for humans/animates, *which* for nonhumans/inanimates.[9] In this case (unlike the introduction of *whom* and *whose* earlier), it was the base forms which adopted referent categorization based on animacy first, implying the strength of this semantic condition. Subsequently, the selection principle expanded to possessive (genitive) relations as well, though not wholly successfully: *of which* is restricted to inanimate referents, while *whose* may denote both animates and inanimates. In CDS terms, these processes illustrate perpetual dynamics, emergentism, and auto-organization (via the establishment of distributional principles shared across sub-systems), and a trend towards the establishment of order.

Overall, a range of different forces and factors play a role here, and they are clearly components of a CDS. As was stated at the beginning of section 3, these forces and factors manifest themselves in the self-reinforcing feedback loops of everyday usage, which steer the interaction between individual human agency and system evolution. Speakers behave linguistically in predictable and systematic ways in order to communicate successfully, though there is also always room for some degree of variability and, hence, the potential for trajectories of change moving in new directions. This variability may even introduce qualitative changes via a butterfly effect (as, for instance, in the occurrence of entirely new forms in a functional paradigm). The factors which had an effect on shaping the choices of relativizers, with possible relevant CDS principles, include the following:

- contact (resulting in massive borrowing from Latin and French) – increasing language-internal complexity and building cross-linguistic networks, possibly triggering chaotic / turbulent developments;
- prestige (associated with Latin-/French-derived forms as a consequence of the high social status of these languages in medieval society) – driving dynamism;
- the marking of grammatical, functional relations (case and its associated semantic roles in a clause) – increasing systemness and order;
- the marking of semantic referent categories (i.e., animacy effects) – increasing systemness and order, and establishing network relations;

**9** By necessity, given space constraints and the overall architecture of my line of argumentation, this very short account disregards many complex factors and details in this diffusion and variation process. For example, Ball (1996) shows that personal and nonpersonal subjects display different strategies of *wh*-diffusion, and that the distinction between restrictive and nonrestrictive clauses and, not surprisingly, variety differences also play a role.

- analogical extension from one sub-system to another (with properties, i.e., relevant semantic distinctions, of personal pronouns being transferred to *wh*-pronouns) – contributing to auto-organization;
- modification and adjustment of such principles in transition (with the male – female – inanimate categorization getting reduced to animate – inanimate) – complexifying, and influencing the tension between order and chaos, and influencing auto-organization; and
- simplicity vs. complexity effects (for example, *who* as a form without case marking is simpler than *who*(*m*), but *of which*, as a multi-word lexeme, is more complex than *whose*, a one-word form) – increasing or decreasing complexity, order, and network relations.

Thus, a constant interaction of and competition between different forces and principles, with varying effects at different points in space and time, in a sense the "meanderings" introduced above, produced an increasingly complex subsystem of relativizer choices and *wh*-pronoun selection, in line with a number of CDS effects shown above and summarized in section 5.

## 4.3 The loss of case marking in *wh*-pronouns: *Whom* > *who*

In non-subject positions, the case-marked form *whom* has increasingly been replaced by unmarked *who* in both interrogative and relative functions, a process commonly seen as a final stage of the general loss of case marking in nominals since ME (Mair 2006: 141–143). While it is difficult to explain this change on the basis of any one principle of linguistic change, CDS illustrates that this change proceeded logically from a combination of perpetual dynamism, network relations, and auto-organization.

Present-day English allows a choice between non-subject *whom* and *who*, as in (2).

(2) (a) ***Whom*** *did you see?*
(b) ***Who*** *did you see?*

The loss of case marking in *wh*-pronouns is a well-known usage problem, discussed in style guides, with the unmarked form often branded as "incorrect" by conservative language observers. Two distribution principles are known to be effective in ModE (Aarts 1994; Aarts and Aarts 2002; de Haan 2002). One is style: *who* is informal and characteristic of speech, while *whom* predominates

in formal, written language.[10] Secondly, the influence of the position of a preposition is strong: the inflected form is the norm with so-called "pied piping", where the preposition immediately precedes the relativizer, as in (3a), but the uninflected form is preferred with preposition stranding, as in (3b), which is an informal construction.

(3) (a) ***To whom*** *did you give it?* (pied piping)
(b) ***Who*** *did you give it* ***to****?* (preposition stranding)

Historically, the case-marked form, attested as an interrogative since late OE and as a relative a little later (see *OED*, s.v. *whom* I.1., III.), results from a systemic simplification, the coalescence of the old dative *whæm* and accusative *whone* pronouns between OE and ME, producing a single object case form *whom*. Non-subject *who* appeared in the fifteenth century, first as a relative and even later as an interrogative pronoun (*OED*, s.v. *who* 5., 13.). Structurally, this is in line with the loss of endings in many word classes and the change of character from a synthetic to an analytic language. Sapir (1921: 167) used the loss of case in *whom* as a prototypical manifestation of the "drift" of English towards complete absence of inflectional endings. He predicted the future disappearance of *whom* and hypothesized it would be gone "within a couple of hundred years". But clearly the loss of the final nasal in *whom* is not a modern fallacy, as Strang (1970: 143) observed: "[t]he present uncertainties about the use of *who* and *whom* are perfectly familiar in the 16th century – indeed they are almost as old as the use of the forms as relatives".

Schneider (1996) traces the relationship between both forms roughly through the last four hundred years. In EModE, represented by the Helsinki Corpus (HC), *whom* clearly predominates. Non-subject *who* begins to appear in noteworthy numbers after 1570 (3% of the sum total of non-subject *whom*/*who*) and increases in the sub-period after 1640–1710 to 8%. By text types in the HC data from 1570–1710, relatively high proportions of the uninflected form are found mainly in informal genres (drama and comedy, private correspondence, less so fiction and diaries; Schneider 1996: 490). Thus, *who* for *whom* is an innovation which spreads successfully in late EModE and represents a change from below in sociolinguistic terms.

Given that this pattern of variation started some five hundred years ago, it is possible to carry out a long-term comparison based on real-time evidence

**10** Based on ICE-GB data, Aarts and Aarts (2002: 126) report a normalized frequency of *whom* of 26.5 per million words (pmw) in speech as opposed to 96.8 pmw in writing.

(though methodologically it is not so easy to identify fully comparable text types). This comparison, a rare chance to trace the perpetual dynamics and network adjustment processes within a sub-system of a CDS, is based on the 1570–1710 texts of the HC and Shakespeare's prose sections of dramas only, the style closest to colloquial speech, for written and spoken EModE, respectively (from Schneider 1992b and 1996), and on the written and spoken parts of the British component of the International Corpus of English (ICE-GB), newly investigated for this paper. In writing, the proportion of non-subject *who* out of all forms in that function is 4% (13/293), as opposed to 11% in the written part of ICE-GB from the early 1990s. As to representations of speech, the proportion of *who* is 35% (24/92) in Shakespeare as opposed to 56% in the spoken component of ICE-GB. So what we find is an increase of *who* over 400 years, but a fairly moderate one, and especially in speech, but not so much in writing.

The Brown quartet of corpora, covering BrE (LOB/FLOB) and AmE (Brown/Frown) from 1961 and 1991, allows another, more recent real-time developmental comparison (though based on written, even printed texts only).[11] It shows that the proportion of *who* in formal writing (of the sum of non-subject *who* plus *whom*) is still low but that it increased slightly in both varieties (more so in AmE, from 1.4% to 10.8%, as opposed to 3.5% to 9.4% in BrE). In absolute numbers, *whom* also increased in AmE (from 144 to 166 tokens) but decreased in BrE (219 to 174) during that thirty-year interval, which is not compensated by the increase of *who* (8 to 18). It seems then that the overall construction type of animate *wh*-objects and complements is used less frequently.

It is interesting to ask how the variation between non-subject *who* and *whom* has diffused into second-language World Englishes (Schneider 2007, 2011a). Possible issues and questions are whether we see effects of archaic retention (the so-called colonial lag, with conservative *whom*) or innovation (e.g., simplification in L2, promoting uninflected *who*), or whether the continuing impact of BrE or the more recent influence of AmE can be observed. With the latter goal in mind, here I focus on ICE data from IndE and PhilE.

Overall, in India non-subject *who* is rare (7 out of 104 instances; 6.7%), whereas in the Philippines it occurs in 38.8% of all tokens (66/170). IndE usage thus appears to be comparatively conservative, with hardly any non-subject *who*, while this form is fairly common in the Philippines, perhaps following the model of less formal American usage (see Schneider 1992a). Figure 1 compares the corresponding style distribution in each variety. It documents a remarkable style sensitivity in PhilE (which is almost completely lacking in IndE): *who*

**11** The 1961 data are from Schneider (1992a); the 1991 ones are new.

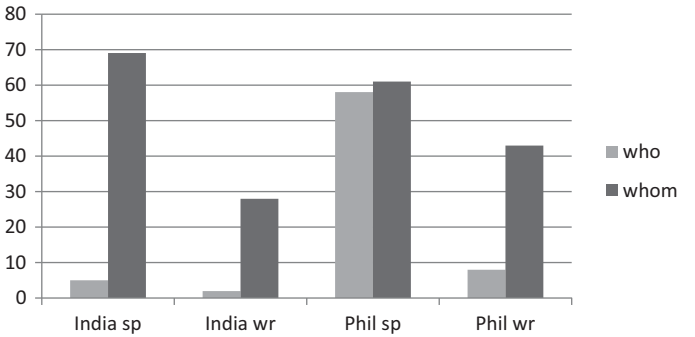


**Figure 1:** Non-subject *who* vs. *whom* by styles in IndE and PhilE (sp = speech; wr = writing).

takes almost half of the tally in speech but remains rare in writing. We see the perpetual dynamics of a CDS operating here: some relationships remain similar; others are reshuffled. It could be argued that this development reflects an embryonic stage of auto-organization, with different varieties developing different style associations with the same form.

IndE and PhilE do share one significant similarity that distinguishes them from other Englishes: *whom* with a stranded preposition, as in (4)–(7), is frequent in both. By contrast, this construction is extremely rare in BrE and AmE, maybe because of a perceived stylistic incompatibility (the form being formal but the pattern informal). It appears that style associations have not been passed on firmly from first- to second-language varieties. In a CDS perspective, this manifests intra-systemic dynamism involving network relations, and possibly emergentism.

(4) *when they surrendered,* ***whom*** *did they surrender* ***to****?* (ICE-Ind S1b-059)

(5) *party member who* ***whom*** [sic] *I could not say no* ***to****.* (ICE-Phil S1b-041)[12]

(6) *these people* ***whom*** *you are now so concerned* ***about*** (ICE-Phil S1b-059)

(7) *please call those* ***whom*** *you are in contact* ***with*** (ICE-Phil W1b-014)

Overall, the developments of *who* and *whom* in space and time again show a number of meanderings, with varying causes and constraints, illustrating, as argued

**12** The sequence *who whom* is a self-initiated repair, indicating a high degree of linguistic insecurity concerning which form is required.

above, CDS principles such as perpetual dynamics, network relations, and developments towards new orderly patterns, i.e., auto-organization. As to causes, *whom* was introduced to satisfy a need for case-marked clause binding, while non-subject *who* originally seems to have been triggered by the effect of an apparent subject position, perhaps somewhere between analogy and simplification (for further discussion, see Schneider 1996). Over time, a rather weak, gradual change, a moderate increase of non-subject *who* over more than 400 years in speech, can be observed. A rather persistent effect (i.e., a pocket of order and stability in the sub-system) is the style distribution, with *who* preferred in speech and *whom* predominant in writing at all times and in all varieties. In the twentieth century, however, we find a partial reversal of a trajectory, and also splitting trajectories, movements in different directions. While *whom* increases in written AmE, it decreases in BrE. In speech *who* becomes well established in the Philippines but not at all in India.

## 4.4 Competition of *wh*-pronoun forms in the genitive: *Whose* vs. *of which*

The third, most interesting example of complex developments in *wh*-pronouns concerns the variation between the *wh*-pronouns *whose* and *of which* in genitival position (as in 8a and 8b).[13] Again, this example of meanderings shows varying phenomena, causes and constraints in different times and varieties. In this case, in addition to dynamics and network relations, there is an increase of complexity, and there are oscillations of systemic relations between moderate versions of chaos (via the breakdown of older relations) and the emergence of new patterns of order.

*Whose* and *of which* stand in a direct paradigmatic relationship of competition, and again this is a competition which has been around since the ME (Johansson 2002) and EModE periods.

(8) (a) *a presbyterian* ***whose*** *name they gave out* (HC, E3 NN BIA FOX)
(b) *two hands, the fingers* ***of which*** *are intermixed* (ARCHER, 1787mark_f4a)

**13** This section largely builds upon and summarizes findings from Schneider (1993). Data from the Helsinki Corpus stem from that source, while all other data (from ARCHER, the Brown quartet, and the ICE corpora) are new.

The focus here is on occurrences of these forms in possessive relative function, excluding other constructions, e.g., interrogatives (where *of which* is very unlikely), absolutes, and chance co-occurrences. Historically, possessive relatives are secondary and came later (with interrogative uses being first). The earliest relative uses of *whose* for persons are attested from ca. 1200, for inanimates in 1382, and of *of which* from 1423 (*OED*, s.vv. *whose*, *which*). In terms of frequencies (based on HC data; cf. Schneider 1993: 242), *whose* predominated throughout EModE (occurring between 72% and 76% of the time in the three sub-periods of EModE as represented in the HC), and *of which* gained ground slowly (with 12% each in the first two periods, rising to 17% in the third one, 1640–1710).

Interestingly enough, a third form, *whereof*, in (9) and (10), is also found in a paradigmatic relation with the others in this slot for a while.

(9) *Euxton chapel* [. . .] *the key* ***whereof*** *was in the hands of* [. . .] (ARCHER, 1687carty2b)

(10) *goods or chattels,* ***whereof*** *they shall be reputed owners* (ARCHER, 1790prce_l4a)

Its earliest relative uses stem from the fifteenth century. During the sixteenth and seventeenth centuries, it was common, and then it disappeared – but left notable grammatical traces (see below). This waxing and waning is also reflected in the EModE HC data: in the three sub-periods it occupies 12%, 17%, and 10% of all genitival relativizers, respectively (Schneider 1993: 242). In ARCHER, the loss of *whereof* shows in clearly decreasing token frequencies: between the seventeenth and the twentieth centuries we find 35, 17, 5, and zero tokens, respectively. In modern data, no examples of *whereof* appear in Frown, FLOB, ICE-India, and ICE-Phil (i.e., in all the 1990s corpora). However, there are eight examples of *whereof* in Brown, all in genre H, in frozen formulae in legal language (*in testimony whereof*, 7 times in text H08; *in witness whereof*, once in H22), and three in LOB, also all in H and in varying formulaic contexts (*in witness / by consequence / the directors whereof*). In other words, the form *whereof* survived as a partly frozen legalese term up to the 1960s but has vanished completely since.

Two major constraints have been in competition in the choice between *whose* and *of which*, which reinforce alternative trajectories of change. One is animacy: for inanimate referents both *of which* and *whose* are available, but for human referents, only *whose* can be chosen. This uneven distribution implies that the animacy principle has been effective but only partially so – an instance of auto-organization which has established order only half-way through, as it

were. Second, there is complexity (meant here in the non-CDS-technical understanding often discussed in linguistics and referred to in footnote 5): *whose* is a pre-nominal one-word unit which stands in the determiner position, so it has canonical syntactic properties, while *of which* is a two-word unit in the post-nominal position outside of the determiner slot which it is supposed to fill functionally. Thus, there is substantial difference in complexity levels, and to the extent that simplicity contributes to the dynamics towards orderly systemic relations this difference may hinder the advancement of *of which*. On the other hand, this represents another instance of the major typological change of English during that period, the long-term dynamism of an analytic form encroaching upon the territory of an older synthetic form. Functionally, it was needed because it carries the animacy distinction earlier established in the base form (*who* vs. *which*) into the possessive relation. This development thus strengthens network relations and increases system-internal order.

The grammaticalization of *of which* as a possessive relative and the acquisition of its grammatical properties can be observed to have proceeded systematically through a number of intermediate steps. Schneider (1993) hypothesized that the combination of both words started as cases of chance coincidence (as in 11a), went through a firm partitive reading with a numeral as antecedent noun (11b), and finally broadened semantically to signal a possessive relation with any antecedent noun (11c).[14] These steps can be viewed as higher complexity levels and increased systemic order emerging in the course of time.

(11) (a) *poverte shal been in foure thynges: In defaute of tresor,* ***of which*** *that David* ***seith***, [. . .] (Chaucer, The Parson`s Tale; c1390)
(b) *to which Ilands we came the 29. day of March, 1609.* ***two of which*** *Ilands were within a league one of another,* (Coverte, Report of An Englishman; 1612)
(c) *blood in the guts, the peristaltick* ***motion of which*** *was scarce discernable;* (Hooke, Micrographia; 1665)

The acquisition of the grammatical properties of *of which* as a possessive relative is also interesting and can be observed through transitional stages. This process again illustrates complex system auto-organization in action. Possessive relatives obligatorily express three distinctive semantic properties (which are not formally expressed but need to be understood by the recipient): definiteness (the ability to uniquely identify the referent), possessiveness, and relative junction (the

**14** Examples (11)–(13) are from Schneider (1993), ultimately based on the HC.

grammatical co-reference relation between antecedent and relative pronoun). *Of which* shows some notable characteristics in this regard. In early stages these three properties may be signaled by features that co-occur with *of which* rather than by *of which* itself. This suggests that the properties are not yet clearly associated with the form *of which*, and thus the relevant information needs to be flagged explicitly by additional means during a transitional stage. Definiteness is often marked explicitly by adding the definite article *the*, yielding the form *of the which* (12a).[15] Possessiveness can be additionally expressed by a possessive pronoun (12b). Relativization can be highlighted by three strategies of expression: by adding the old relativizer *that*, yielding *of which that* (12c), by repeating the referent noun (12d), or by supplementing a near-synonym of the referent noun(s).

(12) (a) *in a full gret palays* [. . .], ***Of the whiche*** *the walles ben in circuyt more than .ij. myle* (Mandeville`s Travels; c1400)
(b) *thei ben lyk to briddes* ***of whiche*** *the nyght lightneth* ***hir*** *lokynge* (Chaucer, Boethius; c1380)
(c) *for of thinges* ***of whiche that*** *the effect nys nat naturely divers, nedes the substaunce moot be* [. . .] (Chaucer, Boethius; c1380)
(d) *and incloased it with a strong* ***wall*** *of stone for a place of buriall,* ***of which walle*** *many partes continue at this tyme* (Hayward, Annals; a1627)

While these additional markers may seem redundant and stylistically awkward from a modern perspective, these syntactically complex structures appear to have been needed to express the relationship in question, which at that stage is not yet clearly signaled by *of which* itself. We see the dynamic system auto-organizing itself, which increases complexity and builds network ties to other structures. Consequently, these features are transitional, as Schneider (1993: 247) documents: the strategy of repeating the antecedent (as in 12d) consistently decreases throughout late ME and EModE and ultimately disappears completely after 1640.

Similarly, the fixing of the syntactic position of *of which* follows a systematic developmental trajectory. Possessive relativizers can occupy the pre-nominal position (as in 13 and 8a) or the post-nominal slot (in 14 and 8b).

(13) [. . .] *suche as ben ygnoraunt.* ***Of which the nombre*** *is infenyte /* (Caxton, Prologues and Epilogues; 1483)

**15** One reviewer points out that this is similar to other instances of grammaticalization and that such "redundancy", motivated by an urge for being noticed and properly understood before the novel construction is properly semanticized, contributes to making the system dynamic.

(14) [. . .] *apparell, the due* ***consideration whereof*** *is referred* [. . .], (HC, E2 XX CORO WCECIL)

Of these, *whose* always occupies the pre-nominal position, in line with its role as a determiner in the noun phrase. *Whereof* appears predominantly (in 79%, i.e., 30/38 of all instances in the HC) post-nominally. The most notable development, however, affects *of which*, a competition between and switch of pre-nominal and post-nominal positions. Figure 2 shows its clear change in positional preferences.

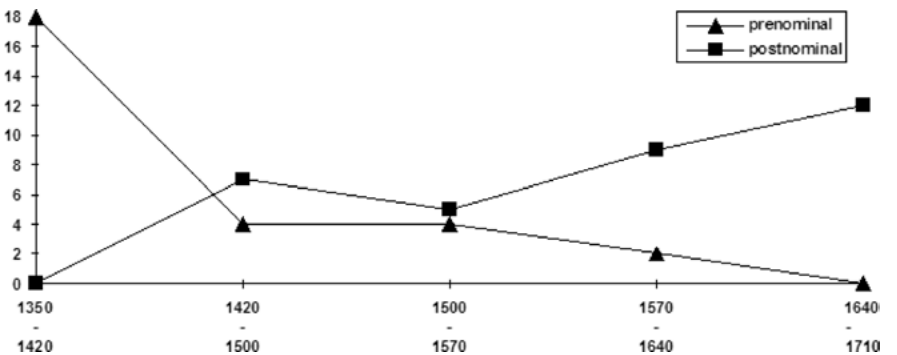


**Figure 2:** Syntactic position of *of which* by periods (Helsinki Corpus, token numbers; from Schneider 1993: 250).

In ME, *of which* starts out exclusively pre-nominally, in the same slot as *whose*, which it supplements semantically (for explicitly inanimate reference). In the fifteenth century, it encroaches on the post-nominal slot which had earlier been occupied (and established for this function as an option) by *whereof*. Throughout EModE, both forms compete there, until by the second half of the seventeenth century *of which* occupies the post-nominal position exclusively. So today's post-nominal position of *of which*, which increases its syntactic complexity, took over and retains an element of structural heritage of its EModE competitor *whereof*.[16] As one reviewer comments, the fact that replacement cycles preserve structure also fits well into the CDS perspective: the change is neither entirely random nor predictable.

**16** Johansson (1997) documents some persistent variability in ModE, depending upon style and syntactic context.

This developmental trend observed in HC data is confirmed and shown to be continuing in texts from ARCHER, i.e., the dynamism, mutual network rearrangements, and system rearrangement can be shown to continue throughout the next few centuries. While *whose* continues pre-nominally throughout, *of which* occupies the post-nominal position inherited from *whereof* by the mid-seventeenth century (by which time *whereof* itself is largely gone): pre-nominal *of which* is still noticeably strong (2 out of 3 instances) in the first ARCHER subperiod (1600–1649) but almost completely absent after 1650. As to overall frequencies (Figure 3), *of which* keeps gaining ground until about 1900 but recedes in the twentieth century. Conversely, *whose*, the form which predominates throughout, gradually loses some ground to *of which* 1600–1900, but rises again thereafter.

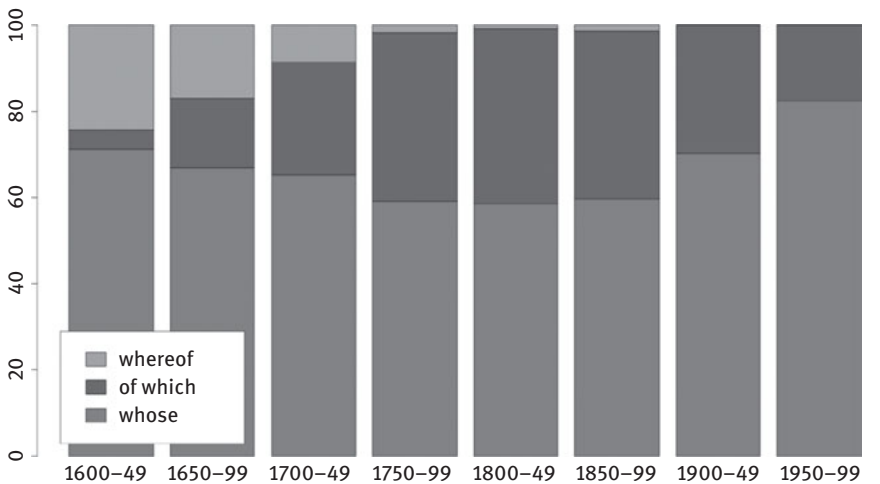


**Figure 3:** Frequency changes of *whereof*, *of which*, and *whose* 1600–1999 (by ARCHER periods; percentages).

As Figure 4, with data from the Brown quartet, shows, this trend continues in the twentieth century. As before, *whose* predominates and keeps increasing. *Of which*, which is considerably less common anyhow, decreases even further in frequency between 1961 and 1991, both in BrE (from 19% down to 15%) and in AmE (15% to 8%), and it is even less frequent in AmE than in BrE. The trend towards *whose* replacing *of which* can be interpreted as a case of complexity avoidance, with the two-word item in the marked post-nominal position increasingly dispreferred.

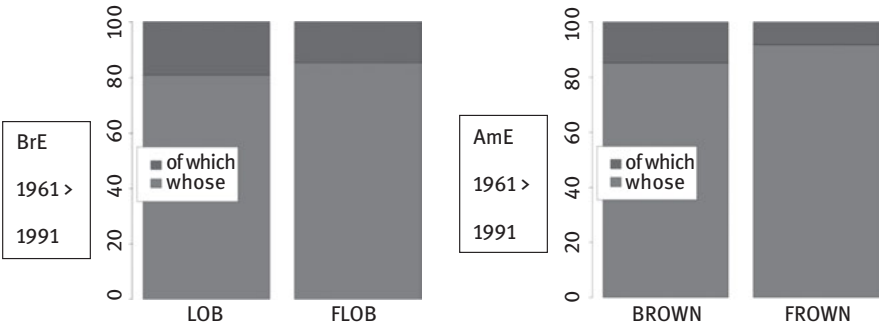


**Figure 4:** Frequency changes of *whose* and *of which* in the later twentieth century (Brown quartet; percentages).

Finally, what happens when these forms diffuse into New Englishes? Basically, the recessive trend of *of which* in AmE and BrE is not as strong there, and it seems to have taken on a life of its own. Figure 5 shows the relative frequencies. In absolute numbers, genitival relativizers are moderately more common in writing than in speech (188 vs. 122 tokens overall in both varieties), which is in line with Johansson's (1995) findings.

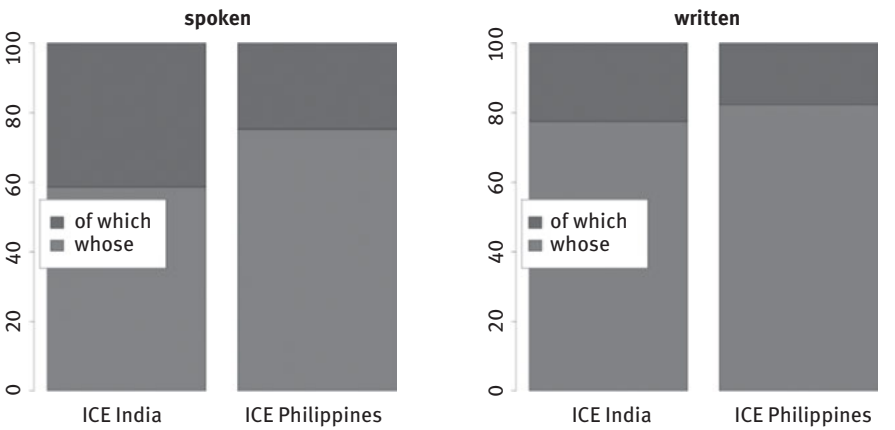


**Figure 5:** Frequencies of *whose* and *of which* in IndE and PhilE by speech and writing (percentages).

Overall, the proportions of *of which* are moderately higher than in BrE and AmE, at about 30% in IndE and 20% in PhilE. This may reflect their origin in

historically earlier periods when the proportions of *of which* were still higher, as well as the fact that IndE is older than PhilE (cf. Figure 3). They are also in line with diffusion from the respective source varieties: *of which* is less common in PhilE than in IndE, similar to what was observed for AmE vs. BrE. *Of which* is relatively more common in speech than in writing, and this difference is more pronounced in IndE. In this variety, perhaps because of its greater proximity to stages of second language acquisition, which favors regularity and simplicity (Schneider 2011a: 192–194), analyticity may come more naturally. Inflection, by contrast, possibly emphasized in the education system, is associated more strongly with formality. These relationships show that sub-domains within a CDS, in our case different varieties of English, display slightly varying effects and trajectories of change in their perpetual dynamics and re-organization. These variable developments depend in these cases on persistent effects such as input time or relationship to other varieties as well as innovation.

In a similar vein, some special, qualitative observations on *of which* in New Englishes can be drawn from the ICE corpora. These observations illustrate the differential speed of development, the potential for different pockets of orderliness, and also the potential for the emergence of qualitatively new properties within a CDS. There are a few cases of archaic retentions, phenomena discussed above and found in earlier times but lost in AmE and BrE: ICE-Ind has one instance of the repetition of the referent noun to make the relationship explicit (15), and also four examples of *of which* in pre-nominal position (as in 16). In addition, there are also patterns which appear innovative, with *of which* used for extended (clausal) reference, by necessity somewhat vaguely. Examples are found in (17) and (18), one from each variety, of reference to an antecedent clause, not a noun. The other examples along this line, all from IndE, show collocational chunking tendencies: one instance of *as a consequence of which* (ICE-Ind S2b-035) and three instances of *as a result of which* (ICE-Ind W1a-008 (twice), ICE-Ind W1a-013).

(15) *we have in all seven* ***units***, ***of which*** *the first* ***unit*** *is called informative writing* (ICE-Ind S1b-072.txt)

(16) *blood file* [. . .] ***of which the result*** *shows inconclusive* (ICE-Ind S2a-067.txt)

(17) *I had registered for it on April 22 1982* ***of which*** *the receipt is attached to this cover.* (ICE-Ind W1b-022

(18) *Sulu slaves spent a life of captivity vestiges* ***of which*** *may be observed in* [. . .] (ICE-Phil W2a-011)

# 5 Meanderings of *wh*-pronouns: The CDS perspective

Adopting a bird's eye perspective, practically all CDS principles, discussed above in section 3 and throughout the analysis in section 4, and recapitulated in this section, turn out to have been effective in the evolution of *wh*-pronouns; some evolutionary processes actually are motivated by and can be subsumed under several CDS headings. Fundamentally, the system turns out to be perpetually in motion, with new forms appearing (*wh*-pronouns in general, *of which*) and occupying specific functional positions, and others disappearing at some point in time (*whereof*). It tends to get increasingly complex, with new forms (case-marked *wh*-pronouns, *of which* as a complex lexeme) evolving. New systematic relationships (increasing order) emerge; for example, semantic categorizations of referents are established and spread to varying degrees. Some qualities and relations are persistent (e.g., style associations of relativizers); some constraints are modified (e.g., whether zero subjects occur or not); and there are also some fairly radical innovations, i.e., qualitative changes (e.g., the emergence of *of which*). The establishment of *of which* can be regarded as complexification: it is a new multiple-word unit as opposed to a one-word unit, and it is unstable in its structural properties, as shown in 4.4. Effects of network relations surface in the transfer of animacy from personal to relative pronouns or in the position competition between *of which*, *whose*, and *whereof*. Auto-organization is to be identified in animacy transfer and also the reduction of animacy categories from three (male – female – neuter) to two (animate – inanimate).

Manifestations of *systemness*, structured relationships between units, are evident all across the board: in the paradigmatic relationship between *wh*-forms, *that*, and zero, or between *whose* and *of which* (and, in between, *whereof*); in the loss of inflection in *whom* – analogously to other nominals; in the transfer of animacy as a principle underlying formal choices from personal to relative pronouns and, later, from base to genitival forms (only partially in the latter case, and hence we witness varying degrees of systematicity).

Emerging *complexity* produces the combination of *of* + *which*, both jointly becoming established as a single unit, and then also the development of the properties of *of which* (its increasingly tightly defined reference potential, which transitionally needs to be marked explicitly; fixing of position options), and also in the transitional competition between the three genitival *wh*-relativizers in EModE.

*Perpetual dynamics* reigns throughout: we observe ongoing processes and changes, both principled and random, across time and space, from OE through

the periods of English to the present day, from England to North America and to the world, ongoing in the younger World Englishes. The persistent transmission of forms and relationships secures a basic degree of continuity but also allows the potential of change – changing frequencies, changing structural relationships, and innovations (like the establishment of *of which* as a new player). This shows most readily in the varying degrees of diffusion of the relative pronoun forms themselves in various contexts, the appearance of *wh*-forms like *who*, *whose*, etc., in relativizer functions, or the introduction and disappearance of *whereof*, down to recognizable differences in developmental trajectories (such as the differing acceptance of non-subject *who* in PhilE and IndE discussed in 4.3).

Clearly many of these processes can also be viewed as *increasing* or *decreasing order*, manifestations of the paths from order to chaos and back. Processes increasing or restoring orderly relations are the evolution of a principled set of forms of relative pronouns and the paradigmatic relations between them, or the partial restoration of order via effects of some motivated distributions, e.g., along functional lines (interrogative pronouns generally becoming relative pronouns) or along semantic lines (with animacy in personal pronouns spreading to regulate relative pronoun choices). However, we also witness *turbulence* and noise, decreasing orderliness and systematicity, e.g., in the introduction of contact-induced borrowings (such as *wh*-pronouns from Latin/French), in random fluctuations of frequencies across periods and varieties (through historical phases, or from BrE to AmE to World Englishes).

Several of these processes clearly also show ongoing processes of *emergence* and *auto-organization*, with the system itself developing and establishing new relationships and functional associations and entities. A case in point is the competition between forms in the same functional position, e.g., *whose* and *of which* (with semantic distinctions being gradually negotiated and established), or *whereof* and *of which* in EModE (with the latter adopting functional properties, e.g., the post-nominal position, of the former). Processes of re-allocation, i.e., old forms (interrogatives) adopting new functions (as relatives), can also be seen in this light. The same applies to processes such as the initial appearance of *wh*-pronouns with case marking (motivated by a functional need of expressing increasingly complex relationships of clause binding), the stepwise acquisition of specific properties (definiteness, possessiveness, reference) of *of which*, or developments to settle positional choices (in the relationship between *whereof* and postnominal *of which*). The mechanisms which implement these principles are essentially the causes of language change: for instance, similarity relations or forces of analogy are important driving forces behind the ongoing dynamic changes in the system.

The *butterfly effect* typical of chaotic systems can be identified in late stages of a chain of influences: the fixing of stress in Germanic on the first syllable caused a progressive loss of endings, which removed grammatical gender and resulted in the rearrangement of the semantics of personal pronouns based on natural gender – and these processes, in turn, affected relative pronouns in later developmental stages (with the rearrangement of the semantics of *wh*-pronouns or the loss of the old case ending in *whom*).

## 6 Conclusion: World Englishes as CDS

Given the above pieces of evidence and discussions, it makes perfect sense to view Englishes, the set of varieties of English both diachronic and dialectal, including World Englishes, as manifestations of a Complex Dynamic System, driven by the evolutionary and emergentist principles which characterize such systems. We see varying but partly overlapping forms and functions of *wh*-pronouns in different varieties of English and in their respective grammatical sub-domains, but those forms and functions are the result of broader linguistic forces that cut across those varieties. Studying *wh*-pronouns in those varieties in isolation fails to capture such dynamics, while, as I have demonstrated in this chapter, applying the CDS framework opens up new possibilities to understand language variation and change as a more comprehensive system, encompassing both linguistic and extralinguistic factors. In other words, CDS helps us cross variety boundaries as we study language. With Kretzschmar (2015, 2018) and others, thus, this is a call for a change of framework and perspective, giving preference to a holistic and processual theory in line with new tendencies in the sciences (e.g., connectionist, holistic approaches) rather than reductionist, segmenting, and categorial thinking. Accepting CDS as an underlying framework of understanding language variability and change implies that former boundaries are crossed in significant ways.

## References

Aarts, Flor. 1994. Relative *who* and *whom*: Prescriptive rules and linguistic reality. *American Speech* 69(1). 71–79.

Aarts, Flor & Bas Aarts. 2002. Relative *whom*: "A mischief-maker". In Andreas Fischer, Gunnel Tottie & Hans Martin Lehman (eds.), *Text types and corpora. Studies in honour of Udo Fries*, 123–130. Tübingen: Narr.

Ball, Catherine. 1996. A diachronic study of relative markers in spoken and written English. *Language Variation and Change* 8(2). 227–258.
Beal, Joan & Karen Corrigan. 2002. Relatives in Tyneside and Northumbrian English. In Patricia Poussa (ed.), *Relativisation on the North Sea littoral*, 125–134. Munich: Lincom Europa.
Beckner, Clay, Richard Blythe, Joan Bybee, Morten H. Christiansen, William Croft, Nick C. Ellis, John Holland, Jinyun Ke, Diane Larsen-Freeman & Tom Schoenemann. 2009. Language is a Complex Adaptive System: Position paper. In Nick C. Ellis & Diane Larsen-Freeman (eds.), *Language as a complex adaptive system*, 1–26. Malden, MA: Wiley.
Bergs, Alexander. 2012. The Uniformitarian Principle and the risk of anachronisms in language and social history. In Juan-Manuel Hernández-Campoy & J. Camilo Conde-Silvestre (eds.) *The handbook of historical sociolinguistics*, 83–101. Malden, MA: Wiley-Blackwell.
Bossomaier, Terry R. J. & David G. Green (eds.). 2000. *Complex systems*. Cambridge: Cambridge University Press.
Brunner, Karl. 1960/1962. *Die englische Sprache. Ihre geschichtliche Entwicklung*. Vols. 1–2. Tübingen: Niemeyer.
Burkette, Alison & William Kretzschmar. 2018. *Exploring linguistic science. Language use, complexity and interaction*. Cambridge: Cambridge University Press.
Bybee, Joan. 2010. *Language, usage and cognition*. Cambridge: Cambridge University Press.
Chambers, J. K. 2006–2007. Geolinguistic patterns in a vast speech community. *Linguistica Atlantica* 27–28. 27–36.
Cheshire, Jenny. 1994. Standardization and the English irregular verbs. In Dieter Stein & Ingrid Tieken-Boon van Ostade (eds.), *Towards a standard English 1600–1800*, 115–133. Berlin: Mouton de Gruyter.
De Haan, Pieter. 2002. *Whom* is not dead? In Pam Peters, Peter Collins & Adam Smith (eds.), *New frontiers of corpus research: Papers from the twenty first international conference on English language research on computerized corpora, Sydney 2000*, 215–228. Amsterdam: Rodopi.
Ellegård, Alvar. 1953. *The auxiliary* do*: The establishment and regulation of its use in English*. Stockholm: Almqvist & Wiksell.
Ellis, Nick C. 2008. The dynamics of second language emergence: Cycles of language use, language change, and language acquisition. *The Modern Language Journal* 92(2). 232–249.
Ellis, Nick C. 2011. The emergence of language as a complex adaptive system. In James Simpson (ed.), *The Routledge handbook of applied linguistics*, 666–679. London: Routledge.
Ellis, Nick C. & Diane Larsen-Freeman (eds.). 2009. *Language as a complex adaptive system*. Malden, MA: Wiley.
Fennell, Barbara. 2001. *A history of English: A sociolinguistic approach*. Oxford: Blackwell.
Gleick, James. 1987. *Chaos. Making a new science*. New York: Viking.
Görlach, Manfred. 1996. And is it English? *English World-Wide* 17(2). 153–174.
Graband, Gerhard. 1965. *Die Entwicklung der frühneuenglischen Nominalflexion*. Tübingen: Niemeyer.
Grund, Peter. 2011. The science of pronominal usage: *He* and *it* in co-reference to inanimate objects in Late Middle English texts on alchemy. *Journal of English Linguistics* 39(4). 335–358.
Hoffmann, Thomas & Graeme Trousdale (eds.). 2013. *The Oxford handbook of construction grammar*. Oxford: Oxford University Press.

Holland, John H. 2014. *Complexity. A very short introduction*. Oxford: Oxford University Press.
Hopper, Paul J. 1987. Emergent grammar. *Berkeley Linguistic Society* 13. 139–157.
Johansson, Christine. 1995. *The relativizers* whose *and* of which *in present-day English. Description and theory*. Uppsala: Acta Universitatis Upsaliensis.
Johansson, Christine. 1997. The positional variation of the possessive relativizer *of which*. In Udo Fries, Viviane Müller & Peter Schneider (eds.), *From Aelfric to the New York Times. Studies in English corpus linguistics*, 51–64. Amsterdam: Rodopi.
Johansson, Christine. 2002. The possessive relativizers *whose* and *of which* in Middle English. In Patricia Poussa (ed.), *Relativisation on the North Sea littoral*, 37–49. Munich: Lincom Europa.
Johnson, Neil. 2009. *Simply complexity. A clear guide to complexity theory*. Oxford: Oneworld.
Keller, Rudi. 1994. *On language change: The invisible hand in language*. London: Routledge.
Kortmann, Bernd & Verena Schröter. 2020. Linguistic complexity. https://www.oxfordbibliographies.com/view/document/obo-9780199772810/. DOI: 10.1093/OBO/9780199772810-0254 (As accessed in April 2020.).
Kretzschmar, William R., Jr. 2009. *The linguistics of speech*. Cambridge: Cambridge University Press.
Kretzschmar, William R., Jr. 2015. *Language and complex systems*. Cambridge: Cambridge University Press.
Kretzschmar, William R., Jr. 2018. *The emergence and development of English*. Cambridge: Cambridge University Press.
Kroch, Anthony. 1989. Reflexes of grammar in patterns of language change. *Language Variation and Change* 1(3). 199–244.
Labov, William. 1972. *Sociolinguistic patterns*. Oxford: Blackwell.
Larsen-Freeman, Diane. 1997. Chaos/complexity science and second language acquisition. *Applied Linguistics* 18(2). 141–165.
Larsen-Freeman, Diane & Lynne Cameron. 2008. *Complex systems and applied linguistics*. Oxford: Oxford University Press.
Lass, Roger. 1992. Phonology and morphology. In Norman Blake (ed.), *The Cambridge history of the English language*, vol. 2, *1066–1476*, 23–155. Cambridge: Cambridge University Press.
Levey, Stephen. 2006. Visiting London relatives. *English World-Wide* 27(1). 45–70.
Lightfoot, David. 1991. *How to set parameters: Arguments from language change*. Cambridge, MA: MIT Press.
Lindblom, Björn, Peter MacNeilage & Michael Studdert-Kennedy. 1984. Self-organizing processes and the explanation of language universals. In Brian Butterworth, Bernard Comrie & Östen Dahl (eds.), *Explanations for language universals*, 181–203. Berlin: Mouton de Gruyter.
Mair, Christian. 2006. *Twentieth-century English. History, variation and standardization*. Cambridge: Cambridge University Press.
Massip-Bonet, Àngels, Gemma Bel-Enguix & Albert Bastards-Boada (eds.). 2019. *Complexity applications in language and communication sciences*. Cham: Springer.
McColl Millar, Robert. 2000. *System collapse, system rebirth: The demonstrative pronouns of English 900–1350 and the birth of the definite article*. Bern: Peter Lang.
Mesthrie, Rajend & Rakesh Bhatt. 2008. *World Englishes*. Cambridge: Cambridge University Press.

Milroy, James & Lesley Milroy. 1985. Linguistic change, social network and speaker innovation. *Journal of Linguistics* 21(2). 339–384.
Mobus, George E. & Michael C. Kalton. 2015. *Principles of systems science*. New York: Springer.
Mufwene, Salikoko S. 2001. *The ecology of language evolution*. Cambridge: Cambridge University Press.
Nevalainen, Terttu. 2015. Descriptive adequacy of the S-curve model in diachronic studies of language change. In Christina Sanchez-Stockhammer (ed.), *Can we predict linguistic change?* http://www.helsinki.fi/varieng/series/volumes/16/
Nevalainen, Terttu & Helena Raumolin-Brunberg. 2002. The rise of relative *who* in Early Modern English. In Patricia Poussa (ed.), *Relativisation on the North Sea littoral*, 109–121. Munich: Lincom Europa.
*OED = Oxford English Dictionary*. https://www.oed.com/ (Accessed June 2019.).
Romaine, Suzanne. 1980. The relative clause marker in Scots English: Diffusion, complexity and style as dimensions of syntactic change. *Language in Society* 9(2). 221–249.
Romaine, Suzanne. 1982. *Socio-historical linguistics. Its status and methodology*. Cambridge: Cambridge University Press.
Sapir, Edward. 1921. *Language*. New York: Harcourt Brace.
Schmid, Hans-Jörg. 2020. *The dynamics of the linguistic system. Usage, conventionalization, and entrenchment*. Oxford: Oxford University Press.
Schneider, Edgar W. 1989. *American Earlier Black English. Morphological and syntactic variables*. Tuscaloosa: University of Alabama Press.
Schneider, Edgar W. 1992a. Who(m)? Case marking of *wh*-pronouns in written British and American English. In Gerhard Leitner (ed.), *New directions in English language corpora. Methodology, results, software developments*, 231–245. Berlin: Mouton de Gruyter.
Schneider, Edgar W. 1992b. Who(m)? Constraints on the loss of case marking of *wh*-pronouns in the English of Shakespeare and other poets of the Early Modern English period. In Matti Rissanen, Ossi Ihalainen, Terttu Nevalainen & Irma Taavitsainen (eds.), *History of Englishes. New methods and interpretations in historical linguistics*, 437–452. Berlin: Mouton de Gruyter.
Schneider, Edgar W. 1993. The grammaticalization of possessive *of which* in Middle English and Early Modern English. *Folia Linguistic Historica* 14(1–2). 239–257.
Schneider, Edgar W. 1996. Constraints on the loss of case marking in English *wh*-pronouns: Four hundred years of real-time evidence. In Jennifer Arnold, Renée Blake, Brad Davidson, Scott Schwenter & Julie Solomon (eds.), *Sociolinguistic variation: Data, theory, and analysis. Selected papers from NWAV23 at Stanford*, 487–499. Stanford, CA: CSLI Publications.
Schneider, Edgar W. 1997. Chaos theory as a model for dialect variability and change? In Alan R. Thomas (ed.), *Issues and methods in dialectology*, 22–36. Bangor: Department of Linguistics, University of Wales.
Schneider, Edgar W. 2007. *Postcolonial English: Varieties around the world*. Cambridge: Cambridge University Press.
Schneider, Edgar W. 2011a. *English around the world: An introduction*. Cambridge: Cambridge University Press.
Schneider, Edgar W. 2011b. English into Asia: From Singaporean ubiquity to Chinese learners' features. In Anne Curzan & Michael Adams (eds.) *Contours of English and English language studies*, 135–156. Ann Arbor, MI: University of Michigan Press.

Schneider, Edgar W. 2020. *Calling* Englishes *as* Complex Dynamic Systems: Diffusion and restructuring. In Anna Mauranen & Svetlana Vetchinnikova (eds.), *Language change: The impact of English as a Lingua Franca*, 15–43. Cambridge: Cambridge University Press.

Siemund, Peter. 2008. *Pronominal gender in English. A study of English varieties from a cross-linguistic perspective*. New York: Routledge.

Stenroos, Merja. 2008. Order out of chaos? The English gender change in the Southwest Midlands as a process of semantically based reorganization. *English Language and Linguistics* 12(3). 445–473.

Strang, Barbara. 1970. *A history of English*. London: Methuen.

Tagliamonte, Sali. 2002. Variation and change in the British relative marker system. In Patricia Poussa (ed.), *Relativisation on the North Sea littoral*, 147–165. Munich: Lincom Europa.

# Section 2: **Linguistic boundaries**

Donka Minkova

# 4 First or best, last not least: Domain edges in the history of English

## 1 The premise: Speech segmentation and boundary signals

Speech comprehension implies a shared understanding of the boundaries of discrete units in a continuous sound stream.[1] Languages differ in the ways in which they mark boundaries (Lehiste 1965; Keating et al. 2004), yet boundary marking is an inalienable property of any language. Boundary marking is also a concern in the matching of written and spoken language and a recognizable obstacle in the reconstruction of speech properties from exclusively written and often imperfect historical documentation. The present study seeks to establish whether some of the testable prosodic and segmental boundary signals of Present-Day English (PDE) can be projected back to Old English (OE).[2]

The opening section lays out the philological, metrical, and phonological premises on which the examination of specific boundary manifestations can be conducted. The merits of two sources of diachronic evidence, orthography and meter, are also assessed in the introductory part. Section 2 describes and juxtaposes the Prosodic Hierarchy schema and the basic schema of OE meter. Next, some relevant empirical findings in English historical phonology are discussed with regard to these premises: Section 3 surveys the behavior of consonants at the left edge of prosodic domains, while Section 4 focuses on vowels, specifically at the right edge of prosodic and metrical units. It takes up the question of Kaluza's Law, a long-time crux in the history of English verse in the light of the experimental and theoretical

**1** I am grateful to the organizers of SHEL 10 for inviting me to present my research and for their generous hosting of that exciting decahedral event. This study has benefitted from comments by audiences at SHEL 10, The Phonetics and Phonology Research Group (Linguistics and English Language, University of Edinburgh), and Frontiers in Comparative Metrics III, Tallinn University. Thanks also to the anonymous reviewers and this volume's eagle-eyed editors for their valuable comments.

**2** The term *prosody*/*prosodic* as used here refers to the rhythmic organization of speech, and the term *meter*/*metrical* refers to the organization of verse.

**Donka Minkova,** Department of English, UCLA, Los Angeles, CA, USA,
e-mail: minkova@humnet.ucla.edu

https://doi.org/10.1515/9783110643282-005

positions on boundary signals in speech and verse. A summary, tentative conclusions, and some directions for future inquiry are offered in Section 5.

## 1.1 The orthographic evidential basis

Any written representation of spoken language falls short of capturing the phonological nuances of actual speech. In speech, the edges of words and phrases are easily perceived, but in writing edge-marking follows a learned set of conventions unavailable to illiterate speakers. In OE, word separation was an innovation introduced by Irish scribes "for the benefit of Irish readers, for whom Latin was an alien language" (Parkes 1991: xvii; also Cook 2016: 18); that practice was adopted by the Anglo-Saxon scribes, though not in its modern form. Parkes (1991) also reports that the system of graded pause-marks dependent on the importance of the break go back to at least the first century AD. Hierarchized punctuation marks, the *distinctiones* system, were used in Latin documents written on the continent. That system's insular variant was based on a series according to the number of points used: one for a minor pause, two for a medium one, and three for a major pause, arranged in a triangle. The scribes also used *punctus elevatus* to mark medial pauses, and a $^{7}$-shaped mark (*simplex ductus*) "within a verse to separate matters erroneously run together" (Parkes 1993: 304–307).

Familiarity with these signs did not make punctuation in the Anglo-Saxon records systematic or infallible (see Donoghue 1987; Parkes 1991, 1993; Burnley 1995; Roberts 2005). Word division in the manuscripts is inconsistent, especially with potentially cliticized monosyllables, e.g., *semunuc*, *þadunstan*, *anman*, *7swa*.[3] The use of points and other marks to represent grammatical structure or pauses varies from non-existent to more ample, though the intentionality behind the marking is often unclear and is "frequently of uncertain significance" (Mitchell 1995: 21).

## 1.2 Verse as evidence for reconstructing boundaries

The interpretation of punctuation as an indication of morphological and prosodic boundaries appears to be more reliable in verse than in prose. From the time of the earliest Latin manuscripts in *scriptio continua* ('words written together'), the verse line was often identified as the "primary unit of textual analysis" (Parkes

**3** Examples from London, British Library Cotton MS Julius E; vii, f. 203r, Ælfric, *Life of St. Martin* and opening of *Life of St. Edmund*, transcribed in Roberts (2005: 82).

1993: 11; Burnley 1995: 48). Against the more unsystematic practice in prose, in verse the use of points coincides more consistently with the edges of metrical units, as noted in the comments on *Guthlac* and *Genesis* B in Roberts (2005: 60, 68). Table 1 offers a comparison between the manuscript's transcription and the edited version of an excerpt from the tenth-century *Guthlac* A.

**Table 1:** Manuscript punctuation and editorial punctuation illustrated.

| Scribal punctuation (Exeter, Exeter Cathedral Library, MS 3501, f. 32v, *Guthlac* A, ll. 5–8; transcription from Roberts 2005: 60) | Editorial punctuation (from *The Complete Corpus of Anglo-Saxon Poetry*, http://www.sacred-texts.com/neu/ascp/) |
|---|---|
| engel hafað yldran hád. greteð gæst oþerne.<br>abeodeð hi(m)<br>godes ærende. Nu þu most feran þider þu<br>fundadest. lon<-><br>ge 7gelome Ic þec lædan sceal wegas þe<br>sindon weþe 7wul<-><br>dres leoht torht ontyned | (ðonne cwið se) engel, (hafað yldran had),<br>5 greteð gæst oþerne, abeodeð him godes<br>ærende:<br>"Nu þu most feran þider þu fundadest<br>longe ond gelome. Ic þec lædan sceal.<br>Wegas þe sindon weþe, ond wuldres leoht<br>torht ontyned. |

The scribal marking of verse units in the sample is incomplete, but where it exists, it is justified both metrically and syntactically. The editorial decision of omitting punctuation at the end of l. 6, after *þu fundadest*, is just that: an editorial decision to show the syntactic link to the next phrase. In the modern text the spacing of the verse line on the page is sufficient to mark the metrical edge, while the scribe was careful to signal that same edge with a point. Note that two of the three capitals (*Nu*, l. 6, and *Ic*, l. 7) are also scribal; this is yet another way that edges could be highlighted.

The ASPR (Anglo-Saxon Poetic Records) corpus shows a wide range of practices; comparing the Exeter and the Junius manuscripts, Dobbie (ASPR: xxi) finds punctuation in the former more sporadic, while in the Junius it is intended as a "metrical punctuation". The rate of pointing by the *Beowulf* scribes is at 19.6% (O'Brien O'Keeffe 1990: 175), and in 93% of those attestations the marking is at the right edge of the line, whereas in the Exeter poems it is the edge of the on-verse, i.e., the mid-line caesura, that is most consistently marked.[4] Either way, the baseline of what we now call *domains* and *boundaries* was nothing strange to the early

**4** Nevertheless, as in prose, verse punctuation has to be treated with caution. In *Beowulf*, for example, "[a] point is supplied, on average, once in nearly five verse pairs [. . .] but the frequency varies widely [. . .]. The point is usually placed at the end of a verse pair, occasionally at the end of an on-verse, and a few times, surprisingly, within a verse" (Fulk et al. 2014: xxxii).

grammarians, poets, and scribes, though neither the old, nor the modern system, can reflect speech adequately.

The earlier and stricter use of visual boundary signals is just one aspect of the poetic records that makes them a good source of information about the prosodic organization of the poets' language. The punctuation marks used by the scribes are "guides to potential pauses in reading aloud" (Cook 2016: 17), an observation which highlights the oral elements in the composition and non-silent transmission and consumption of OE verse. These elements have always been recognized, though not always equally weighed. In a series of studies Orton (1994, 1999, 2014) defends convincingly the position that alliterative verse was "originally, or even essentially, oral" (Orton 2014: 235); so "orality" offers another argument in favor of reliance on verse as a mirror of linguistic properties.

The oral nature of the OE poetic compositions makes them the closest approximation to what can be viewed as "internalized" language reflecting the grammatical competence underlying entirely written source material. Reconstructing the hidden and otherwise untestable structures of early verse is based on the premise that the selection and placement of lexical items in verse is an externalization of the speakers' internalized categories, structures, and preferences. The necessary and indubitable link between the conventions and templates of poetic meter and the prosody of the spoken language is an analytical postulate above and beyond a specific language, period, or verse tradition (Minkova 2003: 22–24). The rest of this project therefore draws on the principles of linguistic grounding of meter in formulating hypotheses about the interaction of the two systems in earlier English.

# 2 The prosodic hierarchy

## 2.1 The prosodic hierarchy in PDE

The units that we identify as the domains of phonological and prosodic processes – the phrase, the word, the syllable – are hierarchically organized. A basic prosodic hierarchy is shown in Figure 1 and Figure 2.[5]

**5** For a useful tutorial on various approaches/interpretations of the prosodic hierarchy, see Shattuck-Hufnagel and Turk (1996). I am not including higher levels, such as the Intonational Phrase (IP) for the domain of the utterance, because so far the IP prosodic contours for OE remain an empirically undiscoverable area.

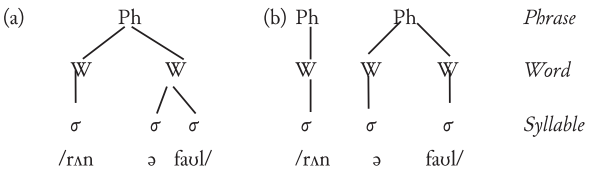


**Figure 1:** A simplified prosodic hierarchy for PDE.

Figure 1(a) shows the noun-adverb phrase *run afoul*, and Figure 1(b) shows an imperative verb phrase plus a noun phrase: *Run! A fowl!*.

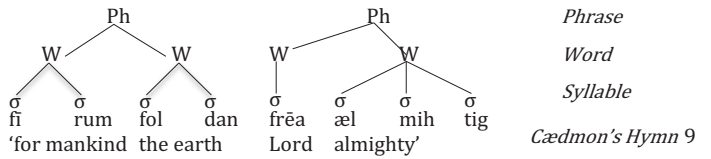


**Figure 2:** A simplified prosodic hierarchy for OE.

Additional layers between the Word and the Phrase are skipped in Figure 2 because they will not be relevant to the discussion here, but see Minkova (2008) for a more detailed hierarchy needed in the account of OE prefixal stress. The syntactic bracketing of a string may or may not coincide with the prosodic bracketing; the co-extensiveness of the prosodic domains and the syntactic strings increases at the higher levels of the hierarchy. In all versions of the prosodic hierarchy a lower domain can be co-extensive with the domain above it (Shattuck-Hufnagel and Turk 1996). Strict Layering (e.g., Hayes 1989) is observed: if a phonological process occurs at one boundary-edge, then it also affects the unit at a higher domain boundary. The domains are of increasing size, and there is a well-established correlation between domain size and the strength of the boundary marking: the larger the domain, the clearer the boundary signals. The smaller the constituent, the harder it is to demarcate its edges.

While prosodic domains are set off at both edges, the way that boundaries are phonetically marked depends on the edge's direction. Fougeron and Keating (1997: 3736) found that "[t]he *beginning* edge of prosodic domains is marked by *lengthening* and increasing contact *for the consonant*. *The final edge* of prosodic domains *is marked by lengthening and decreasing contact for the vowel*" (emphasis added).

Boundary marking is a large and vibrant area of research and one can hardly do it full justice here (e.g., Cutler and Butterfield 1992; Wightman et al. 1992; Byrd and Saltzman 1998; Keating et al. 2004; Cho 2005; Cho et al. 2007; Turk and Shattuck-Hufnagel 2007; Yoon et al. 2007; Cho and Keating 2009; Mo et al. 2009; Byrd and Choi 2010; Katsika et al. 2014; Katsika 2016; Shih 2018). Building on this, the next sections seek to explore if and how the phonetic findings observed synchronically for PDE could be manifested in early English meter. In addition to the search for panchronic edge-related patterns in OE verse, this project promises to open up yet another window on the hypothesis of the essentially oral nature of the surviving alliterative corpus noted in Section 1.2.

## 2.2 Matching the prosodic hierarchy to OE metrical structure

Figure 3 is a schematic representation of the hierarchy of metrical units reconstructed for OE verse.

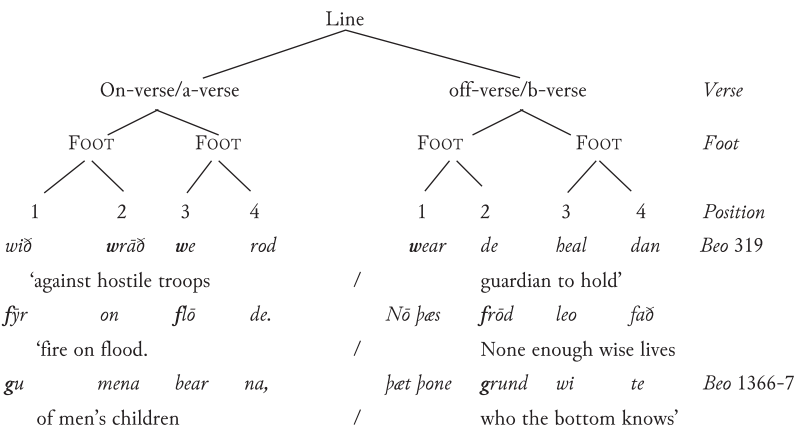


**Figure 3:** An idealized metrical structure of the OE verse line.

In this idealized representation, the OE alliterative line is the largest constituent, branching into two verses, the a-verse and the b-verse, which are linked by (boldfaced) alliteration. A further binary division splits the material into feet; the foot-division here follows Sievers (1893). Each foot comprises two positions: strong (alliterating S, non-alliterating s) or ictic, and weak (w) or non-ictic; the total number of positions for a line is four. Matching this hierarchy to the hierarchies in Figures 1 and 2 is mostly straightforward. The line is the largest break, and it is regularly

coextensive with a phrasal break, possibly a clause break. Although the verses of OE are more additive than prose, "synthetic", to use Cable's label (Cable 1991: 46), they are nevertheless "self-contained and discrete". The edges of the verses are treated here as regularly matching phrasal edges, with all the caveats on the different phrase structure in verse and prose addressed in Cable (1991: 41–65).

The matching of the verse foot to the prosodic word domain in OE is so close that one influential account of English alliterative verse (Russom 1987, 2017) is based on the principle that "[m]etrical feet are abstracted from words, and any word-counting meter is a foot-counting meter" (Russom 2017: 28).[6]

At the position level in verse there is no one-to-one correspondence between syllable and position. Weak verse positions can be filled by one or more syllables, strong positions can be filled by one or sometimes two syllables – more on this in Section 4.2.

The focus of this chapter will be on segmental effects, so details will not be addressed here, but a metrical feature directly applicable to the matching of prosody and OE meter is the so-called *displacement* (Kendall 1991: 43–59), whereby items normally placed in unstressed positions verse-initially, such as *was* 'was' receive metrical ictus at the right edge/the phrase or clause boundary, as illustrated in Table 2.

**Table 2:** Examples of metrical prominence at phrase and clause boundaries in *Beowulf*.[a]

| Non-ictic | Ictic |
|---|---|
| *wæs* þǣm **h**æftmēce / **H**runting nama; (1457)<br>'was the hilted sword / Hrunting name' | **h**erestrǣl **h**earda; / hē on **h**olme *wæs* (1435)<br>'war-arrow hard; / it on the sea was' |
| þæt *wæs* **ā**n foran / **ea**ldgestrēona; (1458)<br>'that was one in forefront / of old treasures;' | swā hē ne **m**ihte nō / hē þǣm **m**ōdig *wæs* (1508)<br>'so he could not / however brave he was' |
| **e**cg *wæs* īren, / **ā**tertānum fāh, (1459)<br>'edge was iron, / with venomous pattern marked' | þæt hē [in] **n**īðsele / **n**āthwylcum *wæs*, (1513)<br>'that he in hostile hall' / of some sort was' |

[a]All examples are from Fulk et al. (2014). Line-final/ictic placement of finite forms of *beon/wesan* occurs in 26.8% of the attestations, while the preferred weak position placement is at 72.2% (Getty 2002: 65).

**6** There is no general consensus on the "optimal" account of OE verse structure. A discussion of the various theories of OE meter is beyond the scope of this study. Here I follow the Sievers-Bliss system, the most widely used model of scansion in the literature. Russom's (1987) word-foot theory and the Sievers-derived metrical foot theories make the same lexical stress-based predictions with respect to alliteration, though the grouping into feet may differ.

This well-studied promotion of otherwise prosodically non-prominent items is another confirmation of the relevance of boundaries in the interaction between the natural prominences in speech and meter, in line with other arguments supporting the oral grounding of OE metrical techniques.

# 3 Consonants at the left edge

## 3.1 Alliteration and the left edge

Synchronically, domain-initial position, lexical stress, and phrasal accent are potentially independent factors in consonantal strengthening (Cho 2005; Cho and Keating 2009). Strengthening due to positional prominence and the strengthening effect of lexical stress can nevertheless be phonetically cumulative, resulting in what has been described as "local hyperarticulation" (Fougeron and Keating 1997; Keating et al. 2004).[7] On the other hand, right-edge phrasal prominence is in tension with word-initial strengthening and that can decrease or even completely obscure left-edge effects.

With that preamble, a reminder of the basics of stress-placement in OE is in order. OE unprefixed lexical words, including compounds, and some prefixed nouns are left prominent: ***wé**rod*, ***w**éarde*, ***fl**ŏde*, ***g**úmena* in Figure 3. Phrasal stress is arguably right prominent (Minkova 2014: 294–302).

In the prosody of early Germanic, the articulatory and perceptual salience of consonants at the left edge of lexical words provides the linguistic foundation of early Germanic alliterative verse (Lehman 1956). Since the whole idea of verse structure implies recurrence of patterns, the demarcation of these patterns requires boundary signals; alliteration is one of the left-edge signals guiding the identification and ordering of metrical domains.

Within each verse, alliteration falls on the first metrical ictus which normally coincides with the stressed syllable of the first foot in each verse.[8] The ictic position of the second foot of the on-verse can also alliterate, and the last ictus/the last

**7** This is a much-abridged account of a very rich research area where the data are continuously tested and reevaluated; see Cho and Keating (2009: 482) for an overview of earlier findings and results, indicating that "[d]omain boundary effects are cumulative while prominence [=stress] effects are not".

**8** This statement excludes Sievers' verse type A3, where there is only one alliterating lift in the right periphery of the a-verse, allowing the option of a preceding, possibly stressed, but non-alliterating lexical item in the left foot, as in *Beowulf* 1963a: Gewāt him ðā se **h**earda / mid his **h**ondscole 'Turned then the hardy one / with his hand-picked retinue'. I count 1,302 such

stressed syllable of the line does not alliterate. These are very familiar facts, yet adding the dimension of left-edge boundary marking helps elucidate some of the alliterative facts in the OE poetic corpus: the so-called "Rule of Precedence" (Sievers 1893), the cohesive alliteration of [s + stop] onsets, and vowel alliteration.

The Rule of Precedence describes the obligatory alliteration on the first noun, adjective, or non-finite verb form in the verse, irrespective of which element heads the syntactic phrase. This rarely violated Rule is illustrated in (1)–(4).

(1) *Be**b**eorh þē ðone **b**ealonīð, / **B**ēowulf lēofa* (*Beowulf*, l. 1758)
'Guard you from wickedness, / Beowulf beloved'

(2) *līcað **l**eng swā wēl, / **l**ēofa Bēowulf.* (*Beowulf*, l. 1854)
'pleases longer thus well / beloved Beowulf'

(3) ***d**ryhten Gēata / **d**racan scēawian;* (*Beowulf*, l. 2402)
'lord of Geats / dragon to face'

(4) ***G**ēata dryhten, / **g**ryrefāhne slōh* (*Beowulf*, l. 2576)
'Geats' lord, / the terribly hostile one slew'

Switching the phrase-internal order of ***B**ēowulf lēofa* 'Beowulf dear' in (1) vs. ***l**ēofa Bēowulf* 'dear Beowulf' in (2) and of ***d**ryhten Gēata* 'lord of the Geats' in (3) vs. ***G**ēata dryhten* 'the Geats' lord' in (4) is fully grammatical in OE. By choosing one order over the other in the verse line, the poet avails himself of the optionality of the syntax. In line with the cumulativity of domain-boundary effects, the first element of the phrase gets an additional boost; post-pausal, phrase-initial strengthening is even more perceptible than word-initial strengthening. The metrical Rule of Precedence has generated controversial interpretations in the reconstruction of the prosody of earlier English (see Minkova and Stockwell [1997] for a survey of earlier claims and a proposal for a pan-historic reconstruction of phrasal prominence in English). Associating the Rule with the articulatory and perceptual enhancement at the left edge of prosodic boundaries found in English and other languages (Keating et al. 2004) provides an independent criterion bridging the metrical structure of alliterative verse and the reconstructed ambient language.

verses in Hutcheson's (1995: 289) list of types, amounting to about 4% of the entire database of about 32,000 verses.

The premise of consonantal strengthening at the left edge of words applies to another aspect of OE alliteration: the cohesive alliteration of [s + stop] initial clusters. It has always been known that <sp-, st-, sc-> alliterate only as clusters, disallowing, e.g., [st-] in *stande* to alliterate with *secgað* in (5), *scyld* with *sweord* in (6), *spell* with *sōð* in (7), i.e., no alliteration on singleton [s-] or [s + sonorant].

(5) ***s**ecgað **s**ælīðend, / þæt þæs **s**ele **stande*** (*Beowulf*, l. 411)
'say seafarers / that this hall stands'

(6) *þæt ic **s**weord bere / oþðe **s**īdne **scyld*** (*Beowulf*, l. 437)
'that I sword bear / or broad shield'

(7) ***s**ōð ond **s**ārlīc, / hwīlum **s**yllīc **spell*** (*Beowulf*, l. 2109)
'true and lamentable / sometimes wondrous spell'

Since adjacent alliterative practices, such as Latin or Old Irish, did not follow the same rules of alliteration, the choice of treating <sp-, st-, sk-> differently in Germanic and OE verse poses a particularly difficult problem for matching OE verse techniques to the reconstructed linguistic realities of the older language. Minkova (2003: 192–237) examined the ways in which the OE patterns of cohesiveness in verse can be analyzed in terms of phonetic and phonological properties of the respective segments. The emphasis in the analysis is the lack of a perceptual break within these clusters; this sets apart onset clusters in which [s] is followed by a more sonorous segment such as a nasal or a liquid.

An approach seeking to establish the phonetic and phonological correlates of boundaries opens a further linguistic angle to the cohesiveness account. The [s + stop] clusters have special phonotactic properties: lack of voicing contrast, no aspiration of the stop, and a single gesture release. Following Byrd and Choi (2010: 41–42), we can add that the duration of these clusters produces a measurable boundary effect.[9] The effect is significant both for the sibilant, and for the stop, though the sibilant duration increase is more substantial. Crucially, the gestural overlap in [sp-, st-, sk-] is most pronounced in word onsets, less so in other s-initial clusters. This new phonetic perspective on the robust lengthening of the [s + stop] initial clusters provides articulatory and perceptual grounding for the

**9** See also Cho et al. (2014: 144) whose results lead to the conclusion that "the temporal structure of the sCV sequences is systematically modulated by prosodic strengthening that serves dual functions of marking prosodic structure – i.e., boundary marking (delimitative function) and prominence marking (culminative function)." They report on similar findings for Modern German.

cohesiveness and strength of <sp-, st-, sk-> at the left edge and adds a new and informative factor in the analysis of OE metrical practices.[10]

Another possible correlation between the Germanic alliterative practices and the speech-based enhancement of the left-edge signal is suggested by the so called vocalic alliteration. Contrary to the strict observance of phonemic identity of consonants in alliteration, at first glance the poets appear to have been oblivious to the differences between vowels, as in (8)–(10).

(8) ***eo****tenas ond* ***y****lfe / ond* ***o****rcneas,* (*Beowulf*, l. 112)
'ogres and elves / and monsters'

(9) ***o****mbeht* ***u****nforht: 'Æghwæpres sceal* (*Beowulf*, l. 287)
'officer unafraid: / "Of each shall'

(10) ***ea****llum* ***æ****þellingum / tō* ***a****ldorceare;* (*Beowulf*, l. 906)
'to all men of valor / life-long grief'

One widely accepted account of why these very distinct vowels were judged identical for the purpose of alliteration is that OE, like Modern German, and optionally like PDE, had prevocalic glottal stop insertion, so that stressed vowel-initial syllables were realized as [ʔ + vowel]. The presence of the glottal stop provided the filler for the onset position, rendering the word initial [ʔV-] string structurally identical to a [CV-] string; [ʔ-] insertion avoids the violation of a syllabic well-formedness constraint favoring a filled onset.[11]

Once again, current phonetic measurements of [ʔV-] at the prosodic left edge offer an empirical validation of the plausibility of a hypothesis which has been famously challenged in the literature.[12] A study of the word-initial realization of complete and incomplete glottal stops in English (Garellek 2012) records testable graded pre-vocalic glottalization as a function of both boundary adjacency and

**10** In the similar, but not identical Old Irish practice (Russom 2005) initial cluster alliteration disallows the splitting of <sm->, in which the conditions of gestural overlap do not obtain, while <sn-, sl-, sr-> can alliterate on <s-> + a vowel. The different treatment of <sm-> is an analytical challenge: however, Old Irish also allows non-domain initial alliteration (Sproule 1987). The two traditions drew not only on different cue strengths, with OE verse using a higher perceptibility threshold, but also on different rules of recurrence.

**11** For full details on the various proposals regarding the rationale of vowel alliteration and some independent metrical confirmation of the glottal stop theory, see Minkova (2003: 135–191).

**12** Jakobson (1963) characterized the glottal stop hypothesis as ad hoc and circular.

lexical stress. Adding the boundary factor thus relates the glottal stop account of vowel alliteration to phonetic events testable in a laboratory setting today.

The additional empirical motivation enables us to come closer to an "explanation" of why Anglo-Saxon speakers would find it natural to use a verse schema where only the first ictus in each verse requires obligatory alliteration, [s + stop] initial clusters are perceived as the only candidates for cluster alliteration, and glottal stop insertion provides the expected onset identity for vowel-initial words. That's the left edge. The next section discusses right-edge vowel length enhancement in the context of a famous metrical crux in OE verse.

# 4 Vowels at the right edge

As adumbrated in Section 2.1, in synchronic experimental studies the right edge of a prosodic domain is associated with vowel lengthening. Prosodic final lengthening is identifiable at every prosodic level, including the syllable (Beckman and Edwards 1990): "[t]ypically, the last vowel before a large phrasal boundary is lengthened, but other units such as final consonants, final VC's, final syllables, and final words have been identified as subject to domain-final lengthening as well" (Byrd and Saltzman 1998: 174). Below I explore vowel lengthening in various contexts in OE.

## 4.1 Word-final lengthening in monosyllables

A familiar diachronic process which fits the behavior of vowels at the right edge is the vowel lengthening in monosyllables best attested in OE and Old Norse (not so well in Old High German), either with or without compensation for the loss of a final consonant, as illustrated in Table 3.

**Table 3:** Lengthening of monosyllables in North-west Germanic.

| Meaning | Gothic | OE |
|---|---|---|
| 'now' | *nu* | *nū* |
| 'so' | *swa* | *swā* |
| 'who' | *hwas* | *hwā* |
| 'me' | *mis* | *mē* |

The relevant items are monosyllabic adverbs and pronominal forms, i.e., candidates for prosodic weakness; predictably, this creates tension between the perception of the vowel as short or long. This is where we can project the stronger effect of domain-final lengthening at phrase boundaries back to pre-OE. The length of the vowel, especially in some of the pronominal monosyllables in OE (*he* 'he', *me* 'me', *þu* 'thou') is reconstructed as fluctuating depending on the position of the pronoun in the phrase or clause: the long vowel goes with stressed position, but when the pronoun was unstressed, the vowel was short.[13]

In the verse, the distribution of monosyllables with "variable" length reflects this fluctuation perfectly. There is no *swā* in verse-final final position (Bessinger and Smith 1978); there are, however, some examples with *me* and *nu*, which are normally in weak positions unless they provide the verse-final ictus, as in (11)–(14).[14]

(11) ***Nŭ hie d****rihtne synt* (*Genesis B*, l. 421b)[15]
'now they to the lord are'

(12) *þeah wit* ***h****earmas* ***nū*** (*Genesis*, l. 736b)
'though we injuries now'

(13) *Swa* ***mĕ æ****fter wearð* (*Daniel*, l. 139b)
'so to me after became'

(14) *þurh þæt* ***w****if on* ***mē*** (Genesis, l. 2685a)
'through that woman in me'

Adding the criterion of domain-finality thus takes us a little closer to a phonetically grounded reconstruction of the vowel length in such items in earlier English and removes some of the guesswork in teaching OE pronunciation.

**13** Compare the *OED* transcription for *me*, pron: /m**i**ː/, /m**i**/, /m**ɪ**/, U.S. /m**i**/, /m**ɪ**/. Clearly, the lengthened forms were dominant in acquisition and transition to Middle English and PDE, since their vowels have undergone the post-OE Long-Vowel Shift.

**14** Campbell (1959: §125) made the decision that "[i]n this book these words are usually printed in the unaccented form, with short vowel, e.g. *me*, *hwa*." The shared assumption is that they get lengthened under stress. Lewis (1987: 79) observed this distribution, but he did not detect or comment on the specific verse-final effect since he was focusing on what can fill the lift and not on why this practice was linguistically licensed.

**15** The source for non-*Beowulf* examples throughout is *The complete corpus of Anglo-Saxon poetry*, http://www.sacred-texts.com/neu/ascp/.

## 4.2 Right-edge vowel lengthening and Kaluza's Law

Two right-edge boundary effects, slowing down of the tempo and absolute duration, are experimentally established in the phonetic literature. One strategy used to control the tempo at the right edge is "decreasing intragestural stiffness to slow down some part of the syllable" (see Edwards et al. 1991: 369, who also found that "the final [ə] is significantly longer than its nonfinal counterpart for all four subjects" [1991: 378]). Phrase-final increased duration affects the boundary-adjacent segment most, and the phonetic effects "tend to spread beyond the primary-stressed syllable to secondary-stressed syllables" (Cho and Keating 2009: 482). These measurable speech-based events have their counterparts in music and in the account of metrical forms (Hayes and MacEachern 1998; Blumenfeld 2016).[16] An attempt to explore a possible connection between vowel lengthening and the account of OE metrical structure is warranted.

As noted in 2.2, strong/ictic metrical positions in OE verse can be filled by one or two syllables. More specifically, if the position is filled by a single syllable (*σ*), it has to be heavy, i.e., of the shape (C)VV or (C)VC(C). This is the "unmarked", typical match of syllable-to-metrical position. This one-to-one correspondence may be disregarded under special metrical conditions for S-positions if the stressed syllable of the word is light, (C)V-; this convention is known as "resolution":

Metrical resolution: H(eavy) = L(ight) + σ (any σ, H or L)

As shown in the definition above, resolution is the metrical equivalence of a stressed heavy syllable and a stressed light syllable + another syllable, hereafter marked as S-w (alliterating ictus) or s-w (non-alliterating ictus).[17] Unlike resolution in the early Latin tradition where "either a heavy syllable or two light may carry the peak-and-cadence complex of stress" (Allen 1969: 197), resolution in OE is possible both for L+H and L+L words. The examples in (15) and (16) show how this works. The bracketed scansions show the syllabic correspondences and resolution (S-w and s-w), and the equivalences to the right are the verse positions.

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| (15) | ond | ***féo*** | ***rum*** | ***gú*** | ***me*** | na | → [w **S**-**w** / **s**-**w** w] = W S / S W |
| | | \| | \| | \| | \| | | (*Beowulf*, l. 73b) |
| | | L | H | L | L | | |
| | 'and | lives | | of men' | | | |

**16** Blumenfeld (2016) also argues for "short-last" effects in forms requiring parallelism from line to line, which is not a requirement in OE alliterative verse.

**17** The scansion conventions follow Stockwell and Minkova (1997): w = non-ictic syllable, S = an ictic syllable bearing primary stress, s = an ictic syllable bearing secondary stress.

(16) ***hǽ*** ***leð*** *under* ***héo*** ***fe*** *num* → [**S-w** w w / **S-w** w] = S W / S W
| | | | | (*Beowulf*, l. 52a)
L H L L
'heroes under heavens'

The intricate technicalities of resolution in relation to the weight of the preceding syllable, the type of metrical foot, the type of verse, are set aside for the moment in the interest of focusing on one particular issue, the *suspension* of resolution in relation to verse-ends. Suspension of resolution means that a disyllabic sequence L + σ is associated with two metrical positions, i.e., the light syllable furnishes the lift by itself. The mechanism of suspension of resolution is illustrated in (17) and (18); notice that both syllables in *monig* and *werod* in (17) and (18) have to count – they occupy separate metrical positions.

(17) *gúð* *rìnc* *mó.* ***nig*** → [S s / s w] = S W / S W, **not** *[S s / s-w]
| | (*Beowulf*, l. 838b)
L **H**
'battle-man many'

(18) *wið* *wrāð* *wé.* ***rod*** → [w S / S w] = W S / S W, not *[w S / S-w]
| | (*Beowulf*, l. 319a)
L **H**
'against hostile troops'

Examples (17) and (18) are a preamble to a much knottier problem, known as Kaluza's Law (e.g., Kaluza 1896; Cable 1991: 6–40, 132–153; Fulk 1992: §271–275; Russom 1995; Hutcheson 1995: 68–96). In a very simplified form, one aspect of Kaluza's Law is a regularity based on the weight distinction in the syllable following the "resolvable" light syllable: resolution is avoided if the post-tonic syllable is heavy. Examples (17) and (18) are the "vanilla" case of suspension of resolution: the syllable following the stressed syllable is unambiguously heavy: resolution does not apply. This is the "rule"; the suspension of resolution is a function of the structure of the verse as a whole, above all the observance of the four-position principle.[18]

**18** The weight of the preceding syllable also matters: verse-final resolution of what looks like HL sequences – where the light final syllable is an inflectional ending – is suspended "in Sievers' types C, D1, and A2K, where it follows a long stressed syllable. [. . .] the distinction is observed in the first foot of types A and E as well" (Fulk 1992: §171). Fulk (1992: §175) also states that "[. . .] it does appear to be true that Kaluza's Law applies only when the resolvable syllables immediately follow a lift in the same foot", but he goes on to say that "[y]et there is

The metrical hot potato is suspension of resolution when a light syllable is followed by a light syllable, where the etymological weight of the second syllable is supposed to dictate the poet's choice: certain historically long vowels (in, e.g., neuter *–a* stem plurals, feminine *-ō* stems nom., acc. sg., masculine *n*-stems nom. sg., etc.; see lists in Bliss 1962: 27–30; Fulk 1992: Appendix C; Hutcheson 1995: 68–96) appear to be sufficient to block resolution. Suspension of resolution in -(C)V̆. CV# items are illustrated in (19)–(22).

(19) *hē ne lēag **fela*** [w w S / **s w**] (*Beowulf*, l. 3029)
'he did not lie much'

(20) *þonne hē **fela** murne* [www **S-w** / s w] (*Beowulf*, l. 1385b)
'than he much should mourn'

(21) *Þær bið hlūd **wudu*** [w w S / **s w**] (*Riddle 3*, l. 24)
'there is loud wood'

(22) *fȳr **wudu** meltan* [S **s-w** / s w] (*Max I*, l. 71b)
'fire wood to melt'

The verse-final unresolved items in (19) and (21) clearly do not behave in this way elsewhere in the poetic corpus, i.e., they can be subject to resolution, as in (20) and (22). The highly controversial explanation of this apparent metrical ambiguity, as noted above, was attributed by Kaluza to the preservation of the historical length of the second syllable. This hypothetical preservation was referred to by Bliss (1962: 35) as a factor which "might provide a new criterion on the relative chronology of Old English verse". Most famously Fulk (1992), whose early dating of *Beowulf* rests on other substantive and convincing grounds, used it as one more argument for the epic's early dating: "*Beowulf* is unique in respect to the great ease and regularity of the poet's ability to distinguish long and short endings" (Fulk 1992: 164); "[. . .] [t]he *Beowulf* poet observes Kaluza's law in a remarkable 106 out of 108 instances" (Fulk 1992: 55).

However, subsequent researchers, unable to refute the strength of the statistical findings, felt unconvinced that this is a valid criterion on the basis of (a) apparent persistence of Kaluza's Law in late verse, (b) the conflict with other

room for doubt here", and identifies the problem of incorporating such a limitation into metrical theory (1992: §175, fn. 12). The discussion of suspension of resolution below will focus on the resolvable sequence in terms of its position with respect to the type of prosodic boundary at its right edge.

early compositions where the etymological length distinctions are not maintained in the same way as in *Beowulf*. Some alternative proposals exist: Hutcheson (1995, 2004) and Suzuki (1996)[19] sought to account for the facts through stability of formulaic poetic language and morphological analogy, but both accounts are addressed independently by Fulk (2007) and shown to be less informative than the length-based account.

One remarkable observation in Fulk's (2007) rebuttal has to do with the exact position of application of Kaluza's Law in the early and the later verse. He separates the Law's evidence in two parts:

> Under a variously formulated principle now referred to as Kaluza's Law, resolution under secondary stress in *Beowulf* occurs only when the second of the resolved syllables was short in early Old English [. . .]. For present purposes, this principle may be referred to as part one of the rule. In complementary fashion, under part two, nonresolution under secondary stress is accompanied by etymologically long final syllables, those that either end in a consonant or bore circumflex accent in early Germanic (the latter comprising mostly the ending *-a*, though *-e* is also found). (Fulk 2007: 317–323)

Part One of the Law follows the general metrical schema in the sense that verse-internal resolution under secondary stress, i.e., in compounds, would be subject to the four-position principle.[20] The additional consideration I propose to bring into the picture has to do specifically with Part Two: the suspension of resolution in verse-final position.

This is related to a regularity which Fulk (1992: 201) labeled "the Rule of the Coda":[21] "[s]uffixes like *–scipe* and *–sume*, with an etymologically short high vowel in the penultimate syllable, count as two metrical positions *in the coda of a verse*;

**19** Suzuki's idea is an elaboration of a suggestion by A. J. Bliss: the Beowulf poet based his length distinctions on morphological analogy. Note that Suzuki's (1996: 305–306) argument dates the composition of *Beowulf* prior to ca. 750, since it establishes a terminus based on distinctions of vocalic quality rather than quantity.

**20** I set aside the discussion of resolution and suspension of resolution of (C)V.CV specifically under secondary stress, noting that here too, verse-finally resolution is suspended. Compare single-position *–wudu* in "holtwudu sēce" 'forest seeks' [S **s-w** / S w] *Beowulf*, l. 1369b to *Beowulf*, l. 2340: "þæt him holtwudu" 'that him forest' [w w S / s w]. See also *wudu* in (22). Verse-internal resolution is not at issue: "Yet Hutcheson's group of late poems, dating from approximately 937 to 1104, offers no appreciable evidence in regard to part one of the rule-that is, in regard to etymologically short endings verse-internally" (Fulk 2007: 318). See Terasawa (1989, 2011: 71–79) for a detailed discussion of compounds and metrical structure in OE.

**21** Coda is defined as "the last full lift and all subsequent syllables". "The rule of the coda applies also to verse in other early Germanic languages" (Fulk 1992: 210).

otherwise they count as one" (emphasis added).[22] He continues: "[t]here seems *no linguistic motivation* for the rule of the coda, since it is metrically conditioned" (Fulk 1992: 211; emphasis added). The *Rule* covers examples such as in (23)–(26).[23]

(23) *on ðām lēod**scipe*** (*Beowulf*, l. 2197a)
'on that country'

(24) *ond se fēond**scipe*** (*Beowulf*, l. 2999b)
'and the enmity'

(25) *līf ond lēods**cipe*** (*Beowulf*, l. 2751a)
'life and country'

(26) *eahtodan eorl**scipe*** (*Beowulf*, l. 3173a)
'praised lordly power'

The account proposed here is that right-edge enhancement provides a linguistic basis for the Rule of the Coda – it is the domain-final perception of a vowel which is not phonologically long, but which is perceptually sufficiently extended in that position to license the metrical usage. The parallel between the unresolved words in (17) and (18) and the suffixes in (23)–(26) with disyllabic underived words, supports the idea that any phrase-final –CV# is potentially subject to lengthening/perceptual enhancement.

As for Kaluza's Law, its apparent preservation in the later verse in verse-final position does not require an appeal to etymological length in vowel-final

**22** It is hard to pinpoint the motivation for restricting the Rule to short high vowels in the penultimate syllable: high vowels are shorter than low vowels; in principle, this would enhance the chances or resolution (see Russom's [1995: 151] "subrule"; fn. 25 below), whereby a resolved sequence, an "ultralong" syllable would be a desirable match for an ictus. On the other hand, as one of the editors observed, if the *–scipe*, *–sume*, suffixes are not at the right edge, they are non-ictic and the intrinsic shortness of the high vowels may be a factor in their perceived "lightness", as in "frēondscipe fæstne" 'friendship firm', *Beowulf*, l. 2069a. Put differently, the increased prosodic perceptibility in the coda overrides the shortness of the high vowels in these suffixes. To what extent the (very few) exceptions, e.g. "eard and eorlscipe" 'land and nobility', *Beowulf*, l 1727a, can be attributed to inherent vowel shortness licensing resolution in the coda is an open question.

**23** For further examples in the corpus, see Fulk (1992: 201, n. 63). On *–ode*, see Hutcheson (1995: 31, n. 115).

inflections[24] – all we need is an appeal to right-edge variable perception of length. The examples of suspended resolution in disyllabic sequences in (19) and (21) may or may not preserve the weak syllable's etymological length in the scop's language. If they did, suspension of resolution in items in -CV would be more systematic, as it is indeed in *Beowulf*. Even if etymological vowel length in the post-tonic syllable was variable, or even if the vowel was short elsewhere in the utterance, the probability of verse-final slowing down and positional lengthening can generate variable results. The crucial point here is that the so-called "exceptional suspension of resolution", both according to Kaluza's Law and the Rule of the Coda, applies verse-finally where positional vowel lengthening boosts the vowel's perceptibility, making a "light"-CV# more similar to a heavy syllable CVC or CVV, as in (27)–(29).[25, 26]

(27) *godes gold**fatu*** [S-w / S s w] (*Daniel*, l. 754a, similarly 532a, 742a)
'God's golden vessels'

(28) *stēap stān**hliðo*** [S / S s w] (*Beowulf*, l. 1409a)
'steep stone-slopes'

(29) *beran beorht **searo*** [s-w / S s w] (*Exodus*, l. 219a)
'to bear bright arms'[27]

**24** "[. . .] it is by no means impossible that certain of the etymologically circumflected vowels retained length into the eleventh century, notably the low vowel –a" (Fulk 2007: 324).

**25** Russom's (1995: 151) "subrule" which states that resolution involving a heavy syllable "would [. . .] be equivalent to an 'ultralong' closed syllable containing a long vowel" does not contradict the proposed reasons for suspended resolution. Domain-final vowel lengthening is a variable phonetic event; occasional resolution is to be expected, so "hyra fyrngeflitu" [w w S / w s-w] in *Judith*, l 264a is not a problem, especially in a relatively late poem. The most detailed account of the attestations of the Law in the later verse with statistical analysis is offered in Hutcheson (2004). His not-for-the-faint-of-heart analysis does not reject the earliness of *Beowulf*, though he finds Kaluza's Law non-predictive with respect to dating. His focus on the persistence of the Law in later verse, which would be odd in an etymologically-based account, leads him to appeal to "oral poetic tradition" as the explanation for that practice. My proposal also ultimately corroborates the oral nature of the poetic composition and transmission, but the specific reference to phonetic grounding makes it less speculative. For an extensive and rhetorically intense survey of the theories and controversies surrounding Kaluza's Law, see Neidorf and Pascual (2014); they ultimately come down on the side of purely etymological conditioning of the Law and defend Fulk's original position.

**26** The list follows the time-line in Fulk (1992: 161–2); I am excluding –VC second syllables – they are unambiguously heavy.

**27** See further *Elene*, ll. 252a, 673a; *Christ II*, ll. 660a, 860a, 794a; *Andreas*, ll. 370a, 467a, 561a, 803a, 1221b, 1657a; *Meters of Boethius*, l. 5.10a.

Commenting on "stēap stānhliðo" (*Beowulf*, l. 1409), Fulk (2007: 318, note 28) labels it a "clear exception" to Kaluza's Law, and identifies a "peculiar" pattern: "[i]t does seem peculiar that [in verse-final position] poets other than the *Beowulf* poet should have been able to recognize long desinences but not, for the most part, short ones" (Fulk 2007: 322).

The situation is less peculiar if we consider it in the light of right-edge vowel-length effects discussed earlier in this section. The right edge of an OE verse characteristically coincides with a syntactic break, as noted in 2.2. OE verses are end-stopped; there is no enjambment. The absence of enjambment in OE verse confirms the co-extensiveness of phrase boundaries with verse ends. As Fulk (1992: 248, note 13) reports "Alfred B. Lord correlates enjambment to literacy and general sophistication of composition". While the level of sophistication of OE verse varies, the regularity of end-stopping argues against an account based exclusively on literacy. End-stopping is thus a necessary component of the proposed interpretation of the non-conformity to Kaluza's Law at line closures in some early texts and the application of the Law in late texts.

Again, the proposal is that the probability of non-resolution at the right edge is a boundary phenomenon, not just, or primarily, an etymological length or stress issue. This is a parameter independent of other factors involved in judgements on the chronology of the poems. It applies variably: *Beowulf* is the most conservative of the poems, while the Law is violated by the Cynewulf and the *Andreas* poet (Fulk 1992: 166–167). In either case the existence of lingering length in -CV inflections is not precluded, but it is not the only factor in play. There is nothing in this approach to dispute Fulk's (2007: 319) dating of Beowulf: "no later than about A.D. 725 if the poem is not Northumbrian in origin, otherwise no later than about A.D. 825". Why the *Beowulf* poet was strikingly more sensitive to the phonetic options for boundary marking, while other early compositors were less attuned to the same signals is not a matter of the absolute fixed value of the vowel in –CV# syllables at domain edges; it is quite possibly a matter of different levels of versification skills and the different styles of the poems' delivery, transmission, and consumption.

# 5 Summary, conclusions, future work

This study builds on the general hypothesis that "speakers signal prosodic structure via systematic phonetic details, and listeners use these cues to prosodic structure in decoding continuous speech" (Cho et al. 2007: 235). Tracing the manifestation of these cues in OE verse enriches the range of factors we can appeal to

in the reconstruction and account of the linguistic and cultural circumstances of poetic composition.

The phonological patterns forming the core of OE verse form can be rethought in terms of boundary signals. Punctuation and word division in the older manuscripts cannot always be relied on, even in verse, but verse structure can offer a closer, possibly more dependable information on internalized language for which there are no sound recordings. The speech-based, phonetically grounded perspective of initial consonant strengthening, applied to left-edge alliteration, throws light on the constraints on alliterating metrical positions, cluster alliteration, and vowel alliteration. The additional parameters behind the choice of [sp- st- sk-] as the only cohesive clusters in OE verse bear on the question of similarities and differences between OE and the verse forms in adjacent traditions. A reference to phonetic boundary effects in positing a pre-vocalic [ʔ-] as the basis of vowel alliteration is neither a blind transfer of hidden phonological structure, nor circular, which addresses Jakobson's (1963) reservations about the hypothesis.

The prosodic right-edge effect of vowel lengthening provides a link to observed variation of vowel duration in monosyllabic vowel-final function words and pronouns. It is also fully in accord with the historical persistence of the minimal-word requirement rendering stressed monosyllable of the type #(C)V# underlyingly ill-formed throughout the entire history of English.

Much of the argumentation regarding Kaluza's Law (4.2) in the voluminous literature on the subject revolves around the stress levels of the resolvable unit and the weight of the preceding syllable. Highlighting the significance of the verse-final position where the most salient suspension of resolution occurs opens up a new line of inquiry.

The proposal here is that suspension of resolution is a case of the metrical four-position principle acting in conjunction with optional verse-final / phrase-final vowel lengthening. Final lengthening is "to a first approximation the final syllable" (Ryan 2018: 6.1); it is almost entirely confined to the phrase-final syllable (Lindblom 1968). Positional lengthening at the right edge, bringing -CV final syllables perceptually closer to the "ideal" weight for regular suspension of resolution offers the poet the option of overriding the phonemic length of the verse-final syllable. The likelihood of neutralization of vowel length in this position is supported in other studies: Ryan (2017: 593) labels this phenomenon "Final indifference". It "refers to the near-universal by which quantity is ignored line finally or prepausally in quantitative metres, as also seen in Latin".

Some of the "oddities" of OE meter are thus argued to be phonetically realistic. This is not the place to elaborate on this, but the so called "happ-y tensing" in English offers a parallel to the lengthening factor adduced in the discussion of Kaluza's Law. Cruttenden (2014) treats word-final [ɪ] as the default for *CGB*

(Conspicuous British English), but recognizes its *GB* (General British) allophone [i] in word-final position as a norm, "a vowel nearer in quality to the long /i:/, rather than /ɪ/" (Cruttenden 2014: 113–116). The process is not new: Jones (2012: 829) cites early eighteenth-century recommended respellings such as *hunnee* 'honey', *munnee* 'money', and *munkee* 'monkey' (on the variability of unstressed word-final <y, i>, see also Beal 2000; Thomas 2007; Gordon 2012; Bekker 2014).

The implications of the claim that edge marking is an important factor in regulating metrical form go beyond the concerns of phonologists and address the concerns of literary historians. The references to the phonetic speech signals for boundary marking enrich the evidence for the basically oral nature of OE verse composition. Allowing for the effects of edge prosodic strengthening in verse, especially in a more emphatic delivery style, makes a view of the poet as a literate archaizer, someone who consciously used old-fashioned pronunciation forms culled from his reading of manuscripts of older poetry, less credible.[28] The Anglo-Saxon oral poetic tradition "was not killed by the advent of literacy, but continued to exercise a powerful influence on the poetic texts that have survived from the period in both inscriptions and manuscripts" (Orton 2014: 250).

The line of research suggested here is promising and invites further inquiry. We still do not know whether there is a match between the probability of suspension of resolution in OE and the type of vowel involved, given the inherent longer duration of low vowels compared to high vowels. I am not aware of a systematic study of orthographic marking in relation to verse-final CV# syllables. An area that deserves a separate paper has to do with the special properties of verse closure. The principle of "beginnings free – endings strict" (Hayes 1989) is shown to be operative in many verse traditions, but it is not immediately obvious how it fits the proposed account here. Stepping outside OE alliterative verse, it is worth exploring how the Middle English alliterative poets reacted to verse-final signals as a factor in the notoriously problematic metrical form at the right edge of the b-verse.

## References

Allen, W. S. 1969. The Latin accent: a restatement. *Journal of Linguistics* 5(2). 193–203.

*ASPR* = Krapp, George Philip & Elliot Van Kirk Dobbie (eds.). 1936. *The Exeter Book. The Anglo-Saxon poetic records*. Vol. 3. New York: Columbia University Press.

**28** This does not imply that some other archaic features, more specifically syntactic features, did not make it into verse composition (see Russom 2017: 55–56).

Beal, Joan C. 2000. HappY-tensing: A recent innovation? In Ricardo Bermùdez-Otero, David Denison, Richard M. Hogg & C.B. McCully (eds.), *Generative theory and corpus studies: A dialogue from 10 ICEHL*, 483–497. Berlin: Mouton de Gruyter.

Beckman, Mary & Jan Edwards. 1990. Lengthenings and shortenings and the nature of prosodic constituency. In John Kingston & Mary Beckman (eds.), *Papers in laboratory phonology 1: Between the grammar and physics of speech*, 152–178. Cambridge: Cambridge University Press.

Bekker, Ian. 2014. The weak vowels of South African English: A critical review and comparative acoustic analysis. *Southern African Linguistics and Applied Language Studies* 32(1). 133–147.

Bessinger, Jess B. Jr. & Philip H Smith. 1978. *A concordance to the Anglo-Saxon poetic records*. Ithaca, NY: Cornell University Press.

Bliss, A. J. 1962. *The metre of Beowulf*. Oxford: Blackwell.

Blumenfeld, Lev. 2016. End-weight effects in verse and language. *Studia Metrica et Poetica* 3(1). 7–32.

Burnley, David. 1995. Scribes and Hypertext. *The yearbook of English studies* 25. 41–62.

Byrd, Dani & Elliot Saltzman. 1998. Intragestural dynamics of multiple prosodic boundaries. *Journal of Phonetics* 26(2). 173–199.

Byrd, Dani & Susie Choi. 2010. At the juncture of prosody, phonology, and phonetics – The interaction of phrasal and syllable structure in shaping the timing of consonant gestures. *Laboratory Phonology* 10. 31–59.

Cable, Thomas. 1991. *The English alliterative tradition*. Philadelphia: University of Pennsylvania Press.

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

Cho, Taehong. 2005. Prosodic strengthening and featural enhancement: Evidence from acoustic and articulatory realizations of /a, i/ in English. *Journal of the Acoustical Society of America* 117(2). 3867–3878.

Cho, Taehong, Yoonjeong Lee, and Sahyang Kim. 2014. Prosodic strengthening on the/s/-stop cluster and the phonetic implementation of an allophonic rule in English. *Journal of Phonetics* 46(1). 128–146.

Cho, Taehong & Patricia Keating. 2009. Effects of initial position versus prominence in English. *Journal of Phonetics* 37(4). 466–485.

Cho, Taehong, James M. McQueen & Ethan Cox. 2007. Prosodically driven phonetic detail in speech processing: The case of domain-initial strengthening in English. *Journal of Phonetics* 35(2). 210–243.

Cook, Vivian. 2016. Background to the English writing system. In Vivian Cook & Des Ryan (eds.), *The Routledge handbook of the English writing system*, 25–44. London: Routledge.

Cruttenden, Alan. 2014. *Gimson's pronunciation of English*. 8th edn. London: Routledge.

Cutler, Anne & Sally Butterfield. 1992. Rhythmic cues to speech segmentation: Evidence from juncture misperception. *Journal of Memory and Language* 31(2). 218–236.

Donoghue, Daniel. 1987. *Style in Old English poetry: The test of the auxiliary*. New Haven, CT: Yale University Press.

Edwards, Jan, Mary E. Beckman & Janet Fletcher. 1991. The articulatory kinematics of final lengthening. *Journal of the Acoustical Society of America* 89(1). 369–382.

Fougeron, Cécile & Patricia A. Keating. 1997. Articulatory strengthening at edges of prosodic domains. *The Journal of the Acoustical Society of America* 101(6). 3728–3740.

Fulk, Robert D. 1992. *A history of Old English meter*. Philadelphia: University of Pennsylvania Press.
Fulk, Robert D. 2007. Old English meter and oral tradition: Three issues bearing on poetic chronology. *The Journal of English and Germanic Philology* 106(3). 304–324.
Fulk, Robert D., Robert E. Bjork & John D. Niles (eds.). 2014. *Klaeber's Beowulf and the Fight at Finnsburg*. 4th edn. Toronto: University of Toronto Press.
Garellek, Marc. 2012. Glottal stops before word-initial vowels in American English: Distribution and acoustic characteristics. *UCLA Working Papers in Phonetics* 110. 1–23.
Getty, Michael. 2002. *The metre of Beowulf*. Berlin: Mouton de Gruyter.
Gordon, Elizabeth. 2012. Standard New Zealand English. In Raymond Hickey (ed.), *Standards of English: Codified varieties around the world*, 318–336. Cambridge University Press.
Hayes, Bruce. 1989. The prosodic hierarchy in meter. In Paul Kiparsky & Gilbert Youmans (eds.), *Phonetics and phonology*, vol. 1, *Rhythm and meter*, 201–260. San Diego: Academic Press.
Hayes, Bruce & Margaret MacEachern. 1998. Quatrain form in English folk verse. *Language* 74(3). 473–507.
Hutcheson, B. R. 1995. *Old English poetic metre*. Cambridge: D. S. Brewer.
Hutcheson, B. R. 2004. Kaluza's law, the dating of Beowulf, and the Old English poetic tradition. *Journal of English and Germanic Philology* 103(3). 297–322.
Jakobson, Roman. 1963. On the so-called vowel alliteration in Germanic verse. *Zeitschrift für Phonetic, Sprachwissenschaft und Kommunikation* 16(14). 85–92.
Jones, Charles. 2012. Phonology. In Alexander Bergs & Laurel J. Brinton (eds.), *English historical linguistics. An International Handbook*, vol. 1, 827–842. Berlin: Mouton de Gruyter.
Kaluza, Max. 1896. Zur Betonungs- und Verslehre des Altenglischen. In *Festschrift zum siebzigsten Geburtstage Oskar Schade*, 101–134. Königsberg: Hartung.
Katsika, Argyro. 2016. The role of prominence in determining the scope of boundary lengthening in Greek. *Journal of Phonetics* 55. 149–181.
Katsika, Argyro, Jelena Krivokapić, Christine Mooshammer, Mark Tiede & Louis Goldstein. 2014. The scope of boundary lengthening as a function of lexical stress and pitch accent. *Journal of the Acoustical Society of America* 135(4). 2195.
Keating, Patricia, Taehong Cho, Cécile Fougeron & Chai-shune Hsu. 2004. Domain-initial articulatory strengthening in four languages. In John Local, Richard Ogden, & Rosalind Temple (eds.), *Phonetic interpretation*: *Papers in laboratory phonology* VI, 143–161. Cambridge: Cambridge University Press.
Kendall, Calvin B. 1991. *The metrical grammar of 'Beowulf'*. Cambridge: Cambridge University Press.
Lehiste, I. 1965. Juncture. In Eberhard Zwirner & Wolfgang Bethge (eds.), *Proceedings of the 5th International Congress of Phonetic Sciences*, 172–200. Basel & New York: S. Karger.
Lehmann, Winfred P. 1956. *The development of Germanic verse form*. Austin, TX: University of Texas Press.
Lewis, David. 1987. The metre of *Genesis B*. In Peter Clemoes, Simon Keynes & Michael Lapidge (eds.), *Anglo-Saxon England*, vol. 16, 67–127. Cambridge: Cambridge University Press.
Lindblom, Björn. 1968. Temporal organization of syllable production. *STL-QPSR* 9(2–3). 1–13.
Minkova, Donka. 2003. *Alliteration and sound change in early English*. Cambridge: Cambridge University Press.

Minkova, Donka. 2008. Prefixation and stress in Old English. *Word Structure* 1(1). 21–52.
Minkova, Donka. 2014. *A historical phonology of English*. Edinburgh: Edinburgh University Press.
Minkova, Donka & Robert Stockwell. 1997. Against the *emergence* of the nuclear stress rule in Middle English. In Jacek Fisiak (ed.), *Studies in Middle English* linguistics, 301–335. Berlin: Mouton de Gruyter.
Mitchell, Bruce. 1995. *An invitation to Old English and Anglo-Saxon England*. Oxford: Blackwell.
Mo, Yoonsook, Jennifer Cole & Mark Hasegawa-Johnson. 2009. Prosodic effects on vowel production: Evidence from formant structure. In *Proceedings of the 10th Annual Conference of the International Speech Communication Association 2009 (INTERSPEECH 2009)*, 2535–2538. International Speech Communication Association.
Neidorf, Leonard & Rafael J. Pascual. 2014. The language of Beowulf and the conditioning of Kaluza's Law. *Neophilologus* 98(4). 657–673.
O'Brien O'Keeffe, Katherine. 1990. *Visible song: Transitional literacy in Old English verse*. Cambridge: Cambridge University Press.
*OED = Oxford English Dictionary*. https://www.oed.com/ (Accessed 2019).
Orton, Peter R. 1994. On the transmission and phonology of the Battle of Brunanburh, *Leeds Studies in English*, n.s. 25. 1–27.
Orton, Peter. 1999. Anglo-Saxon attitudes to Kuhn's Laws. *The Review of English Studies* 50(199). 287–303.
Orton, Peter. 2014. *Writing in a speaking world. The pragmatics of literacy in Anglo-Saxon inscriptions and Old English poetry*. Tempe, AZ: Arizona Center for Medieval & Renaissance Studies.
Parkes, Malcolm B. 1991. *Scribes, scripts, and readers*. London: Hambledon Press.
Parkes, Malcolm B. 1993. *Pause and effect: An introduction to the history of punctuation in the West*. Berkeley and Los Angeles: University of California Press.
Pascual, Rafael J. 2016. Sievers, Bliss, Fulk, and Old English metrical theory. In Leonard Neidorf, Rafael J. Pascual & Tom Shippey (eds.), *Old English philology: Studies in honour of RD Fulk*, 17–33. Cambridge: D. S. Brewer.
Roberts, Jane. 2005. *Guide to scripts used in English writings up to 1500*. London: British Library.
Russom, Geoffrey. 1987. *Old English meter and linguistic theory*. Cambridge: Cambridge University Press.
Russom, Geoffrey. 1995. Constraints on resolution in *Beowulf*. In M. J. Toswell (ed.), *Prosody and poetics in the early Middle Ages: Essays in honour of C.B. Hieatt*, 147–163. Toronto: University of Toronto Press.
Russom, Geoffrey. 2005. Review of Donka Minkova, *Alliteration and sound change in early English*. *Language* 81(3). 745–748.
Russom, Geoffrey. 2017. *The evolution of verse structure in Old and Middle English poetry*. Cambridge: Cambridge University Press.
Ryan, Kevin M. 2017. The stress–weight interface in metre. *Phonology* 34(3). 581–613.
Ryan, Kevin M. 2018. Prosodic end-weight reflects phrasal stress. *Natural Language & Linguistic Theory* 37. 315–356.
Shattuck-Hufnagel, Stefanie & Alice E. Turk. 1996. A prosody tutorial for investigators of auditory sentence processing. *Journal of Psycholinguistic Research* 25(2). 193–247.

Shih, Shu-hao. 2018. On the interaction between phrase-final lengthening and prosodic structure. In Daniel Edmiston, Marina Ermolaeva, Emre Hakgüder, Jackie Lai, Kathryn Montemurro, Brandon Rhodes, Amara Sankhagowit & Michael Tabatowski (eds.), *Proceedings of the 53rd Meeting of the Chicago Linguistic Society*, 353–364. Chicago: Chicago Linguistic Society.

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

Sproule, David. 1987. Complex alliteration, full and unstressed rhyme, and the origin of deibide. *Ériu* 38. 185–200.

Stockwell, Robert P. & Donka Minkova. 1997. Prosody. In Robert Bjork & John D. Niles (eds.), *A* Beowulf *handbook*, 55–85. Lincoln: Nebraska University Press.

Suzuki, Seiichi. 1996. Preference conditions for resolution in the meter of "Beowulf": Kaluza's Law reconsidered. *Modern philology* 93(3). 281–306.

Terasawa, Jun. 1989. Metrical constraints on Old English compounds. *Poetica* 31. 1–16.

Terasawa, Jun. 2011. *Old English metre: An introduction*. Toronto: University of Toronto Press.

Thomas, Erik R. 2007. Phonological and phonetic characteristics of African American vernacular English. *Language and Linguistics Compass* 1(5). 450–475.

Turk, Alice E. & Stefanie Shattuck-Hufnagel. 2007. Multiple targets of phrase-final lengthening in American English words. *Journal of Phonetics* 35(4). 445–472.

Wightman, Colin W., Stefanie Shattuck-Hufnagel, Mari Ostendorf & Patti J. Price. 1992. Segmental durations in the vicinity of prosodic phrase boundaries. *Journal of the Acoustical Society of America* 91. 1707–1717.

Yoon, Tae-Jin, Jennifer Cole & Mark Hasegawa-Johnson. 2007. On the edge: Acoustic cues to layered prosodic domains. In *Proceedings of the International Conference on Phonetic Sciences*, 1017–1020. Saarbrucken, Germany.

Edgar W. Schneider and Sarah Buschfeld

# 5 Expanding boundaries of a function word: Uses of *one* in Early Modern and Modern English

## 1 Introduction

The word *one* is a linguistically most interesting form, because in the course of time it has transgressed many boundaries, i.e., it has adopted new functions and started to appear in new word class and other usage contexts, thus potentially illustrating a number of principles of language variation and change. In Rissanen's (1997: 87) words, the form offers "a fascinating story of enrichment and regularization of forms". It has existed in English since the Old English (OE) period, but since then, at different time periods, it has systematically acquired a wide range of new functions and has grammaticalized in some respects. Hence, as a lexeme it is difficult to categorize and grasp holistically. Thus, it represents an exceptionally strong instance of re-allocation, i.e., the process of an existing form adopting new functions (Lass 1990) – certainly one interesting way of crossing boundaries in linguistic evolution.

The versatility of *one* is illustrated best by the fact that even the plain word class (or part of speech) assignment in Modern English is fuzzy and partly unclear. For example, the online version of the *Longman dictionary of contemporary English* offers five quite distinct word class types which *one* may realize. In their terminology, *one* may be (1) a "number" ("they had **one** daughter"; interestingly, the term is not "numeral"); (2) pronoun ("you should buy **one**"; "**one** was crying"; "**one** of the books"; "a later **one**"); (3) determiner ("**one** person I can't stand"; "**one** thing I like"); (4) adjective ("her **one** concern"); or (5) noun ("Do you have any **ones**?").[1] However, criteria for assignment are not always clear and are not discussed.

Incidentally, it is worth noting that this versatility of *one* continues into modern varieties of English including Postcolonial Englishes (Schneider 2007),

1 All examples from the *Longman dictionary of contemporary English online*.

**Edgar W. Schneider,** Department of English and American Studies, University of Regensburg, Regensburg, Germany, e-mail: edgar.schneider@ur.de
**Sarah Buschfeld,** TU Dortmund University, Dortmund, Germany, e-mail: sarah.buschfeld@tu-dortmund.de

https://doi.org/10.1515/9783110643282-006

and the use becomes even more creative there. For example, *one* may function as a clause-final relativizer in Singapore English (e.g., "those wear black **one**" 'the ones who wear black'; Bao 2015: 104), or as an indefinite article in Cyprus English (e.g., "My daughter is **one** art teacher"; Buschfeld 2013: 116) or in Namibian English (e.g., "Namibia is **one** country where [. . .]"; cf. Kautzsch 2019). These new developments substantiate the basic picture of the form's grammatical and semantic flexibility and its readiness to adopt new functions and cross into new usage domains. These observations are important to keep in mind for the general developmental framework of the form, though in the present article the main focus of investigation will be on developments of *one* in British and American English (BrE and AmE) from 1600 onwards.

The early phase of the broadening of the functions of *one* has been rather thoroughly investigated in work by Kruisinga (1934) and, in particular, Rissanen (1967, 1997), who have provided descriptions of the uses of *one* in OE and Middle English (ME). In addition, of course, modern grammars and dictionaries of English document and discuss uses of *one* extensively (although their classifications and approaches turn out to be far from uniform). However, limited information is available for the period in between. The purpose of the present paper is to bridge this gap and investigate and document the development of *one* between the seventeenth and twentieth centuries and to look into the early differentiation of BrE and AmE, based on the ARCHER corpus (well known in historical linguistics and described below). In these periods and varieties we consider and document quantitative changes not only of word class and sub-class assignment but also with reference to syntactic functions, construction contexts, and the semantic categories of the referents of *one*. We argue that *one* is a linguistically interesting and challenging form that lends itself to an evolutionary study because of its manifold and partly fuzzy functional properties and its readiness to cross word-class boundaries.

## 2 Early developmental trajectories and changing functions of *one*

Initially, in OE *one* (with the form *an*) simply functioned as a numeral, as in (1).[2]

**2** All examples in this section are taken from Rissanen (1997).

(1) *þæt oft in gefeohte **an** feseð tyne* (Wulfstan, *Homilies*)
'that often in a battle one puts to flight ten'

The numeral then undergoes a process of *subjectification*, with no longer the numerical contrast (one as opposed to more) highlighted but the focus on an individual. In subsequent centuries this is then followed by ongoing processes of "grammaticalization and pronominalization" (Rissanen 1997: 87), thoroughly documented and illustrated in Rissanen (1997). He identifies six functional sub-types which constitute the basis of our own categorization in Section 3.2 (cf. Table 1). He concedes, however, that these types often overlap in meaning and are difficult to distinguish (e.g., Rissanen 1997: 100) – and this is something we can basically confirm on the basis of our observations.

In the earliest developmental step in OE *an*/*one* becomes an individualizing pronoun: it broadens to "single out a single individual from a set" (Los 2015: 48), one out of a set of potential referents in contrast to others, as in (2); Rissanen (1997) labels this type "personal-specific". Thus, *an*/*one* becomes a marker of indefiniteness and a "presentative" marker, a procedural signal to the hearer "denoting first mentions of referents that are going to play an important role in the discourse" (Los 2015: 48).

(2) *inn eode **an** þæs cyninges þegna* (Ælfric, *Lives of Saints*)
'in came one of the king's servants'

In the next step of expansion, individualizing *one* "loses its connection [. . .] to a set of referents from which the referent is singled out" (Rissanen 1997: 100), so its reference potential broadens from the "personal-specific" type (e.g., "**one** of my brothers") to the "personal non-specific" (e.g., "**one** whom she will love forever"). Thus, *one* moves to becoming an indefinite[3] pronoun, with first attestations occurring in ME and with its reference potential being either specific (meaning 'a certain') or non-specific ('anyone'), as in (3).

(3) *and yet mighte **one** goo from **one** to another* (Caxton, *Eneydos*)

The earliest instances of this pronominal usage refer to persons, and this pattern is well attested in Early Modern English (EModE). Non-personal referents occur a little later, in the early fourteenth century (Rissanen 1997: 101). During the seventeenth

**3** See Raumolin-Brunberg and Kahlas-Tarkka (1997: 21–25) for a thorough discussion of the notion of (in)definiteness.

century both specific and non-specific pronominal *one* become gradually less common, being replaced by *some* or *any* (Rissanen 1997: 119–121).

Via only a small step of semantic extension, the non-specific personal reference leads to the generic indefinite subject pronoun *one*, denoting what 'people in general' do (Rissanen 1997: 100–101, 116–117; cf. Raumolin-Brunberg 1994). This broadening was caused by the loss of the corresponding form *man* (still present as such in German) in ME, so *one* fills a linguistic gap. This type of usage of *one* developed during the fifteenth century and was common from the sixteenth century onwards, also, and increasingly so during the EModE period, as indefinite compounds, e.g., *some one*, *any one* (Raumolin-Brunberg and Kahlas-Tarkka 1997).

Finally, *one* adopted a novel grammatical role as what Rissanen (1997) (as well as others, e.g. Jespersen 1961: 245–271; Nevalainen and Raumolin-Brunberg 2003: 64–65) calls a "propword", a substitution form for noun phrase (NP) heads. It is noteworthy that unlike a pronoun, which replaces an entire NP, the propword, typically co-occurring with a premodifying adjective, substitutes the NP head only. It thus plays an important role in building cohesion in texts in English to the present day and entails fairly specific semantic properties. For example, this use of *one* obligatorily implies that the class of a referent noun mentioned immediately before remains the same but that a quality associated with this noun, expressed by a premodifying adjective, is explicitly rejected and replaced by another one, i.e., another adjective preceding *one* (e.g., "a boring book and an interesting **one**"; Halliday and Hasan 1976: 92–95). This pattern is already to be found in Chaucer's writings, as in (4).

(4) *I have the mooste steadfast wyf, and eek the mekeste* ***oon*** (Chaucer, *Canterbury tales*)

The propword usage type, called "anaphoric classifying *one*" by Kruisinga (1934: 333), clearly represents a case of grammaticalization. Formal traces of this pattern exist already in OE but are rare; it became systematically established much later. The earliest instances of the use with adjectives appear in the thirteenth century (Rissanen 1997: 102). The structure develops more broadly in late ME, caused by a syntactic tendency of that period to avoid adjectives as heads of NPs. Substantial growth of the construction can then be observed in the sixteenth and seventeenth centuries (Kruisinga 1934: 335; Raumolin-Brunberg and Nurmi 1997; Rissanen 1999: 198; Nevalainen and Raumolin-Brunberg 2003: 64–65).

The EModE period, 1500–1710, as based on the Helsinki Corpus, sees some significant changes (after Rissanen 1997: 120–121). The personal-specific use decreases and falls into "gradual obsolescence". The personal-nonspecific use

becomes a bit more restricted. In contrast, the personal-generic use undergoes a rapid increase (especially after 1640), and the propword use is reported to experience "an explosive increase in [. . .] popularity [. . .] in the latter half of the 17th century" (Rissanen 1997: 133).

Extrapolating from these developments, in the main, empirical part of this paper we investigate usage patterns of *one* from the seventeenth through the twentieth centuries. Similarly to Rissanen (1997), we focus on the word class developments of *one*, but we go beyond his investigations in three important aspects. First, we complement the diachronic picture by investigating the subsequent development of *one* beyond the ME and EModE periods, analyzing occurrences of *one* in the ARCHER corpus between 1600 and 1999. Secondly, as part of the ARCHER corpus from 1750 onwards we add AmE to the overall picture and ask whether any interesting differences between the two main varieties of English can be detected. Thus, in terms of both diachronic coverage, filling the gap left between EModE and Modern English, and with respect to the regional expansion from BrE to AmE, this paper extends the boundary of our knowledge of the development of this form. Thirdly, we broaden the analytical context by not only looking into word class but also into related formal parameters which may be of interest for the development of *one* and can be observed on the basis of a corpus. These comprise the different syntactic functions in which *one* occurs, its construction contexts (i.e., specific surface constituent sequences), and the semantic class of its referents. The decision to consider these parameters as well is partly motivated by the obvious interrelationships between them and partly inspired by Rissanen (1997). Despite his main focus on subtle word class changes, he repeatedly but only in passing refers to other parameters as well, e.g., the semantics of referents (human vs. nonhuman referents of new pronoun uses, for instance; e.g., Rissanen 1997: 112) or the syntactic functions *one* performs (e.g., Rissanen 1997: 94, 106–107, 114, 116).

Our specific research questions are as follows:

1. In terms of the word class of *one*, do we witness a continuation of the developments described by Rissanen (1997) for the time frame up to 1700, or do we find alternative trajectories of change?
2. Can any conspicuous differences between BrE and AmE be observed in the ARCHER data?
3. Are there any noteworthy correlations between the uses of *one* by word class with other syntactic or semantic factors (syntactic functions, constructional patterns, semantic referent categories)?

It seems obvious that the form *one*, with its constantly evolving developmental trajectories, lends itself particularly well to the investigation of boundaries

being crossed. Throughout its history it has transgressed word class assignments as well as its referential potential and the constructions in which it tended to appear. Documenting and understanding these processes of recategorization and restructuring is thus also one of the main goals of this paper.

# 3 Methodology

## 3.1 Data

Our analyses are based on the ARCHER corpus, version 3.2 (A Representative Corpus of Historical English Registers). ARCHER is a multi-genre historical corpus of English designed to continue the systematic historical corpus coverage of the early phases of the language provided in the Helsinki Corpus into the present time (when ARCHER was designed, i.e. in the late 1990s). It was originally compiled by Douglas Biber and Edward Finegan and has been expanded to become an international collaborative corpus project, with fourteen universities in seven countries involved in this effort. After obtaining a User Agreement,[4] ARCHER can be searched online via a CQP web server (http://www.projects.alc.manchester.ac.uk/archer/).

ARCHER covers the period from 1600 to 1999, subdivided into 50-year subperiods. Project-internally they are labelled periods 1–8, i.e., period 1 equals the time frame of 1600–1649; period 2 is 1650–1699; etc. However, for brevity and convenience of reference in the following graphs and in our text we denote these sub-periods by their beginning year and symbolize the extension by a dash, i.e., 1600– refers to 1600–1649, 1650– means 1650–1699, etc. For BrE, all eight periods are included. Given that AmE split off and gained a distinctive character only later, its coverage in ARCHER sets in later, from period 4, i.e., 1750–, only. Thus, the entire corpus is divided into thirteen sub-corpora for half-century periods per variety, eight for BrE and five for AmE. In addition, twelve different registers (genres) are systematically covered, a quality which guarantees the stylistic representativeness of the corpus but is not considered here.

The ARCHER corpus contains as many as 8,622 tokens of *one*. Obviously, for practicality reasons (cf. Leech et al. 2009: 47) it is necessary to draw a

4 Thanks to David Denison, Nuria Yáñez-Bouza, and Sebastian Hoffmann.

sample from these attestations. Our goal was to draw a random sample of 200 tokens per sub-corpus,[5] i.e., to analyze ca. 2,600 tokens, about 30 percent of the entire population. Since a few tokens in the sample were incomplete or unclear for classification and categorization and could therefore not be analyzed, we ultimately investigated and categorized 2,571 tokens of *one* overall.

## 3.2 Analysis

We coded all sampled tokens of *one* for the four different parameters mentioned above which might have an impact on usage conditions of *one* and might change in the course of time: word class, syntactic function, construction context, and referent meaning. Some of these (e.g., word class, syntactic function) constitute conventional descriptive categories of the grammar of English and were thus determined beforehand; others (especially the categories of construction context) were expanded during the process of coding, based on the occurrence of specific patterns found in the sample. With a small number of tokens context was insufficient to determine specific category assignments unambiguously (eighty-five times for syntactic function, eleven times for construction context, thirteen times for meaning – but not at all for word class). These were marked as "na" in the database and disregarded from the analysis of the respective category.

For the category of word class we investigate two different levels of granularity. A "broad" word class analysis looks into the frequencies of *one* as a numeral, determiner, adjective, or pronoun, respectively, as well as a category of phraseologisms, idioms, and the like which we labelled "lexical chunk". For two word classes we also looked into more finely graded subdivisions, namely subtypes of pronouns (which are obviously relevant, given Rissanen's [1997] description of developmental stages) and subtypes of determiner usage. Table 1 documents the word class types and subtypes considered, with examples and explanatory notes.

**5** Randomization was carried out by the simple but effective strategy of selecting every *n*-th token in a given subsample, with *n* determined by the sub-sample size so that the tokens picked spread evenly across the entire sub-corpus. For example, the period 1700– in BrE shows 655 tokens of *one*, so for *n*=3, every third token selected, we get a random sample of 200 tokens.

**Table 1:** Categories analyzed for word class.

| Word class | Examples | Further explanations |
|---|---|---|
| Numeral | "I have **one** child"; "**one** or two" | numerical meaning '1' |
| Determiner (particular) | "**one** thing I can't stand" | meaning 'a particular' |
| Determiner (with names) | "He worked for **one** Mr. Smith." | meaning 'a certain' |
| Determiner (time reference) | "**one** day" | time unit as noun; meaning 'some' |
| Determiner (emphatic) | "**one** hell of a game" | emphatic function |
| Determiner (article) | "**one** art teacher" | as indefinite article |
| Adjective | "her **one** concern was [. . .]" | preceding the noun after a determiner; typically means 'only'[6] |
| Pronoun (propword) | "A red apple?–A green **one**." "He is no cuckold, you are **one**." | substitutive (personal or non-personal referent, previously mentioned, with or without modifier) |
| Pronoun (personal specific) | "**one** of the servants" "**one** who was sick" | a particular human individual, usually identifiable; closed or open reference group |
| Pronoun (personal non-specific) | "There's **one** at the door who wants [. . .]" | reference to individual but not a specific one known to the speaker: ~ 'someone' |
| Pronoun (personal generic) | "**One** shouldn't do that." | typically alone; meaning indefinite (cf. German *man*): 'people in general', 'any individual' |

**6** Both the position and the meaning justify the categorization of these forms as adjective (rather than post-determiner, which might also be conceivable on the basis of the position alone). Substitution tests in paradigmatic relationship with other adjectives confirm this assignment: "her one/only/sole/deepest/ [. . .] concern".

**Table 1** (continued)

| Word class | Examples | Further explanations |
|---|---|---|
| Pronoun/ numeral | "**one** of . . . " (e.g., "**one** of my friends"; "**one** of my children") | head of NP, meaning '1', often partitive; closed, delimited reference group |
| Lexchunk | "**one** hundred"<br>"**one** another" | lexicalized chunk / firm collocation in which *one* occurs |
| Na | | any other or unclear word class; incomplete tokens which cannot be syntactically categorized |

Second, we coded the syntactic functions of the constituent in which *one* occurred,[7] based on canonical syntactic functions as spelled out in Table 2.

**Table 2:** Categories analyzed for syntactic function.

| Syntactic function | Examples | Further explanations |
|---|---|---|
| Subject | "**One** never knows."<br>"My **one** concern is this." | |
| Object | "I know **one** thing."<br>"Give me **one**." | |
| Subject complement | "This is **one** concern of mine."<br>"He's a nice **one**." | |
| Object complement | "I considered him the only **one**." | |
| Prepositional complement | "with **one** person"; "with this **one**"; "In **one** of her houses [. . .]" | *one* in PP in complement function |

**7** This comprises instances in which *one* constitutes different constituents within a noun phrase: the complete noun phrase, the noun phrase head, or a determiner or a modifier within the noun phrase. We decided not to break this down any further here since this would introduce another level of complication that would not seem motivated by any syntactic hypothesis. Of course, the role of *one* within noun phrases (but not combined with syntactic functions) is indirectly categorized in the word class and construction context frames.

**Table 2** (continued)

| Syntactic function | Examples | Further explanations |
|---|---|---|
| Adverbial | "**One** day [. . .]" | if not part of a PP; e.g., adverbial of time or place |
| Na | | any other or unclear syntactic functions; incomplete tokens which cannot be syntactically categorized |

Thirdly, we probed into the surface construction contexts of *one*, classifying relevant constituent sequence strings. Not surprisingly, this turned out to be a more complex type of subdivision, with its options identified in Table 3.

**Table 3:** Categories analyzed for construction context.

| Construction context | Examples | Further explanations |
|---|---|---|
| *One* | "**One** never knows."<br>"I had **one**." | head, unmodified |
| *One's* | "**one's** property",<br>"it is **one's** duty to [. . .]" | *one* + genitive 's + N |
| Det+*one* | "the **one**",<br>"this **one**",<br>"which **one**" | head, preceded by determiner |
| Det+mod+*one* | "the old **one**" | head, preceded by determiner + adjective (or other modifier) |
| Det+*one*+N | "the **one** thing" | *one* preceded by determiner and followed by head noun |
| Det+*one*+ mod+N | "the **one** good thing",<br>"the **one** best friend" | *one* preceded by determiner and followed by adjective (or other modifier) preceding head noun |
| Det+*one*+rel | "the **one** who/which [. . .]" | head, preceded by determiner and followed by postmodifying relative clause |
| *One*+rel | "**one** who [. . .]" | head, followed by relative clause |

**Table 3** (continued)

| Construction context | Examples | Further explanations |
|---|---|---|
| Det+*one*+PP | "the **one** with [. . .]" | head, preceded by determiner and followed by postmodifying PP |
| *One*+PP | "**one** of [. . .]" | head, followed by postmodifying PP |
| *One*+PP (coord.) | "**one** or the other of [. . .]" "**one** or two of [. . .]" | head, followed by coordinated head preceding postmodifying PP |
| *One*+N | "**one** man" | *one* immediately preceding head noun |
| *One*+N (coord.) | "**one** or two years", "**one** or more men" | *one* + coordinated modifier + head noun |
| *One*+mod.+N | "**one** nice guy" | *one* + adjective (or other modifier) preceding head noun |
| N+*one* | "(from) day **one**" | *one* as postmodifier after a N |
| Na | | any other or unclear; context insufficient for classification |

Finally, we inquired into the distribution and manifestations of semantic referent classes of *one* and whether these have been changing in the course of time. Since it is well known that animacy is not a clear-cut binary category but is better understood as a continuum (e.g., Rosenbach 2008) we allowed for a few intermediate types, explained and exemplified in Table 4.

**Table 4:** Categories analyzed for meaning of referent noun.

| Meaning of referent N | Examples | Further explanations |
|---|---|---|
| Human | *gentleman*, *hunter* | names, denomination / characterization of people, human roles, etc. |
| Human (group) | *team*, *company* | collective nouns for humans: group designations, countries, etc. |
| Nonhuman (animate) | *horse*, *devil* | animals, plants, gods, etc. |

**Table 4** (continued)

| Meaning of referent N | Examples | Further explanations |
|---|---|---|
| Concrete (inanimate) | *lamp*, *ship* | things, objects, etc. |
| Time unit[8] | *day*, *year* | time periods |
| Abstract | *reason*, *trouble* | abstract notions except time |
| Na | | any other or unclear; context insufficient for classification |

Of course, these categories of units of analysis are all interrelated. For example, constituent structure is always closely connected with word class and syntactic function. Depending on the precise construction *one* occurs in, it can take on several of them. For example, *one* as head of NPs occurs in characteristic patterns: e.g., *one* preceded by just a determiner and possibly a modifier, as in (5), *one* followed by a relative clause (6), or *one* followed by a prepositional phrase (7). *One* as adjective is often preceded by a determiner and followed by a noun, as in (8).

(5) *a cold **one*** (1665head_f2b)[9]

(6) ***one** who could obtain me that letter* (1845poe_f5a)

(7) ***one** of the Visitors* (1728gran_y3b)

(8) *the **one** bar grubby enough* [. . .] (1971mich_f8a)

**8** We decided to treat time units separately from other abstract nouns since they tend to co-occur with *one* fairly frequently, so there is reason to hypothesize that the time nouns might behave somewhat differently. Table 8 below confirms that this is indeed the case. Pooling all abstract nouns might have had a slightly skewing effect, giving a disproportionately high impact to time nouns.

**9** Examples are identified by ARCHER filenames, which report the year when the token was produced, a four-letter abbreviation of the author's name, and, after the underscore, one-letter symbols for genre, period, and variety.

Similarly, some word classes are associated with semantic categories (e.g., personal generic pronouns refer to humans) or with syntactic functions (e.g., Rissanen [1997: 116] claims that "generic *one* [. . .] is mostly used as the subject"). Investigating these parameters as well may allow us to identify or test such correlations.

In the following section, we report the results of these analyses, viz., raw token frequencies of the respective categories per sub-period and per variety, in form of line diagrams created on the basis of Excel pivot tables. Beyond that, the data set has been submitted to statistical significance testing.[10]

Word class manifestations and their potential changes through space (variety) and time (period) are statistically modelled on the basis of a conditional inference trees analysis, using the package "partykit" in R (cf. Hothorn et al. 2006; R Development Core Team 2014). A Conditional inference trees (ctrees) analysis (for previous applications of this technique, see, e.g., Tagliamonte and Baayen 2012; Bernaisch et al. 2014; Koch et al. 2016) applies recursive partitioning algorithms to the data set in order to "classify / compute predicted outcomes / values on the basis of multiple binary splits of the data" (Bernaisch et al. 2014: 14). In other words, an inference tree analysis investigates a data set in a recursive fashion "to determine according to which (categorical or numeric) independent variable the data should be split up into two groups to classify / predict best the known outcomes of the dependent variable" (Bernaisch et al. 2014: 14). In our study, WORD CLASS is the dependent variable; PERIOD and VARIETY are the independent variables. As Baayen (2013: 364) summarizes:

> The ctree algorithm begins with testing the global null hypothesis of independence between any of the predictors and the response variable. The algorithm terminates if this hypothesis cannot be rejected. Otherwise, the predictor with the strongest association to the response is selected, where strength is measured by a p-value corresponding to a test for the partial null hypothesis of a single input variable and the response. A binary split in the selected input variable is carried out. These steps are recursively repeated until no further splits are supported.

The final result is a readily and easily accessible and interpretable decision tree (as presented in Figure 4 and interpreted and further described in Section 4). For this statistical analysis we decided to slightly trim our data for reasons of better accessibility and explanatory power. We did so on the basis

**10** We would like to thank Sali Tagliamonte, Stefan Th. Gries, and Tobias Bernaisch for their advice on statistical questions, as well as Sven Leuckert for providing us with the script for creating ctrees.

of our findings from the descriptive analysis (see Figures 1–3). We conflated all low-frequency word class distinctions in the determiner and pronoun sets and deleted the category lexchunk as this latter one was of no further interest for the present study. This yielded eight remaining factor levels (word class types) for analysis, viz., adjective, determiner_particular, determiner_other, numeral, pronoun_numeral, pronoun_personal (pooling specific and nonspecific uses), pronoun_personal_generic (the latter two kept separate because of the interesting developmental tendencies of personal generic pronouns according to Figure 2), and pronoun_propword.

We started out with a full model including all potential predictors, viz., period and variety as well as the language internal variables syntactic function, construction context, and meaning. The analysis returned multiple branching trees in which all three language-internal variables turned out to be indeed important predictors for word class. The tree, however, was far too complex to interpret, and some of these interrelations come naturally, e.g., construction contexts being closely connected with word classes, or human referents preferably appearing as subjects. Of course, these are still statistically relevant findings bearing on the realization of word class of *one*, but they are not of immediate interest for the present analysis. We therefore focus on WORD CLASS as the dependent variable and the two independent variables PERIOD and VARIETY only (see Figure 4).

# 4 Results: The development of *one* through time and space

Turning towards the results, we look into diachronic and geographical dimensions of the spread and development of *one*. We begin our analysis by investigating the broad word classes *one* has taken on during the time span investigated, viz., 1600–1999 (see Figure 1). Subsequently, we will zoom in on subtypes of its uses as pronoun (see Figure 2) and determiner (see Figure 3), since these are the domains where, based on the historical records discussed above, further developments and changes may be expected. We proceed by investigating the syntactic functions it has occupied within this particular time span, the different construction contexts it occurs in as well as the diverse meanings it has taken on (as described in Section 3.2). Given the nature of the data, the first three periods always cover the development of *one* in BrE only.

## 4.1 Word class changes

As Figure 1 illustrates, an analysis of the broad word classes in which we find *one* used during the study period reveals quite a clear picture. Between 1600 and 1650, the corpus reveals a frequency of below 80 tokens for all word classes in BrE, with the clearly highest token frequency for *one* as pronoun (77 tokens), followed by *one* as determiner (51), numeral (27), lexchunk (22), and adjective (9). Except for *one* as adjective, which remains mostly stable in its low frequencies (and AmE here behaves in very similar ways to BrE, displaying even slightly lower token frequencies), most of the word classes experience both slight increases and decreases in token frequencies throughout the period. Still, when comparing their initial states after 1600 with their frequencies in the 1950– period nothing of substance has changed in their word class frequencies, and the two varieties seem to behave in largely similar ways. The use of *one* as lexical chunk goes down slightly towards the end of the period of investigation. In the 1800– and 1850– periods we can observe reverse trajectories of behavior between BrE and AmE in their use of *one* as numeral (green lines), determiner (blue lines), and pronoun (red lines). In adopting pronoun and numeral functions AmE can be viewed as trailing BrE by one period, with peaks 50 years later. It seems that at this stage structural innovations originated in BrE and took some time to spread to AmE – which is perhaps in

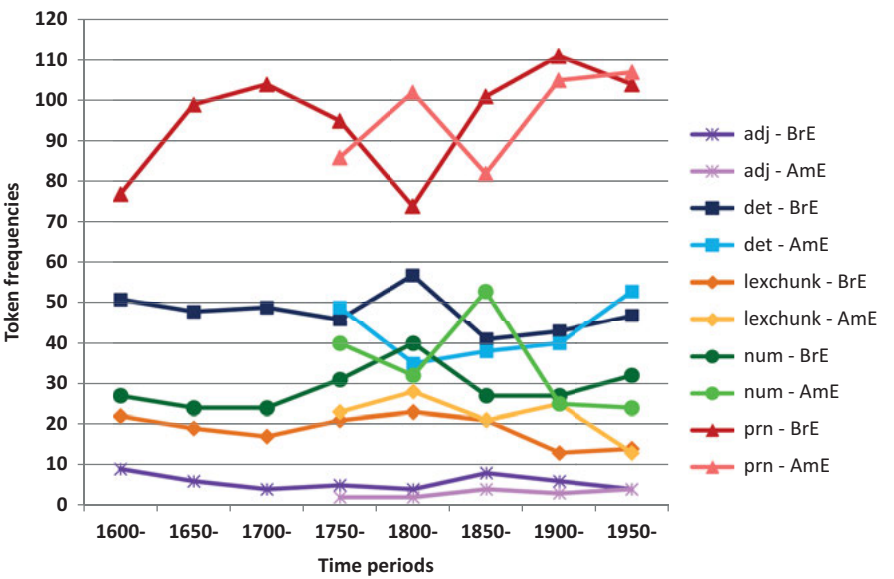


**Figure 1:** Word class of *one* by period and variety.

line with the notion of *colonial lag*, the assumption that AmE (like, possibly, colonial varieties in general) retains a more conservative character (Görlach 1987). Hundt (2009) offers a thorough discussion and corpus-based analysis of this notion. She identifies a few structures in which "BrE is [. . .] ahead of AmE" (Hundt 2009: 17) – though this is clearly not the only possible scenario of "differential change" (2009: 34).

Within the overall picture, it is especially the development of the pronoun type, illustrated in (9), which is remarkable and deserves some further attention. Our results confirm Rissanen's (1997) observations (reported in Section 2) that the story of *one* is essentially a story of pronominalization. Except for some dips in use in both BrE and AmE in the 1800– and 1850– periods (some of which, though not all, may be due to chance), its general use has increased in BrE in the initial period under investigation (1600–) to the last period observed (1950–).

(9) *even if* ***one*** *isn't employed by them* [. . .] (1971tyna_y8b)

Since Rissanen's (1997) analysis has revealed particularly interesting changes among pronoun subtypes, we employ a more fine-grained analysis of subtypes, too. Figure 2 provides the more detailed frequencies of a variety of subcategories of *one* as pronoun (see Section 3.2 for further details). As Figure 2

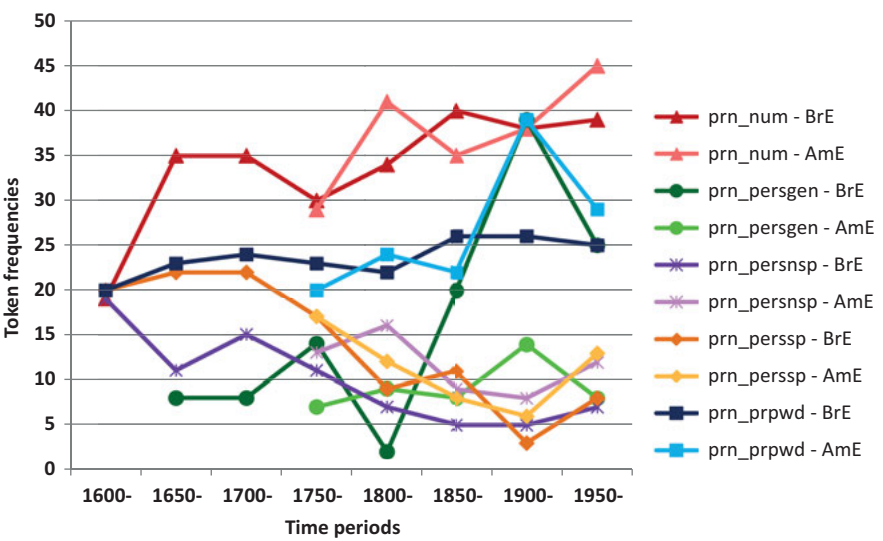


**Figure 2:** Type of pronoun by period and variety.

reveals, *one* as pronoun started out with approximately equal numbers of subtypes between 1600 and 1650 (all have nineteen or twenty tokens). *One* as personal generic pronoun could not be recorded for this time in ARCHER and entered the scene in the 1650–1700 period only. As discussed above, *one* as pronoun experienced a general increase in the course of time. However, a closer look at the data shows that this is not true for all subtypes. There are two types of developmental trajectories to be distinguished here: (1) some pronoun types of *one* have generally gone up in frequencies (with some minor dips and rises) throughout the time periods investigated, viz., *one* as propword (as in 10) and in particular *one* as pronominal NP head with strongly numeral meaning (see 11); (2) other types have experienced a decrease (again with some minor dips and rises) in frequencies, viz., *one* as personal specific pronoun and as personal non-specific pronoun (mainly after 1750– in BrE). These latter changes confirm Rissanen's (1997: 120–121, 125–126) description of developments in the early phase of this period (see Section 2).

(10) [. . .] *gave him a great mind, me only an honest* ***one*** (1844bouc_d5b)

(11) *It is* ***one*** *of the wonders of the world* (16xxbaxt_h2b)

In general there do not seem to be any considerable differences between the varieties, except for *one* as personal generic (or indefinite) pronoun (green lines). It is different from the other patterns observed in two interesting ways. First of all, the occurrence of *one* as personal generic pronoun is first recorded for the 1650– period in the corpus and at much lower frequencies than the other subtypes observed. Secondly, the two varieties behave in very different ways. Whereas the frequencies for *one* as a personal generic pronoun remain comparatively stable in AmE (when comparing the 1750– period to the 1950– period, with an intermediate rise between 1900 and 1950), the numbers for personal generic *one* in BrE experience a drastic increase from the 1800– period to the 1850– and 1900– periods. There is a dip in the latter half of the twentieth century, but the comparatively high frequency of this pronoun subtype in BrE remains stable (especially when compared to the few tokens in AmE). Hence, generic *one* seems to have been emerging as a predominantly British type of usage for the last two hundred years. We can only speculate whether AmE possibly prefers alternative strategies such as the passive to camouflage individual agency (as may be suggested by the observation that during the nineteenth century the frequency of the *get*-passive of AmE overtook that of BrE; cf. Hundt 2009: 18–19).

In a similar vein, Figure 3 illustrates the distributions of subtypes of *one* as a determiner. Here we find one noteworthy outlier, viz., a remarkable growth of *one* as "particular" determiner, as in (12).

(12) [. . .] *it is* [. . .] ***one*** *continued vindication* [. . .] (1850robe_h6b)

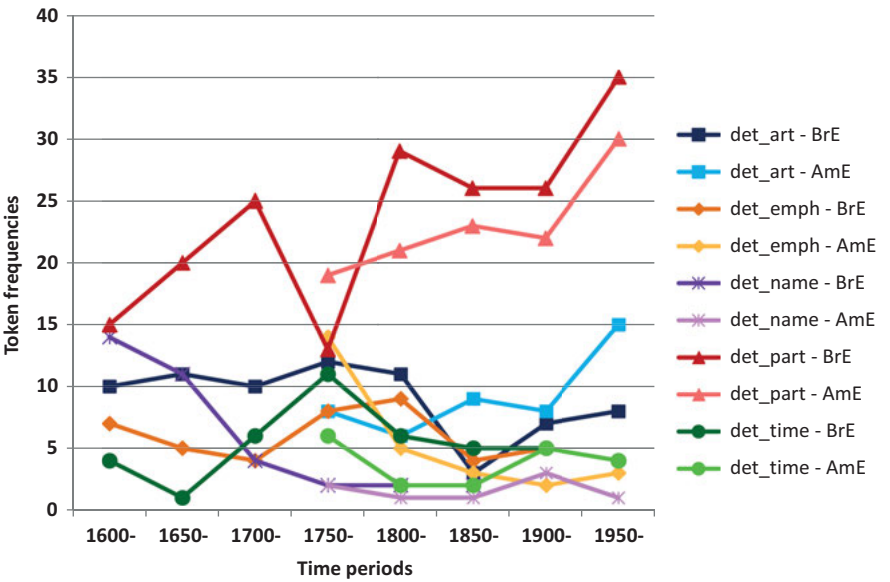


**Figure 3:** Type of determiner by period and variety.

Again, the different subtypes of *one* as determiner start out at comparatively similar frequencies – though not as similar as for the pronoun uses. *One* as a determiner with names starts out as one of the higher frequency determiner types in BrE but experiences a somewhat drastic decrease during the 1700– period, remains rare in 1800–, and disappears thereafter. The AmE data reveal a similar pattern: *one* as a determiner with names comes in at a low token frequency in the 1750– period and never rises above 3 tokens (between 1900 and 1949). *One* as article, emphatic determiner, and with time reference are all stable at low frequencies, with a slight increase up to the 1750– period. In BrE, these three types never get beyond a maximum of 12 tokens; in AmE, they start out at moderate frequencies in 1750–. Subsequently, they experience a further decline in both varieties,

which is particularly steep in *one* as emphatic determiner.[11] Only *one* as article in AmE shows an increase, which is most prominent in the last, 1950– sub-period.

As already pointed out above, *one* as "particular" determiner, illustrated in (13), shows a remarkable increase overall in both varieties, by far the highest and most stable one among the different subtypes (with a conspicuous dip, however, in the 1750– period).

(13) ***One*** *Quality peculiar to this Ink is, that* [. . .] (1772cens_a4a)

AmE comes in at nineteen tokens in the 1750– period, and then this use of *one* also increases quite consistently. Both varieties show a very similar pattern, especially in the last three phases of development under consideration here, with AmE frequencies slightly and consistently ranging below the BrE ones. In the 1950– period, frequencies are as high as thirty tokens in AmE and thirty-five tokens in BrE, respectively. No evident reason is discernible for this development. Tokens are not concentrated in individual texts. Several examples come from the drama category, where perhaps the need to single out one particular referent from several other contextually available ones is most evident.

Figure 4 illustrates the findings from the conditional inference tree analysis, which, as explained above, was just modelled on WORD CLASS as dependent variable and PERIOD and VARIETY. As illustrated in the ctree structure in Figure 4, both independent variables included in the model are important predictors for the word class of *one*. Overall the tree reveals that PERIOD has the stronger influence on word class than VARIETY as it is responsible for the first major split in the data ($p < 0.001$), which divides the data set into the time periods before 1850 and after (as indicated by the branch labels). This shows that there is a significant difference in word class manifestations of *one* between the periods 1600– to 1800– and 1850– to 1950–.

A further significant split shows in the before-1850 data set, once more along the lines of PERIOD, indicating a significant difference in word class manifestations in the 1600– period as opposed to the four following ones (see Node 2; $p = 0.003$). The two terminal nodes (Node 3 and Node 4) provide us with the following information: 1) the number of realizations of *one* (164 and 1058 respectively) which belong to the respective categories defined by the splits, viz., *one* as realized by word class in the 1600– period and *one* as realized by word class in the 1650– to 1800– periods; 2) a bar plot of observed percentages of the possible word class manifestations under observation (see Section 3.2; for a similar description of how to read ctrees,

**11** In BrE in 1950–, this form is not documented at all.

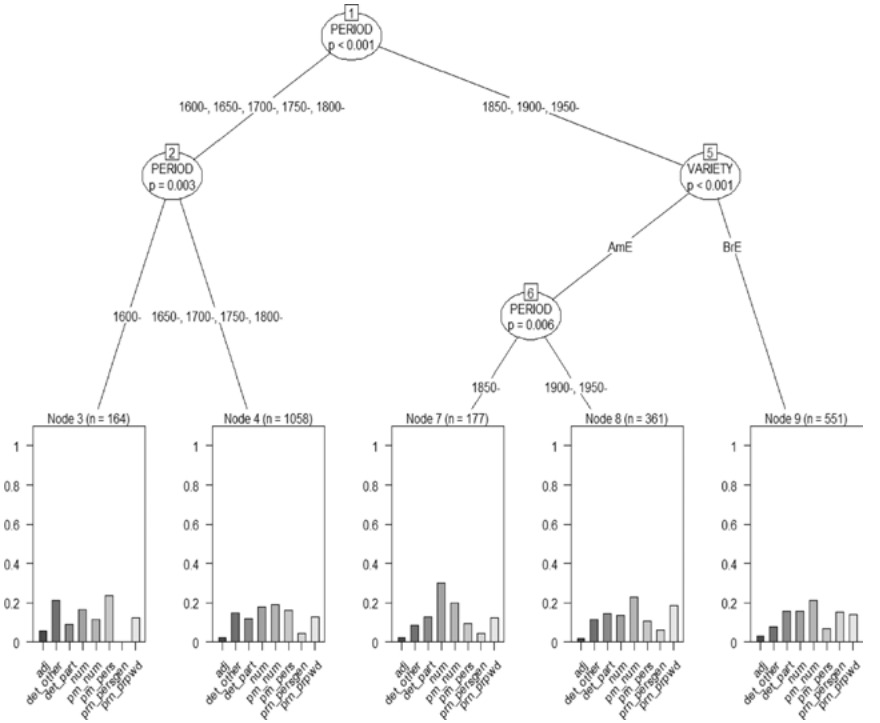


**Figure 4:** Ctree for WORD CLASS, VARIETY, and PERIOD.

see Bernaisch et al. 2014: 16). As a comparison of the two bar plots reveals, the changes are manifold and involve quite a few of the word class subtypes, e.g., the percentages of personal generic pronouns, *one* as pronoun/numeral, and *one* as numeral go up in frequency, *one* as determiner_other goes down, *one* as personal generic pronoun, which does not manifest in the 1600– period at all, generally rises in the 1650– to 1800– periods (and further increases in the course of time, as the other bar plots, i.e., Nodes 7–9, will reveal). The general tendency seems to be that word classes of *one* level out somewhat, i.e., that they approach each other in frequencies (some rise and some drop in their overall percentages).

The right branch of the initial period split (viz., the data from the 1850– to 1950– periods) shows another interesting split, this time with VARIETY as significant predictor ($p < 0.001$), indicating that word class manifestations of *one* are significantly different between BrE and AmE. Within the AmE data set, another split manifests between the early AmE data (in the 1850– period) and AmE in the 1900– and 1950– periods. Compared to especially the bar plot in Node 4, word class manifestations of *one* appear to differentiate again. *One* as pronoun/

numeral goes up, *one* as personal-generic pronoun increases in frequency in particular in BrE (as noted above), and *one* as numeral is particularly strong in AmE in the 1850– period.

Summing up, the tree suggests that both period and variety are significant predictors for realizations of *one* in terms of word class, which is an interesting and relevant finding clearly going beyond Rissanen's (1997) earlier observations. The findings suggest that important changes in word class manifestations and distributions occur once AmE had branched off and started to developed distinctive characteristics (see Schneider 2007: 273–282). The ongoing dynamism especially of AmE manifests itself furthermore in Node 6, which illustrates the fact that another split divides the 1850– and post-1900 phases of AmE with respect to the word class distributions of *one*. However, cross-tabulation of the predicted probabilities with the observed outcomes has returned a comparatively low accuracy rate (0.2146257), which implies that the model does not fit the data very well. This implies that we clearly are confronted also with a fairly high (and ultimately unclear) amount of random variability.

## 4.2 Changes in syntactic functions

Next, we take a more detailed look into the different syntactic functions *one* (or the NPs which contain it) can perform at the sentence level. As Figure 5 illustrates, different syntactic functions of *one* come at different frequencies, though the distribution shows fewer outliers and no completely exceptional developments compared to the one for word classes. (There is a possible exception for *one* in object complements, which shows nearly no occurrences at all, but then, object complements are generally a rarely appearing syntactic function.)

In general, *one* occurs most often as or in a prepositional complement, as in (14), starting out at fifty-nine tokens in BrE between 1600 and 1650.

(14) [. . .] *cover them all with* ***one*** *general slough* (1769bard_m4a)

The use experiences some mostly moderate ups and downs in its development and one more noticeable decline in the 1850– period in both BrE and in AmE, then goes up again, closely approaching the value initially taken by BrE in the final period under observation (1950–). *One* as prepositional complement is closely followed by *one* as subject, *one* as/in objects and *one* as/in subject complements, in this order. Certainly these relationships mirror the frequency distributions of the different syntactic functions in English sentences in general, viz., subjects occur generally more often than objects and, in

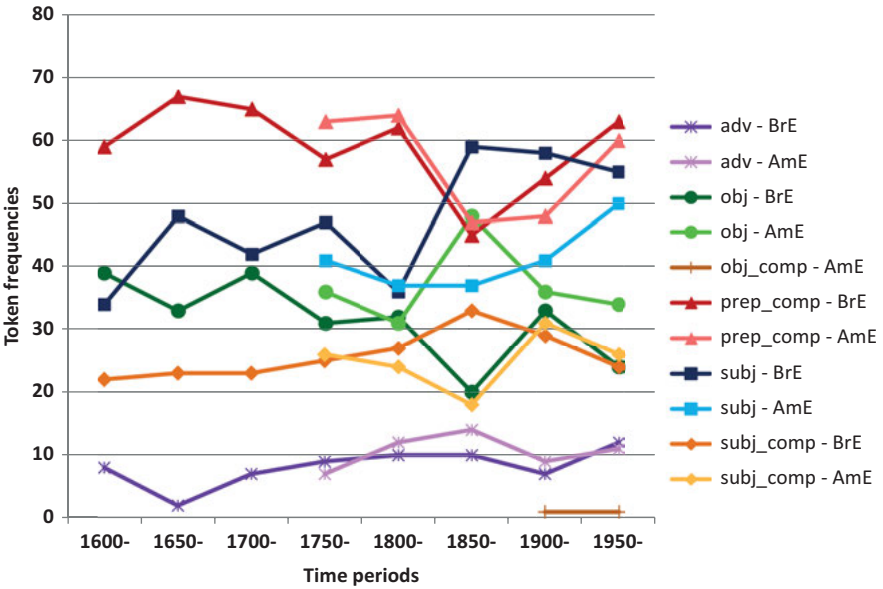


**Figure 5:** Syntactic function of *one* by period and variety.

turn, subject complements. *One* in subject function experiences a clear upward trend in the course of time, with BrE in the lead in 1850– and 1900–. *One* in object function experiences a slight overall decrease in frequencies, with one divergent development between the two varieties in the 1850– period, in which AmE experiences an increase in token frequency and BrE a decrease. The numbers for *one* in subject complement function are fairly stable. Only BrE shows minimally higher token frequencies when comparing initial and final ones (which, once more, might be easily due to chance). What is interesting (but difficult to interpret and, once more, possibly a chance effect) is that the two varieties again diverge in their frequencies in the 1850– period, with BrE going slightly up and AmE slightly down. Note that a similarly divergent development can be observed for *one* in subject function. *One* as adverbial comes in at comparatively low frequencies, which remain mostly stable over time. This, again, is certainly not too surprising when considering the general distribution and frequencies of the syntactic functions as such and the often optional character of adverbials. We will return to a more general assessment of patterns of change, including some of these changes in syntactic functions, in 4.3 and 5.

## 4.3 Changes in construction contexts

When considering the different construction contexts in which *one* occurs (see Section 3.2), the picture is less stable and clear, with many constituent sequences experiencing apparently unpredictable up and down developments over time. We split the construction contexts up into two sets for better illustration and accessibility. We start with all construction contexts which have a determiner as initial, left-hand element followed by *one*, either in isolation or preceded or followed by other grammatical material, such as modifiers, nouns, relative clauses, or prepositional phrases. Since token frequencies are often low and any findings tentative, we comment on those with higher frequencies only and gloss over the others, reporting their general tendencies only.

As Figure 6 reveals, in the Det-preceded construction contexts *one* occurs most frequently in the Det + mod + *one* construction, as in (15), which corresponds to Rissanen's (1997) innovative propword type.

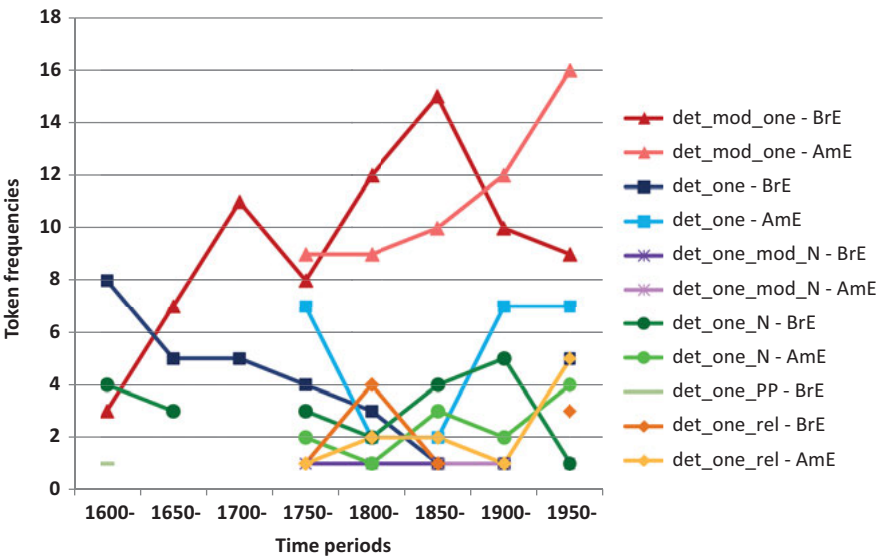


**Figure 6:** Construction context of *one* (constructions containing *one* preceded by determiner) by period and variety.[12]

**12** Interrupted lines are caused by missing instances at particular periods.

(15) [. . .] *an old tale: I have now a new* ***one*** (1628ford_d1b)

It appears to be generally on the rise in our data, too, especially and consistently in AmE. In BrE the pattern experiences a decline from the 1850– period onwards; in general it develops from a low frequency in the 1600– period via some seemingly erratic ups and downs and especially a peak in 1850– to a moderate token number in the 1950– period, (but note once again that token numbers are far from robust and frequency differences might be due to chance). The other construction pattern which is still somewhat frequent is Det + *one* (blue lines), as in (16).

(16) *the* ***one*** *traversing the room* [. . .] (1797fost_f4a)

The construction experiences a sharp overall decrease in BrE until the 1900– period and then goes up again in the 1950– period. AmE follows a similar pattern, with a decrease between 1750 and 1899, followed by an increase in the two final periods. The remaining constructions come at extremely low frequencies, none of them going beyond a 5-token-benchmark.

Token frequencies are generally much higher when looking into the second set of construction contexts investigated, viz., all structures employing *one* as the first element, followed by different grammatical material such as nouns, prepositional phrases, modifiers, genitive-*s*, and coordinated structures (PP and N). This set also includes the construction context N + *one*, shown in (17), which is very rare.

(17) *Week* ***One*** (1971wesk_y8b)

We focus on construction contexts with a frequency beyond a twenty-token threshold; all other patterns are too rare to permit reliable conclusions and do not display any particularly interesting developmental behavior or relevant differences between the two varieties.

As Figure 7 illustrates, the three higher-frequency structures are plain *one*, as in (18), *one* + PP (19), and *one* + N (20), in ascending order.

(18) [. . .] *praying all the estate shall be reduced into* ***one****, and the warranty shall come* (1612cinq_l1b)

(19) *the question was* ***one*** *of public policy*; (1928gree_l7a)

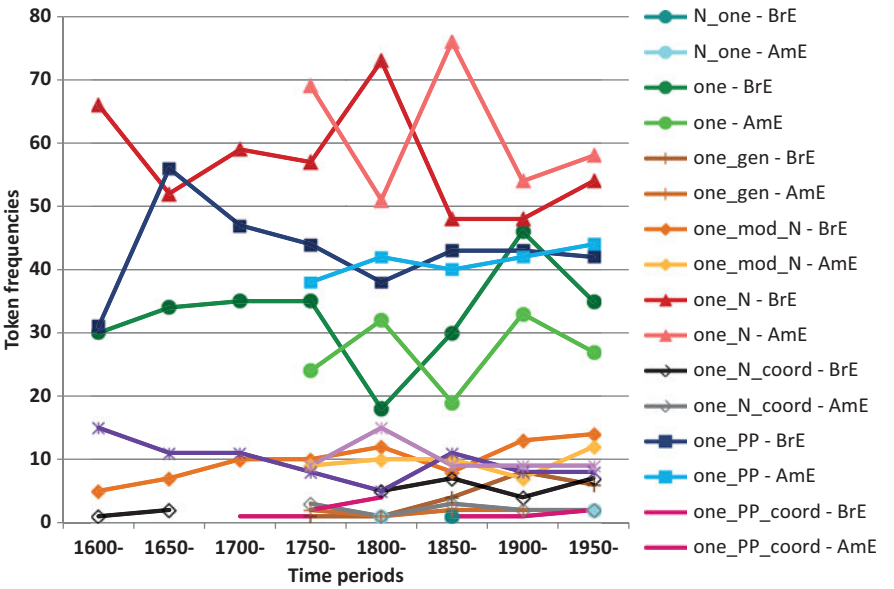


**Figure 7:** Construction context of *one* (constructions starting in *one* or N) by period and variety.

(20) [. . .] *which was, that at the End of* ***one*** *Journey, you had begun another*; (1653osbn_x2b)

Plain *one* as an NP undergoes a slight overall increase in both varieties, which is slightly stronger in BrE than in AmE. In both varieties *one* experiences a dip in this construction, which occurs slightly lagged, viz., in the 1800– period in BrE and in the 1850– period in AmE. Both varieties experience another slight decrease in the 1950– period.

*One* + PP, the second strongest construction context *one* occurs in and also a manifestation of *one* as an NP, shows a considerable increase during the 1600– to 1650– periods in BrE, followed by a continuous decrease until the 1800– period, and then experiences another slight increase, but token numbers remain comparatively stable. The latter also applies to *one* in this construction in AmE, with token frequencies being overall very similar to the frequencies observed in BrE.

The strongest construction context overall in which *one* occurs in both varieties under investigation is *one* + N, i.e., typically as a determiner, even though the construction appears to experience a slight general decrease in frequency,

accompanied by some probably erratic, but interestingly once more divergent developmental swings in both varieties (see the red lines with triangles in the 1800– and 1850– periods in Figure 7).

What we are witnessing here, thus, is an increased tendency to employ *one* either in nominal functions (as NP head in the prop-word pattern or as a full NP realization) or, as observed above, as a particular determiner. This seems in line with the developments in syntactic functions observed above, with prepositional complements and subjects in the lead, closely followed by direct objects – all of these being functions in which these construction types are likely realizations. It may be hypothesized that these are steps towards an increasing grammaticalization of *one*, at the expense of its original lexical meaning – but any case for such a hypothesis would need to be built more strongly and comprehensively.

## 4.4 Changes in referent semantics

When turning towards our last variable, the meaning of *one*, we see an even clearer difference between a few high-frequency variables and low frequency ones. Once more, we do not consider the latter ones in any detail. Some of the relatively rare meanings undergo slight increases in frequency (e.g., time), and others go even further down (e.g., non-human animate). There do not seem to be any conspicuous differences between the two varieties here.

As Figure 8 illustrates, three meanings occur at higher frequencies, viz. concrete inanimate referent, exemplified in (21), abstract referent (22), and human referent (23).

(21) [. . .] ***ships**, of which **one** seemed to a great **one*** (1654whit_j2b)

(22) [. . .] *old **method** of proceeding seems to me the most sensible **one*** (1825wood_s5b)

(23) [. . .] *& my **classmate** Packard; the only **one** now in town* (1787adaq_y4a)

After an increase in both varieties (in BrE between 1600 and 1799, in AmE between 1750 and 1899), *one* as denoting a concrete inanimate object has gradually dropped in frequency in more recent times; this decline is slightly stronger in BrE than in AmE. The use of *one* denoting abstract concepts, on the other hand, has clearly increased in the course of time. BrE starts out with thirty-seven tokens with abstract meaning in the 1600– period and shows a frequency of sixty-five

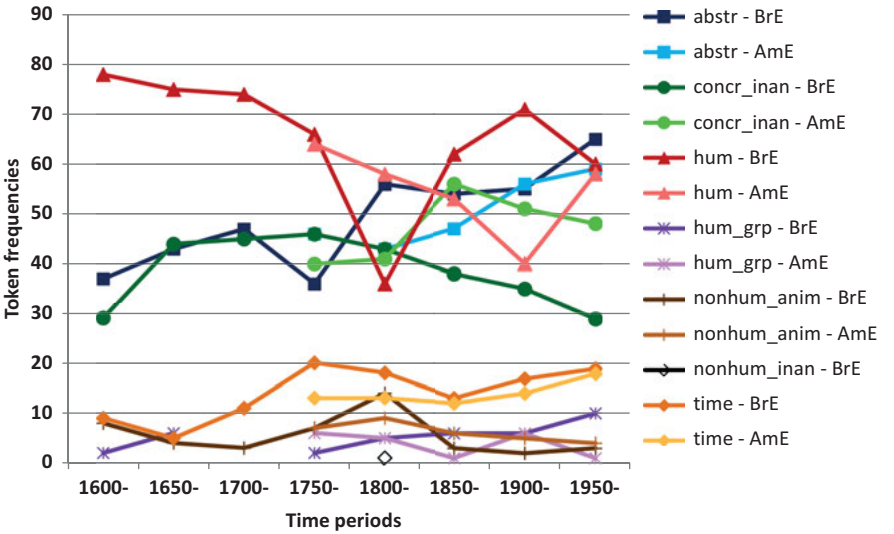


**Figure 8:** Meaning of *one* according to period and variety.

in the 1950– period (with a dip in the 1750– period). AmE *one* with abstract meaning increases more or less steadily between 1750 and 1999 (with a slight dip in the 1800– period which might easily be due to chance). As in many of the earlier observations, both varieties seem to follow the same general developmental trend. *One* with human referents slightly decreases in usage frequency, even though it is generally the most frequent meaning of *one*. In BrE it declines consistently for more than two centuries after 1600–, with a strong dip in the 1800– period, followed by a slight increase in the 1900– period. AmE *one* with human referents goes down from 1750– to the 1900– period and subsequently slightly increases again to roughly the same frequency as found in BrE, although it does not reach its initial value from the 1750– period again. In the 1900– period the two varieties once more show a short period of divergent development.

These overall changes may be regarded as broadly corresponding to the somewhat speculative observations made at the end of 4.3, on changing construction preferences. An increase in abstract nouns at the expense of human and especially non-human referents can be seen as corresponding to the growing tendency to use *one* for grammatical functions rather than for specific lexical references to objects or persons.

# 5 Discussion and conclusion

Our analysis has shown quite a number of detailed developmental trajectories of properties of *one* over a four hundred year period. Obviously, there are a number of minor ups and downs which are always to be found in such a quantitative analysis. Such constant frequency fluctuations characterize languages as complex systems (Kretzschmar 2015; Schneider, this volume) and may represent chance oscillations but also possibly lead to more firmly established systemic changes (like the ones hinted at at the end of 4.3 and 4.4). We have also identified some interesting developments over time and variety, such as the slow growth of *one* as a propword (and the Det + mod + *one* pattern – consistently in AmE but only up until the end of the nineteenth century in BrE) or, more recently, its increasing use as a pronoun or a "particular" determiner with a distinctly numerical meaning (which re-establishes a meaning close to the original, OE one). These are clearly instances of word-class boundaries being redefined. In the course of time *one* comes to be used slightly more frequently as a subject and less commonly as an object. Using *one* for abstract referents has gained and for concrete and (less so) human referents has lost ground. We have hypothesized that these may be indicators of the word moving along a grammaticalization path, adopting increasingly abstract and grammatical rather than concrete lexical referential functions.

As to the relationship between BrE and AmE, the two varieties appear to follow very similar general developmental routes (e.g., either a general decline or increase or just stable frequencies in particular for the low frequency types, structures, and meanings), although divergent tendencies can also be found at times. Occasionally (e.g., with respect to the *one* + N or plain *one* patterns, or in the "particular" determiner function, or as a prepositional complement) AmE seems to be trailing BrE by about half a century or less (and we hypothesized above, in 4.1 on word class changes, that this might represent an instance of colonial lag); in one instance (generic *one*) a BrE preference seems to have emerged in the recent past; and the Det + mod + *one* (propword) construction suggests divergence during the last century (continuously rising in AmE but decreasing in BrE). The ctree analysis suggests that variety differentiation can indeed be identified and plays a subtle but statistically significant role: a major time split before and after 1850 as well as another split between 1850– and the twentieth century in AmE only implies an ongoing, perhaps increasing tendency for both varieties to drift apart from each other in terms of subtle quantitative preferences, with AmE perhaps showing more dynamic developments. This may be seen as corresponding to a recent tendency to view differences between varieties not as mainly qualitative but rather as quantitative: while

earlier writings on BrE versus AmE tended to highlight select differences (e.g., *autumn – fall*, or nonrhotic vs. rhotic pronunciation; Strevens 1972), it is now understood that differences are characterized by a very large set of preferences, tendencies, and consistently different quantitative choices (Algeo 2006 for lexis; Rohdenburg and Schlüter 2009 for grammar; Leech et al. 2009).

There are obviously a number of interactions between the syntactic and semantic properties considered in this study. Some such interrelationships may reflect the fact that cognitive structures call for specific syntactic representations (e.g. agents are typically human and subjects; subject complements often give information about the subject's qualities and are hence represented by adjectives); others are more complex and may be viewed as precursors of schematic constructions in a Construction Grammar framework (Hilpert 2014). For example, *one* as a particular determiner shows a tendency to co-occur with prepositional complement function and abstract referents and to collocate with nouns like *side*, *point*, *case*, or *respect*. Hence, the prepositional phrases with *one* in examples (24) to (27) may be seen as increasingly firm phraseological expressions on their way to becoming schematic constructions.

(24) *Furniture has been crowded to* ***one*** *side by the courtroom* (1961simp_d8b)

(25) *the burning carbon arc lamps which at* ***one*** *point turn the actors into . . .* (1971tyna_y8b)

(26) *viz. that in* ***one*** *case irritation of these regions caused the eyes to be* [. . .] (1873ferr_s6b)

(27) *though in* ***one*** *respect this* [. . .] (1810enqu_s5a)

We find the syntactic descriptions and predictions of Rissanen's (1997) thorough study largely borne out: continuing EModE trajectories of change, certain pronominal and determinative uses of *one* are in use with gradually changing frequencies, with the overall pattern of change representing structured variability more than any systematic development strengthening one particular usage type. This needs to be attested more thoroughly for modern varieties of English, however. As the World Englishes examples quoted at the end of Section 1 promise, the story of *one* continues, transgressing new boundaries, and it may offer further insights into the nature of linguistic change.

While the boundary-crossing propensity of *one* is obvious after all the evidence outlined so far, one question that remains is how unique the form appears to be in that respect. Clearly, with this property in the history of English it seems

to be at least a very special and unusual case – not many other forms have adopted so many functions, crossed so many category boundaries; but there are parallels as well. First, there are two roughly comparable cases, words which also have been developing vibrantly and are used in widely different ways: *like* (compare, e.g., the functions as preposition in *like others*, as adjective in *like colors*, as verb in *I like you* and as a "new quotative" in *She's like " . . . "*; cf. D'Arcy 2017) and *so* (a premodifying adverb in *so happy*, a clause substitute in *I think so*, and, increasingly, a premodifier of verbs as in *I so wish*). Second, if the hypothesis spelled out above, i.e., that *one* gradually increases in its grammatical roles and decreases in its lexical-referential functions, it may be seen as tying in with general tendencies towards grammaticalization in English and other languages (Hopper and Traugott 1993).

# References

## Primary source

ARCHER = A Representative Corpus of Historical English Registers, 1600–1999. For information, see http://www.projects.alc.manchester.ac.uk/archer/

## Secondary sources

Algeo, John. 2006. *British or American English? A handbook of word and grammar patterns*. Cambridge: Cambridge University Press.

Baayen, R. Harald. 2013. Multivariate statistics. In Robert J. Podesva & Devyani Sharma (eds.), *Research methods in linguistics*, 337–372. Cambridge: Cambridge University Press.

Bao, Zhiming. 2015. *The making of Vernacular Singapore English. System, transfer and filter*. Cambridge: Cambridge University Press.

Bernaisch, Tobias, Stefan Th. Gries & Joybrato Mukherjee. 2014. The dative alternation in South Asian Englishes: modelling predictors and predicting prototypes. *English World-Wide* 35(1). 7–31.

Buschfeld, Sarah. 2013. *English in Cyprus or Cyprus English? An empirical investigation of variety status*. Amsterdam: John Benjamins.

D'Arcy, Alexandra. 2017. *Discourse-pragmatic variation in context: Eight hundred years of LIKE*. Amsterdam: John Benjamins.

Görlach, Manfred. 1987. Colonial lag? The alleged conservative character of American English and other "colonial" varieties. *English World-Wide* 8(1). 41–60.

Gries, Stefan T. 2009. *Statistics for linguistics with R: A practical introduction*. Berlin: Mouton de Gruyter.

Halliday, M. A. K. & Ruquaia Hasan. 1976. *Cohesion in English*. London: Longman.

Hilpert, Martin. 2014. *Construction Grammar and its application to English*. Edinburgh: Edinburgh University Press.

Hopper, Paul J. & Elizabeth Closs Traugott. 1993. *Grammaticalization*. Cambridge: Cambridge University Press.
Hothorn, Torsten, Kurth Hornik & Achim Zeileis. 2006. Unbiased recursive partitioning: A conditional inference framework. *Journal of Computational and Graphical Statistics* 15(3). 651–674.
Hundt. Marianne. 2009. Colonial lag, colonial innovation, or simply language change? In Günter Rohdenburg & Julia Schlüter (eds.), *One language, two grammars? Differences between British and American English*, 13–37. Cambridge: Cambridge University Press.
Jespersen, Otto. 1961. *A Modern English grammar on historical principles. Part II: Syntax (first volume)*. London: Allen & Unwin; Copenhagen: Munksgaard.
Kautzsch, Alexander. 2019. Namibian English on the web: Lexical and morpho-syntactic features in a Corpus of Namibian Online Newspapers (CNamON). In Alexandra Esimaje, Ulrike Gut & Bassey Antia (eds.), *Corpus linguistics and African Englishes*, 231–258. Amsterdam: John Benjamins.
Koch, Christopher, Claudia Lange & Sven Leuckert. 2016. "This hair-style called as 'duck tail'": The "intrusive *as*"-construction in South Asian varieties of English and Learner Englishes. *International Journal of Learner Corpus Research* 2(2). 151–176.
Kretzschmar, William R., Jr. 2015. *Language and complex systems*. Cambridge: Cambridge University Press.
Kruisinga, E. 1934. On some uses of *one*. *The Review of English Studies* 10(39). 331–337.
Lass, Roger. 1990. How to do things with junk: Exaptation in language evolution. *Journal of Linguistics* 26(1). 79–102.
Leech, Geoffrey, Marianne Hundt, Christian Mair & Nicolas Smith. 2009. *Change in contemporary English*. Cambridge: Cambridge University Press.
Los, Bettelou. 2015. *A historical syntax of English*. Edinburgh: Edinburgh University Press.
*Longman dictionary of contemporary English online*. <www.ldoceonline.com/dictionary/one> (Accessed Aug. 10, 2018).
Nevalainen, Terttu & Helena Raumolin-Brunberg. 2003. *Historical sociolinguistics*. Harlow: Pearson.
Raumolin-Brunberg, Helena. 1994. The development of the compound pronouns in *–body* and *–one* in Early Modern English. In Dieter Kastovsky (ed.), *Studies in Early Modern English*, 301–324. Berlin: Mouton de Gruyter.
Raumolin-Brunberg, Helena & Leena Kahlas-Tarkka. 1997. Indefinite pronouns with singular human reference. In Matti Rissanen, Merja Kytö & Kirsi Heikkonen (eds.), *Grammaticalization at work. Studies of long-term developments in English*, 17–85. Berlin: Mouton de Gruyter.
Raumolin-Brunberg, Helena & Arja Nurmi. 1997. Dummies on the move: Prop-*one* and affirmative *do* in the 17th century. In Terttu Nevalainen & Leena Kahlas-Tarkka (eds.), *To explain the present: Studies in the changing English language in honour of Matti Rissanen*, 395–417. Helsinki: Société Néophilologique.
R Development Core Team. 2014. R: A language and environment for statistical computing. R Foundation for Statistical Computing. <http://www.R-project.org/>. (Accessed Aug. 28, 2018.).
Rissanen, Matti. 1967. *The uses of* one *in Old and Early Middle English*. Helsinki: Société Néophilologique.

Rissanen, Matti. 1997. The pronominalization of *one*. In Matti Rissanen, Merja Kytö & Kirsi Heikkonen (eds.), *Grammaticalization at work. Studies of long-term developments in English*, 87–143. Berlin: Mouton de Gruyter.

Rissanen, Matti. 1999. Syntax. In Roger Lass (ed.), *The Cambridge history of the English language*, vol. III, *1476–1776*, 187–331. Cambridge: Cambridge University Press.

Rohdenburg, Günter & Julia Schlüter (eds.). 2009. *One language, two grammars? Differences between British and American English*. Cambridge: Cambridge University Press.

Rosenbach, Anette. 2008. Animacy and grammatical variation – findings from English genitive variation. *Lingua* 118(2). 151–171.

Schneider, Edgar W. 2007. *Postcolonial English. Varieties around the world*. Cambridge: Cambridge University Press.

Strevens, Peter. 1972. *British and American English*. London: Macmillan.

Tagliamonte, Sali & R. Harald Baayen. 2012. Models, forests, and trees of York English: Variation as a case study for statistical practice. *Language Variation and Change* 24(2). 135–178.

Erik Smitterberg

# 6 Non-correlative commas between subjects and verbs in Early and Late Modern English sermons and scientific texts

## 1 Introduction

### 1.1 Aim and scope

In this study, I examine the distribution of non-correlative commas between subjects and verbs in Early and Late Modern English. Example (1) illustrates this usage, which should be contrasted with (2), an instance of a correlative comma.[1]

(1) [. . .] *but that which inflam'd their* ***Zeal, was*** *Traditions of their own, the Commandments of Men, and not of God.* (Sermons, Birch, 1650–1699)

(2) *Thirdly, Niter, which is made by the affusion of an Acid Spirit upon an* ***Alcali, may*** *be almost totally distill'd into an Acid Spirit,* [. . .] (Science, Coxe, 1650–1699)

A non-correlative comma between a subject and a verb marks a linguistic boundary – prosodic and/or syntactic – in this place independently of previous punctuation, as in (1). A correlative comma, in contrast, signals the end of a unit that is included within the subject and whose beginning is also typically marked by punctuation; for instance, the correlative comma in (2) marks the end of a non-restrictive relative clause inside the subject noun phrase (see Quirk et al. 1985: 1610; Smitterberg 2013: 367–368). In the interest of brevity, I shall refer to commas like the one in example (1) as *non-correlative S|V commas*.[2]

**1** To facilitate identification of the relevant comma in corpus examples, the comma as well as the immediately preceding and following words appears in bold face. In examples that do not feature a comma, the linguistically relevant components of the clause have been highlighted.
**2** This term also subsumes the rare cases where another punctuation mark, e.g., a dash or a semi-colon, is used instead of a comma; see (8) in section 2.2 for an example.

**Erik Smitterberg,** Department of English, Uppsala University, Sweden,
e-mail: erik.smitterberg@engelska.uu.se

https://doi.org/10.1515/9783110643282-007

Correlative and non-correlative commas also differ in acceptability: only correlative punctuation is typically allowed between the subject and the verb in Present-day English. However, this has not always been the case; as Crystal (2015: 243) notes, the non-use of commas to mark boundaries between clause elements "is one of the big differences between modern punctuation and earlier practice". Salmon ([1962] 1998) argues that non-correlative S|V commas were part of a recognized system of punctuation in early seventeenth-century English. In addition, while Quirk et al. (1985: 1606) claim that non-correlative punctuation marks between subjects and verbs have been proscribed "[s]ince the early nineteenth century", the results in Smitterberg (2013) indicate that non-correlative S|V commas were not uniformly proscribed in Late Modern English grammars, usage guides, etc., and that they still occurred in nineteenth-century newspaper texts. To understand how and why the current proscription was established, we need more research on the incidence and status of non-correlative S|V commas in English texts from the seventeenth to the nineteenth centuries.

The aim of the present diachronic study is to examine the use of non-correlative S|V commas in sermons and scientific texts from A Representative Corpus of Historical English Registers (ARCHER). First, in a quantitative analysis, seventeenth-century and nineteenth-century texts from these genres are examined with regard to the frequency of non-correlative S|V commas as well as to two factors that may be hypothesized to influence that frequency, viz., genre and subject length. In this part of the study, I apply a variationist framework and contrast clauses with and without non-correlative S|V commas. Secondly, a qualitative analysis that focusses on clauses with non-correlative S|V commas aims at providing clues to the function of such punctuation in the material. This analysis takes as its point of departure the distinction between rhetorical punctuation, which connects writing to speech, and grammatical punctuation, which signals syntactic relationships in a written text (Baron 2001: 22–23; Schou 2007: 195). It is hoped that the inclusion of one speech-related genre (Sermons) and one written expository genre (Science) will shed light on the relationship between these two main functions that punctuation has been argued to fill.

## 1.2 Background

Smitterberg (2013) is one of the few existing quantitative analyses of non-correlative S|V commas. This study also includes a brief survey of Late Modern English normative works. While authors such as Lowth (1762: 161) argue that non-correlative S|V commas are not acceptable, this general statement is qualified in several sources, e.g., Robertson (1785: 72–73), Lennie (1864: 160), and

Brown (1869: 774–775), which mention linguistic contexts where such commas should be allowed. These contexts include subjects with long postmodifiers, clausal subjects, and subjects that consist of lists of conjoined units; clauses where the verb BE is followed by an adjective phrase or a *to*-infinitive clause are also mentioned (Smitterberg 2013: 369–370; see also section 3.2).

As regards the quantitative distribution of non-correlative S|V commas, Smitterberg (2013) focusses on nineteenth-century newspaper language. While such commas become increasingly rare during the 1800s, they are not completely absent even at the end of the century. These results are corroborated by an analysis of non-correlative S|V commas in nineteenth-century private letters and scientific texts (Smitterberg 2020). While such commas are far more frequent in scientific language and near-absent in private letters from the last three decades of the 1800s, the feature is attested in both genres at the beginning as well as the end of the century. A functional analysis of the commas in the material indicates that their occurrence is increasingly restricted to a shrinking set of linguistic contexts, e.g., before main-verb BE and after subjects that comprise lists.

While there are thus valuable indications of language change in previous work, less is known about the development of non-correlative S|V commas before 1800. In particular, we need to examine printed texts which pre-date the codification of English that characterizes the second half of the eighteenth century (Auer 2012: 940) as a complement to data on this linguistic feature after it had begun to be proscribed. The present study is a first step towards filling this gap in research.

# 2 Method

## 2.1 Material

The primary material used for this study was taken from ARCHER. Punctuation is a linguistic feature that may potentially not be perfectly preserved when texts undergo transformations between formats (e.g., from a manuscript to a printed edition or from a printed text to a corpus file). However, basing the study on original documents would have been impossible owing to limitations of time, and given that ARCHER has been subject to several rounds of revision and expansion since its original compilation in the 1990s, it is hoped that it provides a comparatively reliable window on Modern English punctuation.[3] This

**3** See the entry for ARCHER on the Corpus Resource Database (CoRD) for information on the history and development of the corpus (http://www.helsinki.fi/varieng/CoRD/index.html).

assumption of reliability is strengthened by the results, which imply the existence of clear trends over time in the dataset. ARCHER has also formed the basis for previous research on punctuation; Denison and Hundt (2013: 149) look at relative clauses in ARCHER and report that punctuation "turned out to be somewhat more consistent than [they] expected" in terms of separating restrictive and non-restrictive clauses.

In order to maximize the diachronic spread of the results while focussing on periods during which non-correlative S|V commas can be expected to occur in texts, I drew on the periods 1650–1699 and 1850–1899 for data. The analysis covers two genres, viz., Science and Sermons. For the Sermons genre, all ten British texts available from the two relevant periods were used. In the Science genre, five of the ten British texts from each period were randomly selected for this exploratory study.[4] This restriction was necessary owing to limitations of time: the manual identification of relevant tokens is very time-consuming, and it was methodologically preferable to include more Sermons texts than Science texts, as the Science genre yielded a larger number of relevant tokens per text (see section 2.2).

Science and Sermons were selected based on their genre characteristics. The Science genre is one of the most "literate" genres in Biber and Finegan's (1997) factor-score analysis of the corpus and comprises formal, expository prose written to be read silently. Sermons are an interesting contrast to Science since they are what Culpeper and Kytö (2010: 17) call "speech-purposed", i.e., written to be spoken. As noted by Rodríguez Puente (2014: 179), this – in combination with their relative formality – may cause them to contain a mixture of oral and literate features. The juxtaposition of speech-related and non-speech-related material is of particular interest in the functional analysis of non-correlative S|V commas. Scientific texts are likely to feature complex syntax, and grammatically motivated commas may then be used in order to clarify sentence structure. In contrast, a Sermons text may contain punctuation that mirrors prosodic features of the speech event when the sermon is delivered, which increases the likelihood that rhetorically motivated punctuation will occur. Both genres can thus be expected to feature a comparatively high frequency of non-correlative S|V commas, but for different reasons; in addition, it is an interesting research question which of these genre characteristics results in higher proportions of non-correlative S|V commas.

4 For the period 1650–1699, the texts used were 1675hook, 1675ray-, 1676boyl, 1676coxe, and 1676newt; for the period 1850–1899, I drew for data on 1873ferr, 1874gunt, 1874klei, 1874lank, and 1874lass.

## 2.2 Data

Non-correlative S|V commas may occur in both main and subordinate clauses. However, several features of many subordinate clauses may be hypothesized to influence the distribution of non-correlative S|V commas; these include, but are not restricted to, the presence of a subordinator, the absence of a finite verb, and restrictions on word order in clauses with relative markers and WH-elements.[5] In order to keep the scope of the study manageable and to avoid genre differences in the distribution of main and subordinate clauses influencing the results, the present study focusses on main clauses. The first step in the selection of data was thus to identify all main clauses in the material through manual analysis.[6] These 2,845 main clauses were then examined one by one so that irrelevant tokens could be excluded from the counts.

My aim was to include in the counts only tokens where a non-correlative comma between subject and verb was either present or possible, given the linguistic context. This led to the exclusion of a large number of main clauses: imperative clauses, which typically lack a subject, as in (3); clauses where all or part of the subject followed the finite verb, as in (4); and clauses that already had correlative punctuation between subject and verb, such as (5).

(3) ***Help*** *us afresh to put up the standard!* (Sermons, Booth, 1850–1899)

(4) *Well, then,* ***might the inspired Psalmist*** *break forth in wonder, at the contrast between what man is intended for, and what he effects:* [. . .] (Sermons, Wilberforce, 1850–1899)

(5) *The Motion of the two disciples, and the design of our Adversaires of* ***Rome, were*** *both of them for destruction, and for destruction by that merciless*

**5** Main clauses may of course also contain linguistic material that precedes the subject, notably co-ordinators and adverbials. In Smitterberg (2020), I demonstrate that non-clause-initial subjects of main clauses are less likely than clause-initial subjects to be followed by non-correlative S|V commas in nineteenth-century English. However, subordinate clauses still differ from main clauses in this respect: finite subordinate clauses regularly contain explicit clause links (subordinators or WH-words) preceding the subject, and non-finite subordinate clauses lack finite verbs, in which case the subject is also often absent (Biber et al. 1999: 193, 198). These differences justify treating subordinate clauses separately.

**6** Clauses beginning with the conjunction *for* were counted as main clauses in this analysis, as such clauses were frequently used as full sentences in my material. See Quirk et al. (1985: 920–928) for an analysis where *for*-clauses are instead treated as subordinate, which may be more appropriate in Present-day English.

*element of Fire, which acts as they do, ad extremum virium.* (Sermons, Lamplugh, 1650–1699)

In addition, I excluded clauses where the subject and the finite verb were separated by other linguistic units, such as the adverb *generally* in (6), since intervening units may affect punctuation.

(6) *Throughout the several figures the same* ***letters generally indicate*** *the same parts of the machine.* (Science, Lassell, 1850–1899)

A small number of tokens that were quotations from other sources were also excluded; this category comprises mainly quotations from the Bible, which are quite frequent in Sermons texts. As many of these quotes are from the King James Bible and thus represent early-seventeenth-century usage from another genre, it was felt that including them would have reduced the representativity of the dataset.

Finally, tokens where a comma between the subject and the finite verb could be interpreted either as correlative or non-correlative were excluded, such as (7).

(7) *But, whether it is the Volatil or Acid Salt, which combines with the Oyl or* ***Sulphur, is*** *now the subject of our Inquiry*. (Science, Coxe, 1650–1699)

On the one hand, the comma after *Sulphur* seems non-correlative, given that the *which*-clause, which is part of an *it*-cleft, would not normally be enclosed in punctuation in Present-day English. On the other hand, as there is already a comma immediately preceding that *which*-clause, a correlative reading is also possible. After these exclusions, the dataset comprised 837 main clauses.

However, I also applied one data-driven restriction to the remaining tokens. This restriction was based on the fact that non-correlative S|V commas do not appear to be a genuine option after all types of subject. In particular, short subjects such as noun phrases consisting of a single pronoun are never followed by commas in the texts examined. The shortest subject preceding non-correlative punctuation in my material, exemplified in (8), comprises four words. (As the example shows, other punctuation marks are occasionally used instead of a comma in this position. As made clear in note 2, such tokens were also included.)

(8) [. . .] *And, To reject their* ***Authority; was*** *to Resist the Ordinance of God, who had said, The Priests Lips shall preserve Knowledge, and the People shall seek it at their Mouths.* (Sermons, Birch, 1650–1699)

Owing to this distribution of data, I excluded main clauses with subjects of less than four words in length from the counts. Although this cut-off point is somewhat imprecise in that it neglects other potentially important characteristics of the subject (e.g., length in syllables or syntactic make-up), it is important not least from a genre perspective: genres where a large proportion of main clauses have simple subjects would have very low proportions of non-correlative S|V commas if such clauses, where commas are not possible, were included in the counts. It is also a fundamental component of variationist studies that "[c]ontexts that do not vary but are categorically encoded with one or other variant are not included in the analysis of variation" (Tagliamonte 2012: 10).

After the restriction on subject length had also been implemented, the database contained a total of 338 relevant tokens, corresponding to 11.9% of all main clauses in the material. The percentage of relevant tokens is lower in Sermons (175 out of 2,088, or 8.4%) than in Science (163 out of 757, or 21.5%), probably due to the higher frequency of short pronominal subjects and imperative clauses in the former genre. The quantitative and qualitative analyses in section 3 are based on these 338 tokens.

# 3 Results

## 3.1 Quantitative analyses

I begin by accounting for overall trends in the data. Table 1 presents the distribution of main clauses with and without non-correlative commas in the two genres examined.

**Table 1:** The proportion of non-correlative S|V commas by period and genre.

| Period | Science | | | Sermons | | | Science + Sermons | | |
|---|---|---|---|---|---|---|---|---|---|
| | Comma | No comma | Total | Comma | No comma | Total | Comma | No comma | Total |
| 1650–1699 | 7 (15.6%) | 38 (84.4%) | 45 | 22 (26.2%) | 62 (73.8%) | 84 | 29 (22.5%) | 100 (77.5%) | 129 |
| 1850–1899 | 2 (1.7%) | 116 (98.3%) | 118 | 12 (13.2%) | 79 (86.8%) | 91 | 14 (6.7%) | 195 (93.3%) | 209 |
| Total | 9 (5.5%) | 154 (94.5%) | 163 | 34 (19.4%) | 141 (80.6%) | 175 | 43 (12.7%) | 295 (87.3%) | 338 |

It is clear from Table 1 that there is an overall decrease in the proportion of non-correlative S|V commas, which tallies with previous work. The period difference is statistically significant according to the chi-square test (d.f. = 1; $\chi^2 = 17.9$; $p < 0.001$). The decrease is also present in each genre examined, though the period difference in Sermons only borders on significance if allowance is made for several tests being carried out on the same dataset (d.f. = 1; $\chi^2 = 4.72$; $p = 0.030$), and the difference in Science does not bear testing owing to one expected frequency falling below 5.

Science and Sermons also turn out to differ from each other, the similar diachronic trend evinced in the two genres notwithstanding. In the overall figures, non-correlative S|V commas are nearly four times more frequent in Sermons, a difference that reaches statistical significance (d.f. = 1; $\chi^2 = 14.7$; $p < 0.001$). The results also indicate that the difference increases over time, chiefly because non-correlative S|V commas are very rare in late-nineteenth-century Science texts. While the genre difference for the seventeenth-century data is not significant (d.f. = 1; $\chi^2 = 1.90$; $p = 0.168$), Sermons are significantly more likely to feature non-correlative S|V commas in the latter half of the 1800s (d.f. = 1; $\chi^2 = 10.9$; $p = 0.001$).

The consideration of two extralinguistic features – time and genre – thus demonstrates that explicitly marking this prosodic and/or syntactic boundary by means of punctuation became a less frequent choice during the Modern English period, but also that this boundary was more frequently signalled in sermons than in scientific texts. In section 3.2, I shall link these statistical tendencies to possible differences in the function fulfilled by such boundary-marking in the two genres. First, however, the influence of a linguistic feature on punctuation will be considered, viz., subject length.

I have shown elsewhere (Smitterberg 2013, 2020) that subject length influences non-correlative S|V comma usage beyond the cut-off point at four words: longer subjects are more likely to be followed by commas. To test whether this also holds for the two genres considered here, I divided all relevant tokens into four classes with respect to subject length: short subjects (4 words), medium-length subjects (5–6 words), long subjects (7–9 words), and very long subjects (10 words or more). The cut-off points were decided on mostly on the basis of the distribution of the data: it was considered important to categorize subjects of unusual length separately as they were hypothesized to behave differently; at the same time, each category needed to have a sufficiently large number of members for the counts to be comparatively robust. The results of this categorization are given in Table 2.

The effect of subject length is clear from Table 2: Science does not evince any non-correlative S|V commas after short and medium-length subjects; in Sermons, such commas are attested in all length categories, but both genres feature increasing

**Table 2:** The proportion of non-correlative S|V commas by subject length and genre.

| Subject | Science | | | Sermons | | | Science + Sermons | | |
|---|---|---|---|---|---|---|---|---|---|
| | Comma | No comma | Total | Comma | No comma | Total | Comma | No comma | Total |
| Short | – | 47 (100%) | 47 | 1 (2.0%) | 50 (98.0%) | 51 | 1 (1.0%) | 97 (99.0%) | 98 |
| Medium | – | 52 (100%) | 52 | 5 (8.8%) | 52 (91.2%) | 57 | 5 (4.6%) | 104 (95.4%) | 109 |
| Long | 5 (12.5%) | 35 (87.5%) | 40 | 9 (25.7%) | 26 (74.3%) | 35 | 14 (18.7%) | 61 (81.3%) | 75 |
| Very long | 4 (16.7%) | 20 (83.3%) | 24 | 19 (59.4%) | 13 (40.6%) | 32 | 23 (41.1%) | 33 (58.9%) | 56 |
| Total | 9 (5.5%) | 154 (94.5%) | 163 | 34 (19.4%) | 141 (80.6%) | 175 | 43 (12.7%) | 295 (87.3%) | 338 |

proportions of comma use with longer subjects. The Science data does not bear testing owing to low expected frequencies, but the differences in the Sermons dataset are statistically significant (d.f. = 3; $\chi^2 = 47.6$; $p < 0.001$). In fact, with very long subjects, non-correlative S|V commas are actually the norm in the Sermons data, where they account for nearly 60% of tokens.

It is also noteworthy that the difference between Science and Sermons shown in Table 1 cannot be due simply to average subject length in these genres. Given that the overall percentage of non-correlative S|V commas is proportional to subject length, a genre that features a high proportion of such commas might do so simply because main-clause subjects are longer than average in that genre. But since Sermons feature higher proportions of commas than Science does in every length class (see Table 2), differences in average subject length cannot account for the genre differences in Table 1. In other words, while the tendency for writers to mark a prosodic and/or syntactic boundary between subject and verb by means of punctuation clearly increases with the length of the subject, genre-specific differences in this regard still remain.

## 3.2 Contextual analysis

In this section, I focus on the function and linguistic environment of the forty-three tokens of non-correlative S|V commas in my material. In previous research, punctuation is often argued to fill two main functions, which I refer to as *rhetorical*

and *grammatical*.[7] The function of rhetorical punctuation is to connect writing to a – real or imagined – speech event by signalling, for instance, prosodic breaks. Grammatical punctuation, in contrast, marks syntactic relationships between linguistic units in a written text regardless of what the spoken version sounded – or would have sounded – like (see Baron 2001: 22–23; Schou 2007: 195). The distinction between these functions is of considerable importance in textual analysis, as it relates to the degree of independence a written text has from spoken language.

Some sources suggest that rhetorical punctuation became rarer during the Modern English period (e.g., Ronberg 1995: 55; Schou 2007: 198; but cf. Salmon 1988: 295); moreover, Quirk et al. (1985: 1611) claim that punctuation practice in late-twentieth-century English was "governed primarily by grammatical considerations". However, while Quirk et al. (1985: 1606) appear to argue that non-correlative S|V commas mainly have rhetorical functions, in my analysis they can also fill grammatical functions. On the one hand, the comma may mirror a prosodic break in speech. As Quirk et al. (1985: 1606) note, long subjects are often followed by "a prosodic break, usually realized by the end of a tone unit, often by a pause as well". It is this spoken boundary that a comma with a rhetorical function corresponds to. On the other hand, regardless of whether a prosodic break would (be likely to) occur in speech, the comma also marks the boundary that separates two units of grammar (i.e., the clause elements subject and verb).

As the non-correlative S|V comma may potentially fill both of the main functions of punctuation, the question arises whether the genre differences attested in section 3.1 may in part be due to the functions of the commas in the material. Specifically, it may be the case that sermons, which are speech-purposed texts "designed to produce monologue" (Culpeper and Kytö 2010: 17), feature more rhetorical punctuation than scientific texts do owing to the relatively close relationship between a written sermon and its spoken delivery. In contrast, scientific texts often contain complex linguistic structures, which may increase the likelihood of a grammatical S|V comma being used to signal the syntactic boundary between subject and verb. This factor may in turn account wholly or partly for the greater tendency towards comma use in Sermons: for various reasons, rhetorical commas in Sermons might outnumber grammatical commas in Science (see section 2.1).

7 While these functions seem fairly well established in the scholarly literature, other names are sometimes used for them; for instance, grammatical punctuation may instead be called *syntactic*. In the present study, I use *syntactic* and *prosodic* to refer to the boundary marked by a comma, while the punctuation itself is labelled *grammatical* or *rhetorical*, respectively.

In order to investigate the function of the non-correlative S|V commas in my material, I carried out a qualitative analysis of the linguistic environment of these commas. The departure point of this analysis was the linguistic features that, according to some Late Modern English normative sources, may license a comma; however, I also drew on my previous work and on insights from a corpus-driven examination of the relevant tokens (see Smitterberg [2013, 2020] for details). To the extent that the presence of one or several of those features may constitute a sufficient criterion for either a grammatical or a rhetorical use of a comma, the relative distribution of the two functions in Science and Sermons can be compared.

However, there are several problems associated with identifying grammatical and rhetorical punctuation in historical texts. First, rhetorical punctuation is related to prosodic breaks, and it cannot be determined with certainty where such a break would have occurred in past speech. Secondly, the two reasons for marking a boundary by means of punctuation may of course co-occur in the same main clause, as I have tried to indicate with the phrase "prosodic and/or syntactic boundary" above. That is, a subject that is both long and complex may well cause the writer to use a comma both to mirror a prosodic break before the verb phrase and to clarify the syntactic structure of the clause. The contextual analysis is thus necessarily tentative and inherently subjective, but it may nevertheless contribute to a better understanding of why and how Modern English writers used punctuation. In particular, if commas are used after subjects that appear not to be syntactically complex despite their length, that may potentially identify rhetorical punctuation. In contrast, as most syntactically complex subjects are also long enough to be likely to be followed by a prosodic break in speech, rhetorical functions of commas can rarely if ever be ruled out.

As a first step, contexts where a grammatical non-correlative S|V comma may have been thought convenient in order to clarify sentence structure were identified. Clausal subjects, as in example (9), belong to this category: if a subject clause with its own verb phrase occurs inside the main clause, writers may have felt the need to indicate the boundary of the smaller verbal unit.

(9) *Whatever bright cloud of hope and prophecy had formerly floated about his* ***cradle, has*** *long been scattered and forgotten;* [. . .] (Sermons, Beard, 1850–1899)

Similarly, a subject that consists of a noun phrase may contain a postmodifier that comprises or includes clausal material, such as the appositive clause in (10). Such constructions may also make it desirable to signal the grammatical boundary between the completed subject structure and the predicate.

(10) *The fact that the development of the eggs of this mollusk takes place within a pair of brood-cavities formed at the root of the inner gill-lamella on each* ***side, enables*** *the observer very readily to obtain embryos in different stages of development*. (Science, Lankester, 1850–1899)

There are also a few "pseudo-correlative" cases such as (11), where a non-correlative S|V comma looks like a closing correlative comma without an opening comma.

(11) *But especially our Country men who are satisfied in the experience of* ***it, should*** *seriously bethink themselves, If there may not be an easier and cheaper way of Conveyance, for a greater quantity thereof to be brought up into the middle of the Country*. (Science, Hooke, 1650–1699)

The comma in (11) bears a surface similarity to that in (10): the subject ends in a clausal postmodifier – an appositive clause in (10) and a relative clause in (11) – and a comma precedes the verb phrase. However, if the relative clause in (11) is analysed as non-restrictive, which is a possible reading, a comma would be expected both before *who* and after *it*;[8] in contrast, a comma would be unlikely between *fact* and *that* in (10), where there is no clear prosodic boundary. It is thus difficult to know whether the comma in (11) is a non-correlative S|V comma that fills the grammatical function of signalling the end of the subject (possibly in conjunction with the rhetorical function of mirroring a pause), or whether it is a correlative comma that marks the end of a non-restrictive relative clause within the subject, whose beginning for some reason lacks parallel marking.

Another context discussed in contemporary sources may indicate either a grammatical or a rhetorical function. Several sources mention that the verb BE followed by a complement may be preceded by a non-correlative comma, as in (12). I included in this category all cases where a form of BE was the only part of the verb phrase, regardless of whether it was followed by an adjective phrase, a *to*-infinitive clause, etc. (cf. section 1.2).

(12) *Yea garden herbs, and fruits, in those* ***places, are*** *more, and those better in their kind*. (Science, Hooke, 1650–1699)

**8** Denison and Hundt's (2013: 149) analysis of ARCHER reveals that "[n]on-restrictive relative clauses have always strongly tended to be preceded by some punctuation mark". Denison and Hundt (2013: 157–158) discuss (11) as an ambiguous token.

Without further clarification, it is difficult to know whether the motivation for such a comma is chiefly grammatical or rhetorical. The difficulty is compounded by the fact that all tokens in the material where BE makes up the verb phrase except (12) meet one more criterion that makes a non-correlative S|V comma likely, i.e., they have subjects that are clauses, contain clausal material in a postmodifier, or are lists (see below). This makes it impossible to know which criterion may have "triggered" the comma (and punctuation may of course also be due to the two criteria combined).

A comma was also sometimes prescribed to signal the end of a subject that lists more than two items, like that in (13).

(13) *What depth of tenderness, what steadiness of judgment, what a majestic and yet winning purity, what a faculty of self-devotion (not yet too hardly tried), what a simple intensity of* ***devoutness, must*** *have watched and helped the child, as he grew and blossomed into man!* (Sermons, Beard, 1850–1899)

These commas are perhaps those in the data that are the least likely to fulfil grammatical functions. Despite the length of the subject, the structure of the main clause is arguably not difficult to parse even without a comma marking the subject off from the remainder of the clause, given that there are no clausal units before the verb; in contrast, a prosodic break between the last item in the list and the following verb phrase would be very likely, especially perhaps in a sermon, where pauses at significant points in the delivery may make the sermon easier to take in and emphasize its rhetorical force.

Using the environments exemplified above as criteria for grammatical or rhetorical punctuation is a difficult endeavour. As mentioned above, rhetorical functions can rarely be ruled out, and subjects followed by a verb phrase consisting only of a form of BE resist easy functional classification. In addition, the environments are of course not mutually exclusive, and several of them often co-occur in individual tokens. In spite of this, and although raw frequencies are low, the results may help to reveal what kinds of linguistic and genre-related constraints affect non-correlative S|V comma use across the Modern English period.

Like tokens that occur after lists as subjects, tokens that met none of the criteria listed above were hypothesized to be more likely to fill a rhetorical function, where the comma chiefly mirrors what would have been a spoken prosodic break after a long subject, but where the structure itself is not difficult to parse. However, in the present dataset, this turned out to concern only one single token, viz. (14), and even there the syntax can be argued to be complicated, for a reason not mentioned in the contemporary sources I have consulted.

(14) [. . .] *and, on the other hand, many actions externally mean and* ***lowly, may****, because of the state of his heart who does them, be truly exalted and honorable.* (Sermons, Caird, 1850–1899)

The adjective phrase in (14) is postposed, a marked usage which may have occasioned a comma to clarify sentence structure; alternatively, the comma may mirror a spoken prosodic boundary. Although the only relevant token for the present study is from the 1800s, Smitterberg (2020) demonstrates that commas that meet none of the criteria listed above became rarer across the nineteenth century, which may indicate that rhetorical punctuation became less frequent over time.

One context whose frequency decreases in diachrony is that where the comma is pseudo-correlative, as in (11), which tallies with the results in Smitterberg (2020). While there are five relevant tokens in the 1600s, (15) is the only potential token in the nineteenth-century sample.

(15) *This being engaged with the wheel of 77* ***teeth, causes*** *the speculum to revolve on its axis,* [. . .] (Science, Lassell, 1850–1899)

In (15), the comma is pseudo-correlative if *this* is analysed as the head of a subject noun phrase in which *being . . . teeth* is a non-restrictive participle-clause postmodifier;[9] but an alternative analysis in which all of *this . . . teeth* is a non-finite subject clause would make this a clausal subject instead. Regardless of how (15) is analysed, the dearth of late-nineteenth-century tokens of this feature both in this material and in the texts examined in Smitterberg (2020) is most likely indicative of standardization: it became the norm to signal the inclusion of some units, e.g., non-restrictive postmodifiers, within larger units by means of commas on either side of the unit.

In contrast, tokens with clausal subjects or clausal postmodifiers occur in both periods and both genres, though tokens where the entire subject is a clause are quite rare in the late-nineteenth-century material both here and in the material used in Smitterberg (2020). Raw frequencies are of course too low

**9** If *being . . . teeth* is analysed as a participle clause, there is also a possibility of an adverbial interpretation (= *Being engaged with the wheel of 77 teeth, this causes . . .*) (Quirk et al. 1985: 1270–1271). That underlying structure would make the token irrelevant, as the adverbial clause then separates the subject and the verb. Example (15) is the only token in the dataset that features such ambiguity; as the token is relevant in two out of three interpretations – viz. the gerundial-clause reading and the *this* + postmodifier reading – it was nevertheless kept in the dataset.

to allow safe conclusions, especially as most of these tokens also meet other criteria, but it seems that clausal material in the subject is a relatively resilient feature that co-occurs with non-correlative S|V commas. This may indicate the persistence of grammatical punctuation in the material, though, as mentioned above, the possibility of rhetorical punctuation after these typically long subjects cannot be ruled out either.

The association with grammatical punctuation is potentially strengthened by the relative prevalence of non-correlative S|V commas after subjects incorporating clausal material in the non-speech-related Science genre. With the exception of one seventeenth-century token, all relevant commas in Science follow subjects that comprise or contain clauses. However, 27 of the non-correlative S|V commas in Sermons also meet this criterion; it is thus not exclusive to scientific texts.

In contrast, a few contexts are strongly associated with Sermons and may thus help to explain the genre differences. Most importantly, lists as subjects, as in (13) above and as in (16), occur exclusively in Sermons.

(16) [. . .] *the Abrahamic covenant, the separation of Israel, all the rites and all the* ***prophecies, are*** *but the unfoldings of its precious meaning.* (Sermons, Punshon, 1850–1899)

As mentioned above, the function of these non-correlative S|V commas seems to be primarily rhetorical. It is easy to imagine a pause for effect being made after each item in the list in a corresponding speech event, which results in a final comma that also marks the prosodic boundary between subject and verb. In contrast, there seems to be less need to clarify the structure of the written sentence itself with a non-correlative S|V comma: notably, there is no verb phrase in any of the conjoins making up the subject that could be interpreted as a main-clause verb. Three tokens of commas after lists were also found in Smitterberg's (2020) examination of nineteenth-century private letters and scientific texts, which may indicate that commas in this position can fulfil rhetorical functions even in non-speech-related genres such as academic writing.

Secondly, tokens where the non-correlative S|V comma is followed by a verb phrase that consists only of a form of BE, as in (17), occur in both genres in the period 1650–1699, but only in Sermons in the nineteenth-century data.

(17) *For all those who share in Adam's* ***nature, are*** *partakers also in the fellowship of his sin.* (Sermons, Wilberforce, 1850–1899)

Although this interpretation is speculative, a rhetorical function can perhaps be inferred here as well; for instance, a brief break after *nature* may help to emphasize the parallelism between Adam's nature and his sin. However, the relative clause *who . . . nature* also contains verbal material, which may complicate the syntactic parsing of the sentence, so a grammatical function is a clear possibility as well.

The last context that is characteristic of Sermons (there are two seventeenth-century examples from Science, but none from the 1800s) is what I will call "semi-clausal" structures. Unlike the preceding contexts, this environment emerged as a significant category during the examination of the corpus data. In semi-clausal tokens, the subject is a noun phrase consisting of a pronoun with generic meaning postmodified by an adnominal relative clause, as in (18) and (19).

(18) *They who have with the greatest judgement and care searched into the nature and first principles of humane* ***Societies, have*** *all agreed that the chief end and design of men in joyning together was, for the mutual benefit and advantage of each other*, [. . .] (Sermons, Stillingfleet, 1650–1699)

(19) *That which was restored by the Sanctions of* ***Christianity, was*** *this very Law*, [. . .] (Sermons, Birch, 1650–1699)

Given that most of the meaning expressed by the subject is contained in the relative clause, tokens such as (18) and (19) are quite similar to nominal relative clauses (cf. *Whoever has with . . .*; *What was restored . . .*). It is difficult to say whether the commas are rhetorically or grammatically motivated: on the one hand, a prosodic break after the subject for rhetorical effect seems likely; on the other hand, the subject is syntactically complex and includes a verb phrase, which may trigger grammatical punctuation to indicate the boundary between subject and verb.

In sum, the contextual analysis has demonstrated that virtually all tokens in the dataset meet at least one criterion that has either been listed in Late Modern English sources or been identified in my previous work (Smitterberg 2013, 2020) as encouraging the use of non-correlative S|V commas. The dearth of tokens that meet no criteria indicates that the occurrence of non-correlative S|V commas was restricted to a fairly small set of possible contexts. At the same time, the decrease in the frequency of non-correlative S|V commas (see section 3.1) shows that they became an increasingly marked feature even in their typical contexts of occurrence. The genre difference in the proportion of non-correlative S|V commas noted in section 3.1, where the Sermons genre was shown to feature a higher proportion of such commas than the Science genre, is due in part to a small number of specific environments. Since these environments probably (lists as subjects) or

potentially (verb phrases consisting only of a form of BE; semi-clausal structures as subjects) encourage the use of rhetorically motivated punctuation, this finding may help to explain why sermons favour non-correlative S|V commas to a greater extent than scientific texts do. Sermons is a speech-purposed – and thus speech-related – genre, and this relationship between the written text and a corresponding speech event may tally with a higher frequency of non-correlative S|V commas whose main function is to mirror spoken prosody in Sermons than in a non-speech-related genre like Science. However, as will be discussed in section 4, the fact that Science texts feature more non-correlative S|V commas than private letters (Smitterberg 2020) indicates that grammatically motivated punctuation is also present in the nineteenth-century material.

# 4 Conclusion

The present study has shown that non-correlative S|V commas became an increasingly rare feature of Modern English texts. This result tallies with previous work such as Smitterberg (2013, 2020). In addition, both the genre parameter and the length of the subject were shown to influence the distribution of such commas: the proportion was higher in Sermons than in Science and higher in main clauses with long subjects. When the main clauses that featured non-correlative S|V commas were considered separately in a contextual analysis, it became clear that their occurrence was closely linked to a limited number of linguistic contexts, most of which had been identified in contemporary works dealing with punctuation. Links between these contexts and the function of punctuation must remain tentative, especially since rhetorical punctuation after long subjects can virtually never be ruled out. However, most commas in Science were potentially grammatical in that they could be linked to syntactic reasons for inserting punctuation. In contrast, the Sermons genre evinced a number of commas, particularly after subjects comprising lists, that seemed to be rhetorically motivated; the comma would have corresponded to a spoken prosodic break, but the subject did not appear to have such a complex structure that it would have been felt necessary to signal its end by means of punctuation.

The genre differences in the material are interesting in combination with Smitterberg's (2020) analysis of private letters and scientific texts, in which the latter were more likely to feature non-correlative S|V commas. Provided that the different sets of scientific texts used in the two studies are comparable, these results imply that the occurrence of non-correlative S|V commas cannot be easily linked to the most important genre distinction in Modern English, viz,. that of

oral vs. literate genres (see, for instance, Biber and Finegan 1997). Scientific writing is more literate than both private letters (which may contain oral features owing to informal production circumstances) and sermons (which are speech-related); yet Science occupies the middle position between Letters and Sermons as regards the proportion of commas. This is most likely because non-correlative S|V commas are multi-functional, with both rhetorical and grammatical uses. The proportion of such commas may thus be particularly low in texts like private letters, which feature neither an overt relationship with a speech event (which might encourage rhetorical punctuation) nor a large number of syntactically complex subjects (which may favour grammatically motivated commas).

The decrease in the proportion of main clauses with non-correlative S|V commas, which has been attested in all genres studied, and which means that Modern English was gradually changing towards norms for Present-day English, may be due to a combination of factors. As mentioned above, several scholars argue that rhetorical punctuation becomes rarer during the Modern English period, which can be expected to affect the incidence of rhetorically motivated commas. One factor in this development is probably that writing became increasingly independent of speech in general, which may decrease the incidence of written commas that mirror spoken prosodic breaks. At the same time, previous studies such as Schou (2007) indicate that Modern English readers became increasingly able to process written texts. One consequence of such a development may be that the number of contexts in which writers felt that a comma was necessary to mirror syntactic boundaries decreased.

The present study has far from exhausted the topic of non-correlative S|V commas. Their development should be traced through the entire period covered by ARCHER, in more genres, and in both British and American texts. More work is also needed on the functions of non-correlative S|V commas. Nevertheless, the results reached in this study have important implications for how we interpret historical punctuation as regards, for example, to what extent non-correlative S|V commas were part of common practice at different times, in different genres, and in different linguistic contexts, and in what way their occurrence can be related to syntactic and prosodic boundaries. The combination of quantitative analysis and close readings of individual tokens has shed light on an underexplored field. As indicated by Salmon (1999: 50), by the mid-seventeenth century “nearly all the punctuation marks in common use now were known, but there were some differences in form and function”. The present study has shown that these differences are well worth exploring.

# References

## Primary source

ARCHER = A Representative Corpus of Historical English Registers. 1990– 1993/2002/2007/2010/2013. Originally compiled under the supervision of Douglas Biber and Edward Finegan at Northern Arizona University and University of Southern California; modified and expanded by subsequent members of a consortium of universities. Current member universities are Bamberg, Freiburg, Heidelberg, Helsinki, Lancaster, Leicester, Manchester, Michigan, Northern Arizona, Santiago de Compostela, Southern California, Trier, Uppsala, Zurich.

## Secondary sources

Auer, Anita. 2012. Late Modern English: Standardization. In Alexander Bergs & Laurel J. Brinton (eds.), *English historical linguistics: An international handbook*, 939–952. Berlin: Mouton de Gruyter.

Baron, Naomi S. 2001. Commas and canaries: The role of punctuation in speech and writing. *Language Sciences* 23(1). 15–67.

Biber, Douglas & Edward Finegan. 1997. Diachronic relations among speech-based and written registers in English. In Terttu Nevalainen & Leena Kahlas-Tarkka (eds.), *To explain the present: Studies in the changing English language in honour of Matti Rissanen*, 253–275. Helsinki: Société Néophilologique.

Biber, Douglas, Stig Johansson, Geoffrey Leech, Susan Conrad & Edward Finegan. 1999. *Longman grammar of spoken and written English*. Harlow: Pearson.

Brown, Goold. 1869. *The grammar of English grammars*, 10th edn. New York: William Wood.

Crystal, David. 2015. *Making a point: The persnickety story of English punctuation*. New York: St. Martin's Press.

Culpeper, Jonathan & Merja Kytö. 2010. *Early Modern English dialogues: Spoken interaction as writing*. Cambridge: Cambridge University Press.

Denison, David & Marianne Hundt. 2013. Defining relatives. *Journal of English Linguistics* 41(2). 135–167.

Lennie, William. 1864. *The principles of English grammar*, 57th edn. Edinburgh: Oliver & Boyd.

Lowth, Robert. 1762. *A short introduction to English grammar*. London: A. Millar & R. & J. Dodsley.

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

Robertson, Joseph. 1785. *An essay on punctuation*. London: J. Walter.

Rodríguez Puente, Paula. 2014. Colloquialization and "decolloquialization": Phrasal verbs in formal contexts, 1650–1990. In Simone E. Pfenninger, Olga Timofeeva, Anne-Christine Gardner, Alpo Honkapohja, Marianne Hundt & Daniel Schreier (eds.), *Contact, variation, and change in the history of English*, 163–186. Amsterdam: John Benjamins.

Ronberg, Gert. 1995. They had their points: Punctuation and interpretation in English Renaissance literature. In Andreas H. Jucker (ed.), *Historical pragmatics: Pragmatic developments in the history of English*, 55–63. Amsterdam: John Benjamins.

Salmon, Vivian. 1988. English punctuation theory 1500–1800. *Anglia* 106. 285–314.
Salmon, Vivian. 1998 [1962]. Early seventeenth-century punctuation as a guide to sentence structure. In Mats Rydén, Ingrid Tieken-Boon van Ostade & Merja Kytö (eds.), *A reader in Early Modern English*, 47–62. Frankfurt am Main: Peter Lang.
Salmon, Vivian. 1999. Orthography and punctuation. In Roger Lass (ed.), *The Cambridge history of the English language*. vol. 3, *1476–1776*, 13–55. Cambridge: Cambridge University Press.
Schou, Karsten. 2007. The syntactic status of English punctuation. *English Studies* 88(2). 195–216.
Smitterberg, Erik. 2013. Non-correlative commas between subjects and verbs in nineteenth-century newspaper English. In Nils-Lennart Johannesson, Gunnel Melchers & Beyza Björkman (eds.), *Of butterflies and birds, of dialects and genres: Essays in honour of Philip Shaw*, 367–382. Stockholm: Acta Universitatis Stockholmiensis.
Smitterberg, Erik. 2020. Non-correlative commas between subjects and verbs in nineteenth-century private letters and scientific texts. In Claudia Claridge & Merja Kytö (eds.), *Punctuation in context – past and present perspectives*, 23–45. Bern: Peter Lang.
Tagliamonte, Sali A. 2012. *Variationist sociolinguistics: Change, observation, interpretation*. Chichester: John Wiley.

Raquel Vea Escarza

# 7 Old English verbs of envy: Class membership and grammatical behaviour

## 1 Introduction

This study addresses the grammatical behaviour and class membership of Old English verbs expressing 'envy', *æfestian*, *andian*, *niðan*, and *ofunnan*, so as to determine the (in)consistency of their syntactic behaviour.[1] The theoretical basis of this research is provided, on the one hand, by Levin's (1993) framework of verb classes and alternations, according to which meaning determines the syntactic behaviour of a verb, and, on the other, by Role and Reference Grammar (Foley and Van Valin 1984; Van Valin and LaPolla 1997; Van Valin 2005, 2014). The starting point of this research is the inventory of verbs provided by *A thesaurus of Old English* (*TOE*) (Roberts et al. 2017), which classifies the lexicon on a conceptual basis. Verbs of envy share the expression of a common meaning, largely associated with the experiencer's covetous desire to own a quality or a possession that belongs to another person or entity, but also with the experiencer's ill will towards somebody or something. Providing that alternations[2] alone cannot be insightful enough to define class membership, it has been deemed convenient to resort to other criteria, mostly influenced by the Role and Reference Grammar framework, to analyse the grammatical behaviour of these verbs and to ascertain whether or not they constitute a homogeneus group both semantically and syntactically. These criteria are semantic valence (macrorole assignment), syntactic valence (number and type of arguments a predicate can take), morphological case of arguments, prepositional government in oblique core arguments, thematic relations, participation in alternations, and nexus and juncture (see section 2.2 for further discussion).

The purpose of this research is, therefore, to describe and analyse the syntactic behaviour of this group of Old English verbs that constitute a class according

**1** This research has been funded through the grant FFI2014-59110-P, which is gratefully acknowledged.

**2** According to Levin (1993: 2), verbs participate in different Diathesis alternations or "alternations in the expressions of arguments, sometimes accompanied by changes of meaning" (see 2.1).

**Raquel Vea Escarza,** University of La Rioja, Departmento de Filologías Modernas, Logroño, Spain, e-mail: raquel.vea@unirioja.es

https://doi.org/10.1515/9783110643282-008

to the *TOE*'s meaning-based classification. This research reappraises class boundaries as established by the *TOE* with a focus on the verbs of envy by applying a syntactic-semantic perspective. One of the contributions of this work is thus the introduction of a syntactic component into the analysis of this class and the reexamination of the *TOE*'s classification of these verbs according to a well-defined set of parameters that synthesise grammar and meaning. The results of this investigation demonstrate that the uniformity of this meaning-based class is not maintained when syntactic aspects are taken into consideration too.

The remainder of this chapter is organised as follows. The second section defines the theoretical model that supports the analytical part of the study. Section 3 surveys previous work on Old English verb classes and on the application of Role and Reference Grammar to Old English. Section 4 explains the obtaining of the data and the procedural steps of the study. Section 5 presents the analytical part of the research with an emphasis on the different criteria of analysis applied to the Old English verbs and a discussion of the convergent and divergent patterns of behaviour identified for them. Finally, section 6 presents some concluding remarks.

# 2 Theoretical framework

The analytical part of this study requires a general framework that can guide a syntactic-semantic description of verbs and provide a justification for class membership. In order to investigate the connection between the meaning of the verbs and their syntactic behaviour, this study brings together the two theoretical approaches that will be explained in detail in this section: Levin's (1993) model of verb classes and alternations, and Role and Reference Grammar. I have adopted these frameworks since one framework alone is not sufficient to explain the complex interdependence of syntax and semantics in the case of the verbs under study.

## 2.1 Levin's (1993) framework of verb classes and alternations

Levin's (1993) framework of verb classes and alternations is grounded in the assumption that the meaning of a verb largely determines its syntactic behaviour. This theory is underpinned by two descriptive notions that are interrelated: alternations and verb classes. According to this double-sided theory, the semantics of a verb determines the type and number of arguments that are associated with it, whereas members of the same verb class must share at least one

meaning component and present a similar syntactic behaviour. An alternation allows for the expression of a verb's arguments in at least two different ways, which may lead to an order and function alteration of these elements within the sentence and also to occasional changes of meaning (as in *I envy her courage* vs. *I envy her for her courage*; see discussion of examples 1 and 2 below). In this regard it must be noted that alternations affect arguments or predicate elements (see 2.2), i.e., constituents of the clause core. In syntactic terms, non-arguments are part of the periphery and include, for instance, temporal or locative elements.

Levin (1993) proposes a total of forty-nine classes, and many other subclasses, including, to name but a few, Verbs of Change of Possession (e.g., *feed*, *give*, *lease*, *lend*), Psych Verbs (e.g., *amuse*, *admire*, *marvel*, *appeal*), and Judgement Verbs (e.g., *congratulate*, *punish*, *recompense*, *ridicule*). The boundaries and edges that exist among these groups can be explained on the grounds of the semantic and syntactic properties that define each of them. If we take Psych Verbs and Judgement Verbs as an example, it must be noted that both types express a reaction to something; however, the former involve a particular feeling, whereas Judgement Verbs are rather related to the expression of a judgement or opinion. With respect to the syntactic patterns that characterise each group, Judgement Verbs typically take sentential complements as objects, i.e., clauses that are arguments of the verb; the realization of this pattern is rare in the case of Psych Verbs.

The other main idea of Levin's (1993) theory are the so-called Diathesis alternations, which designate the syntactic possibility of a verb to express their arguments in at least two different ways, giving rise in some occasions to a change of meaning. According to Levin, the verb 'to envy' is classified as a negative *admire*-type Psych Verb, a group of verbs which are primarily found in alternating constructions involving arguments in the predicate. Specifically, these verbs take part in the Possessor-Attribute Factoring Alternations, which fall into three different constructions and that allow for a double expression of the possessor and the possessed arguments. The first subtype of alternation, the Possessor Object Alternation, is typical of transitive verbs that offer the possibility to express a) only one constituent, a possessor and attribute in the form of a noun phrase (1), and b) two constituents, one being the possessor and the other the attribute introduced by a prepositional phrase involving *for* (as in 2).

(1) *I envy her courage*.

(2) *I envy her for her courage*.

In the second subtype, the Attribute Object Alternation, the first part of the alternation is expressed by a noun phrase encoding the possessor and the attribute (as in 1). In the alternating counterpart, the attribute is a noun phrase and the first constituent of the predicate, whereas the possessor is represented as a prepositional phrase headed by *in*, as represented in (3).

(3) *I envy the courage in her.*

The third subtype, the Possessor and Attribute Alternation, combines the two previous ones, thus giving rise to alternating pairs (4) and (5).

(4) *I envy her **for** her courage.*

(5) *I envy the courage **in** her.*

## 2.2 Role and Reference Grammar

Together with Levin's (1993) framework of verb classes and alternations, another theory that provides a more general perspective on the interrelation between syntax and semantics and serves as a theoretical foundation for the study is the functionally oriented Role and Reference Grammar (henceforth RRG) (Foley and Van Valin 1984; Van Valin and LaPolla 1997; Van Valin 2005, 2014). RRG describes the clause as a layered structure in which a nucleus (normally a verb) together with its arguments form a core, and a periphery can be optionally appended. This layered (or hierarchical) structure results from the expansion of a predication (a verb plus its arguments) by means of the insertion of operators (grammatical features) and satellites (adverbials). The notion of the layered structure of the clause rests on two key pillars, namely the predicate and the non-predicating elements. Within the non-predicating elements, a distinction must be made between arguments and non-arguments, the presence of the former (noun and adpositional phrases) being entailed by the semantics of the predicate. By way of illustration, in the sentence *Robin gave it to Mary in the park*, there are three arguments, namely *Robin*, *it*, and *to Mary*, since there is a semantic relation between these elements and the predicate, while *in the park* is a non-argument and thus a peripheral element in the sentence.

From an RRG perspective, the structure of complex sentences is explained in terms of nexus relations and juncture, the former alluding to the syntactic relations that hold between units, the latter to the nature of these units in a

complex construction. Regarding juncture, three types of units of different degree of complexity can participate in complex structures: nuclear, core, and clausal juncture. Concerning nexus relations, these include coordination, subordination, and cosubordination. The three types of nexus can be found in the three forms of syntactic relations (or juncture), thus originating a variety of juncture-nexus combinations.

Concerning juncture, in a nuclear juncture it is expected that at least two nuclei form a single nucleus to which a common set of arguments is attached. In (6),[3] two nuclei, namely *make* and *eat* are found in a single core (verbal nucleus and its arguments).

(6) *I will make John eat the cakes.*

At a core level, core junctures contain at least two cores in a clause. The different cores in this structure share a core argument. In (7), *Harry* is the argument shared by the first core (*Fred saw Harry*) and the second core (*Harry leave the room*).

(7) *Fred saw Harry leave the room.*

A clausal juncture is found in complex structures of the type illustrated in (8).

(8) *Anna read for a few minutes, and then she went out.*

As a coordinated sentence in which two units of equal status are linked, creating independent main clauses, example (8) also illustrates a coordination nexus. Examples (9) and (10) illustrate another type of nexus relation, subordination, by means of which one of the clauses is embedded in the other, being either an argument (daughter subordination) or a modifier (peripheral subordination) of the main clause.

(9) ***That she arrived late*** *shocked everyone.*

(10) *Pat went to the party* ***after he talked to Chris***.

In the third type of nexus, cosubordination, the units share at least one operator (tense, aspect, modality, etc.). This is illustrated in (11), whereby *sat* and *reading* share the operator of imperfect aspect.

**3** Examples (6) through (12) have been drawn from Van Valin (2005: 197–199).

(11) *Kim sat reading a book.*

One final central component of RRG theory deserves attention: the semantic interpretation of verbal arguments. This idea is based on two generalised semantic roles or macroroles called *actor* and *undergoer*, which comprise different thematic relations (see below). In a transitive predication, the actor is the first argument and the undergoer the second argument of the verb. In an intransitive predication, the only argument can be an actor or an undergoer, depending on the semantic properties of the predicate. This is illustrated in (12) and (13).

(12) *Jill is sleeping in the car.*

(13) *Jill is swimming.*

In (12), the first argument, *Jill*, is an undergoer because *sleep* is a stative verb, whereas the first argument of (13), *Jill*, is an actor because *swim* is an active verb. The semantic macroroles actor and undergoer are generalisations across argumental structures. That is to say, the actor is the first argument of verbs like *eat*, *put*, and *give*, while the undergoer is the first argument of *be*, *die*, and *have* and the second argument of *eat*, *put*, and *give*.

Apart from the aforementioned semantic roles (that is, actor and undergoer), RRG acknowledges the presence of thematic relations, which are generalisations across verb-specific roles (e.g., agent, experiencer, patient, etc.; Van Valin 2005: 53). Figure 1 is an adaptation of Van Valin's (2005: 58) thematic relations continuum and represents the different thematic relations posited by RRG in relation to the five possible argument positions. This hierarchy is not defined in categorical terms, but rather as a gradient system of relations between syntax and semantics. The actor is located to the left of the continuum, whereas the undergoer appears to the right of the continuum. Those thematic relations closer to the left side are thus more agent-like, whereas those placed on the other side of the continuum are more patient-like.

In RRG, the term *M-transitivity* is used to refer to the number of macroroles a verb can take and which determine the semantics of that verb. On the other hand, the term *S-transitivity* relates to the number of syntactic arguments a verb takes. A verb can take from zero to two macroroles (M0, M1, or M2). M0 would be equal to an atransitive verb (*rain*), which takes no arguments and no macroroles; M1 corresponds to intransitivity (*die*) and takes one argument and one macrorole; M2 represents transitivity (*send*) and takes two arguments and two macroroles. Ditransitive verbs, like *send* or *give*, are M2 verbs with an S3 value, since the number of arguments they take is three but the number of macroroles

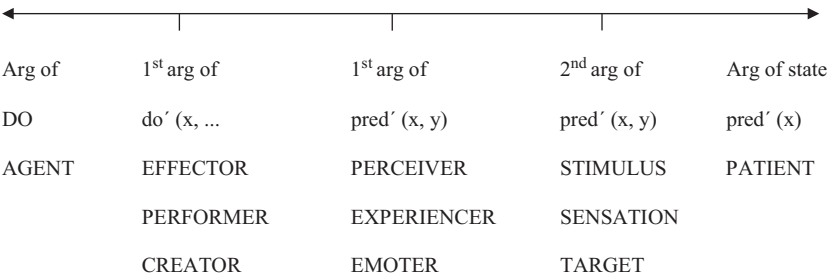


**Figure 1:** Thematic relations continuum and argument positions (adapted from Van Valin 2005: 58).
The abbreviations in the header row represent the possible positions of the arguments and their thematic relations to the predicate. "Arg of DO" refers to the single argument of an agentive verb; "1st arg of do′" signifies the effector of an action and implies the presence of a second argument; the third column includes the first arguments of less agentive verbs, namely verbs of cognition, perception, or emotion; the fourth column lists the type of roles the second argument of a state predicate performs, to wit, the entity that is known, seen, felt, etc.; and, finally, the right-most column refers to the single argument of a state predicate that receives the consequences of an action.

is two. Notice that in the sentence *Robin sent Mary a letter*, the verb *send* has a semantic valence of 3, since we can identify three syntactic arguments, namely *Robin, Mary* and *a letter*, whereas there are only two semantic macroroles: actor for *Robin* and undergoer for *Mary*. As observed in this example, the number of syntactic arguments a verb takes does not necessarily coincide with its number of macroroles.

## 2.3 Bringing syntax and semantics together

Nothwithstanding the validity of Levin's (1993) model in the study of verb classes and alternations and the correlation between semantics and syntax, a classification of the verbal lexicon that is ultimately justified on the grounds of the verbal participation in alternations proves insufficient in a historical language such as Old English. Old English is a synthetic language in which inflectional endings prevail over prepositional government and word order; this seems to be a convincing reason why a complementary model can partly compensate for some of the constraints of Levin's (1993) model. Along these lines, RRG complements Levin's (1993) model inasmuch as it offers a more general view of grammar and, specifically, of the combination of syntatic and semantic aspects. Levin's (1993) model proceeds from syntactic frames and then considers

semantic components to finally draw the conclusion that syntactic behaviour is semantically determined. RRG, by contrast, includes complex clauses, verb complementation, morphological case or semantic roles, creating associations between syntax and semantics.

These two theoretical frameworks have inspired the seven criteria of analysis that have guided the analytical part of the study. They represent key aspects within these theories and capture both semantic and syntactic features of the grammar. The seven criteria are

- *Semantic valence*, which refers to the number of macroroles a verb can take. Semantic valence will be further represented through the abbreviations M1 and M2. If a verb's semantic valence is M1, it means that it is an intransitive verb, and its argument can only be assigned one semantic role, either actor or undergoer. If the semantic valence of a verb is M2, it will be a transitive verb with both macroroles. Finally, if a verb is assigned M1 and M2, this means it can be either transitive or intransitive, and semantic roles will apply accordingly.
- *Syntactic valence*, which points to the number and type of arguments a predicate can take. Regarding the typology of arguments at a core level, there are direct and oblique arguments: nominative, accusative and dative are considered direct cases, and the others, including phrases marked by a preposition, are oblique.
- *Morphological case of arguments*, which includes nominative, accusative, genitive, and dative.
- *Prepositional government in oblique core arguments*, which refers to the type of prepositions a verb can rule and the morphological case of their complements.
- *Thematic relations*, which cover the type of roles the arguments of a verb fulfil on a lexical basis;
- *Alternations* (Levin 1993), which refer to the capacity of a verb to express its arguments in two different ways (at least);
- *Nexus and juncture*, which concern the internal relation structure of clauses which affects complex sentences; more specifically, nexus refers to syntactic relations between units and juncture to the nature of these units.

The combination of two theories that address a wide range of verbal features – namely verbal augmentation, meaning and form assocations, and the connection between syntax and semantics more generally – will facilitate a more detailed analysis than has been previously undertaken through the melding of syntactic and semantic perspectives.

## 3 Review of previous work

Studies conducted from several syntactic-semantic approaches on the verbal lexicon of Old English have helped describe verbal behaviour on the basis of a set of principles that consider the syntactic and semantic configuration of a given verbal class. Several contributions have revolved around different Old English verb classes by focusing on constructions, i.e., systematic patterns of form and meaning associations, on the representation of their logical structure, as well as on alternations, which concern recurrent form and meaning contrasts.

Some of these works are framed within the Lexematic-Functional model (see Geeraerts 2010), a theory that combines Functional Grammar and Structuralist Semantics and applies them to the study of verbal syntax and semantics. A few of the most relevant contributions are Cortés Rodríguez and González Orta's (2006) verbs of sound; Cortés Rodríguez and Torres Medina's (2003) verbs of running; C. D. García Pacheco's (2013) verbs of feeling; L. M. García Pacheco's (2013) verbs of existence; González Orta's (2002a, 2002b, 2003) verbs of warning, speech, and smell perception and emission, respectively; and Vera Díaz's (2005) verbs of colour. Other studies have concentrated on particular verbs, such as the work by Sosa Acevedo (2007) on (*ge*)*sēon* and (*ge*)*lōcian*. A third type of works has addressed constructions, for example, González Orta's (2006) study on the resultative construction in speech verbs or Sosa Acevedo's (2009) on the conative construction.

Martín Arista (2001, 2016) has applied RRG to the study of Old English verbs. He demonstrates that Old English can express by syntactic means almost all the *Aktionsart* types defined in RRG. Even though this verb typology has not been considered a criterion of analysis in the present study inasmuch as the verbs under scrutiny here are all verbs of state, this type of work demonstrates that RRG has its application to the study of a historical language from a syntactic-semantic perspective.

## 4 Data retrieval and methodology

This study draws on *A thesaurus of Old English* (*TOE*) to retrieve the inventory of verbs of envy: *æfestian*, *andian*, and *ofunnan*. Standard Old English dictionaries, mainly Bosworth and Toller (1973), Clark Hall (1996), and Sweet (1976), and the lexical database of Old English Nerthus, have supplied additional information about the aforementioned verbs regarding meaning aspects, alternative spellings, and inflectional forms. As for the citations, *The dictionary of Old English* (*DOE*; Healey 2016) has been consulted for *æfestian* and *andian*.

*The dictionary of Old English Corpus* (Healey et al. 2004) has provided additional citations for these verbs. Searches of the citations using the York Parsed Corpus of Old English Prose (YCOE) (Taylor et al. 2003) have allowed for checking the syntactic information relevant for the study.

Once the initial group of verbs was defined, a search in the Nerthus database made possible the discovery of a fourth verb, *nīðan*, whose meaning coincides with the other members', and its incorporation was justified from this perspective. According to *TOE*, *nīðan* is the single component of the group 'to envy, hate'. However, based simply on the information provided by *TOE*'s defining headings, it is not very clear why these four verbs belong to two different groups.

For *æfestian* and *andian*, the lemmatised forms identified by the *DOE* have been trusted, whereas for *ofunnan* and *nīðan* the database *Freya*[4] has been used, which compiles and systematises information regarding the inflectional forms, lemmas (dictionary headwords), alternative spellings, meaning, etc. as they appear in secondary sources, i.e., glossaries and studies in the Old English language.

At the same time, the study presents certain limitations with respect to the information offered by the main works of reference, mainly related to their partial scope. For *æfestian* and *andian*, their entries in the *DOE* (A-H) have provided accurate semantic and syntactic information as well as textual occurrences with the intended meaning. However, this information is not available for the verbs *nīðan* and *ofunnan*, since these entries are not yet available in the *DOE*. For some textual fragments stored in the *DOEC*, it has not been possible to find a translation or glossary, thus obscuring the analytical process. The *YCOE* has contributed useful syntactic data throughout the syntactic annotation of roughly one half of the texts in the *DOEC*, although for the other half there is not information available in this regard.

## 5 Grammatical behaviour of Old English verbs of envy

This section addresses the characterization of the four verbs (*æfestian*, *andian*, *ofunnan* and *nīðan*) in terms of the criteria selected for analyzing their grammatical behaviour (section 2.3). Table 1 summarises the main information

4 *Freya* is one of the databases created by the Nerthus project that is available for internal use.

**Table 1:** Implementation of syntactic-semantic factors to Old English verbs of envy.

| Criteria of analysis | Old English verbs of envy | | | |
|---|---|---|---|---|
| | ***æfestian*** | ***andian*** | ***ofunnan*** | ***nīðan*** |
| **Semantic valence** | M2 | M1 and M2 | M2 | M2 |
| **Syntactic valence** | ARG 1, ARG 2, ARG 3:<br>– 2 direct core arguments<br>– 1 oblique core argument | ARG 1, ARG 2, ARG 3:<br>– 3 direct core arguments<br>– 1 oblique core argument | ARG 1, ARG 2, ARG 3:<br>– 3 direct core arguments | ARG 1, ARG 2:<br>– 2 direct core arguments |
| **Morphological case of arguments** | ARG 1: nom.<br>ARG 2: acc./gen./dat.<br>ARG 3: dat. | ARG 1: nom.<br>ARG 2: acc./dat./ clause in subordination<br>ARG 3: dat. | ARG 1: nom.<br>ARG 2: gen.<br>ARG 3: dat. | ARG 1: nom.<br>ARG 2: clause in subordination |
| **Prepositional government in oblique core arguments** | *on* + dat/acc<br>*wiþ* + dat | *ofer* + acc<br>*on* + dat/acc<br>*ongean* + dat/acc<br>*togeanes* + dat | – | – |
| **Thematic roles** | ARG 1: experiencer<br>ARG 2: target<br>ARG 3: patient | ARG 1: experiencer, patient<br>ARG 2: target, patient<br>ARG 3: patient | ARG 1: experiencer<br>ARG 2: target<br>ARG 3: patient | ARG 1: experiencer<br>ARG 2: target |
| **Participation in alternations** | Attribute Object alternation | – | – | – |
| **Nexus and juncture** | – | Linked clause in subordination | – | Linked clause in subordination |

derived from the implementation of the parameters of analysis to these verbs and should be interpreted in the following way: an abbreviated description for each criterion and verb has been inserted in each cell. As for the semantic valence, all the possible realizations have been indicated in each case. The syntactic valence illustrates the highest number of arguments of each type that a verb can take. Prepositional government has been described with the suitable preposition(s) and the case(s) they rule. The type of thematic relation has been provided for the fifth criterion. In the sixth place, the name of the alternation has been included where the verb actually evinces an alternating pattern. Finally, the type of complex relation has been provided when a complex construction applies.

In what follows, a description of each verb is accompanied with Old English examples retrieved from the *DOE* corpus and preceded by their Cameron number.[5] This is followed by a summary discussion that returns to the similarities and differences among the verbs sketched in Table 1.

## 5.1 *Æfestian*

An analysis of the complementation patterns of the verb *æfestian* reveals that it can take up to two direct core arguments and one oblique core argument. The nominative is the unmarked[6] case for the first argument, whereas the second argument can be either unmarked for accusative or marked for genitive when referring to an inanimate entity. The dative is the unmarked case for the third argument, which is illustrated in (16) below. In (14), an example is provided of an accusative second argument for this verb: (*oþra manna) goddæde* '(men's) virtues'. In (15), the second argument is genitive-marked: *his godra weorca* 'his good works'.

(14) *hi symble* ***æfæstiað*** *oþra manna goddæde* **(**GD 2 (C); B9.5.4 [0181 (8.117.2)])
'he continuously **envies** other men's virtues'

(15) ***æfæstgende*** *his godra weorca* (GDPref and 3 (C); B9.5.5 [0350 (16.211.11)])
'**envying/envious of** his good works'

**5** The whole list of texts by Cameron number (alphanumeric code that identifies each text and is named after the first compiler of the corpus, Angus Cameron) is available at http://www-users.york.ac.uk/~lang22/YCOE/info/YcoeTextCam.htm

**6** In this context, unmarked refers to the case that is considered customary for the first argument.

The verb *æfestian* also allows for the expression of two constituents in the VP, one being the coveted attribute in the form of a noun phrase, the other the possessor of the coveted attribute, a phrase headed by a preposition, in this case *on*. The attribute is the second argument of the construction, inflected for the accusative, and allocated the thematic role of target; the possessor of the attribute, in turn, constitutes an oblique core argument in the form of a prepositional phrase governing the dative case. In the continuum of thematic relations this third argument would be the most patient-like.

(16) *hi* ***æfstiað*** *on oðrum mannum hyra mægenes god*. (GD 2 (H); B9.5.10.2 [0105 (8.117.6)])
'they **envy** the goodness of their virtue in other men'

Another syntactic pattern for *æfestian* is to have an oblique core argument in second position displayed by a prepositional phrase headed by prepositions *on* or *wiþ* and followed by a dative-case noun phrase. Examples (17) and (18) illustrate each preposition.

(17) *Ðes iunga man ne* ***æfestigað*** ***on*** *nánum ðingum ðe hé hór gesihð* (ApT; B4.1 [0140 (14.30)])
'This young man **envies** nothing that he here sees'

(18) *Ðá geseah hé dæs sácerdes mód byrnan and* ***æfæstigian wiþ*** *his lífe* (GD 2 (C); B9.5.4 [0194 (8.119.5)])
'Then he saw the priest's spirit burn and **feel envy towards/envy** his life'

In terms of the alternations assigned to the *admire*-type Psych Verbs by Levin (1993), the evidence shows that the only alternation that is fully accomplished by this verb is the Attribute Object Alternation. Examples (19) and (20) illustrate this.

(19) *hi symble* ***æfæstiað*** *oþra manna goddæde* (GD 2 (C); B9.5.4 [0181 (8.117.2)])
'He continuously **envies** other men's virtues'

(20) *hi* ***æfstiað*** *on oðrum mannum hyra mægenes god* (GD 2 (H); B9.5.10.2 [0105 (8.117.6)])
'He **envies** in other men the goodness of their virtue'

In (19), the target is a noun phrase with direct core argument status (*oþra manna goddæde*) and inflected for the accusative. In (20), *hyra mægenes god* is

the target, which receives direct core argument status and accusative case, and *on oðrum mannum* is the possessor of the coveted entity, an oblique core argument that consists of a prepositional phrase whose complement is inflected for the dative.

## 5.2 *Andian*

In terms of its S-transitivity, *andian* can take a maximum of three direct core arguments and one oblique core argument, though a third direct core argument is rare. The unmarked case for the first argument is the nominative. The second argument can be either a direct or an oblique core argument. If it is direct, it will be inflected for the accusative or for the dative. If it is an oblique core argument, it will be inflected for the accusative or for the dative too, depending on the preposition. In some cases, as will be explained next, the same preposition can rule both cases. In (21), the second argument is introduced by a prepositional phrase headed by preposition *on*, and is inflected for the accusative.

(21) *ne* ***andgiað on*** *þone welegan* (PPs (prose); B8.2.1 [0724 (48.16)])
'do not **envy** the rich'

Other prepositions the verb *andian* appear with are *ofer*, *ongean*, and *togeanes*. As for case inflection, *on* and *ongean* admit both accusative and dative, *ofer* takes accusative, and *togeanes* rules dative. In (22), *togeanes* is actually a postposition that follows the pronominal complement.

(22) *þa ongann se ungeseowenlice feond him* ***togeanes andigen*** (LS 28 (Neot); B3.3.28 [0023 (45)])
'Then the invisible foe began to **feel envious of** him'

In the reflexive realization of the verb *andian*, exemplified in (23), the second argument is a direct core argument, whereas in (21) and (22) the second argument is marked by adposition (i.e., a preposition or postposition that complements a noun phrase).

(23) *se þe him* ***andað*** *naht ys him wyrse* (LibSc; C15 [0684 (15.8)])
'The one who **envies** himself, nothing is more wretched to him'

Only one example has been found of the verb *andian* in a complex sentence. As shown in (24), the embedded *that*-clause is a direct core argument of the matrix predicate *andian*.

(24) *æfre hi **andedon** arleaslice þæt se hælend wolde þa hæþenan þeoda him to folce geceosan* (ÆHomM 12 (Brot 1); B1.5.12 [0059 (193)])
'they always **envied** wickedly that the Saviour would accept the heathen nations to his people'

The presence of two arguments in the VP is not a common syntactic realisation of the verb *andian*. The sentence in (25) can be taken as an isolated example of this three-argument construction, in which the VP arguments are inflected for dative and accusative respectively.

(25) *þeaw þwyra ys **andian** oþrum mægenes god* (LibSc; C15 [0696 (15.20)])
'it is an evil practice to envy others the goodness of their virtue'

The verb *andian* also has an intransitive realisation. In (26), the only argument, with thematic role experiencer, is inflected for nominative.

(26) *se þe **andaþ** he na lufað* (LibSc; C15 [0689 (15.13)])
'the one who **envies** does not love'

As for the participation in alternations, no systematic alternating patterns have been identified in this Old English verb. The alternations attributed to the verb *envy* in Present Day English seem not to work in this case, given that the verb *andian* allows neither for two oblique core arguments marked by adposition nor for one direct and one oblique core argument – also marked by adposition – in the VP that permit the alternating pattern.

## 5.3 *Ofunnan*

*Ofunnan* is a ditransitive verb that takes three direct core arguments, the first inflected for the nominative, the second for the genitive, and the third for the dative, this last case being the unmarked case for the third argument of ditransitives. On the semantic side, with respect to thematic role assignment, the first argument is designated experiencer – *seo*, in (27) – whose antecedent is *neaht* 'night', an inanimate entity; the second argument is the target, and the last is an argument in dative case with a patient-like thematic role.

(27) *(þonne seo **neaht** becymeð) seo me eðles **ofonn*** (Rim; A3.15 [0023 (70)])
'(then the night approaches,) which **begrudged** me my homeland'

No other syntactic valence has been found for this verb, as a consequence of which it is not possible for *ofunnan* to take part in any of Levin's (1993) alternating patterns.

## 5.4 *Nīðan*

The only syntactic configuration attributed to the verb *nīðan* is that of a transitive. It therefore takes two arguments, the first inflected for the nominative, the second being a clause in subordination. The first argument serves the thematic role experiencer, whereas the second argument corresponds to the target. Due to the fact that there is only one syntactic configuration that applies to this verb, there is no possibility for *niðan* to participate in any alternating pattern. An example is provided in (28).

(28) *(hí) hefelíce **níþað** þá hí selfe nó ðý ær habban willað* (D 2 (C); B9.5.4 [0181 (8.117.2)])
'(they) heavily **envy** what they do not wish to have none the sooner for themselves'

## 5.5 Boundaries and edges in verb class membership

This subsection addresses the converging and diverging aspects that characterise Old English verbs of envy from a syntactic-semantic perspective, including the three verbs constituting a category in the *TOE*, *æfestian*, *andian*, and *ofunnan*, and also the verb *nīðan*, which has been added to the analysis due to its sense relatedness.

As described in subsections 5.1–5.4 and summarised in Table 1, the verbs *æfestian*, *andian*, *ofunnan*, and *nīðan* share certain aspects of their grammatical behaviour. Overall, these verbs are prototypically transitive; only the verb *andian* can appear without an expressed object. As for the number of arguments that can be assigned to these predicates, *æfestian*, *ofunnan*, and *nīðan* take a maximum of two direct core arguments, and *andian* a maximum of three; as for the number of oblique core arguments, *æfestian*, *andian*, and *ofunnan* can take one of this type. The verb *nīðan*, in turn, only rules direct core arguments.

Concerning the thematic roles (see Section 2.2 and Figure 1) associated with these arguments, the first argument in all four cases assumes the role of experiencer. As for the second argument, it is assigned the thematic role of target if it refers to an inanimate entity, whereas patient is preferred when the target is a human entity. Only the verb *andian* can attribute the role of patient to the second argument as it is the only verb that has an animate second argument (see examples 21–23). If there is a third argument, it will adopt the role of patient (verbs *æfestian*, *andian*, and *ofunnan*). Since verbs of envy are verbs of state, the first argument will never be agent-like. Taking Figure 1 as a reference, the first argument would occupy a middle position in the hierarchy. The other arguments tend to be more patient-like and are thus placed to the right in the continuum. *Andian* is the only verb that can have a semantic valence of M1 as it is used intransitively. In terms of thematic roles, the first and only argument of this verb will assume the role of patient.

The morphological case of the first argument is always nominative; no other case has been found for this argument. The second direct core argument of the verb *æfestian* is inflected for either accusative or genitive, although the accusative is preferred when there is a third argument (oblique). When the second argument is an oblique core one, the prepositional phrase will rule the dative case. The dative is also the unmarked case for the third argument. The verb *andian* shows some differences as regards the inflection of the second and third arguments. In the absence of a third argument, the second argument, when it is a direct core one, is inflected for the dative; however, it is inflected for the accusative if a third argument occurs. For those oblique core arguments in second position, the accusative or the dative will be chosen, depending on the preposition heading the phrase. The third argument is always inflected for the dative. The verb *ofunnan* only presents a syntactic pattern in which the second argument is inflected for the genitive and the third for the dative, the second being non-direct case, and the third a direct one. It should be pointed out that this verb, together with *æfestian*, is the only one that has a predicate argument inflected for a non-direct case in a non-prepositional phrase. Finally, the peculiarity of the verb *nīðan* is that the argument of its predicate (or second argument) is exclusively performed through a subordinate clause.

The prepositions found in oblique core arguments of the verb *æfestian* are *on* and *wiþ* (+ dat.) in second position arguments, and only *on* (+ dat.) in third position arguments. For the verb *andian*, prepositional phrases in oblique core arguments are only found in second position; the prepositions that head these phrases are *ofer* (+ acc.), *on* and *ongean* (+ dat./acc.), and *togeanes* (+ dat.).

Regarding their participation in alternating constructions, the analysis has revealed that only the verb *æfestian* takes part in one of them, the Attribute

Object Alternation. An explanation for this modest participation in alternating patterns may be found in the fact that Old English had an elaborate inflective system where prepositional phrases did not have a prominent presence; as the language evolved, this highly inflected system was gradually reduced, paving the way for other syntactic means, such as the proliferation of prepositional phrases.

The last criterion, nexus and juncture, applies to complex constructions. Both *æfestian* and *ofunnan* are found in simple clauses, whereas *andian* and *nīðan* can have, in addition, a clause in subordination as its second argument.

# 6 Conclusion

The starting point of this research has been the inventory of verbs of envy proposed by *TOE*, arranged in ordered categories on the basis of a conceptual criterion. My analysis of these verbs (*æfestian*, *andian*, and *ofunnan*) has revealed the necessity to correlate semantics and syntax in order to evaluate their consistency as a group. A fourth verb, *nīðan*, was added to this list driven by the fact that there was a meaning correspondence between these four verbs.

My chapter has established a number of parameters that have guided the analytical part of the study and that constitute the defining properties of the verbs under analysis. These criteria have been drawn from two main sources: the framework of verb classes and alternations (Levin 1993) and Role and Reference Grammar. After applying these criteria to help determine the (in)consistency of the verb group both in terms of semantic and syntactic behaviour, notable differences have been detected, mainly in their syntactic valence, in the morphological case of the arguments appearing in the VP, in the prepositional government in the oblique core arguments, and in their acceptance of complex configurations. This diversity of grammatical realisations leads to the conclusion that, if both meaning components and grammatical behaviour are considered, the group comprised of *æfestian*, *andian*, *ofunnan*, and *nīðan* is not a homogeneous one, notwithstanding their transitive nature or the partial coincidences in argument realisation.

The analysis of the fourth candidate, *nīðan*, has made it possible not only to complete the inventory of the Old English verbs that expressed the idea of envy but also to reassess the boundaries and edges of the group under scrutiny. This verb has evinced a relevant similarity with *andian*, which is the fact that they can appear in both simple and complex configurations; this does not occur for *æfestian* and *ofunnan* since they are restricted to simple structures. If

the focus is rather on the type of arguments the predicate takes, *nīðan* bears similarities with *ofunnan* given that both verbs exclusively take direct core arguments, while the others can also take an oblique one. The only characteristic that makes *nīðan* unique in the group is its two-argument pattern since no evidence of a third argument has been discovered in this verb. Should the participation in alternations be considered, *æfestian* distances itself from the other verbs, being the only verb that participates in an alternation (Attribute Object Alternation). In terms of semantic valence, there is similarity between *nīðan*, *æfestian*, and *ofunnan* since these verbs are exclusively M2, whereas no M1 semantic realization seems to be possible in any of them. Briefly stated, verbal affinity is especially present when semantically-oriented parameters apply (macrorole assignment and thematic roles); marked divergencies are, however, found when it comes to syntactic configurations.

All things considered, a semantic-syntactic-based analysis substantiated by RRG and the framework of verb classes and alternations has revealed information that a strictly semantic analysis is unable to provide. An analysis of this type has thus enabled us to reconsider the role the members of this class have in the group as a whole. In this case, the verb *nīðan*, which was not a member of the *TOE*'*s* 'to be envious, envy' group but formed a group on its own, has evinced certain similarities with the other verbs as regards semantic traits and syntactic patterns, although none of these coinciding features are solid enough to justify a unified group. The inclusion of this fourth candidate has confirmed that meaning similarity does not necessarily guarantee syntactic analogy.

This study insists on the necessity of merging syntax and semantics in this type of analysis. On one side, *TOE*'s "to be envious, envy" group has proved not to be a uniform class from a syntax-semantics combined perspective. On the other, some caution must be also exercised when dealing with verb classes on the basis of Levin's (1993) conception of them since, as has been observed, these verbs do not share a fully analogous grammatical characterization.

This should not be taken as a drawback but rather as an opportunity to reassess the concept of verb class in connection with their boundaries and edges. Moreover, the parameters selected for the analysis constitute themselves a valuable source of information that can help look into the extent and scope of other Old English verb classes and establish interclass associations. Although much work is still needed, ongoing studies conducted by the Nerthus project predict similar results for different verb classes. The application of a set of principled criteria inspired by both syntactic and semantic traits paves the way for a more eclectic approach to the description and definition of verb classes and to the reconsideration of their boundaries.

# References

## Primary sources

Bosworth, Joseph & T. Northcote Toller. 1973 [1898]. *An Anglo-Saxon dictionary*. Oxford: Oxford University Press.

Clark Hall, John R. 1996 [1896]. *A concise Anglo-Saxon dictionary*. Toronto: University of Toronto Press.

Healey, Antonette diPaolo (ed.). 2016. *The dictionary of Old English in electronic form A-H*. Toronto: Dictionary of Old English Project, Centre for Medieval Studies, University of Toronto (accessed on March and April 2018).

Healey, Antonette diPaolo (ed.), with John Price Wilkin & Xin Xiang. 2004. *The dictionary of Old English web corpus*. Toronto: Dictionary of Old English Project, Centre for Medieval Studies, University of Toronto.

Nerthus: Lexical Database of Old English. http:www.nerthusproject.com.

Roberts, Jane & Christian Kay, with Lynne Grundy. 2017. *A thesaurus of Old English*. Glasgow: University of Glasgow. http://oldenglishthesaurus.arts.gla.ac.uk/ (accessed on February 2018).

Sweet, Henry. 1976 [1896]. *The student's dictionary of Anglo-Saxon*. Cambridge: Cambridge University Press.

Taylor, Ann, Anthony Warner, Susan Pintzuk & Frank Beths. 2003. The York-Toronto-Helsinki parsed corpus of Old English prose. http://www.helsinki.fi/varieng/CoRD/corpora/YCOE/ (accessed in March and April 2018).

## Secondary sources

Cortés Rodríguez, Francisco & Marta González Orta. 2006. Anglo-Saxon verbs of sounds: Semantic architecture, lexical representation and constructions. *Studia Anglica Posnaniensia* 42. 249–284.

Cortés Rodríguez, Francisco & Dolores Torres Medina. 2003. Old English verbs of running: Linking semantic representation and morphosyntactic structure. *Folia Linguistica Historica* 24(1). 153–174.

Foley, William & Robert Van Valin. 1984. *Functional syntax and universal grammar*. Cambridge: Cambridge University Press.

García Pacheco, Carmen Dolores. 2013. *Los verbos de sentimiento en inglés antiguo: Arquitectura léxica e interfaz semántica-gramática*. PhD Dissertation. San Cristóbal de La Laguna, Universidad de La Laguna.

García Pacheco, Luz María. 2013. *El dominio verbal de la existencia en anglosajón. Análisis semántico-sintáctico*. PhD Dissertation. San Cristóbal de La Laguna, Universidad de La Laguna.

Geeraerts, Dirk. 2010. *Theories of lexical semantics*. Oxford: Oxford University Press.

González Orta, Marta. 2002a. Linking syntax and semantics in Old English verbs of warning. *Estudios Ingleses de la Universidad Complutense* 10. 157–182.

González Orta, Marta. 2002b. The syntax and semantics interface of present-day and Old English speech verbs: *Say* and *tell* versus *secgan* and *tellan*. *Journal of English Studies* 3. 81–98.
González Orta, Marta. 2003. The Old English verbs of smell perception and emission: Analysis of the interface of their semantic and syntactic representation. *Selim* 12. 33–48.
González Orta, Marta. 2006. The resultative construction in Old English: Towards a semantic network of verb classes. *Studia Neophilologica* 78(2). 123–137.
Levin, Beth. 1993. *English verb classes and alternations*. Chicago: University of Chicago Press.
Martín Arista, Javier. 2001. Sintaxis medieval inglesa (1): Complementación, caso y sintaxis verbal. In Isabel de la Cruz Cabanillas & Javier Martín Arista (eds.), *Lingüística histórica inglesa*, 224–312. Barcelona: Ariel.
Martín Arista, Javier. 2016. Syntactic and lexical active accomplishments with Old English verbs of motion. Paper presented at the 49th Annual Meeting of the Societas Linguistica Europaea, University of Naples Federico II, 31 August–3 September.
Sosa Acevedo, Eulalia. 2007. The semantic representation of Anglo-Saxon (*ge*)*séon* and (*ge*)*lócian*: Syntactic evidence for meaning decomposition. *RÆL: Revista Electrónica de Lingüística Aplicada* 6. 92–107.
Sosa Acevedo, Eulalia. 2009. Lexical classes and the conative construction in Old English. *Studia Anglica Posnaniensia* 45. 69–90.
Van Valin, Robert. 2005. *Exploring the syntax-semantics interface*. Cambridge: Cambridge University Press.
Van Valin, Robert. 2014. *Some questions concerning accomplishments*. Lecture delivered at the 2014 Symposium on Verbs, Clauses, and Constructions (22th October 2014). University of La Rioja, Spain.
Van Valin, Robert & Randy LaPolla. 1997. *Syntax: Structure, meaning and function*. Cambridge: Cambridge University Press.
Vera Díaz, Elba. 2005. Semántica y sintaxis de los verbos de color en inglés antiguo. *Revista de Filología* 23. 319–334.

# Section 3: **Language and language variety boundaries**

Anatoly Liberman

# 8 Germanic /r/ as an isogloss, rhotacism, and the West Germanic gemination

## 1 The isolated phoneme /r/ as an isogloss and boundary maker

It is the purpose of this paper to trace the history of the phoneme /r/ in Germanic and show how it acquired its present-day realization in British and American English. Many events will have to be discussed along the way: rhotacism, which resulted in a clash between old (inherited) /r/ and new /r/ (from */z/); the enigmatic resistance of /r/ of either origin to lengthening; and its final integration into the system of resonants. Solutions to three riddles will be offered, namely, what triggered rhotacism, what caused West Germanic lengthening, and why /r/ remained short before /j/. It will also be shown how a phonemic approach to sound change may provide answers where attention to purely phonetic factors fails to do so.

To an outsider, no feature in the phonetic makeup of a modern Indo-European language defines its auditory character more strongly than the articulation of [r]. From the phonological point of view it may not be too important whether one has mastered the Parisian uvular [r], the Russian trill, or the British English "liquid" [r], as opposed to American [r], but the resulting sound colors the speaker's accent in a decisive way; it is a classic shibboleth. Although, in theory, phonology studies only distinctive features and leaves realization to phonetics, the production and distribution of /r/ are among the most important and controversial questions of Indo-European historical phonology. One notes that Armenian and Classical Greek lacked initial /r/ (see the discussion of this evergreen question in Lehmann 1951: 14), that /r/ often sides not with /l/, which is close to it in articulation and perception, but with /h/ and /w/ in forming an environment for phonetic change, and that, like perhaps no other phoneme, it tends to disappear after vowels, even if not without a trace.

To complicate matters, /r/ serves as a magnet for other phonemes. Rhotacism (that is, any sound becoming /r/) is a near-universal phenomenon. One need not think of it exclusively in Germanic-Latin terms (/s/ > /r/), for a look at the languages of the world shows that almost any consonant can change to /r/.

**Anatoly Liberman,** Department of German, Nordic, Slavic, and Dutch, University of Minnesota, Minneapolis, MN, USA, e-mail: aliber@umn.edu

https://doi.org/10.1515/9783110643282-009

Here is a fairly isolated example from Standard English (in dialects, many more such examples will be found). The word *pottage*, remembered today only from the biblical phrase *a mess of pottage*, eventually developed into *porridge*. The intermediate stage, namely *podech* and *podditch*, has been recorded. Of the same origin is /r/ in the "vulgar" forms *geron* 'get on' and the like (*ODEE*).

Apparently, in such words, intervocalic /t/ was weakened into a sound of the type known from American English *writer*, *seated*, and *title*, which are indistinguishable from *rider*, *seeded*, and *tidal*, and then underwent rhotacism. Similar forms occur in a Hessian dialect of German (Howell 1987: 330, note 15). An instructive array of facts can be found in Dzendzelivs'kij (2000). In the Ukrainian dialects he surveyed, /l v d j l n/ and /t/ occasionally yield /r/. Some cases are due to dissimilation and more or less unpredictable factors. Yet /l/ > /r/ is common, the change of /ʒ/ (that is, /ž/) to /r/ in the root /boʒ/ 'God' finds, as we will see, a parallel in Germanic, and /d/ > /r/ is ubiquitous, though Ukrainian /r/ is a trill. In light of such facts, one should beware of describing rhotacism as a simple, even predictable, shift from one distinctive feature to another, as in /s/ to /r/ (this is what Touratier 1975 tried to show for Latin). Isolated examples of the Ukrainian type occurred also in Old Norse, where /l/ sometimes turned into /r/, and perhaps /ð/ did in the attested form *guðrs* 'of god' (Noreen [1923] 1970: sec. 238/3).

In the history of /r/ as a boundary maker, the main event was rhotacism, that is, the change of Germanic /s/ to /r/. It drove a wedge between Gothic, which lacks this phenomenon, and the rest of Germanic, but, even within Germanic, some isoglosses look puzzling and create barriers. Germanic **raisjan*, the causative of *rīsan* 'to rise', developed according to the expected pattern: since *-jan* verbs had stress on the suffix, /s/ acquired voice by Verner's Law: **raisján* > **raizján* (voicing) > **rairjan* (rhotacism) and English *rear* (verb). But for reasons that have never been explained or even discussed in grammars and etymological dictionaries, in Old Norse, this form escaped the influence of Verner's Law or, perhaps most uncharacteristically, changed /r/ to /s/ on the analogy of the past singular. Hence we find Old Icelandic *reisa*, which was borrowed into English as *raise*. Consequently, *rear* and *raise* are etymological doublets. As usual in such cases, their meanings diverged, in order not to be confused, but we can still raise children and rear them. To summarize: the phoneme /r/ carves the linguistic map of Germanic, with serious consequences for English, like a knife, but not always according to expectation, and thereby produces uneven edges and corners. One of the questions confronting the researcher is what makes /r/ constantly side with /h/ and /w/ in providing an environment for phonetic change (particularly, all kinds of breaking in Germanic). Regardless of whether [h] is a pharyngeal consonant or a voiceless

vowel, from the systemic point of view, it is a phoneme devoid of any distinctive features (unless we give "pharyngeal" the status of such a feature). In language history, [h] is the last stage in any spirant's disappearance.

The phoneme /r/ looks better integrated into the system, but this integration is illusory. The illusion owes its origins to the fact that, unlike /h/, /r/ is voiced and fully audible, while /h/ is realized as breath. True, /r/ is a resonant, which means that, given a certain type of environment, it can be syllabic. Other than that, its phonological and phonetic descriptions coincide (for example, a trill). Systemic isolation unites /r/ with /h/, and it is the lack of integration typical of /r/ that causes its instability.

Also, the third member of the triad, /w/, shares the features of /h/ and /r/, or rather the lack of features, because phonemic characteristics are the product of oppositions and correlations. Like /r/, /w/ too can be syllabic (/w/ is usually called a semivowel), but, to perform this function, it has to be vocalized and become /u/. All three phonemes – /h, /r/, and /w/ – are "weak" not because, or at least not only because, of their idiosyncratic realization but on account of their marginal status in the system. They frequently affect preceding vowels and, by doing so, make a decisive step toward disappearance (yield up their characteristics and merge with their neighbors). As usual, the intermediate stage may be retained for a long time during which the conservative and the avant-garde forms coexist. Important phonemic changes sometimes take centuries to realize their full potential. Needless to say, the progress of a change can be reversed.

The phonemes /r/, /h/, and /w/ can be called peripheral (the most incisive work on such "outsiders" has been done by Josef Vachek – see Vachek 1976). Not unexpectedly, they tend to cling to one another. It is customary to distinguish between syntagmatic and paradigmatic changes. A paradigmatic change presupposes the addition or subtraction of some element(s). A syntagmatic change results in the redistribution of the elements present in the system. The foregoing remarks addressed the paradigmatic status of the three phonemes, but /r/, /h/, and /w/ also form syntagmatic unions. Thus, in Old Germanic, we find the initial groups /hr/ and /hw/ in addition to /hl/ and /hn/; /hw/ still exists in the idiolects of many English speakers. Another common group was /wr/, whose traces remain in Modern English spelling (cf. *wright*, *wring*, and so forth). The aspirated Greek *ro* (*ṕ*) is another well-known example of [hr]. Of note is the common substitution of [w] for [r] in children's and in adults' affected speech. The simpering pronunciation of [w] instead of [r] by nineteenth-century British "swells" was the target of sustained ridicule by Dickens, the *Punch*, and many others (see, e.g., Horn and Lehnert 1954, 2: 947–48). In baby talk, [w] replaces [r] even in such languages that have no phoneme /w/ and regardless of the realization of /r/ (Wollock 1982: 202).

The foregoing remarks aimed at showing that /r/, which belongs with such poorly integrated phonemes as /h/ and /w/, is weak (devoid of distinctive features) from the phonological point of view. Its articulation may not always be hard to master. It is its systemic weakness (its lack of integration) that explains why, along with /h/, it so often appears as the last step on other consonants' path toward disappearance and why it itself is often so unstable.

## 2 The phoneme /r/ and Germanic lenition

The history of /r/ in English and elsewhere in Germanic is part of *lenition* (or weakening), a major process that began with the First Consonant Shift. The reconstructed */$b^h$ $d^h$ $g^h$/ yielded fricatives in Germanic, and fricatives are weaker than stops by definition. The phonemically voiceless Indo-European */p t k/ acquired aspiration and, as a result, also lost force (the greater the aperture in the articulation of a sound, the less effort it takes the speaker to produce it). A sound change of such dimensions can be likened to an avalanche or a river winding its way to the valley, or a stone rolling downhill. The direction of the movement is obvious, but the details are not, for many of them are unpredictable. The "river" or the "stone" may encounter obstacles and change its route. Besides, in language history, we are dealing with both mechanical pressure and numerous sociological factors, most often conservative (the intuitive or conscious striving to resist change). Therefore, as already noted, a change can be reversed, contrary to a river that will not flow upwards, even when it is made to choose a circuitous way. On the whole, a historical linguist always faces an interplay of choice and chance. Rhotacism (/s/ > /z/ > /r/) is a classic case of lenition, but the zigzags of the "river" deserve our attention for their own sake and because they too produced new boundaries and set English as part of West Germanic in opposition to Scandinavian.

Sound change, which is a physical process in that it depends on phonation, first affects the positions that offer it the least resistance and may even stall there. All fricatives were affected by lenition, but we also observe so-called exceptions. As is known, in Old English, intervocalic fricatives underwent voicing: hence *rīsan* [-z-]; by contrast, in Old Norse, /s/ remained voiceless: *rísa* [-s-] 'to rise'. I have not been able to find any explanation of this phenomenon in the works I consulted. Even today, intervocalic voicing is sometimes referred to as assimilation: the vowels on both sides of the fricative are voiced, and the consonant allegedly succumbs to their influence. However, the phonetic context, far from being able to produce change, only facilitates or hinders it.

Supposing that an analogy would help: water is necessary for the existence of fish; yet fish need more than water for their appearance and survival.

It was probably Verner (1877) who was the first to recognize the role of lenition in the history of Germanic, but his main discovery made such an impression that his minor observations remained partly unappreciated. Since Verner's Law deals with the voicing of fricatives, it is natural that Verner referred to the weakening of articulation. Although later some scholars recognized lenition as a dominant trend in the history of Germanic, they used it only as background for their reasoning.

One of such scholars was A. I. Smirnitskij. In 1959, his long article on rhotacism and the loss of final */z/ in West Germanic appeared posthumously. Since the article was written in Russian, it had minimal, if any, readership among Germanic scholars in the West. Smirnitskij's book on Old English (1955, also posthumous), put together from his lecture notes, contains a chapter on the same subject and is available in English (Smirnitskij 1990), but its focus is on the chronology of rhotacism and the proof that the so-called loss of final */z/ in West Germanic was really the loss of */R/, the immediate product of rhotacism, as known from Old Norse. In the 1959 article, the chronology of rhotacism is discussed at great length. In recent times, lenition has been at the center of Kurt Goblirsch's research since the 1990s; he devoted a series of articles on this subject and a book (Goblirsch 2018). To conclude, rhotacism is a product of Germanic lenition.

## 3 Rhotacism

Rhotacism resulted in major changes on the linguistic map of Germanic, but it also shifted boundaries between individual words. The clash between the English verbs *rear* and *raise* has already been discussed. Here is another example, this time of a merger rather than of a split. Two homonyms met in Old English: *werian* 'to defend' (Gothic *warjan*) and *werian* 'to dress' (Gothic *wasjan*). Both have /e/ in the root by *i*-umlaut. But in *werian* 'to dress', /r/ emerged by rhotacism, while /r/ in *werian* 'to defend' is old.

The point to make here is that both steps of rhotacism – */s/ > */z/ and */z/ > /r/ presuppose weakening. This follows from the history of /r/ in Old Norse. In that language, the product of rhotacism was designated by the rune R; /r/ and /R/ coexisted for a long time. The unrecorded existence of the phoneme of the Scandinavian type in West Germanic has always been recognized, and disagreement concerns only the length of the period it maintained its independence from /r/. As we have seen, according to Smirnitskij (1959), it was */R/, not /z/,

that West Germanic lost. His conclusion appears well-argued, but chronology need not concern us here. It will be sufficient to agree with him and accept the fact that the process */s/ > */z/ and */z/ > /r/ should be presented as */s/ > */z/, */z/ > /R/, and /R/ > /r/ for all Germanic, though for West Germanic /R/ needs an asterisk (*/R/).

The phonetic nature of Old Germanic /r/ and of the new /r/ < */z/ has been a matter of protracted debate. Yet Howell's (1987) conclusion seems persuasive enough: common Germanic */r/ was, most likely, realized as an apical (that is, not as a uvular) sound (see also Manganella 1958: esp. 150; also van Haeringen 1962). The loss of apicality, where it did happen, must have occurred in the course of the individual history of the Germanic languages. Also, though Wollock (1982) deals mainly with the development of /r/ in German and French, it contains many stray remarks on other languages, and, as a general rule, uvular /r/ appears to be secondary wherever it occurs.

It won't be an exaggeration to say that /r/, a peripheral phoneme, underwent weakening and tended to disappear or at least come as close to the brink as possible. The result was the creation of additional isoglosses and boundaries. In the context of the present discussion, the details concerning the realization of Germanic */r/ are of prime importance, because we have to determine the character of the product of rhotacism and solve the riddle of the typologically rare Modern English /r/, whether in its British or American variety. Our point of departure is, of necessity, Old Scandinavian.

According to universal belief, /R/ < */z/ resembled Czech /ř/, an affricate combining the elements of /r/ and /ʒ/. Regardless of the truth of this statement (which will be discussed later), it should be repeated that /R/ must have been a weakened variant of old */r/. Equally important is the fact that /R/ must have been a palatalized consonant. Verner (1877) was again the first to say so, though, quite naturally, he did not speak of phonemes and distinctive features (see Steblin-Kamenskij 1963: note 18). (Among other things, that note is important because it contains references to the works on the phonetic realization of old */r/, not included in Howell's [1987] extensive bibliography.) In the days of Verner, phoneticians often used the term *palatal* for *palatalized*, which is a misnomer. The difference in usage goes beyond terminological niceties. Palatal is a nonbinding phonetic label, while, palatalized, at least today, presupposes the existence of a certain distinctive feature in Old Germanic. Yet few researchers are willing to recognize such a feature. We cannot avoid the problem, because in Old Norse, /R/ caused palatal umlaut.

In the second half of the nineteenth century, several leading scholars believed that palatalized consonants had played an important role in the progress of *i*-umlaut, but umlaut occurred in some words that had no intervening consonants

between the root vowel and /i j/. This circumstance was used as proof against the existence of palatalization, though it only meant that *i*-umlaut could not be reduced to a single formula. However, the anti-palatalization school won (its main argument was the alleged absence of palatalized consonants in Modern Germanic), and its victory did great damage to the progress of Germanic studies (see a detailed discussion of this controversy in Liberman 2007).

Curiously, despite the "official" viewpoint, palatalization was smuggled into all major Neo-Grammarian textbooks as a factor in some changes, especially with regard to velars (Luick 1964: 835–841 is a typical example), but nowhere can we find an explanation of where that feature came from. We are simply told that /k/ and /g/ underwent palatalization in Anglo-Frisian. However, if such a distinctive feature did not even exist in Old English and Old Frisian, what made /k/ and /g/, supposedly always palatalized before front vowels, yield affricates? What caused the change at the end of the fourth century (so Luick 1964: sec. 637, note 8)? Were the products of palatalizations new phonemes? But if the larger aspect of palatalization is disregarded, the change becomes a total mystery.

Since the days of the old polemic, modern Dutch and German dialects, including many southern ones, have been explored in great detail. The existence of palatalization in them is an indubitable fact, and so is its history, going back to the oldest period. I devoted a long essay to this question (Liberman 2007) and see no need to retell it here, the more so as, to the best of my knowledge, no one has objected to or refuted my conclusions.

It is instructive to observe how the refusal to accept distinctive palatalization in Old Germanic hampers the effort of some of the best scholars to understand the nature of /R/. Smirnitskij (1959) noted that, since old /r/ caused breaking in Old English, it must have been velar. /R/, he continued, might be "velarized enough", though remaining fricative at the epoch of breaking. In Old Norse, he continued, /R/ was palatalized, as follows from its ability to cause umlaut; allegedly, it acquired palatalization at a later period, for the change of /i/ to /e/, as in Old Icelandic *mér* and Old English *mē* 'me' before a palatalized consonant would be hard to admit. Like everybody else before him, Smirnitskij (1959) took it for granted that breaking had been caused by the velar character of *h* [x], *w*, *r*, and *l*. We do not know whether Germanic /x/ was still velar (rather than /h/) at the time of breaking, but it should be repeated that /h/, /w/, and /r/ can be called velar phonemes only if we recognize that they were opposed to their non-velar, or palatalized, partners. For speakers to become aware of assimilation and reflect its results in spelling, the feature that causes the change has to be distinctive. The term "velarized enough" and reference to becoming palatalized some time later take us nowhere. A distinctive feature cannot be graduated, and it cannot appear suddenly from nowhere.

Steblin-Kamenskij (1963) faced a similar difficulty. He emphasized the fact that, since /z/ and /R/ are pronounced with the tip of the tongue more to the front than happens in the articulation of a non-sibilant /r/, the phonemes /a o u/ before /R/ must have had more advanced allophones. Those, he suggested, were later replaced by the phonemes /æ ø y/. Without referring to Smirnitskij's (1959) statement, but certainly having it in mind, he noted that the change /i/ > /e/ had taken place only in unstressed positions and therefore did not run counter to the direction of *R*-umlaut. The idea of allophones of one phoneme being drawn into the sphere of another phoneme has little to recommend it, because we cannot know anything about the minute details of ancient realizations.

In my view, the following scenario is more realistic. In the process of rhotacism, */z/ yielded /R/. Since an *r*-like rune was chosen for it, [R] must have been perceived as an *r*-like sound: resembling but different from [r]. By the time of rhotacism, Germanic had distinctive palatalization, and /R/, a new phoneme, to support its independence, became palatalized. In its choice, it had little leeway, for, as shown above, /r/ is (and was) peripheral, or featureless. For the first time ever, it acquired a distinctive feature: it became distinctively velar, as opposed to the newcomer /R/, distinctively palatalized.

The existence of two *r*-like phonemes must have been hard to maintain, and they merged (a common situation in language history [cf. the short coexistence of open and closed *ē* in Middle English]). In Old Norse, /r/ won out. I doubt that /R/ resembled Czech /ř/: /R/ need not have been an affricate (just an open, weak [r], as opposed to the inherited trill). I will risk going even further. Not improbably, Modern American /r/ is a continuation of /R/. If so, in English, /R/ ousted its rival, while German went the Scandinavian way. In British English, it has undergone further weakening, obedient to the ever-going process of lenition. The setting up of palatalized /r/ in Old Norse and, by implication, in Old English, will go a long way toward elucidating one of the pivotal moments in the history of the West Germanic *gemination*.

## 4 The phoneme /r/ and the West Germanic gemination

These are some of the best-known examples of the West Germanic gemination, as we know it from Old English: *sellan* 'to give, lend', *scieppan* 'to create', *settan* 'to set', *hliehhan* 'to laugh', and so forth (Gothic *saljan*, *skapjan*, *satjan*, *hlahjan*). Consonants were lengthened not only before /j/ but also before /r/ and /l/, as in *bittor* 'bitter', *æppel* 'apple', etc. Three questions arise in connection with this

change. 1) Why did it happen? 2) Are we dealing with a single process, or did the lengthening before /j/ and before the resonants have different causes? 3) At the beginning of the present paper, Old English *werian* 'to defend' and *werian* 'to dress' were cited. Both are *-jan* verbs (cf. Gothic *warjan* and *wasjan*). Why wasn't /r/ lengthened?

In Liberman (2007), devoted to palatalized and velarized consonants in Germanic, I dealt with verbs like Old English *settan* < **satjan*. In their Old Saxon cognates, both the double letter and <i> appear. I suggested that, at one time, such verbs had palatalized postvocalic consonants and that after the loss of palatalization the consonants underwent compensatory lengthening. In Old Saxon, <i> might have been a graphic sign of the extant or former palatalization. The change that affected Old English and Old High German, though it drove a wedge between West Germanic and Scandinavian and resulted in the appearance of an important isogloss, did have a limited parallel in Old Norse, but there only two velars were affected, as in *liggja* 'to lie, to repose' and *ekkja* 'widow' (this change is usually called *gi- ~ ki*-umlaut). In Modern Icelandic, the spelling remains the same, but <j> designates only palatalization. *Liggja* still retains a geminate, while in *ekkja* part of the consonant length went over to preaspiration: thus, [ligg'a] and [$e^h$k'(:)a] (simplified transcription).

In the vast literature devoted to the West Germanic gemination, palatalization is almost never considered for the reasons given above: the great masters of the Neo-Grammarian school lost interest in this feature, and to this day palatalization remains a "subject of interest" only in dialectological studies. Paul (1880: 106) mentioned *mouillierung*, but did not develop the subject. Other than that, the causes of the gemination remain unknown.

At one time, the influence of circumflex was conjured up to explain the process (Sievers 1878: 161–162). The fashion to reconstruct the Old Germanic acute, grave, and circumflex attracted the greatest specialists, such as Paul, Morsbach, and Luick (to say nothing about their students) and resulted in numerous bold but unprovable hypotheses (see a detailed discussion of this episode in the history of Germanic philology in Liberman 1979). Those accents, whose existence could not be demonstrated on independent grounds, that is, typical ad hoc concepts, were introduced, phlogiston-like, to account for many puzzling phenomena. A similar phantom, also created by Sievers (1878), is the presumed shifting of syllable borders. It too was allowed to appear for the sole reason of explaining the change. Luick shared this view, and it still has some advocates. Braune (2004: 98–99) offers a fairly exhaustive bibliography of the subject, bypassing some obsolete hypotheses rejected already by the time of Sievers and Paul. In his discussion, he is thoroughly non-committal. An important note on consonant lengthening and syllabification is Sievers (1892). In it, Sievers developed his idea

that the clue to gemination should be sought in the details of syllable division, as he saw it.

As regards the resistance of /r/ to lengthening, only two explanations seem to have been offered. Wilmanns dealt with this question many times. However, in the final version of his compendium (Wilmanns 1911: 189), he repeated the suggestion made decades earlier. He believed that after /r/ a weak epenthetic vowel had arisen. Allegedly, an Old High German word *ferio* (pronounced as *ferjo*) 'ferryman' became *ferijo*; consequently /r/ found itself separated from /j/ and stayed short. It is true that at least in Old High German, the realization of /j/ after /r/ may have differed from its realization elsewhere (Braune 2004: sec. 118, note 3), but no evidence for the epenthesis exists. Even if Wilmanns's (1911) idea had merit on independent grounds, it would be necessary to clarify the chronology of the change: did the epenthesis arise before the gemination, or had it always existed? Finally, as regards the treatment of semivowels, Old High German and Old English went different ways.

According to Boer (1918: 201), in words like Old English *nerian* 'to save' (from **nerjan*; cf. Gothic *nasjan*), /j/ after /r/ became /i/ before the onset of the gemination and blocked the change of /r/ to /r:/. To justify his reconstruction, Boer (1918) referred to the influence of Germanic accents, especially the circumflex. Since, as indicated above, such accents are phantoms, there is no need to dwell on this point. This approach alone invalidates his reasoning. But his use of relative chronology is also unacceptable. The Neogrammarian "reified" chronology, to use Henning Andersen's word (personal communication). Historical changes were represented as trains: one train left just before another one arrived. If we set up a certain order of events, everything will look plausible, but when that order cannot be substantiated, the conclusion loses its appeal.

My own hypothesis is much less speculative, but it presupposes the existence of distinctive palatalization in West Germanic. I assume that Old English *sellan*, *scieppan*, *settan*, etc. passed through the stage when they had /l' p' t'/, etc. (palatalized). For physiological reasons [r'] cannot be lengthened without losing its articulatory qualities. The palatalized trill disintegrates at the first attempt to extend its duration. The American [r] can be palatalized (though not without an effort), but, when prolonged, it yields a [ʒ]-like noise. The situation described above is not unique. Sievers (1901: sec. 564) observed that in Sanskrit and Greek, aspirated consonants were never lengthened, because their lengthening would have produced [h:] and destroyed the stop. The special treatment of /r/ created a small distributional island and drew unpredicted boundaries: a small group of words, those with /r/ after a short vowel, resisted a sweeping change, known as the West Germanic gemination.

Quite obviously, the geminates in the words *bittor* and *æppel*, cited at the beginning of this section, that is, words with final resonants, had nothing to do with palatalization. If we disregard references to the phonetic phlogiston and the *deus ex machina*, be it the circumflex or a shift in syllable boundaries, we should agree that no one has yet explained why those stops underwent lengthening. I will attempt to lay bare the cause of the lengthening, fully realizing how much is needed to make my case persuasive.

Possibly, the gemination in question is connected with Germanic lenition. A word like *æppel* could be expected to undergo lenition with the rest of the consonants, and it sometimes did. As a rule, our records of forms like *abl* do not predate late Middle English, but such forms occur in many dialects and cannot be recent (see *able* 'apple', *kedle* 'kettle', *lidle* 'little', *bodm* 'bottom' [*bodm* already existed in Old English!], etc., in Wright [1905: sec. 275 and 283], Luick [1964: sec. 799, note 1], and Horn and Lehnert [1954: 999]). The form *pebble* (< *papel* ~ *popel*) made it into the Standard and belongs with *pottage* > *podech* > *porridge*, discussed in section 1.

We are used to saying, in the spirit of Roman Jakobson and André Martinet, that phonetic change receives impulses from the system (empty slots, push-chains, drag-chains, etc.), but it should also be remembered that every sound change, whatever its cause, disrupts the process of communication, and the system takes on the role of preserving the *status quo*. This is the cause of all compensatory processes.

I would like to suggest that forms like Old English *æppel* and *bittor* resisted lenition and for this reason reinforced the endangered consonants, which, if weakened, might have disappeared altogether. To put it differently: the gemination of stops before resonants seems to have been a weapon against lenition. Such a process could not be fully consistent. That is why we often find variants: *bittor* and *bitor*, and so forth; however, on the whole, the "weapon" worked well. I am venturing this explanation in the hope of resuming the old debate. Even Karl Luick used to put forward tentative solutions of whose value he was not sure. Perhaps we, who owe him such a debt of gratitude, can also be allowed to follow the example of the great scholar and offer risky conjectures without the fear of being laughed out of court.

## 5 Conclusion: Old and modern boundaries

As we can see, rhotacism, most likely, had as its consequence the existence of the typologically rare English resonant /r/ and different realizations of British

and American /r/. For learners of English few boundaries are more noticeable and harder to overcome. Consonant lengthening before /j/ occurred not only in West Germanic, but only there did it have such drastic results. For centuries, until the disappearance of geminate, in all of West Germanic, the West Germanic consonant lengthening served as an important boundary between the West and the North. Phonemic and phonetic boundaries arise and disappear, and their history determines the shape of the modern languages as it did in the remote past.

## References

Boer, R. C. 1918. Syncope en consonantengeminatie. *Tijdschrift voor Nederlandsche Taal- en Letterkunde* 37. 161–222.

Braune, Wilhelm. 2004. *Althochdeutsche Grammatik I*. 15th edn. [Ingo Reiffenstein (ed.).] Tübingen: Max Niemeyer.

Dzendzelivs'kij, J. O. 2000. "Iz sposterezhen' nad rotatsizmom v ukraïns'kij movi" ["From the Observations about Rhotacism in Modern Ukranian"]. *Movoznavstvo* 6. 10–22.

Goblirsch, Kurt. 2018. *Gemination, lenition, and vowel lengthening: On the history of quantity in Germanic*. Cambridge: Cambridge University Press.

van Haeringen, Coenraad B. 1962. Invloed van *r* op klinkers in germaanse talen. In Coenraad B Van Haeringen, *Neerlandica. Verspreide opstellen*, 2nd edn., 84–107. Den Haag: Daamen.

Horn, Wilhelm & Martin Lehnert. 1954. *Laut und Leben. Englische Lautgeschichte der neueren Zeit (1400–1950)*. 2 vols. Berlin: Deutscher Verlag der Wissenschaften.

Howell, Robert B. 1987. Tracing the origin of uvular *r* in the Germanic languages. *Folia Linguistica Historica* 7(2). 317–349.

Lehmann, Winfred P. 1951. The Distribution of Proto-Indo-European /r/. *Language* 27(1). 13–17.

Liberman, Anatoly. 1979. Prephonological views on the history of English syllable accents. In Harry Hollien & Patricia Hollien (eds.), *Current issues in the phonetic sciences. Proceedings of the IPS-77 Congress, Miami Beach, FL, 17–19 Dec. 1977*, 321–329. Amsterdam: John Benjamins.

Liberman, Anatoly. 2007. Palatalized and velarized consonants in English viewed against their Germanic background, with special reference to *i*-umlaut. In Christofer M. Cain & Geoffrey Russom (eds.), *Studies in the history of the English language. Managing chaos: Strategies for identifying change in English*, 5–36. Berlin: Mouton de Gruyter.

Luick, Karl. 1964. *Historische Grammatik der englischen Sprache*. Stuttgart: Bernhard Tauchnitz. [With additions by Friedrich Wild & Herbert Koziol.].

Manganella, Gemma. 1958. Il valore di *h* e *r* nei dialetti antico-germanici. *Annali del'Istituto Universitario Orentale di Napoli [AION]. Sezione Germanica* 1. 139–151.

Noreen, Adolf. 1970 [1923]. *Altnordische Grammatik I*. 5th edn. Tübingen: Max Niemeyer.

*ODEE* = Onions C. T. (ed.). 1966. *The Oxford Dictionary of English Etymology*. Oxford: Clarendon Press.

Paul, Hermann. 1880. Beiträge zur geschichte der lautentwicklung und formenassociation. 4. Die westgermanische consonantendehnung. *Beiträge zur Geschichte der deutschen Sprache und Literatur* 7(1). 105–170.

Sievers, Eduard. 1878. Zur accent- und lautlehre der germanischen sprachen. *Beiträge zur Geschichte der deutschen Sprache und Literature* 5(1). 63–163.
Sievers, Eduard. 1892. Grammatische Miscellen. 4. Zur westgermanischen Gemination. *Beiträge zur Geschichte der deutschen Sprache und Literatur* 16. 263–265.
Sievers, Eduard. 1901. *Grundzüge der Phonetik*. 5th edn. Leipzig: Breitkopf & Härtel.
Smirnitskij, A. I. 1955. *Drevneangliiskii iazyk* [The Old English Language]. Moscow: Izdatel'stvo literatury na inostrannykh iazykakh. [V. V. Passek (ed.)].
Smirnitskij. A. I. 1959. Otpadenie konechnogo *z* v zapadnogermanskikh iazykakh i izmenenie *z* v *r* [The Loss of Final *z* in West Germanic and the Change *z* > *r*]. *Trudy instituta iazykoznaniia* IX. *Voprosy germanistiki*, 115–136. Moscow: Izdatel'stvo Akademii nauk SSSR.
Smirnitskij. A. I. 1990. Rhotacism in Old English and the West Germanic loss of *z. *General Linguistics* 30. 197–206. [Translated by Ian Ritchie & Anatoly Liberman.]
Steblin-Kamenskij, M. I. 1963. The Scandinavian rhotacism and laws governing the change of distinctive features. *Philologica Pragensia* 6. 362–367.
Touratier, Christian. 1975. Rhotacisme Synchronique du latin classique et Rhotacisme Diachronique. *Glotta* 53 (3/4). 246–281.
Vachek, Josef. 1976. On the Integration of the Peripheral Elements into the System of Language. In Josef Vachek, *Selected writings in English and general linguistics*, 77–90. Prague: Academia; The Hague & Paris: Mouton.
Verner, Karl. 1877. Eine Ausnahme der ersten Lautverschiebung. *Zeitschrift für vergleichende Sprachforschung auf dem Gebiete der indogermanischen Sprachen* 23(2). 97–130.
Wilmanns, Wilhelm. 1911. *Deutsche Grammatik*. 3rd edn. Straßburg: K. J. Trübner.
Wollock, Jeffrey. 1982. Views on the decline of apical *r* in Europe: Historical study. *Folia Linguistica Historica* 3(2). 185–238.
Wright, Joseph. 1905. *The English dialect grammar*. Oxford: H. Frowde.

Angela Hoffman and Merja Kytö

# 9 Migration, localities, and discourse: Shifting linguistic boundaries in Swedish-American cookbooks

## 1 Introduction

Migration sets into motion dynamic forces in language contact, influencing the majority language of a new host community as well as the linguistic forms used by the immigrant community (cf. Schneider 2003, 2007).[1] Many immigrant languages – among them, Swedish in America – have survived for generations as *heritage languages*, that is, languages separated from the region of their origins and transplanted to a new location. Heritage languages in the U.S. are typically learned in family settings, rather than through guided, formal instruction in schools in the majority community. Further, heritage languages frequently bear traces of ongoing language contact with the majority language of the host community (see also Wilkerson and Salmons 2008; Rothman 2009; Johannessen and Salmons 2015; Page and Putnam 2015). Quite a number of texts revealing evidence of language contact between heritage languages and American English may be found in local archives of churches and church-related colleges. It is in such archival collections where we located the materials that are the focus for the present study.

Regarding Heritage Swedish in the U.S., we are interested in Swedish-American cookbooks, the vast majority of which contain instances of language mixing between Swedish and English. Before introducing further details about our investigation, we first offer brief explanations of the central terminology in our study. Later in the analytical portion of our chapter, we provide numerous examples of language mixing phenomena. When we use the term *Heritage Swedish* in the

**1** We thank the editors and the anonymous reviewer for their perceptive comments and suggestions on our manuscript. For assistance in locating cookbooks, we thank Darroll Atwood, Lanis Atwood, Dag Blanck, Samantha Crisp, Rebecca Nirider Dunlap, Gunilla Grügiel, Carol Gusenius, Arlis Henry, Barbara Hoffman, Lisa Huntsha, Muriel Johnson, Bruce Karstadt, Greg Mayer, Judy Mayer, Ronald Peterson, Jill Seaholm, and Judit Söderblom. We are grateful to Gudmund Danielson, Renée Lund Danielson, Birgitta Grape, Karin Lindegård, and Lars Åstrand for answering questions about traditions in Swedish congregational cookbooks and kitchens. We thank Åke Eriksson for verifying some of our translations.

**Angela Hoffman, Merja Kytö,** Department of English, Uppsala University, Uppsala, Sweden, e-mail: angela.hoffman@engelska.uu.se; merja.kyto@engelska.uu.se

https://doi.org/10.1515/9783110643282-010

present study focused on printed materials, the heritage phenomena are limited to orthography and semantics that diverge from Standard Swedish. We use the term *matrix language* in a theoretically neutral sense to refer to the language that is dominant in the discourse (cf. Myers-Scotton 2002). American English is the matrix language in the cookbooks we examine, and varieties of Swedish (Heritage and Standard Swedish) comprise the embedded non-English lexical morphemes inserted into the discourse. We use the terms *code-mixing* and *code-switching* to refer to the instances when writers embed Swedish lexical items in the discourse. We use the term *code-mixing*[2] to refer to points at which Swedish lexical items appear in a phrase together with English. *Code-switching* is the term we use to describe longer stretches of Swedish. We use these terms, instead of *borrowing*, to capture the reality that some of the cooks represented in the community cookbooks are/were native speakers of Swedish, and many are/were proficient users of Heritage Swedish. Meanwhile, however, other recipe contributors, that is, third- and fourth-generation Swedish-Americans, had/have comparatively little active proficiency in Swedish, sometimes limited to lexis pertaining to food and Swedish table prayers. For this latter group of Swedish-Americans, it may be relevant to describe Swedish as their *postvernacular variety*; in other words, the Swedish language for them is post-vernacular in the sense that "the instrumental value [. . .] for communicating information [. . .] is narrowing in scope" even while the "symbolic value [. . .] is expanding" (Shandler 2006: 4; see also Shandler 2004; Benor 2015: 218–219). The migration history of Swedes in the communities we investigate extended across several decades, with the consequence that Swedish-born immigrants settled in towns and cities where second- and third-generation Swedish-Americans also were living. We surmise that the language expertise of Swedish-born cooks has helped the second and subsequent generations of Swedish Americans in expressing cultural concepts (cf. Karstadt 2003).

We focus on contexts where the lexis of English and Swedish have cohabitated in texts from Kansas, Minnesota, and Illinois, investigating whether and how proportions of the languages have shifted over time. Exploring these processes is important so that we learn about the chronology and stages of the language shift and of the factors that have influenced such change. From previous socio-historical research of documents produced in communities in these states (Hoffman and Kytö 2018), we have observed that language shift from Swedish to English moved at different speeds depending on the domain (i.e., religious vs.

**2** Regarding one of the functions of code-mixing, Muysken (2000: 278) notes that "[c]ode-mixing provides bilingual speakers with a means to drastically extend their verbal repertoire, and there is evidence that many bilingual speakers take the opportunity".

educational institutions). In this chapter, we aim at presenting a more nuanced picture of language shift development by zooming in on cookbooks, a genre representative of domestic sphere documents, to which we apply the theoretical framework of enregisterment (Agha 2003, 2007; Adams 2009; Johnstone 2016).

The explanatory power of enregisterment helps us understand why certain lexical items appear to gain more visual prominence in the cookbooks. As usual in language contact in immigrant settings, several factors are at play influencing the developments of language shift. Our research questions are therefore the following:

(i) What evidence of language contact can one find in the community cookbooks? Further, to what extent are the boundaries between Swedish and English distinct, and where do the languages overlap?
(ii) What does enregisterment reveal about the use of Swedish and English in the cookbooks?
(iii) What do the naming patterns of recipes in local cookbooks look like across the communities?

Our chapter is organized as follows: we briefly survey the demographics of Swedish settlement in North America, focusing on localities of relevance to our study, in Section 2. We then discuss the language situation from the Swedish-American perspective and introduce and contextualize Swedish-American cookbooks in this sociolinguistic context in Section 3. After surveying the theoretical perspectives of enregisterment to be applied in our analyses in Section 4, we proceed to investigate our data (Section 5). In our concluding section (6), we summarize our main findings and point to prospects of future study.

## 2 Demographics of the early Swedish settlements in the U.S.

Large-scale mass migration from Sweden to the U.S. took place for nearly a century (ca. 1825 to 1930). The Swedish language was sustained by the slightly more than one million Swedish immigrants who formed many Swedish-speaking networks, neighborhoods, and communities across a vast geographical range of the U.S. (Hasselmo 1974; Hedblom 1983; Karstadt 2003; Blanck 2012). Many thousands

of Swedes immigrated to the Midwestern states of Illinois, Minnesota, and Kansas. Table 1 provides a population overview of some of the counties in those states.[3]

**Table 1:** Swedish-born population by county (based upon U.S. Census records reported in Nelson 1943, Rice 1981, and Lindquist 1953).

| **County and *State*** | **1870** | **1880** | **1890** | **1895** | **1900** | **1905** | **1910** | **1920** | **1930** |
|---|---|---|---|---|---|---|---|---|---|
| McPherson *Kansas* | | | 2680 | | 2507 | | 2257 | 1742 | 1246 |
| Hennepin *Minnesota* | | 3524 | | 22480 | | 27126 | 28127 | | 26869 |
| Rock Island *Illinois* | 2061 | 3435 | 4661 | | 5653 | | 5644 | 5191 | 4038 |
| Henry *Illinois* | 5309 | 4769 | 4324 | | 4534 | | 3787 | 2891 | 2106 |
| DeKalb, Kendall, LaSalle *Illinois* | | | | | 2680 | | 2820 | | |

We examine printed materials from McPherson County in Kansas (where Lindsborg and Marquette are located), Hennepin County in Minnesota (where Minneapolis is located), Rock Island County in Illinois (where the city of Rock Island is located), Henry County in Illinois (where Andover is located), and DeKalb/Kendall/LaSalle Counties in Illinois (where the city of Sandwich is located).

The chosen communities vary with respect to the size of their populations. Hennepin County is a large, urban county; two of the counties have rural populations (McPherson County in Kansas and Henry County in Illinois), while medium-sized populations are found in Rock Island, DeKalb, Kendall, and LaSalle Counties. In Hennepin County, Swedish migration was sustained by the many job-market prospects in Minneapolis (and also in its "twin city" of St. Paul, Minnesota; see Rice 1981) and was characterized by chain migration, as many Swedish young adults joined relatives and friends who had previously settled in Minnesota (Ostergren 1988; Karstadt 2003). Tens of thousands of Swedish-born speakers were living in the area, and the Swedish-born population in Hennepin County continued to rise until approximately 1910.

**3** Some of the cells in Table 1 are empty, due to different years of Census records available for different states.

By contrast, the Swedish population in most of the other counties we investigate peaked many decades earlier. McPherson County in Kansas and Henry County in Illinois are rural areas with economies highly dependent on agricultural production (Nelson 1943). In McPherson County, the number of Swedish-born residents decreased after 1890. By the end of the nineteenth century, few new immigrants from Sweden settled in south-central Kansas. The tracts of Kansas prairie farmland previously available through the Homestead Act of 1862 had been claimed by pioneers in the late 1860s; some decades later, the children of the Swedish pioneers were taking over the farms, further reducing the possibility for incoming immigrants to establish farms (Nelson 1943).

## 3 Swedish and English in contact in cookbooks

In some domains of language use in the communities mentioned above, English moved into the textual space occupied previously by Swedish as new generations of language users were losing contact with the mother country of their ancestors (Hoffman and Kytö 2018). Augustana College, in Rock Island, Illinois, founded by Swedish Lutherans, published its college catalog in Swedish for about 25 years, until the language shift to English took hold in 1876. Elsewhere, language shift from Swedish to English took place quickly in the official spheres of educational institutions. Some individual congregations, among them Bethany Lutheran Church in Lindsborg, Kansas, and Salem Lutheran Church in Sandwich, Illinois, continued using the Swedish language to record the minutes of their annual congregational meetings well into in the 1930s (Hoffman and Kytö 2018).[4] Thereafter, however, English took hold in the official sphere.

The language shift situation in the official sphere described above, with English having moved into domains previously held by Swedish, would perhaps give the impression that English had successively taken over all areas previously occupied by Swedish. Yet when we analyze printed texts in the domestic sphere of the same Swedish-American groups, we note that a different pattern holds for community cookbooks, an everyday genre (see further Section 5). These cookbooks were typically produced, printed, and circulated by volunteer committees whose members were influential women in the community. Community cookbooks typically contain an introduction (or preface) in which the compilers situate the contents of the cookbook and comment on the charitable aims of the

4 In the early 1900s, children in the Lindsborg area could attend Swedish-language classes that were offered in the summer (see Hedblom 1983; cf. Ostergren 1988).

publication. The recipes then follow, organized in sections such as appetizers, meat, fish, fowl, baking, and desserts. In many cases, the contributor of a recipe is included as an attribution. In addition to recipes, early cookbooks contained a miscellaneous section in which home remedies (e.g., for treating illnesses) and helpful hints (for managing households) were described. English is the matrix language of the cookbooks we have located (see Primary sources), but certain features of these books are nonetheless characterized by varying proportions of English and Swedish that continue into the most recent editions of cookbooks.

Cookbooks are interesting in that they encode cultural values, by, for instance, informing readers in introductions to the books about notions such as *smörgåsbord* (the traditional Swedish buffet). Such information helps readers understand what creates a dining experience that is "Swedish".[5] In addition to the didactic explanations by the compilers of the Swedish-American cookbooks, the cookbook committees often blur the boundaries of Swedish and English, quite commonly in the hybrid names assigned to recipe names. Examples of recipe names which open the boundaries between English and Swedish are *Swedish Smörbakels* (Standard Swedish *Smörbakelser* = 'butter pastry') and *Limpa Bread* (Standard Swedish: *Limpa*; *limpa* + *bread* is a doublet in the sense that *limpa* is the Swedish word for 'loaf' and combines with *bread* from English) in *On Our Way Rejoicing*, 1979, Sandwich, Illinois, as well as *'Vridna Studenter' Swedish Cookies* ('Twisted Students Swedish Cookies') and *Emma's Apple Kaka* (Standard Swedish: *Emmas Äppelkaka*) in *Swedish Recipes*, 1956, Minneapolis. Code-mixing and code-shifting are key ways that "Swedish"-style cooking and entertaining are signaled to readers.

We suggest that socio-cultural phenomena linked with food have preserved Swedish lexis, even though English has been the main vehicle in communicating Swedish cooking practices to persons within and outside the borders of Swedish-American social networks. Cooking and entertainment practices have been among the most important socio-cultural phenomena for maintaining a sense of closeness to the ancestral home for former immigrants and their children and grandchildren (cf. Tellström 2015). Descriptions of these cooking and entertainment practices simultaneously invite people in the majority host community to cross the linguistic boundaries and cultural borders that previously separated the Swedish-speaking population from the other inhabitants of the area.

We examine seventeen cookbooks from the localities listed in Table 1 from various time points in order to understand how Swedish-American groups have

5 Double quotation marks around the word "Swedish" connote Swedish identity rather than the Swedish language.

used patterns of code-switching and code-mixing. We explore names of suggested menu items for *smörgåsbord*, names of recipes (ca. 7,000), and instructions for preparation.[6] Within the data pertaining to recipe names for "Swedish"-style cooking, we are interested in lexis ranging from English (e.g., *Swedish Meatballs*), to Heritage Swedish (*Peppar Kakor*), and to forms in Standard Swedish (*Ostkaka*); see further 5.4. We apply the framework of *enregisterment* to help us understand not only the cohabitation of Swedish lexis with English discourse in the cookbooks but also the para-textual dimensions of the pages in the books.

# 4 Enregisterment through Swedish-American lexis

Food is a prime cultural domain in which lexis from an immigrant group intertwines with the majority language. In cookbooks, the lexis of the immigrant food culture is not only inserted into the matrix language of the host community to connote names of settler food dishes but is mobilized to carry symbolic value (cf. Remlinger 2009). The framework of enregisterment (Agha 2003, 2007; Adams 2009; Remlinger 2009; Johnstone 2011, 2013, 2016; Jaffe 2015) provides an explanatory account of the association of linguistic forms with local meanings. According to Johnstone (2013: 110), enregisterment is "the linking of linguistic form and social meaning in the context of a schema of ideas that allows the linkage to make sense". Jaffe (2015: 559) explains that dialect formation involves "a named, distinct, 'enregistered' dialect [being] freighted with local meanings and with meanings about locality". In applying the framework of enregisterment to the cookbook data, we have the opportunity to interpret the symbolic meaning of mixed-language data in a locality.

Figure 1 illustrates some ways Swedish-American lexis in cookbooks is accompanied by contextualization to convey local meaning. The figure features a cookie recipe entitled *Carlotta Kakor* (the literal translation to English is 'Charlotte Cookies') scanned from two editions of *Measure for Pleasure*. The version of the recipe appearing in Edition 1[7] (1961: 76) identifies the contributor of the recipe as Mrs. Hagbard Brase. In Edition 4 of the cookbook (2005: 208), the

**6** In the community cookbooks, English is nearly without exception the language used in the instrumental portions of the texts. Of the thousands of recipes in our material, only two contain instructions written in Swedish.

**7** We use the term *edition* to refer to the updated versions of the cookbooks published across time in a community. Some of these versions, especially the early ones, lack any mention made of the edition in question. Table 2 provides a key to our abbreviations for identifying the chronology of the cookbooks in a given community.

CARLOTTA KAKOR
(Tulip Cookies)

1/2 lb. sugar (1 c.)
1 lb. butter or margarine
1 1/2 lb. flour (7 sifted cups)
1 egg white
almond flavoring and almonds

Place half of the flour on board. Place all sugar and butter on flour and work in well. Work in remaining flour. Roll out with stockinged rolling pin, or pat down to 1/4 inch thickness with edge of hand. Brush with slightly beaten egg. Sprinkle with sugar and finely chopped nuts (blanched almonds are best). Bake in hot oven, 375°--8 minutes or until done. Bake next batch on a cool sheet. Let stand after baking until cooled slightly.

. . .Mrs. Hagbard Brase

76

## Carlotta Kakor

*(Charlotte Cookies)*

"My grandmother, Mrs. Hagbard Brase, made these cookies for Bethany Lutheran Church bazaars many years ago. They were sold for a penny each. Someone suggested to her that if she would make them smaller, they could make more money for the church. The only other cookie cutter she had was tulip-shaped." – *Kristin Johnson Lindsberg*

½ pound sugar (1 cup)
1 pound butter or margarine
7 cups flour, sifted
1 egg white
½ cup almonds, chop
¼ teaspoon almond extract

Place 3 ½ cups flour on board. Put sugar and butter on the flour. Work in very well. Knead in remaining flour. Roll out to ¼" thickness. Cut with any cookie cutter. Brush with slightly beaten egg white. Sprinkle with sugar and finely chopped nuts. Bake each batch on a cold cookie sheet. Bake 8 minutes at 375 degrees.

*...Minna (Dr. Hagbard) Brase*
*Wife of Bethany Organ and*
*Harmony Professor, 1908-1953*

208

**Figure 1:** Versions of a recipe for Carlotta Kakor from *Measure for Pleasure*: Edition 1 (1961: 76) and Edition 4 (2005: 208). Images reproduced with permission from the Bethany College Auxiliary.

list of ingredients and the instructions for preparing the cookies are virtually identical to the presentation in Edition 1. In Edition 4, however, some notable expansions have been made in explaining the local meaning of the recipe. In an introductory paragraph, Kristin Johnson, a granddaughter of Minna Brase, contextualizes the meaning of the cookies and links them with the history of Bethany Church, the congregation that founded Bethany College. At the foot of the recipe in Edition 4, the attribution to Minna Brase has been expanded to include details about her husband's professorship at the college, giving readers further information about Minna Brase's status in Lindsborg. Some para-textual features, especially the symbol of a Swedish Dala horse in the upper left-hand corner, further "freight" the recipe with Swedish-American meaning, which we view as having become heavier over time.

An inference one can draw from the scanned recipe on the left-hand side of Figure 1 is that Minna Brase signaled Swedish baking culture with the lexis she used in line 1 in naming the cookies;[8] many decades later, as exemplified by Edition 4 of the cookbook, the cookbook committee accentuated the Swedish association of the recipe by adding visual para-textual resources to the recipe format. The combination of lexical and para-textual patterns allow the recipe contributors and cookbook committees to create local meanings in the cookbooks (Remlinger 2009).

# 5 Shifting content and changing language boundaries in cookbooks

## 5.1 Proportions of "Swedish" recipes in community cookbooks

As noted in Section 4, we examine data from seventeen cookbooks in the selected communities. In total, the number of recipes printed in the cookbooks surpasses 7,000 recipes, with slightly more than 1,200 of the recipes overall designated as being "Swedish". An overview is given in Table 2.

**8** It is unclear who provided the information in line 2 of the recipes which contain English renderings of the recipe name.

**Table 2:** Distribution of "Swedish" recipes in the cookbooks.

| Locality | Abbreviation | Year cookbook published | Number of all recipes | Number of "Swedish" recipes | Percentage of recipes that are "Swedish" |
|---|---|---|---|---|---|
| Andover, Illinois | AND1 | 1950 | 377 | 35 | 9% |
| | AND2 | 1957 | 441 | 35 | 8% |
| | AND3 | 1974 | 405 | 52 | 13% |
| Subtotal | | | 1223 | 122 | 10% |
| Rock Island, Illinois | ROCK1 | 1979–1980 | 358 | 18 | 5% |
| Subtotal | | | 358 | 18 | 5% |
| Sandwich, Illinois | SAND1 | 1979 | 192 | 24 | 13% |
| | SAND2 | 1994 | 414 | 59 | 14% |
| Subtotal | | | 606 | 83 | 14% |
| Lindsborg, Kansas | LIND1 | 1961 | 354 | 76 | 21% |
| | LIND2 | 1970 | 441 | 80 | 18% |
| | LIND3 | 1991 | 535 | 97 | 18% |
| | LIND4 | 2005 | 520 | 104 | 20% |
| Subtotal | | | 1850 | 357 | 19% |
| Marquette, Kansas | MARQ1 | 1910 | 400 | 6 | 2% |
| | MARQ2 | 1978 | 672 | 40 | 6% |
| | MARQ3 | 1998 | 637 | 45 | 7% |
| Subtotal | | | 1709 | 91 | 5% |
| Minneapolis, Minnesota | MINN1 | 1942 | 161 | 72 | 45% |
| | MINN2 | 1956 | 594 | 95 | 16% |
| | MINN3 | 1980 | 524 | 197 | 38% |

**Table 2** (continued)

| Locality | Abbreviation | Year cookbook published | Number of all recipes | Number of "Swedish" recipes | Percentage of recipes that are "Swedish" |
|---|---|---|---|---|---|
| | MINN4 | 1996 | 524 | 197 | 38% |
| Subtotal | | | 1803 | 561 | 31% |
| Total | | | 7549 | 1232 | 16% |

Table 2 shows that the proportion of "Swedish" recipes (i.e., recipes for dishes from Swedish cooking traditions but rendered in the English language; see also note 5) in relation to the total number of recipes in cookbooks changes over time in a community. The proportions of "Swedish" recipes in three of the Minneapolis cookbooks have been the highest of all of the cookbooks we investigated. Given the very large Swedish-born population in the area (see Table 1), one could reasonably anticipate that the books printed there would contain a high proportion of "Swedish" recipes. Yet the results shown in Table 2 also show that the percentages of "Swedish" recipes in these cookbooks have also varied over time, with 45% of "Swedish" recipes in the cookbook from 1942, a steep decline in the 1956 cookbook (16%), later followed by an increase to 38% in books printed in 1980 and 1996. It is interesting that an increased proportion of "Swedish" recipes in the Minneapolis cookbooks appears many decades after the peak in the Swedish-born population. Relative stability appears in the cookbooks published in Lindsborg, Kansas, where the books have contained a relatively constant proportion of "Swedish" recipes, hovering around 20%, over time in the four editions of cookbooks. The cookbook with the lowest percentage of "Swedish" recipes was produced in Marquette, Kansas, in 1910 (2% of the recipes are "Swedish"). Later in the century, however, cookbook committees in Marquette slightly increased the proportion of "Swedish" recipes. A similar trend has emerged in the cookbooks produced in Illinois, where cookbook committees in churches in Andover and Sandwich increased the proportion of 'Swedish' recipes over time. We return to the demographic factors in 5.4.

## 5.2 Ways that Swedish lexis cohabitates with English

Titles of the cookbooks containing Swedish and English have been produced in each of the states we have investigated. In Illinois, the cookbook committee in

Andover produced the books *Vårt Dagliga Bröd* [9] ['Our Daily Bread'] *The Andover Lutheran Church Centennial Cook Book 1850–1950* and *Smör och bröd gör kinden röd* ['Butter and Bread Make the Cheek(s) Red'] *Andover Lutheran Church Cook Book 1850–1957*. In Kansas, the Bethany College Auxiliary has produced four editions of a cookbook bearing the title *Measure for Pleasure. Featuring Hyllningsfest Smörgåsbord, Lindsborg, Kansas*.

As mentioned above, Swedish and English lexis co-exists on the covers of some books. In editions of cookbooks produced by a group affiliated with the American Swedish Institute in Minneapolis, Swedish and English are not mixed phrase-internally in the titles but appear in separate visual fields; in other words, the main title is *Var Så God* (usually translated as 'Help yourself') and the subtitle is *Heritage and Favorite Recipes & Handbook of Swedish Traditions*. The same pattern by which code-switching appears on the covers of cookbooks can be seen in two (of the three) cookbooks from Andover, Illinois. Figure 2 shows the covers of three books printed by Andover Lutheran Church (also known as Augustana Lutheran Church) in 1950, 1957, and 1974.


**Figure 2:** Covers of three cookbooks produced by Andover Lutheran Church, 1950, 1957, 1974. Images reproduced with permission from Andover Lutheran Church.

**9** In the Andover cookbook titles containing Swedish lexis, we use the capitalization patterns as they appear on the covers of the books in Figure 2.

On the covers of the cookbooks in Figure 2, Swedish-American identity is enregistered through bilingual lexis and the use of para-textual space. As apparent in the books printed in 1950 and 1957, Swedish and English appear to abide by a linguistic boundary (for example, there is no language mixing within phrases). Further reinforcing the boundary between languages is the use of the visual space on the covers of the books, with Swedish lexis appearing at the top and English at the bottom. The cookbook from 1950 has the main title *Vårt Dagliga Bröd*; and the cookbook from 1957 has the main title *Smör och bröd gör kinden röd*. In the third cookbook, produced by the congregation in 1974, the cookbook committee featured a portrait of the renowned Swedish soprano Jenny Lind and a facsimile of her signature. Prospective readers of the cookbooks, regardless of whether they understand Swedish or not, recognize the message indexed by the main titles of the Andover cookbooks printed in 1950 and 1957: the featured style of cooking is "Swedish". If the Swedish lexis in the main titles is opaque to some prospective readers, the image of the girl in a stylized folk costume on the cover of the 1957 book hints at the kind of recipes collected by the congregation. Furthermore, the first chapter of the recipe book is dedicated to showcasing Swedish food and recipes for a traditional *smörgåsbord*.

A cookbook committee in another Swedish-American community makes use of some para-textual space in enregistering Swedish culinary traditions along with linguistic clues from Swedish (cf. Remlinger 2009 on enregisterment as a marketing tool on postcards and bumper stickers in Michigan in the late 1900s). Swedish lexis appears on covers of the four editions of cookbooks produced in Lindsborg, Kansas. In these books, English and Swedish co-habitate in the sub-title, and the cookbook committee devotes many pages in each edition of their cookbook to explaining how to plan and serve *smörgåsbord*. Figure 3, scanned from pages in Edition 4 (2005) of *Measure for Pleasure*, provides an example of such a page. On page 7 of the cookbook, the committee displays how to arrange the dishes on a buffet table. The suggested menu, on page 6 of the scanned image, is presented in two visual fields: the Swedish names of food dishes appear on the left-hand side of the menu, while English-language renderings are presented on the right-hand side (Hoffman and Kytö 2017).

Figure 3 shows how Swedish and English lexical items are separated by the boundaries established by the columns and how para-textual clues (e.g., the folkloristic drawings, including the Swedish Dala horse on the table to the top right and the woman wearing a Swedish folk costume) highlight the Swedish connections. Yet the Swedish connections are established in a mix of orthographic variation. Some of the Swedish lexis is rendered in Heritage Swedish orthography, e.g.,

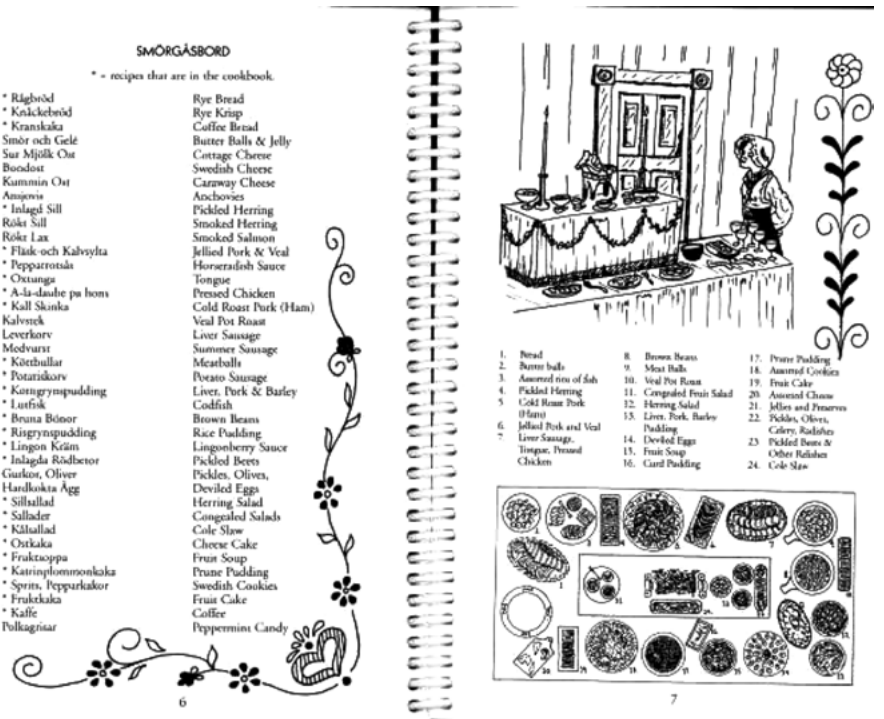


**Figure 3:** Pages from LIND4, pages 6–7. Images reproduced with permission from the Bethany College Auxiliary.

*A-la-daube pa hons* ('Pressed Chicken'; the recipe name in Standard Swedish would be *À la daube på höns* or *Aladåb på höns*)[10] and *Hardkokta Ägg* ('Hard-boiled eggs'; the recipe name would be *Hårdkokta ägg* in Standard Swedish). The examples discussed in this section demonstrate the care committee members devoted to conveying the values of the Swedish smörgåsbord as a specimen of Swedish culture introduced in the American context (see Hoffman and Kytö 2017 for further analysis of Heritage Swedish forms).

**10** From French *à la daube* ('made in a pot').

## 5.3 Shifting language boundaries in the names of recipes

As evident above in Table 2, the proportion of "Swedish" recipes has increased in editions of the cookbooks over time. It is unclear, however, whether the expanded coverage of "Swedish"-style cooking in the twentieth century involved adding recipes for a popular dish such as *Swedish Meatballs* and assigning the same name to every recipe for this dish. We collected a sample of recipe names and prepared an inventory of all of the names used to designate three dishes appearing on the menus for *smörgåsbord* in all of the cookbooks, namely *ostkaka* ('cheesecake'), *köttbullar* ('meatballs'), and *rågbröd* ('rye bread'). Table 3 presents a list of the names of these food dishes for the communities in which we have access to three editions of their cookbooks. To allow a maximum time span for observations, we sampled data from Editions 1 and 3 of the respective cookbooks.[11] We give the number of recipes attested, but as the figures are small, we do not use them as basis for drawing quantitative conclusions, apart from observing that an item that in the beginning was not represented at all gains visibility in a later edition.

**Table 3:** Chronology of names of recipes for *ostkaka*, *köttbullar*, and *rågbröd* in selected communities.

| | **Ostkaka ('cheesecake')** | **Köttbullar ('meatballs')** | **Rågbröd ('rye bread')** |
|---|---|---|---|
| LIND1 (1961) | *Ostkaka* (3)<br>*Wendla & Edla Wahlin's Ostkaka*<br>*Talsh*[a] *Ostkaka (Imitation Cheese Cake)* | *Swedish Meatballs with Sour Cream*<br>*Köttbullar (Swedish Meat Balls)* | *Limpa (Swedish Loaf)*<br>*Swedish Rye Bread* (4) |
| LIND3 (1991) | *Ostkaka (Swedish Cheese cake)* (2)<br>*Ostkaka* (2)<br>*Quick Ostkaka* | *Köttbullar (Swedish Meatballs)* (2)<br>*Swedishmeatballs with Sour Cream* [sic] | *Limpa (Swedish Loaf)*<br>*Swedish Rye Bread* (5)<br>*Svensk Rägbröd* [sic] |

**11** Communities for which we have not located a third edition (i.e., Rock Island and Sandwich, Illinois) are not included in this table. Recipe names presented in the table are those which appear in line 1 of the recipe. See Figure 1, two versions of *Carlotta Kakor*, for examples of line 1 and line 2 in the name of a recipe.

**Table 3** (continued)

| | **Ostkaka ('cheesecake')** | **Köttbullar ('meatballs')** | **Rågbröd ('rye bread')** |
|---|---|---|---|
| MARQ1 (1910) | *Ost kaka* (2) | [none] | [none] |
| MARQ3 (1998) | *Easy Ostkaka*<br>*Ostkaka* (4) | *Swedish Meatballs* | *Delia Lindh's Swedish Rye Bread*<br>*Edna's Swedish Rye Bread*<br>*Mom's Swedish Rye Bread*<br>*Onie's Swedish Rye Bread*<br>*Swedish Rye Bread* |
| AND1 (1950) | *Ost-Kaka* (2)<br>*Osta Kaka* (2) | *Swedish Meat Balls* | *Rye Bread* (4) |
| AND3 (1974) | *Osta Kaka* (3) | *Swedish Meat Balls* (2)<br>*Kott Bulla (Meat Balls)* | *Swedish Rye Bread* (2)<br>*Rye Bread* (4) |
| MINN1 (1942) | *Cheese Cake (Ostkaka)* | *Appetizer Meat Balls (for Smorgashord)* [sic]<br>*Meat Balls (Kottbullar)* | *Rye Bread*<br>*Orange Rye Bread (Limpor)*<br>*Swedish Rye Bread* |
| MINN3 (1980) | *Cheese Cake* (2) | *Royal Swedish Meat Balls*<br>*Swedish Smörgåsbord Meat Balls*<br>*Swedish Meat Balls* (4)<br>*Swedish Meatballs* | *Swedish Rye Bread* (4)<br>*Rye Bread* (2)<br>*Easy Bake Rye Bread*<br>*Swedish Limpa* |

[a] *Talsh* is very likely a misinterpretation of the Swedish word *Falsk* ('false').

Overall, the numbers of recipes for the popular food dishes *ostkaka*, *köttbullar*, and *rågbröd* in relation to the hundreds of recipes in each of the community cookbooks are rather small, but by tracking the lexis used in a given community over time, comparing Editions 1 and 3, we locate some interesting movement. Table 3 shows that the cookbooks include a range of names for the popular recipes in Standard Swedish (*Ostkaka*, *Köttbullar*, *Limpa*), Heritage Swedish renderings (*Talsh Ostkaka*, *Kott Bulla*, *Svensk Rägbröd*), English translations (*Swedish Meatballs*, *Swedish Rye Bread*), and names of recipes which combine lexis from two varieties (*Wendla & Edla Wahlin's Ostkaka*, *Appetizer Meat Balls (for Smorgashord)* [sic]). The rich variation in the use of

the recipe names points to ample linguistic resources mobilized by the cookbook compilers and the contributors of the recipes.

We return to an observation about recipe data reported in the 1910 cookbooks published by a cookbook committee in Marquette, Kansas. The cookbook featured relatively few Swedish recipes, only six (equivalent to 2%) of the total number of recipes being "Swedish", but in the cookbook produced in 1998, the number of "Swedish" recipes had risen to forty-five (equivalent to 7%) of the total number of recipes. In Edition 3, the number of recipes for *ostkaka* and *rågbröd* had increased (only one recipe for meatballs appears). English lexis is used to name the rye bread recipes, and we note a fascinating development in that proper nouns with possessive case are used to differentiate one rye bread recipe from another, such that *Delia Lindh's Swedish Rye Bread* and *Edna's Swedish Rye Bread* are prepared in different ways.

The recipe contributors in Marquette use Swedish lexis to denote *ostkaka*. The food dish *ostkaka* maintains its special lexical properties, while recipes for *rågbröd* are denoted with proper nouns and English lexis. We acknowledge an important factor at play here, namely that *ostkaka* contrasts significantly from American-style cheesecake, the latter prepared with cream cheese. Even so, the increasing frequency of the lexis for *ostkaka* (and a variant *ost kaka*) in the Marquette cookbooks suggests that a Swedish word has gained semantic weight. We return to the the notion of semantic weight in our concluding remarks.

## 5.4 Language contact in different communities

Table 4 illustrates the range of the lexical composition of recipe names in line 1. We lean on the (creole) continuum model to show the dimensions of lexical combinations for recipe names in bilingual speech communities (cf. Rickford 1987; Valdés 2001). Further, we link the various lexical compositions of recipe names to points along a continuum, and we provide examples from the material to illustrate the range of the recipe naming patterns. Naming pattern *Language A*, which heads the far left-hand column of the table, designates recipe names that are realized in Standard Swedish, while *Language B*, far right-hand column, designates recipe names that are realized in English. As a wide range of "Swedish" recipe names in the cookbooks have lexis comprised of naming patterns that blur boundaries between language varieties (Swedish, Heritage Swedish, English), we set up three columns in the middle zone of the Swedish/English continuum in Table 4. Some naming patterns blur the edges between Swedish and English and involve code-switching (for further explanation of how we differentiate between Standard and Heritage Swedish, see Section 1; see also Hoffman and Kytö 2017:

278–279). The column designated as *Ab* presents examples of recipe names that begin with Swedish lexis but continue in English; column *Ba* shows examples of recipe names that begin in English but switch to Swedish. A further category of hybrid naming patterns, classified below the heading *AB*, contains lexis from Heritage Swedish as well as proper nouns teamed with common nouns.

The range of patterns shown in Table 4 provides an overall picture of the lexical flexibility cooks and cookbook committees have used in naming recipes. We now analyze the recipe naming patterns in the two localities that have printed the highest proportion of "Swedish" recipes: Lindsborg and Minneapolis.

**Table 4:** Lexical composition of recipe naming patterns.

| Naming pattern A | Ab | AB | Ba | Naming pattern B |
|---|---|---|---|---|
| **Standard Swedish** *Ostkaka. Skorpor.* | **Swedish + English** *Vridna Studenter Swedish Cookies. Svenska Sausage Casserole.* | **Heritage Swedish** *Blixten Torten. Ragmunkar. Peppar Kakor. Kalu-Cylta. Potatis Curv. Osta Kaka.* | **English + Swedish** *Swedish Pepparkakor. Swedish Limpa.* | **English names for Swedish recipes** *Swedish Meatballs. Swedish Dreams. Swedish Rye Bread.* |
| | **Swedish + Heritage Swedish** *Bruna Boner*.[a] | **Heritage Swedish + English** *Lute Fish. Agg-Med Ost Lada Casserole.* | **English + Heritage Swedish** *Swedish Salt Sill Salad. Swedish Kringler. Fine Kringles.* | |
| | | **Swedish Proper Nouns + English** *Mrs. Hjalmar Wetterström's Ice Box Dessert.* | | |

[a]In response to a reviewer's query, we ascribe the recipe name *Bruna Boner* to this category as the orthography of *Bruna* is Standard Swedish, and that of *Boner* is clearly Heritage Swedish. *Bruna* could, of course, also appear as e.g., *Brown*, which would have led to our classification of the recipe name as English + Heritage Swedish.

Table 5, which presents the most-frequent naming patterns for "Swedish" recipes in Lindsborg, shows apparent stability in the use of lexis. The dominant pattern used by cookbook committees in four editions is the use of Standard Swedish. English names for "Swedish" recipes have consistently held second place, in terms of the naming patterns. In Editions 3 and 4 of the Lindsborg cookbooks, a slight reduction of recipe names rendered in Standard Swedish (41% in 1991; 39% in 2005) is visible, and this slight decline is accompanied by a slight rise in the proportion of English names for "Swedish" recipes.

**Table 5:** Most frequent naming patterns for "Swedish" recipes in Lindsborg cookbooks.

| Year and number of "Swedish" recipes in the edition | Most-frequent naming pattern of "Swedish" recipes | Next most-frequent naming pattern of 'Swedish' recipes |
|---|---|---|
| LIND1, 1961 (N=76) | Standard Swedish (43%) | English (29%) |
| LIND2, 1970 (N=80) | Standard Swedish (44%) | English (29%) |
| LIND3, 1991 (N=97) | Standard Swedish (41%) | English (36%) |
| LIND4, 2005 (N=104) | Standard Swedish (39%) | English (38%) |

The overall patterning in the Minneapolis data (Table 6) suggests that cookbook committees have made different choices, as compared to Lindsborg, in the ways "Swedish" recipes are named.

**Table 6:** Most frequent naming patterns for "Swedish" recipes in Minneapolis cookbooks.

| Year and number of "Swedish" recipes in the edition | Most-frequent naming pattern of "Swedish" recipes | Next most-frequent naming pattern of "Swedish" recipes |
|---|---|---|
| MINN1, 1942 (N=72) | English (60%) | English + Swedish (18%) |
| MINN2, 1956 (N=95) | English + Swedish (28%) | English (27%) |
| MINN3, 1980 (N=197) | English (79%) | Swedish proper nouns + English (11%) |
| MINN4, 1996 (N=197) | English (79%) | Swedish proper nouns + English (11%) |

Table 6 shows that "Swedish" recipes have been expressed with English renderings with a high proportion in three of their cookbooks (in Editions 1, 3, and 4; code-mixing between English and Swedish was the most frequent pattern in Edition 2). We note, however, that the relatively frequent use of English names holds true for line 1 of recipe names (cookbook committees in Minneapolis tend to include a Swedish version of a recipe name in line 2).[12] Table 6 also shows that the next most-frequent naming pattern of recipe names in Editions 1, 3, and 4 is one from the middle zone of the continuum as described in Table 4.

The high proportion of English renderings for names of "Swedish" recipes coupled with evidence of combined patterns (the next most-frequent naming patterns) in the Minneapolis data is interesting in light of the demographic patterns presented in Table 1. As reported in the table, Hennepin County was the adopted home for more Swedish-born persons (at least 26,000 persons in the first decades of the century) than the combined Swedish-born population in all of the other localities we investigated. The Swedish-speaking population in Hennepin County had been renewed by multiple waves of new Swedish immigrants over the years (Rice 1981; Karstadt 2003). Intense language contact between Swedish and English in Hennepin County continued for a much longer period of time than was the case in McPherson (Kansas) and Henry (Illinois) counties, whose Swedish-born population in the first third of the 1900s was less than half of its peak population in the 1800s. We therefore note an interesting paradox in the naming patterns of "Swedish" recipes in the cookbooks from Minneapolis and Lindsborg: Lindsborg cookbook committees use Standard Swedish in line 1 of their recipes at a higher proportion than do the Minneapolis cooks.[13] The use of Standard Swedish in line 1 of recipe names in Lindsborg books gives added emphasis to the source of the culinary culture. Meanwhile, when Minneapolis cookbook committees place English translations in line 1 for the names of Swedish dishes (followed in line 2 with the Swedish names), they use English as the linguistic gateway to Swedish culinary culture. The Minneapolis cookbook committees had contact with Swedish immigrant recipe contributors over many decades in the 1900s. While we cannot know how these committees negotiated their preferences for using Swedish or English names, it is clear from the data that the majority language principle played a role in their decision-making. So demographic patterns were likely not the only factor at play when the boundaries between English and Swedish were negotiated by cookbook committees.

**12** The scope of the present study did not allow analysis of the second lines of recipe names.

**13** When we analyzed the Andover cookbook data, we found that the most-frequent naming pattern for "Swedish" recipes was English, and that the next most-frequent naming pattern was Heritage Swedish.

# 6 Concluding remarks

Regarding our first research question (Section 1), we found that the proportion of recipes representing "Swedish" food increases over time (see 5.1), and that the recipe names for these dishes are characterized by shifting proportions of English and Swedish. Some of the cookbooks, particularly MINN2 (see Table 6 in 5.4), show relatively open boundaries between the two languages in the sense that code-mixing with English and Swedish was the most obvious naming pattern for the "Swedish" recipes.

We gained further insights into the language contact between English and Swedish by applying the theoretical framework of enregisterment, which pertains to our second research question (see especially Section 4 and 5.2). The covers of many of the community cookbooks showed evidence of code-switching and code-mixing. Such linguistic phenomena are often accompanied by para-textual features, clues which accentuated the "Swedish" style of the cookbook. Observation of such para-textual features, when teamed with linguistic analysis of code-switching and mixing, provides a record of how the American cookbooks, with English as the matrix language, nonetheless are enriched by new lexis from an immigrant community.

Addressing our third research question, we presented evidence supporting the notion that different Swedish immigration patterns in the localities played a role in the practices that cookbook committees adopted in naming the featured recipes. On the basis of the broader picture provided by the data, we see that the lexical patterns manifest in the cookbook data reflect the social identity work manifest in the Swedish-American communities across the 1900s and beyond. Importantly, these cookbooks act as vehicles for carrying Swedish lexis across time and space, even when the English language is otherwise used in the vast majority of all other domains. Further with respect to lexical patterning, we found that certain recipes, namely for *ostkaka* ('cheesecake'), *köttbullar* ('meatballs'), and *rågbröd* ('rye bread'), did not disappear but have gained more presence in the cookbooks over time (see Table 3 in 5.3). Thus, unlike the findings of Bunin Benor (2015: 229), in which heritage lexis in a postvernacular Yiddish community waned before it returned (what she identifies as a "boomerang effect in ethnic language use"), we see evidence that Swedish lexis has increased over time, gaining prominence and semantic weight. Such evidence gives us reason to pursue further analysis of semantic weight (cf. Schneider 2003, 2007) in the future and compare patterning in Swedish-American cookbooks with those in other immigrant communities.

As code-switching and code-mixing phenomena are visible in the cookbooks printed in all of the Swedish-American localities we investigated, we see

that the linguistic boundaries between the Swedish and English languages have been permeable across the entire twentieth century (and beyond, for the Lindsborg, Kansas community). When we team this evidence together with the para-textual features in the books, we find strong evidence that recipe contributors and cookbook committees have maintained Swedish cultural expression, even when producing books mainly in English. The books serve as a record of how Swedish culture and linguistic heritage have been preserved in the domestic sphere and in a range of Swedish-American communities. Cookbooks are a particularly good vehicle for preserving heritage because they are so clearly linked to dinner tables and the words and traditions used by family members. We are keen to understand more about the interaction in Swedish-American communities between persons for whom Swedish is a postvernacular variety and those for whom Swedish is their native language. Our future study will also extend the scope of our research beyond community cookbooks so as to make comparisons concerning language contact phenomena and enregisterment with usage in national cookbooks published during the same period.

# References

## Primary sources

AND1: *Vårt dagliga bröd* [Our daily bread]. *The Andover Lutheran Church centennial cook book 1850–1950*. 1950. Andover, IL: Andover Lutheran Church.

AND2: *Smör och bröd gör kinden röd* [Butter and bread make the cheeks red]. *Andover Lutheran Church cook book 1850–1957*. 1957 Andover, IL: Andover Lutheran Church.

AND3: *Jenny Lind cook book*. 1974. Andover Lutheran Church. Kansas City, KS: Bev-Ron Publishing Company.

LIND1: Bethany Teachers' Wives. 1961. *Measure for pleasure. Featuring Hyllnings Fest smörgåsbord, Lindsborg, Kansas*. Lindsborg, KS: Bethany College.

LIND2: The Bethany Dames Club. 1970. *The new measure for pleasure*. Lindsborg, KS: Bethany College.

LIND3: The Bethany College Auxiliary. 1991. *Measure for pleasure. Featuring Hyllningsfest smörgåsbord, Lindsborg, Kansas*. Lindsborg, KS: Bethany College.

LIND4: The Bethany College Auxiliary. 2005. *Measure for pleasure. Featuring Hyllningsfest smörgåsbord, Lindsborg, Kansas*. Lindsborg, KS: Bethany College.

MARQ1: The Ladies of the Elim Lutheran Church and Their Friends. 1910. *The Marquette cook book: Four hundred choice recipes*. Marquette, KS: Elim Lutheran Church.

MARQ2: Marquette, *Elim cookbook* (1978) = *1878–1978 Elim cook book*. 1978. Marquette, KS: Elim Lutheran Church.

MARQ3: *Elim Treasures*. *Elim Lutheran Church, Marquette, Kansas. 120th anniversary, 1878–1998*. 1998. Kearney, NE: Cookbooks by Morris Press.

MINN1: Ladies Auxiliary 1942. *Swedish recipes. Compiled by Ladies Auxiliary Swedish Institute Male Chorus*. 1942. Minneapolis, MN: American Swedish Institute of Swedish Arts Literature and Science.
MINN2: Women's Club of the American Swedish Institute. 1956. *Swedish recipes. Compiled by the Women's Club of the American Swedish Institute, Minneapolis, Minn*. 1956. Minneapolis, MN: The Women's Club of American Swedish Institute.
MINN3: *Var så god* [Help yourself]. *Heritage and favorite recipes & handbook of Swedish traditions*. 1980 [Third printing 1983]. Minneapolis, MN: American Swedish Institute.
MINN4: *Var så god. Heritage and favorite recipes & handbook of Swedish traditions*. 1980 [Sixth printing 1996, with new preface]. Minneapolis, MN: American Swedish Institute.
ROCK1: *Favorites from First*. 1979–1980. Rock Island, IL: First Lutheran Church.
SAND1: *On our way rejoicing. A history of Salem Lutheran Church, Sandwich, Illinois, 1904–1979* [1979]. Sandwich, IL:. Salem Lutheran Church.
SAND2: *Feast and celebration. Salem Evangelical Lutheran Church. Sandwich, Illinois anniversary cookbook 1904–1994*. 1994. Hills, MN: Crescent Publishing.

## Secondary sources

Adams, Michael. 2009. Enregisterment: A special issue. *American Speech* 84(2). 115–117.
Agha, Asif. 2003. The social life of cultural value. *Language and Communication* 23(3–4). 231–273.
Agha, Asif. 2007. *Language and social relations*. Cambridge: Cambridge University Press.
Benor, Sarah Bunin. 2015. How synagogues became shuls. The boomerang effect in Yiddish-influenced English, 1895–2010. In Janne Bondi Johannessen & Joseph C. Salmons (eds.), *Germanic heritage languages in North America: Acquisition, attrition and change*, 217–233. Amsterdam: John Benjamins.
Blanck, Dag. 2012. Friends and neighbors? Patterns of Norwegian-Swedish interaction in the United States. In Philip J. Anderson & Dag Blanck (eds.), *Norwegians and Swedes in the United States. Friends and neighbors*, 5–20. St. Paul, MN: Minnesota Historical Society Press.
Hasselmo, Nils. 1974. *Amerikasvenska* [American Swedish]. Stockholm: Esselte Studium.
Hedblom, Folke. 1983. Katalog över ljudband inspelade i Amerika med arkivets inspelningsbil åren 1962, 1964, 1966 [Catalogue of sound recordings made in America with the archives' recording van in the years 1962, 1964, and 1966]. Uppsala: Uppsala Dialekt- och Folkminnesarkivet.
Hoffman, Angela & Merja Kytö. 2017. The linguistic landscapes of Swedish heritage cookbooks in the American Midwest, 1895–2005. *Studia Neophilologica* 89(2). 261–286.
Hoffman, Angela & Merja Kytö. 2018. Heritage Swedish, English, and textual space in rural communities of practice. In Jan Heegård Petersen & Karoline Kühl (eds.), *Selected proceedings of the 8th Workshop on Immigrant Languages in the Americas (WILA 8)*, 44–54. Somerville, MA: Cascadilla Press.
Jaffe, Alexandra. 2015. Book review of Barbara Johnstone. *Speaking Pittsburghese. The story of a dialect. Journal of Sociolinguistics* 19(4). 559–563.
Johannessen, Janne Bondi & Joseph C. Salmons (eds.). 2015. *Germanic heritage languages in North America: Acquisition, attrition and change*. Amsterdam: John Benjamins.

Johnstone, Barbara. 2011. Dialect enregisterment in performance. *Journal of Sociolinguistics* 15(5). 657–679.
Johnstone, Barbara. 2013. *Speaking Pittsburghese: The story of a dialect*. New York: Oxford University Press.
Johnstone, Barbara. 2016. Language theory in contemporary sociolinguistics: Beyond Dell Hymes? In Nikolas Coupland (ed.), *Sociolinguistics: Theoretical debates*, 417–432. Cambridge: Cambridge University Press.
Karstadt, Angela Hoffman. 2003. *Tracking Swedish-American English. A longitudinal study of linguistic variation and identity*. Uppsala: Acta Universitatis Upsaliensis.
Lindquist, Emory Kempton. 1953. *Smoky Valley people. A history of Lindsborg, Kansas*. Lindsborg, KS: Bethany College.
Muysken, Pieter. 2000. *Bilingual speech: A typology of code-mixing*. Cambridge: Cambridge University Press.
Myers-Scotton, Carol. 2002. *Contact linguistics*: *Bilingual encounters and grammatical outcomes*. Oxford: Oxford University Press.
Nelson, Helge. 1943. *The Swedes and the Swedish settlements in North America*. 2 vols. Lund: Gleerup.
Ostergren, Robert Clifford. 1988. *A community transplanted. The Trans-Atlantic experience of a Swedish immigrant settlement in the Upper Middle West, 1835–1915*. Uppsala: Acta Universitatis Upsaliensis.
Page, B. Richard & Michael T. Putnam (eds.). 2015. *Moribund Germanic heritage languages in North America: Theoretical perspectives and empirical findings*. Leiden: Brill.
Remlinger, Kathryn. 2009. Everyone up here: Enregisterment and identity in Michigan's Keweenaw Peninsula. *American Speech* 84(2). 118–137.
Rice, John G. 1981. The Swedes. In June Drenning Holmquist (ed.), *They chose Minnesota: A survey of the state's ethnic groups*, 248–256. St. Paul: Minnesota Historical Society Press.
Rickford, John R. 1987. *Dimensions of a creole continuum. History, texts and linguistic analysis of Guyanese Creole*. Stanford, CA: Stanford University Press.
Rothman, Jason. 2009. Understanding the nature and outcomes of early bilingualism: Romance languages as heritage languages. *International Journal of Bilingualism* 13(2). 155–163.
Schneider, Edgar W. 2003. The dynamics of new Englishes: From identity construction to dialect birth. *Language* 79(2). 233–281.
Schneider, Edgar W. 2007. *Postcolonial English. Varieties around the world*. Cambridge: Cambridge University Press.
Shandler, Jeffrey. 2004. Postvernacular Yiddish: Language as a performance art. *The Drama Review* 48(1). 19–43.
Shandler, Jeffrey. 2006. *Adventures in Yiddishland: Postvernacular language and culture*. Berkeley: University of California Press.
Tellström, Richard. 2015. *Hunger och törst. Svensk måltidshistoria. Från överlevnad till statusmarkör* [Hunger and thirst. Swedish mealtime history. From survival to a marker of status]. Stockholm: Forum.
Valdés, Guadalupe. 2001. Heritage language students: Profiles and possibilities. In Joy Kreeft Peyton, Donald A. Ranard & Scott McGinnis (eds.), *Heritage languages in America: Preserving a national resource*, 37–77. McHenry, IL: Center for Applied Linguistics.
Wilkerson, Miranda E. & Joseph Salmons. 2008. "Good old immigrants of yesteryear" who didn't learn English: Germans in Wisconsin. *American Speech* 83(3). 259–283.

Don Chapman

# 10 Specimen texts and boundaries in the history of the English language

## 1 Introduction

Boundaries are important in the history of the English language, as the very name of the field suggests. *The* history of *the* English language (HEL) calls for at least two idealized entities – the English language and the history of that language – and textbooks inevitably contribute to establishing and maintaining the boundaries defining these entities simply by choosing the language varieties, language changes, language features, language periods, and historical events to focus on. The boundaries of the discipline are reinforced as textbooks converge on their choices. Important historical events recounted in most textbooks, for example, include the invasion of Germanic tribes, followed by the coming of Christianity, the Viking invasions, and the Norman invasion (Chapman 2019). Important language changes include Grimm's law, the loss of nominal inflections, the Great Vowel shift, and the adoption of French and Latin loanwords.

A less-noticed, yet still important feature of textbooks for establishing and maintaining boundaries is the representative or specimen texts. These are texts like the Lord's prayer to exemplify all periods of English, Chaucer's first lines of the *General prologue to the Canterbury Tales* to exemplify Middle English, and Shakespeare's portrayal of Edgar in *King Lear* to exemplify Early Modern English. If the textbook itself is meant to represent a certain portion of the wide history of the English language, the specimen texts represent certain stages and varieties. Do we see convergence in choices of specimen texts in HEL textbooks, somewhat like a literary canon? Do students and teachers have a greater knowledge of some core of texts than others? If so, which texts and authors? And what does the convergence tell us about the boundaries within the discipline and surrounding the discipline itself? These are the questions this paper will address. As this paper will detail later, HEL textbooks indeed converge on a small group of specimen texts, and this core says much about the importance of boundaries in HEL accounts – boundaries between English and other languages, between periods of English, between dialects of English, and between registers of English. The

**Don Chapman,** Department of Linguistics, Brigham Young University, Provo, Utah, USA, e-mail: don_chapman@byu.edu

https://doi.org/10.1515/9783110643282-011

specimen texts that are most repeated from textbook to textbook highlight the boundaries inherent in the discipline.

Specimen texts have been used in HEL accounts for as long as we have had such accounts. In 1605 William Camden published a chapter on the history of English within a book he called *Remains concerning Britain*, and he included five translations of the Lord's prayer, from the Lindisfarne glosses to the Wycliffe translation, to illustrate the changes in English over that time (Camden [1605] 1984: 25–27; Sevenker 2016: 23–24). HEL writers have been using at least some of the same texts as their predecessors for almost as long (Sevenker 2016: 30). Sevenker (2016: 42, 48) lists at least two writers (Martin 1749 and Greenwood 1711) who borrowed Camden's texts wholesale. Writers, of course, have added many other texts along the way, so that by now a large number of specimen texts have shown up in HEL accounts in many different uses. Sometimes specimen texts illustrate whole language periods and sometimes single language changes. Sometimes they are given as exercises and sometimes as metalinguistic commentary from earlier periods. Sometimes they are given as the "best" writing and sometimes as the most representative writing. Sometimes they recount a historical event (like the Norman invasion) and sometimes instances from daily life. Specimens have become a staple in HEL accounts, and while some accounts use more specimens than others, it is the rare account that includes no specimens.[1]

## 2 Method

To find the most repeated specimen passages, I examined fifty-one HEL accounts ranging from Camden (1605) to van Gelderen (2014). I included nearly all the HEL titles listed by Crowley (2003) and Blockley (2012). I also included a few accounts listed by Sevenker (2016) and a few other books that happened to be on my bookshelf. A list of the HEL accounts examined is given in the references (Primary sources).[2] Most accounts are book-length, though the earliest ones are only chapter-length accounts embedded in larger works. Among the book-length accounts, most could fairly be called textbooks, though the appearance of textbooks has changed over time. Books written in the nineteenth century usually lack the panels, bullet points, study questions, and summaries that typify textbooks today, but they still present introductory and general information, like textbooks. This study does not include other genres, such as HEL workbooks

**1** Smith (1912) is one such. Numerous word examples are included, but no passages.

**2** Nielsen (1998) and Nielsen (2005) have been treated as a single textbook.

(e.g., Millward 1996), HEL readers (e.g., Fisher and Bornstein 1984; Burnley 1992), popularizing books (e.g., McCrum et al. 1986; Bragg 2003), and period textbooks (e.g., Mitchell and Robinson 2011; Beal 2004), mainly because the aims of such genres are sufficiently distinct from HEL textbooks to introduce genre as a complication in analyzing the convergence and variation of specimen texts.

Identifying specimen texts presents its own problems, especially given the mass of material to examine. I tabulated passages that were easy to perceive, being set apart from the prose by indentation, marginal labels, panels, exercises, or appendices. Passages set off only by quotations marks or italics were not included, partly because they would have been harder to find, but also because writers seldom used only those devices for marking specimen texts. Similarly, at least two lines had to be cited for the passage to be tabulated, again, partly because single-line quotations are harder to spot, but, more importantly, because they tended to be less contextualized illustrations of a particular word or grammatical construction. Future studies might profitably look at this kind of specimen text as well, but they should probably be separated from those meant to give more context. Facsimiles were not included if the text was not otherwise elaborated, since facsimiles usually emphasized the appearance and setting of the text more than the text itself. Nor did I include the mere mention of a work or author, even when included in a list of important authors and texts, since the texts were not provided.

# 3 Results

Most of the 1146 texts (types) found in HEL textbooks were unique: 860 texts occurred only once, and 286 texts occurred more than once. There were 1494 tokens of those texts occurring more than once, so that in all, 2354 tokens were tabulated. As these numbers show, the specimen texts do not constitute a canon in a strict sense, since nearly all writers include many unique specimen texts. The percentage of unique specimen texts per textbook ranges from 0% (four textbooks, all with fewer than ten specimen texts overall) to 63%. There doesn't seem to be any obvious trend over time: high percentages show up early in Johnson's (1755) history (56%), late in Millward and Hayes (2012) (56%), and in several textbooks in between (nine textbooks with percentages greater than 50%). Similarly, low percentages show up early with Martin (1749) (10%), late with Algeo and Butcher (2014) (25%), and in several histories in between (twenty-six textbooks with percentages less than 30%). The average percentage of unique specimen texts in all the books is 31%.

Yet we also see in the tabulation a convergence on a few texts and authors. The twenty most-quoted authors are given in Table 1. Here the token count of all passages in all textbooks is given in column 3 and a range count of the total number of books citing the author is given in column 4.

**Table 1:** Top-twenty authors of specimen texts in HEL textbooks.

| Rank | Author | Tokens | Range | Rank | Author | Tokens | Range |
|---|---|---|---|---|---|---|---|
| 1 | Geoffrey Chaucer | 153 | 42 | 11 | Aelfric | 25 | 15 |
| 2 | William Shakespeare | 114 | 29 | 12 | Alexander Pope | 23 | 10 |
| 3 | John Trevisa | 50 | 25 | 13 | Dan Michel | 21 | 12 |
| 4 | Samuel Johnson | 37 | 18 | 14 | John Dryden | 21 | 9 |
| 5 | William Caxton | 32 | 22 | 15 | Robert Mannyng | 20 | 8 |
| 6 | John Milton | 29 | 19 | 16 | Edmund Spenser | 19 | 14 |
| 7 | George Puttenham | 29 | 17 | 17 | Thomas Wilson | 19 | 12 |
| 8 | Alfred (& Alfredian Texts) | 27 | 19 | 18 | William Langland | 17 | 9 |
| 9 | Jonathan Swift | 27 | 11 | 19 | Orm | 16 | 13 |
| 10 | Robert of Gloucester | 25 | 19 | 20 | Caedmon | 16 | 11 |

Several patterns stand out in this table that will be discussed later, but for now, we can note how well represented the authors from the traditional literary canon are, as if each period or even century could be represented by its most famous author: Alfred for Old English, Chaucer for Middle English, Shakespeare for the sixteenth century, Milton for the seventeenth century, and Pope for the eighteenth century. Notably, there are no authors on this list from later than the eighteenth century. These canonical authors all have a wide range of quoted specimen texts, whereas several of the lesser-known writers on this list have only one or two oft-included texts. Of course the authors of several popular specimen texts are not on this chart, because the author is unknown, like the *Beowulf* poet.

The range of specimen texts and the contribution of anonymous authors can be seen more clearly when Table 1 is compared with a similar list of most quoted works, given in Table 2. As Tables 1 and 2 show, an author's position in this "canon" usually depends on just one work. Milton is an exception, as no single work of his makes the top twenty, yet together they rank him number six among authors. Shakespeare and Chaucer are slightly exceptional in having

**Table 2:** Top-twenty specimen texts used in HEL textbooks.

| Rank | Work (author) | Tokens | Range | Rank | Work (author) | Tokens | Range |
|---|---|---|---|---|---|---|---|
| 1 | *Bible* (multiple translators) | 258 | 41 | 11 | *A proposal for correcting, improving, and ascertaining the English tongue* (Swift) | 19 | 8 |
| 2 | *Canterbury tales* (Chaucer) | 113 | 40 | 12 | Preface to *Eneydos* (Caxton) | 18 | 16 |
| 3 | *Anglo-Saxon chronicle* (anonymous) | 75 | 25 | 13 | *Piers Plowman* (Langland) | 17 | 9 |
| 4 | *Polychronicon* (Trevisa) | 50 | 25 | 14 | *Ormulum* (Orm) | 16 | 13 |
| 5 | *Arte of English poesie* (Puttenham) | 29 | 17 | 15 | *Arte of rhetorique* (Wilson) | 16 | 12 |
| 6 | *Beowulf* (anonymous) | 29 | 15 | 16 | Caedmon's hymn | 16 | 11 |
| 7 | Chronicle of Robert of Gloucester | 24 | 18 | 17 | *Troilus and Crysede* (Chaucer) | 16 | 9 |
| 8 | *Dictionary of the English language* (Johnson) | 22 | 13 | 18 | *Sir Gawain and the Green Knight* (anonymous) | 14 | 8 |
| 9 | *Ayenbite of inwyt* (Dan Michel) | 21 | 12 | 19 | *King Lear* (Shakespeare) | 14 | 7 |
| 10 | *Cursor mundi* (anonymous) | 20 | 14 | 20 | *Macbeth* (Shakespeare) | 13 | 9 |

more than one specimen text in the top-twenty list. Otherwise, most authors are included mainly because of just one text, like Trevisa, whose translation of Higden's *Polychronicon* is one of the most included. Even when a writer is known for several works, it is largely one particular work, like Samuel Johnson's *Dictionary* or Caxton's preface to *Eneydos*, that is most frequently included. A few

texts like *The Canterbury tales* and the Bible are composite works and benefit from having several passages from several different parts included. I have not systematically tabulated specific passages within texts, but it is clear that certain passages were particularly popular within a work, like the depiction of the Prioress in Chaucer's *General prologue*, and the Lord's prayer from the Bible, which may be the most-included passage of all.

# 4 Discussion

## 4.1 Periods

This canon of specimen texts, as limited as it may be, can still be seen to reinforce several boundaries. Some of the clearest boundaries demarcated by specimen texts are those between periods of English. Nearly all specimen texts are labeled within a clear period, though the exact labels of those periods has changed over time (see Curzan [2012] for a discussion of period labels). In this paper, I have grouped the specimen texts by the periods typically used today – Old English (OE; 500–1100), Middle English (ME; 1100–1500), Early Modern English (EModE; 1500–1700), Late Modern English (LModE; 1700–1900), and Present-Day English (PDE; 1900–present)–and the labels in this paper are not necessarily those used in each of the books.

The most notable specimen text for demarcating period boundaries is the Bible, especially the Lord's prayer, since the Bible was translated at several different times, so that the same passage can be shown from several different dates. That is how the Bible is first used in the earliest examples of specimen passages that have been identified (Camden 1605), and that is one of the chief uses of the Bible still (the use in Görlach 1997 is a good example). Nearly all textbooks include passages from the Bible (41 out of 51), and most include contrastive uses, to some degree. The effect of such tidy comparisons is to reinforce the differences between the periods and to reify the periods themselves.

The use of the Bible points toward another trend in HEL textbooks, and that is the greater exemplification of some periods more than others. Table 3 provides the figures of the Bible passages from different periods. The token and range count are remarkably comparable for OE, ME, and EModE, underscoring the degree that Bible texts are contrastively used to illustrate the periods. But the token and range counts drop off considerably after 1700. Textbook writers apparently see less need to illustrate contemporary English (LModE and PDE), even with a contrastive text. Including contemporary Bible passages at all is a comparatively

**Table 3:** Specimen texts from Bible translations by period.

| Period | Total Tokens | Range of Books |
|---|---|---|
| OE translations | 67 | 34 |
| ME translations | 63 | 27 |
| EModE translations | 59 | 29 |
| LModE translations | 2 | 1 |
| PDE translations | 22 | 9 |

recent practice: Bryant was the first in 1948 and the other nine textbooks to include Biblical passages from post seventeenth-century translations all come after 1968.

This same drop-off is noticeable, though less dramatic, in the distribution of all specimen texts, which is given in Table 4.

**Table 4:** All tokens and types of specimen texts by period.

| Period | Tokens | Types | Ratio of Tokens to Types | Range of Textbooks |
|---|---|---|---|---|
| OE | 296 | 80 | 3.70 | 48 |
| ME | 686 | 197 | 3.48 | 49 |
| EModE | 729 | 383 | 1.90 | 48 |
| LModE | 349 | 228 | 1.53 | 38 |
| PDE | 279 | 245 | 1.14 | 29 |

It is clear that the number of tokens for the ME and EModE periods are far greater than the tokens for LModE and PDE. And even for the LModE texts, there is a distinct preference for older texts. Of the 349 LModE tokens, 222 (64%) are from the eighteenth century. This trend is largely uniform throughout the tradition; five text books did not treat LModE or PDE at all, and four that did treat LModE and PDE did not include any specimen texts. On the other hand, there are a few textbooks that counter that trend: seven textbooks devote at least half their tokens to LModE and PDE. The highest is Graddol et al. (1996), at 80.7%, which reflects the emphasis of that book on variation throughout the world.

The distribution in Table 4 also reveals a pattern not seen in Table 3, namely a lower number of OE specimen texts. This distribution, too, is fairly regular

throughout the tradition. Of those textbooks with more than ten specimens, the highest proportion of OE specimens is 39% from Toller (1900). The proportion of OE specimen texts was less than 10% for twenty-seven of the textbooks. Even among ME specimen texts, far more tokens come from the fourteenth and fifteenth centuries (543) than from the earlier centuries (143). So the largest concentration of specimen texts comes from the fourteenth through eighteenth centuries.

This is the same distribution of the most-included authors and texts in Tables 1 and 2. Of the twenty authors and texts, only three authors and four texts (not counting some translations of the Bible) come from outside this five-century swath: Aelfric, Alfred, and Caedmon and the *Anglo-Saxon Chronicle*, *Beowulf*, and *Caedmon's hymn* from the OE period, and Robert of Gloucester, his *Chronicle*, Orm, and the *Ormulum* from the early ME period. For some reason, the middle part of the history of English is the most popular part to illustrate. Part of the reason for this could be a Goldilocks principle. OE is too unfamiliar, perhaps, and there is little to be gained by multiplying texts that are all equally difficult and impenetrable for beginning students. On the other hand, texts written since 1800 are too familiar, so there is little to be gained by illustrating a stage of English that students are already well familiar with. The texts in the middle are just right. They are strange enough to be worth exhibiting, but familiar enough to be decipherable. As more and more textbooks include and illustrate World Englishes, we may see more specimen texts from PDE, since those texts, too, are interesting enough to illustrate and familiar enough to understand. Of the 279 PDE tokens, sixty-four are specimens of World Englishes, another four are quotations about World Englishes, and sixty are specimens of other types of non-standard English.

Noticeably, no text or author from after 1800 is in the most included lists. This trend is reinforced when we look at the distribution of unique passages by period, as seen in Table 5.

**Table 5:** Percentage of unique texts per period.

| Period | Unique Token | Total Tokens | Percent of Unique Tokens |
|---|---|---|---|
| OE | 45 | 296 | 15.2% |
| ME | 120 | 686 | 17.5% |
| EModE | 279 | 729 | 38.3% |
| LModE | 184 | 349 | 52.7% |
| PDE | 219 | 279 | 78.5% |

We have already seen that unique texts far outnumber repeated texts, but Table 5 shows that the unique passages increase with each chronological period, until they predominate in LModE and PDE. Even texts that are not repeated numerous times are likely to be repeated at least a few times in the OE and ME periods. We seem to have a tighter canon for the earlier periods. This has been so throughout the textbook tradition. Most books have a much lower percentage of unique OE and ME passages than they do for PDE. In fact, all the PDE passages are unique in several textbooks.

In all these measurements – use of the Bible, use of all specimen texts, use of most-included texts and authors, and use of unique passages – the use of illustrative texts in the most recent periods is different from the earlier periods. The specimen texts not only illustrate and reinforce the boundaries between all periods, they also suggest an important boundary between recent English and earlier English, and perhaps a similar, though weaker boundary between OE and later English. At least when measured by specimen texts, the fourteenth through eighteenth centuries are the most important. This boundary between recent and remote may be partially explained by the Goldilocks principle, as already noted, but it may also suggest a more reified view of the past compared to the present. "History", perhaps, is seen as more distant, and the boundary drawn around "the history" is set for an earlier time. Characterizations of contemporary English are perhaps seen more as a separate topic from history, and indeed many earlier textbooks (e.g., Harrison 1848; Trench 1889; Krapp [1909] 1966) explicitly contrast "history" with "present structure". Since the 1970s, textbooks have increasingly crossed that "pastness" boundary, and have included language change up to the present, but they have yet to converge on a set of illustrative texts. Given that the basic narrative of later changes has become one of divergence into national and world varieties of English, as will be discussed later, perhaps such convergence on texts is less likely to happen.

## 4.2 Dialects

Compared to periods, dialect boundaries are emphasized far less by specimen texts, though the dialect boundaries remain important in HEL textbooks. Dialects receive the most attention in discussions of ME. A few texts like Millward and Hayes (2012: 210–212) include side-by-side passages of the same text copied from different dialect regions to show the diversity of ME dialects. More often, separate texts are given to illustrate distinct dialects. The *Ayenbite of inwit* might be included as an example of the Kentish dialect, and *Sir Gawain and the Green Knight* is sometimes included as an example of the West Midlands dialect that stands in

marked contrast to the nearly contemporary poetry of Chaucer. A few passages are also included because they portray dialects in literature, like the Northern dialect in the *Reeve's tale*, and the Southern dialect in the *Second shepherds play*. Furthermore, textbooks often include explicit observations about dialects, like the passage from Trevisa in (1), which is widely repeated in HEL textbooks.

(1) *for men of þe est wiþ men of þe west, as hyt were vndur þe same party of heuene, acordeþ more in sounyng of speche þan men of the norþ wiþ men of the souþ; þerfore hyt ys þat Mercij, þat buþ men of Myddel Engelond, as hyt were parteners of þe endes vndurstondeþ eyþer oþer.* (quoted in Lounsbury 1896: 118)

Several HEL accounts also include Caxton's story about the innkeeper who said she could not understand French when a traveler used a Northern form for *eggs*, in (2).

(2) *the gode wyf answerde that she coude speke no Frenshe. And the marchaunt was angry, for he also coude speke no Frenshe, but wolde haue hadde egges, and she vnderstode hym not. And thenne at laste a nother sayd that he wolde haue eyren. Then the good wyf sayd that she vnderstod wel.* (quoted in Lounsbury 1896: 160)

The illustration of dialects continues to a lesser degree for EModE. Tellingly, few texts are given as exemplars of regional dialects in EModE; with the increasing standardization of English, textbook writers apparently have fewer exemplars to include. But portrayals of regional dialects are still frequently repeated, like Edgar's portrayal of Kentish dialect in *King Lear* or Llewelyn's portrayal of Welsh dialect in *Henry V*. Occasionally textbooks include explicit remarks about regional dialects, too, like the oft-cited claim from George Puttenham that the English within sixty miles of London is the best kind to use in poetry, in (3).

(3) *ye shall thefore take the vsuall speach of the Court, and that of London and the shires lying about London within lx. Myles, and not much aboue* (quoted in Mugglestone 2006: 170)

For LModE and PDE, specimen texts illustrating dialects shift to exemplars of national dialects and World Englishes. It's not uncommon for individual textbooks to include illustrative passages of American English, Australian English, Irish English, or Scottish English. And in more recent years, textbook writers like Graddol et al. (1996), Brinton and Arnovick (2011), and van Gelderen (2014) have

included passages from Indian English, African English, Singapore English, and other World Englishes. Unlike the exemplars of ME dialects, however, there is practically no convergence on texts for these illustrations.

These texts occur regularly throughout the tradition. Trevisa's remarks show up as early as 1851 in Craik's textbook and is regularly included afterwards. Caxton's preface was first included in Emerson (1894). The text is repeated in assorted textbooks throughout the tradition, but may be a little more popular more recently. The illustrations of dialects in the EModE period is less common, with seven books each including passages from *King Lear* and *Henry V*. Illustrating national dialects is a fairly recent phenomenon. Keane (1875) was the first of many to include Barbour's *Bruce*, and Bryant (1948) first used Burns to illustrate Scottish English. As a national language, Scottish is not commonly illustrated until the 1990s. The first specimen of American English shows up in Lounsbury (1896), but not necessarily to illustrate American English. Krapp ([1909] 1966) includes a section on national dialects, but does not provide any specimens. For most of the twentieth century, American specimens remain limited to American textbooks, though by the twenty-first century most textbooks include specimens of American English and several other varieties outside England. For World Englishes, Harrison (1848) could be the first to include a specimen with his passage from a "West Indies" creole, though he includes it to criticize it. Serious illustration of World Englishes does not start until the 1990s.

But the dialect boundaries have never been emphasized as fully as the period boundaries. Even for ME, a relatively small number of specimens are given to illustrate dialectal differences. Undoubtedly, practical considerations partially account for the privileging of period boundaries over dialectal ones: it would be unmanageable to provide the history of each dialect. That much diversity would be difficult for beginning students to grasp. At the same time, as the next two sections will discuss, a central theme of that single history has been the move toward Standard English, which is often assumed to be metonymous for "the English Language". The continued persistence of dialects would have been a nuisance or even embarrassment in such a move. Perhaps the recent inclusion of national dialects and World Englishes into this history shows more confidence in recognizing that English is not limited to one variety in one location.

## 4.3 Literary language

Another trend that stands out is the preponderance of literary texts and authors. In Table 1, Chaucer, Shakespeare, and Milton are in the top six quoted authors, and in Table 2, the *Canterbury tales* is joined by *Beowulf*, *Piers Plowman*, and *Sir*

*Gawain and the Green Knight*. The boundaries around literary works and literary language will be hard to define in detail, but these works and authors have clearly been considered important literary texts for centuries, and early HEL accounts explicitly maintained the superiority of this variety of English. Crowley (2003: 77–105) has detailed how early HEL writers saw "the English language" whose history they were documenting as being literary language. Milroy (2002: 11) similarly writes, "they were attaching the canon of language history to the canon of literary history, of which it was (for them) a branch". Indeed, one boundary that early HEL accounts did not maintain very well was that between the history of the English language and the history of English literature. Johnson's (1755) account, which he initially included with his dictionary, included practically nothing but literary texts, from Alfred's translation of Boethius to Barclay's *Ship of fools*. He even evaluated the texts by the degree that they represented literary craft:

> Nearly about this time, the following pieces of poetry seem to have been written, of which I have inserted only short fragments; the first is a rude attempt at the present measure of eight syllables, and the second a natural introduction to *Robert of Gloucester*, being composed in the same measure, which, however rude or barbarous it may seem, taught the way to the *Alexandrines* of the *French* poetry. (Johnson 1755: 6)

Toller (1900: 246) is explicit about the importance of literary language:

> [printing] consequently tended to subordinate all local forms of speech to the one form that was used in the literature; and so, for our purpose, it is no longer necessary, as was the case for the times before Chaucer, to consider the different forms of English to be found in different localities; the language of literature becomes the representative English language.

The title of Marsh (1871) is telling: *The origin and history of the English language and of the early literature it embodies*. Similarly, Craik wrote a thorough history of the English language in 1851 and then wrote a history of English literature in 1864. These early accounts certainly bear out Milroy's (2002) claim. If anything the boundary separating literary language (and therefore legitimate) from non-literary language (and therefore non-legitimate) was sometimes sharper and more enforced than the boundary between English language and English literature.

The explicit championing of literary language attenuated in the twentieth-century textbooks, yet literary texts have continued to be among the most frequently used as illustrations. Chaucer's writings are used in almost all textbooks, while Shakespeare's writings show up in most. Both authors are included in all textbooks since 2006. Milton shows up in nineteen textbooks; these tend to be from earlier accounts, though he is still included in a few recent textbooks, like Freeborn (2006) and Gramley (2012). Almost as a mirror image, we see *Sir Gawain and the Green Knight* in eight textbooks all coming since McLaughlin (1970).

*Beowulf* has been included fairly regularly throughout the tradition. Today's textbook writers are not necessarily trying to privilege literary language, and perhaps one reason they continue to include literary texts is the greater familiarity of such texts for their readers: Shakespeare and Chaucer will require less explanation than, say, Margaret Paston or John Fisher. Yet the explicit ideological work from the past continues, to a degree, as the literary texts are perpetuated as samples of "the history" of "the English Language". Whether textbook writers continue to subscribe to the superiority of literary language, they may be unintentionally conferring more importance or legitimacy to that variety of English that Toller championed as a national variety, and that has continued to receive much institutional sanction.

Here the important boundary reinforced by the selection of literary texts is not so much limited to "literary" vs. "non-literary" language, but more generally to language that is sanctioned by tradition and education and language that is not. Milroy (2002) speaks of such privileged language as the "legitimate language", which is that variety used in education, business, government, and other powerful institutions. The legitimate language has aspects of literary language, Standard English, and upper-class English combined. It is hard to come up with one name for it, though "Standard English" is most often used. The preference for literary specimen texts is one way that the legitimate language is privileged in most HEL textbooks, whatever the intentions of the textbook writers. The role of HEL for defining and reinforcing this legitimate language has been noted especially by Milroy (2002) – a legitimate language has a history – and is even more evident in discussions of Standard English, which is the topic of the next section.

## 4.4 Standard English

The issues of legitimate language point to one of the chief boundaries that specimen texts have reinforced: the boundary between Standard English and non-standard varieties. The boundary is so strong, in some cases, that Standard English is equated with the English Language. Specimen texts have converged the most to document contemporary opinions regarding the standardizing processes of English, described by Haugen (1966: 929-934) and amplified by Milroy and Milroy (1999: 22) and Nevalainen and Tieken-Boon van Ostade (2006: 273–286) as consisting of Selection, Acceptance, Diffusion, Maintenance, Elaboration of function, Codification, and Prescription. Specimen texts may illustrate the increasing status of English vs. French (Selection, Elaboration of function), the robust dialect differences in ME (Selection), the social dialects and the development of a national

standard in the sixteenth century (Selection, Diffusion, Maintenance), the tensions from borrowing learned Latin and Greek terms in sixteenth century English (Elaboration of function), or the desire to regulate the language that crescendoed in the eighteenth century (Codification and Prescription). These issues account for the high ranking of writers like Trevisa, Caxton, Puttenham, and Wilson in the most-cited lists. Crowley (2003: 23) noticed this basic standardization story and its use of these specimens in nineteenth-century HEL accounts:

> Frequent references are made to specific texts of particular authors in an ordered chronology: Higden's *Polychronicon*, Chaucer's comments on linguistic diversity in *Troilus and Criseyde*, Caxton's Preface to *Eneydos*, Waller's complaint about English orthographical mutability, the Royal Society's desire for simple prose, Swift's *Proposal*, Johnson's *Dictionay* and so on familiarly. The pattern of repetition set out in this taxonomic classification (itself familiar to the nineteenth century), clearly delineates a designated field of knowledge as it traces the outlines of a structure that holds these texts together.

This story has continued in textbooks throughout the twentieth century and into the twenty-first. For the concerns about English's status vis-à-vis French, HEL writers have relied heavily on Trevisa's comments in his translation of Higden's *Polychronicon*, shown in (4).

(4) *Þis manere was moche i-vsed to fore þe firste mereyn and is siþþe sumdel i-chaunged* [. . .] *gentil men haueþ now moche i-left for to teche here children Frensche*. (quoted in Baugh and Cable 2012: 147)

Several textbooks also quote Chaucer's description of the Prioress, in (5).

(5) *And Frensh she spak faire and fetisly,*
*After the scole of Stratford atte Bow*
*For Frensh of Paris was to hir unknowe* (quoted in Baugh & Cable 2012: 138)

These remarks focus on the diminishing prestige of French and the selection of English for public use. Other passages show the increasing number of functions that English performs, such as the Proclamation of Henry III.

These two themes (Elaboration of function and Selection of English) continue in specimens from the sixteenth century, where Latin replaces French as the language that English encroaches upon. When it comes to the stratification of English vocabulary and inkhorn terms, Wilson's remarks from the *Arte of rhetorique*, along with his tongue-in-cheek inkhorn letter, are often included, as in (6).

(6) *I know them that thinke* Rhetorique *to stande wholie upon darke wordes, and hee that can catche an ynke horne terme by the taile, him they coumpt to be a fine Englisheman, and a good* Rhetorician (quoted in Baugh and Cable 2012: 213–214)

A few textbooks also mention the Royal Society's promotion of a particular English style for Scientific writing (Graddol et al. 1996; Freeborn 2006) or Isaac Newton's switch from Latin in *Principia mathematica* to English in *Opticks* (Gramley 2012: 146).

Added to the notion of selecting English instead of French or Latin is the notion of selecting a particular regional dialect. Such selection is occasionally illustrated for ME with the Chancery standard, but it is illustrated much more frequently for EModE. We have already seen Puttenham's advice to use the English spoken within sixty miles (section 4.2, example 3), and Puttenham adds social dialects to the process of selection. In one passage, he first suggests that poets use the language "which is spoken in the kings Court, or in the good townes and Cities within the land" (quoted in Mugglestone 2006: 169–170), but then advises against using other social dialects, as in (7).

(7) *the speach of a craftes man or carter, or other of the inferiour sort, though he be inhabitant or bred in the best towne and Citie in this Realme, for such persons doe abuse good speaches by strange accents or illshapen soundes, and false ortographie*. (quoted in Mugglestone 2006: 169–170)

For the eighteenth century, common specimens speak less of Selection and Elaboration of function and more of Codification. Passages from Johnson and Swift are most common, such as the passage from Swift's *A proposal for correcting, improving, and ascertaining the English tongue* in (8).

(8) *I* [. . .] *complain* [. . .] *that our language is extremely imperfect; that its daily improvements are by no means in proportion to its daily corruptions; that the pretenders to polish and refine it have chiefly multiplied abuses and absurdities; and that in many instances it offends against every part of grammar*. (quoted in Fennell 2001: 150)

In illustrating the standardization process, these specimen texts emphasize several boundaries, such as the boundaries between English and other languages, particularly French and Latin, the boundaries between regional and social dialects, and even the boundaries between prescribed and proscribed variants.

The specimen texts supporting the story of standardization are commonly used throughout the textbook tradition. Trevisa is first included in Craik (1851) and continues to be included in most twenty-first-century textbooks. Similarly, Puttenham's remarks are included in Trench's (1889) textbook and in five twenty-first-century textbooks (Fennel 2001; Nielsen 2005; Mugglestone 2006; Hogg and Denison 2006; Freeborn 2006). Wilson is first included in Johnson's account (1755) and last included in Baugh and Cable (2012).

The particular story that textbook writers tell about the standardization of English varies. There are clearly some triumphalist stories that none-too-subtly validate a single Standard English. Within this narrative, speakers chose English instead of French and Latin. They developed linguistic resources to handle areas that had previously been handled by Latin and French. They made English a national language that would be understood by all speakers, and regulated it enough for it to be useful. But not all textbook authors are as explicitly triumphalist, and some include the same passages more to show how our notions of a single legitimate language have been historically constructed (e.g., Freeborn 2006; Gramley 2012). What hasn't changed, though, is the perceived usefulness of these passages. Clearly the boundary between Standard English and other varieties of English is one of the most important boundaries in HEL textbooks, whether the intent is to reinforce or question those boundaries.

The concern for English standardization in HEL textbooks points to another boundary, namely boundaries between classes of speakers. The story that legitimizes this variety of English typically focuses on the perspective of those most invested in Standard English. For example, the increasing status of English vs. French can be portrayed as a crucial step in the development of English for all speakers: we speak English today, not French, and this article is written in English, not Latin. But in reality, the increase in English's status and its increased functional load mattered more to the upper crust of society; Trevisa's observation explicitly notes that it is gentlemen's children who learn French less after the Black Death, and it is a Prioress that Chaucer satirizes as having provincial French (examples 4 and 5 above). Much of the country would have had no French at all, nor would they have needed any. There would have been little tension between English and French for much of the populace, and only certain segments of society would have been initially affected by the increased status of English. So when a king's proclamation is delivered in English or when courts or parliament begin to use English, it reflects a change mainly for the powerful segments of society. The highlighting of these moments also highlights the perspective of a certain class.

Similarly, the concern for the stratified vocabulary that resulted from French and Latin borrowings applies more to certain segments of society than others.

Yes, the large-scale borrowing did make English more difficult to understand for people who did not know those languages, but initially only for those speakers who needed to use the language characterized by such borrowings. The tension between English and Latin reflected in the inkhorn terms would have been felt more by those seeking to use some kind of language of wider communication more than by those who needed English only for their local uses. For the local English user, the stratified vocabulary and inkhorn terms would have been much less an issue, except as they came into contact with the prestigious language, as in court cases or schools once education became universal.

Yet saying that HEL has traditionally been a justification and legitimation of Standard English – as valuable as that insight is – does not go far enough. The nature of that single "legitimate language"–even if it is called Standard English – is more complex and deserves closer examination. As several scholars have mentioned (discussed well in Crowley 2003), standard can have several meanings. The equation of Standard English with legitimate language keys on the sense of standard as 'excellent' or 'of the highest quality'; another important meaning of standard is 'having widespread currency'. Another way of looking at the single variety that HEL accounts have legitimized in the past is the language of widest coverage or the variety used for wider communication. It is employed in writing for an audience that goes beyond local communication needs. Because it is widely used, perhaps HEL accounts are justified in focusing on that variety.

Seeing the standard as a variety used for wider communication complicates the process in forming the Standard English defined in most HEL accounts or, what's more, in forming the canonical history of the English Language. The history of English can be seen as the appropriation and legitimation of the language variety used by the powerful segments of society; their English becomes *the* English Language. But the history of Standard English can also be seen as a leveling of power, as this variety of English becomes accessible to many classes of people. Perhaps the use of English instead of French by gentlemen should be seen as revolutionary: gentleman sided with English-speaking non-gentleman against the use of French. Similarly, even if inkhorn terms were troublesome for those who needed to use a language of wider communication but who didn't know Latin or French, the reaction against them still owed to their impeding wider understanding, not the confusion they created for the powerful and educated, who already knew Latin. The concern was in communicating with those who didn't know these languages. The discussion of inkhorn terms and stratified vocabulary doesn't entirely draw a boundary between the language of the upper and lower classes (though there was much effort expended on that task, as well), but in some ways it effaces that boundary, as the English of wider communication widens to reach as many people as possible. The

triumph can be seen for the language of wider communication, too, not simply the language of the powerful.

Several scholars have pointed out that the history of the English language has been told as the history of the language of the most powerful segments of society. To be sure, much in HEL accounts, including the selection of specimen texts has historically strengthened the notion that the English worth studying – the English with a history – is the English of the powerful. At the same time, the boundary-work done by HEL accounts and particularly specimen texts reveals another cultural current in which writers from the past show their anxieties not about hegemonic English, but about other hegemonic languages, namely French and Latin. In this current, the use of English – even the use of the variety that itself becomes hegemonic – is seen as the revolutionary language, the disruptive language.

## 5 Conclusion

Specimen texts open an interesting window on to the history of the English language, and they are often the first representatives of early English that readers encounter. With a language history as diverse as English's, it is, perhaps, reassuring that textbook writers employ a wide variety of specimen texts. At the same time, textbooks as a group have converged on certain texts and authors, some of whom would seem surprising, given their otherwise inconspicuous presence in our culture today. We have seen in this paper that these most-repeated texts highlight important boundaries within the history of the English language, namely those boundaries between periods, dialects, registers, and varieties of English. The boundaries between periods appear to be here to stay, and we can expect continued illustration of periods and even reification of those periods by such illustration. The boundaries between dialects have been less stable. One impulse, more popular in earlier textbooks, but still persistent in recent textbooks, is to emphasize a history of convergence, so that dialect differences are noted, but not emphasized. And when dialect differences are so great that they must be recognized, as in ME, there is at least a hint that these dialect differences are a foil for the later convergence upon a standard. Another impulse, however, is to recognize the divergences in English with its spread throughout the world. This impulse first showed up in the recognition of national standard varieties, but it has also been extended to World Englishes. How HEL textbooks negotiate the complexity that will come with including more dialects within the history of the English language will be interesting to watch. The boundary between literary language and

non-literary language is a complexly durable boundary. No longer do textbooks openly champion literary language, but literary language still holds enough attraction to be frequently chosen as illustrations of the history of English. The passages are better known for students and instructors, partly because of traditional literary education and partly because of the HEL textbook tradition itself. No doubt, literary passages will continue to be used as specimen texts, but it will also be interesting to see how conscious textbook writers are in counteracting the reifying and legitimizing of just one variety that such use performs.

The most controversial boundary may well be the one between Standard English and other varieties. Scholars like Crowley (2003), Milroy (2002), Crystal (2005), Bailey (2002), Watts (2011), and Sevenker (2016) have skillfully pointed out that that HEL accounts celebrate mainly just one variety of English – a variety they often call Standard English – so that the English Language is often regarded as this variety of English. At least since the 1970s, textbooks have recognized this tension, but the reification implied in *the history of the English language* will continue to prompt the privileging of some kind of relatively homogenous version of English.

I have not meant to object to these particular boundaries or even the practice of defining boundaries. Defining boundaries is a necessary function for any discipline (see Adams, this volume), and in this case, the boundaries are defining the fundamental element of the history of the English language, namely the English language whose history is narrated and studied in the discipline. Perhaps it should not be surprising that the boundaries that have been drawn and reinforced in HEL textbooks have tended to favor the kind of English we use most in university courses. This is the variety that is valuable to the instructors and presumably to students as they gain surer facility with it. How English came to be used for such prestigious functions is still an interesting story which the specimen texts enliven. Just as interesting are the ways that English has diverged and continues to diverge. Specimen texts increasingly enliven that story, too.

# References

## Primary sources

Algeo, John & Carmen Acevedo Butcher. 2014. *The origins and development of the English Language*. 7th edn. Boston: Wadsworth.

Bambas, Rudolph C. 1980. *The English language: Its origin and history*. Norman, OK: University of Oklahoma Press.

Barber, Charles, Joan C. Beal & Philip A. Shaw. 2009. *The English language: A historical introduction*. 2nd edn. Cambridge: Cambridge University Press.
Baugh, Albert C. & Thomas Cable. 2012. *A History of the English language*. 6th edn. London: Longman.
Blake, Norman F. 1996. *A history of the English language*. London: Macmillan.
Bloomfield, Morton W. & Leonard Newmark. 1963. *A linguistic introduction to the history of English*. New York: Knopf.
Bradley, Henry. 1925 [1904]. *Making of English*. London: MacMillan.
Brinton, Laurel J. & Leslie K. Arnovick. 2011. *The English language: A linguistic history*. 2nd edn. Oxford: Oxford University Press.
Brook, G. L. 1958. *A history of the English language*. London: Deutsch.
Bryant, Margaret. 1948. *Modern English and its heritage*. New York: MacMillan.
Camden, William. 1984 [1605]. *Remains concerning Britain*. [Ed. R. D. Dunn.] Toronto: University of Toronto Press.
Classen, E. 1919. *Outlines of the history of the English language*. New York: Greenwood Press.
Craik, George L. 1851. *Outlines of the history of the English language*. London: Chapman & Hall.
Culpeper, Jonathan. 2005. *History of English*. 2nd edn. London: Routledge.
Emerson, Oliver Farrar. 1894. *The history of the English language*. London: Macmillan.
Emerson, Oliver Farrar. 1906. *An outline history of the English language*. New York: Macmillan.
Fennell, Barbara. 2001. *A history of English: A sociolinguistic approach*. Malden, MA: Blackwell.
Fisiak, Jacek. 2000. *An outline history of English*, vol. 1, *External history*. 3rd edn. Poznan: Wydawnictwo Poznanskie.
Free, John. 1749. *An essay towards an history of the English tongue*. London.
Freeborn, Dennis. 2006. *From Old English to Standard English*. 3rd edn. New York: Palgrave MacMillan.
van Gelderen, Elly. 2014. *A history of the English language*. Amsterdam: John Benjamins.
Görlach, Manfred. 1997. *The linguistic history of English*. London: Macmillan.
Graddol, David, Dick Leith & Joan Swann. 1996. *English: History, diversity and change*. London: Routledge.
Gramley, Stephan. 2012. *The history of English: An introduction*. London: Routledge.
Gray, Austin K. 1915. The history of the English language. In H. C. O'Neill (ed.), *A guide to the English language: Its history, development, and use*, 1–24. London: T. C. & E. C. Jack.
Harrison, Matthew. 1848. *The rise, progress and present structure of the English language*. London: Longman, Brown, Green & Longmans.
Hogg, Richard & David Denison (eds.). 2006. *A history of the English language*. Cambridge: Cambridge University Press.
Johnson, Samuel. 1755. The history of the English language. In Samuel Johnson, *A dictionary of the English language*. London.
Keane, A. H. 1875. *Handbook of the history of the English language*. London: Longmans, Green & Co.
Krapp, George Philip. 1966 [1909]. *Modern English: Its growth and present use*. New York: Frederick Ungar.
Lounsbury, T. R. 1896. *History of the English language*. New York: Henry Holt.
Marsh, George P. 1871 [1862]. *The origin and history of the English language and of the early literature it embodies*. New York: Charles Scribner.

Martin, Benjamin. 1749. Institutions of Language. In Benjamin Martin, *Lingua Brittanica reformata: Or a new English dictionary*, 1–111. London.
McLaughlin, John C. 1970. *Aspects of the history of English*. New York: Holt, Rinehart & Winston.
Meiklejohn, J. M. D. 1906. *The English language: Its grammar, history, and literature*. Boston: D. C. Heath.
Millward, C. M. & Mary Hayes. 2012. *A biography of the English language*. 3rd edn. Fort Worth: Harcourt Brace.
Mugglestone, Lynda (ed.). 2006. *The Oxford history of English*. Oxford: Oxford University Press.
Myers, L. M. 1966. *The roots of modern English*: Boston: Little, Brown.
Nielsen, Hans Frede. 1998. *The continental backgrounds of English and its insular development until 1154*. Odense: Odense University Press.
Nielsen, Hans Frede. 2005. *From dialect to standard: English in England 1154–1776*. Odense: University Press of Southern Denmark.
Pei, Mario. 1952. *The story of English*. Philadelphia: J. B. Lippincott.
Peters, Robert A. 1968. *A linguistic history of English*. New York: Houghton Mifflin.
Peyton, V. J. 1771. *The history of the English language*. London.
Robertson, Stuart & Frederic G. Cassidy. 1954. *The development of modern English*. 2nd ed. Englewood Cliffs, NJ: Prentice-Hall.
Singh, Ishtla. 2005. *The history of English*. London: Arnold.
Smith, Jeremy. 1999. *Essentials of early English*. London: Routledge.
Strang, Barbara. 1970. *A history of English*. London: Routledge.
Thomas, P. G. 1920. *An introduction to the history of the English language*. London: Sidgwick & Jackson.
Toller, T. N. 1900. *Outlines of the history of the English language*. London: MacMillan.
Trench, Richard Chenevix. 1889. *English past and present*. London: Kegan Paul, Trench.
Williams, John M. 1975. *Origins of the English language*. New York: The Free Press.
Wyld, Henry Cecil. 1907. *The growth of English*. London: Murray.

## Secondary sources

Bailey, Richard W. 2002. A thousand years of the history of English. In Donka Minkova & Robert Stockwell (eds.), *Studies in the history of the English Language: A millennial perspective*, 449–472.Berlin: Mouton de Gruyter.
Beal, Joan C. 2004. *English in modern times*. London: Arnold.
Blockley, Mary. 2012. Resources: Textbooks. In Alexander Bergs & Laurel J. Brinton (eds.), *English historical linguistics: An international handbook*, vol. 2, 1178–1190. Berlin: Mouton de Gruyter.
Bragg, Melvyn. 2003. *The adventure of English*. New York: Arcade.
Burnley, David. 1992. *The history of the English language: A source book*. London: New York.
Chapman, Don. 2019. Internal versus external history: Events in HEL. In Colette Moore & Chris C. Palmer (eds.), *Teaching the history of the English language*, 54–62. New York: MLA.
Crowley, Tony. 2003. *Standard English and the politics of language*. 2nd edn. New York: Palgrave Macmillan.

Curzan, Anne. 2012. Periodization in the history of the English language. In Alexander Bergs & Laurel J. Brinton (eds.), *English historical linguistics: An international handbook*, vol. 2, 1233–1256. Berlin: Mouton de Gruyter.
Crystal, David. 2005. *The stories of English*. Woodstock, NY: Overlook Press.
Craik, George L. 1864. *A compendious history of English literature, and of the English language from the Norman Conquest with numerous specimens*. New York: Charles Scribner.
Fisher, John H. & Diane Bornstein. 1984. *In forme of speche is chaunge*. Lanham, MD: University Press of America.
Greenwood, James. 1711. *An essay towards a practical English grammar*. London.
Haugen, Einar. 1966. Dialect, language, nation. *American Anthropologist* 68(4). 922–935.
McCrum, Robert, William Cran & Robert MacNeil. 1986. *The story of English*. New York: Viking.
Millward, C. M. 1996. *A workbook to accompany* A biography of the English language. Fort Worth: Harcourt Brace.
Milroy, Jim. 2002. The legitimate language: Giving a history to English. In Richard Watts & Peter Trudgill (eds.) *Alternative histories of English*, 7–25. London: Routledge.
Milroy, James & Lesley Milroy. 1999. *Authority in language*. 3rd edn. London: Routledge.
Mitchell, Bruce & Fred C. Robinson. 2011. *A guide to Old English*. 8th edn. Oxford: Blackwell.
Nevalainen, Terttu & Ingrid Tieken-Boon van Ostade. 2006. Standardisation. In Richard Hogg & David Denison (eds.), *A history of the English language*, 271–311. Cambridge: Cambridge University Press.
Sevenker, Justin Ross. 2016. *Old English composition: Literacy and the history of English in the United States*. Pittsburgh: University of Pittsburgh dissertation.
Smith, Logan Pearsall. 1912. *The English language*. New York: Henry Holt.
Watts, Richard. 2011. *Language myths and the history of English*. Oxford: Oxford University Press.

R. D. Fulk

# Coda: HEL-bound

By and large, academicians seem to harbor ambivalent feelings about borders. The ambivalence may stem in part from the function of boundaries as devices with dual, conflicting purposes: they promote both the positive value of nurturing and sustaining community and the negative one of exclusiveness and (a term of provocation in the present political climate) elitism. For instance, in the ideal, the professoriate values the aim that the ranks of the educated should be ever more inclusive. In actuality, however, resources are finite, and aptitude is not evenly distributed. As a result, educators are obliged to be gatekeepers. In similar fashion, the carving up of academe into disciplines is essential to the progress of scholarship, since most academic accomplishments depend upon the sort of expertise that can be attained only by limiting the scope of inquiry. Yet some discoveries can be made only when researchers are able to cross disciplinary boundaries: consider, for example, how dependent the study of chemistry is upon a knowledge of the principles of physics. And policing the disciplinary borders has its unsavory aspects as well: ignorance of neighboring disciplines can be detrimental to scientific rigor, and it can lead to distorted views (see, e.g., Winters and Nathan [1992] and Fulk [2016] on relations between linguistics and philology).

Given the potential rewards to be derived from assuming a cross-disciplinary perspective, as well as the dangers that attend excessively isolative thinking, the impulse to transgress boundaries, or to abrogate them altogether, is natural enough. Throughout the humanities, in the latter part of the preceding century a major aim was to undermine the foundations of boundaries and to invalidate the opposed binarisms of either-or thinking.[1] Yet boundedness is a fundamental condition of existence, a prerequisite to identity. Humans, moreover, are predisposed to reinforce boundaries: an example of such an unconscious, involuntary impulse is pointed out by Minkova (this volume), who describes how the right and left edges of prosodic domains are lent salience. With some effort, borders can be redrawn, but rarely can they be erased.

**1** On deconstruction, see, e.g., Culler (1982). Gender has been a particularly intense focus of efforts to destabilize starkly oppositional habits of thought (see, e.g., Butler 1990; Sedgwick 2008), efforts that have grown only more concerted over time with the rise of transgender activism.

**R. D. Fulk,** Department of English, Indiana University, Bloomington, IN, USA
e-mail: fulk@indiana.edu

https://doi.org/10.1515/9783110643282-012

Considerably less labor is required simply to cross borders and act the tourist, and from this much is to be gained. The present volume is a Baedecker in this sense, or, more aptly, a photo album of exotic places visited. The field of inquiry is narrow – the history of the English language – and yet the aims and methods of the studies included are so diverse as to produce a richly informative dialogue of studies in conversation with one another. And so while the topic appears to be circumscribed, with well-marked boundaries, in fact to pass from one study to the next is to travel to where a different language is spoken. The history of the English language, to judge by these studies, is a surprisingly polyglot discipline. And yet it is remarkable how common concerns serve to link aims, methods, and domains as diverse as those of these studies. It may be useful to outline some of the ways in which the contributions to this volume tacitly interact with one another, addressing issues raised in common from widely divergent perspectives.

Adams (this volume) brings an Olympian perspective to the problem of boundary-construction, asking us to reconsider some conventional suppositions so ingrained as to appear inevitable. The question of mode – in this instance, the opposed types of narrative and analysis – goes to the heart of the linguistic enterprise, since it is inseparable from the question of what the purpose of that enterprise is, or should be. Should linguists speak to a popular audience of non-specialists? The desirability of more general relevance for the field and of counteracting impressions of academic exclusivity would appear to be self-evident. Yet concerns about the superficiality that inevitably attends popularization are not to be dismissed lightly (see, e.g., Pullum 1991, but see also Krebs 2019). The problem of superficiality is also relevant to the matter of scope: one reason that no one has written a linguistic history of the entire North Sea littoral over the course of a millennium is the inevitable shallowness such a history would evince. Adams' own preference is for microhistories, but the idea of variety in scope and mode, and of the interplay of internal and external history, points to a way to enrich the field.

Some of these problems of scope are touched on by Schneider (this volume), but from a rather different perspective, since he is an unapologetic advocate for the larger domain of Complex Dynamic Systems (CDS) and is dismissive of the mainstream Saussurean approach to language, which is "too clear and simplifying, too strongly reductive" (p. 73), representing "reductionist, segmenting and categorial thinking" (p. 101). He also descends from the abstract to the concrete, exemplifying the relevance of the concerns of CDS to the history of *wh*-pronouns in English. Perhaps somewhat ironically, since the chapter focuses on issues of complexity, it can be shown that the variety of factors involved in the rise of *wh*-relatives is even more complex than those listed here. *Hw*-pronouns did in fact

have relative uses already in Old English – usually as fused relative pronouns, so called because they fuse the relative and its antecedent in a single form, e.g., *hwæt* in the sense 'that which, what' – but occasionally with an explicit antecedent (see the *Dictionary of Old English* [*DOE*; Cameron et al. 2018] s.v., *hwā*, *hwæt* IV.A.2, IV.B.1). And it is worth noting that, in addition to such relative functions, Old English *hw*-pronouns were not simply interrogative but also indefinite, and all these uses are most likely related developmentally to the use of interrogatives in indirect questions (and some other functions of *hw*-pronouns detailed in the *DOE*). In other words, the factors that very likely played the most significant role in the development of *wh*-relatives are those most prominent in historical linguistics on the Saussurean model: the internal pressure of mutually defining entities and categories in relation to one another. Such interplay of elements is complex and worthy of detailed delineation, and it suggests that, whereas it undoubtedly is the case that historical analysis could benefit from looking beyond the confines of structuralist principles for aspects of causation, dangers do lurk in regarding Saussurean linguistics as simplistic.

Expanding the parameters of historical linguistic research by introducing an augmented range of influences is also the aim of Devitt (this volume), in this instance by taking into account the influence of genre upon the process of language change. Like Schneider, after discussion of the elements that constitute and contribute to the concept, she concretizes the issue, exploring how her 1989 study of several diagnostic variables in the slowly anglicizing Scots of the Early Modern period both confirmed the significance of genre as a regulating factor in the change and failed to take into account all the relevant principles to be derived from regarding genres as social actions fusing linguistic forms and extralinguistic influences. It seems a particularly significant insight to acknowledge the indeterminacy of genres – the difficulties that attend defining them, their characteristic tendency to shift and evolve and maintain only porous borders – thus highlighting from an unrelated perspective the dynamism of linguistic systems on which Schneider rightly lays stress.

The contribution by Liberman (this volume) could hardly be more different, being a study of what became a single phoneme in the course of Germanic language history and of its relation to some historical questions of long standing, but it shares with Devitt's a focus on the consequences of indeterminacy. Liberman argues that it is the indeterminacy of /r/ – its featurelessness and peripherality – that is the chief cause of its phonetic instability. As with so much of Liberman's work, this study does not shy away from staking out bold and unusual positions (in a sense, going peripheral), but in doing so it is rendered capable of offering some striking solutions to problems which, though long studied, still lack consensual explanations. On the unorthodox assumption that

there was palatalization of consonants in early Germanic, it becomes possible to explain the failure of /r/, alone among the consonants, to undergo gemination in West Germanic. More arresting is the argument that of the two sources of /r/ in North and West Germanic – the /r/ inherited from Proto-Indo-European and the /r/ that developed by rhoticism from */s/ > */z/ > */R/ – the former triumphed over the latter everywhere but in English, where the latter became the standard, thus explaining the isolation of English among related languages as regards its idiosyncratic articulation of /r/. Indeterminacy thus comes to be seen as a powerful explanatory tool and a source of dynamic change.

Something of this idea of the reach of indeterminacy puts Liberman into conversation with a study of an entirely different sort, the chapter by Hoffman and Kytö (this volume) on the use of English, Swedish, and Heritage Swedish in Swedish-American cookbooks. The authors highlight the permeability of the borders between the relevant languages as deployed in these cultural artifacts. Hoffman and Kytö also tacitly forge links to the work of Devitt (this volume), not just in regard to her characterization of genres as indeterminate and unstable, but by their application of the concept of enregisterment. Studying how extralinguistic features of the texts reinforce the cultural work performed by code-mixing and code-switching, they illustrate her point that a genre, such as the culinary recipe, is not to be defined solely on the basis of linguistic or contextual considerations, but it must be regarded as a social action entailing both.

Yet if liminal status and indeterminate borders are motives for change, they are also a prompt for reinforcement, with the goal of counteracting instability. Like Hoffman and Kytö, Chapman (this volume) studies a particular textual genre, the HEL textbook, and, unlike them, finds evidence of boundary-reinforcement in the selection of texts illustrating aspects of the history of English. He highlights the continual re-use of particular texts and passages, attesting to something resembling a canon of illustrative texts. The effect of canonicity is to reinforce "boundaries between English and other languages, between periods of English, between dialects of English, and between registers of English" (p. 249), even though recent textbooks tend to signal an awareness of how questionable an aim it is to portray the development of English as leading to a single variety or register. The paradox of boundaries both promoting and inhibiting change may seem more natural when the ambivalent attitudes toward boundaries mentioned at the outset of this coda are recalled.

Other examples of boundary reinforcement are highlighted by Minkova (this volume), whose aim is to show that the strengthening of right and left edges in domains at various levels that is evident in contemporary English is discernible, as well, in Old English. The formal properties of poetry highlight peripheries in various ways. At the syntactic level this is exemplified by the

metrical prominence (presumably correlating to stress) of rightward-shifted sentence elements that would normally be accorded no prominence when unaffected by "displacement" – a principle more commonly known as the *Satzpartikelgesetz*, or Kuhn's first law of sentence particles (Kuhn 1933). At the phonological level the principle is exemplified by the formal requirements of alliteration, or onset-matching, including the preferential treatment of nomina, the matching of initial clusters, and vocalic alliteration. Most striking is that the assumption of right-edge strengthening furnishes a brilliant solution to a puzzle of long standing: why only the poet of *Beowulf* appears to have been able to distinguish between etymologically heavy and light codas verse-internally, whereas later poets seem for the most part to have known to avoid light codas only verse-finally. Minkova proposes that heightened phonological prominence at the right edge of the verse renders length distinctions there more salient.

The selection of representative texts in textbooks, as in Chapman's study, is to be regarded as a conscious choice by individuals, as is the determination of alliterating staves in Minkova's. Yet most of Minkova's examples illustrate a commoner sort of linguistic phenomenon, since the choices involved are involuntary and ingrained. The study by Schneider and Buschfeld (this volume) provides further examples of the latter type, whereby the lexeme ONE "mov[es] along a grammaticalization path, adopting increasingly abstract grammatical rather than concrete lexical referential functions" (p. 162). Redrawing the boundaries between ONE and other lexemes in their function as grammatical markers conveys a sense of the extent to which constructing and deconstructing boundaries is an ineluctable facet of human cognition, and thus of language. The Enlightenment program of "fixing" the language, rendering it unchanging, was doomed to failure not simply because of a perverse lack of resolve among the prescriptivists' contemporaries, which they thought could be surmounted, but because the cognitive boundaries on which linguistic stability depends are rendered porous by the very nature of consciousness.

The distinction between consciously and unconsciously motivated change is itself something of a porous border, since linguistic competence and performance are not always distinguishable with certainty. The study by Smitterberg (this volume) illustrates why. Examining the incidence of comma insertion between subject and predicate, when no correlative construction is involved, in two textual genres in the second halves of the seventeenth and nineteenth centuries, he finds two likely motivations: to mark the close of a lengthy subject and to indicate a rhetorical pause in oral delivery. At this historical remove it is frequently impossible to ascribe with confidence an instantiation to one or the other cause. Yet even a live informant might often have difficulty pinpointing the cause: it could be that both play a role at once, or it might involve yet a

third motive as yet unidentified, or it could be that, although both motives would appear to result from a conscious choice, an informant could not identify the motive, only the need for a comma. The purpose of historical linguistics is surely to identify and explain language change, and yet explanations themselves mark out boundaries, excluding alternatives, when in fact explanatory indeterminacy may at times be the best explanation of all.

As with many of the contributions to this volume, particularly those of Schneider and Devitt, as discussed above, the study of Old English verbs of envying by Vea Escarza (this volume) foregrounds methodology, in this instance with such extensive attention as perhaps to command more attention than the findings themselves. A particularly striking methodological innovation is the co-application of two discretely bounded theories of the relation between verbal semantics and syntactic behavior. Here the examination of boundaries extends into two discrete dimensions: explicitly, with the demarcation of a class of verbs with four members, and implicitly, with the insistence upon the extension of analytic relevance beyond the realm of semantics, upon which the selection of verbs is predicated, to include syntactic properties of the verbs. Troubling the border between syntax and semantics surely returns us to the spirit of the contribution with which this survey of the volume's contents began, Adams' call for a fundamental rethinking of the discriminations routinely taken for granted in the study of English language history.

The contributions to this volume, whether focused on broad issues that touch upon the very foundations of historical linguistics or on matters of fine detail, or on something between the two, collectively pose a challenge. They ask us to rethink the limits on the study of English language history and to consider alternative interpretations, aims, definitions, research questions, sources of data, levels of analysis, kinds of evidence, ways of grouping, methods, scales, and modes. That is to say, they expand the boundaries of the field, inviting new kinds of research and new ways of approaching issues of long standing. They suggest where HEL is ideally bound: unboundedness.

## References

Butler, Judith. 1990. *Gender trouble: Feminism and the subversion of identity*. New York: Routledge.

Cameron, Angus, Ashley Crandell Amos, Antonette diPaolo Healey, et al. 2018. *Dictionary of Old English: A to I* online. Toronto: Dictionary of Old English Project. https://tapor.library.utoronto.ca/doe/. Accessed July 20, 2019.

Culler, Jonathan. 1982. *On deconstruction: Theory and criticism after structuralism*. Ithaca: Cornell University Press.

Fulk, R. D. 2016. A philological tour of HEL. In Don Chapman, Colette Moore & Miranda Wilcox (eds.), *Studies in the history of the English language VII: Generalizing vs. particularizing methodologies in historical linguistic analysis*, 11–27. Berlin: Mouton de Gruyter.

Krebs, Paula M. (ed.). 2019. *Public humanities in action*. Issue of *Profession*, Spring 2019. https://profession.mla.org/public-humanities-in-action/ (Accessed July 2019.).

Kuhn, Hans. 1933. Zur Wortstellung und -betonung im Altgermanischen. *Beiträge zur Geschichte der Deutschen Sprache und Literatur* 57. 1–109.

Pullum, Geoffrey. 1991. *The great Eskimo vocabulary hoax, and other irreverent essays on the study of language*. Chicago: University of Chicago Press.

Sedgwick, Eve Kosofsky. 2008. *Epistemology of the closet*. Berkeley: University of California Press.

Winters, Margaret E. & Geoffrey S. Nathan. 1992. First he called her a philologist and then she insulted him. In Diane Brentari, Gary N. Larson & Lynn A. MacLeod (eds.), *The joy of grammar: A festschrift in honor of James D. McCawley*, 351–367. Amsterdam: John Benjamins.

# Index

actuation 8, 34
Adams, Michael 24, 33–34, 227, 231, 267, 272, 276
adjective 3, 117, 135, 138, 141–142, 146, 148–149, 151, 163
– phrase 169, 178, 180
adverbial 6, 144, 157, 171, 180, 190
*æfestian* 14, 187, 195–199, 202–205
Alfred 128, 252, 256, 260
alliteration 13, 112, 114–120, 122, 129–130, 275
– *see also* vowel
alternation *see* verb
American English (AmE) 13, 89–91, 96–98, 136, 139–140, 149–163, 225–226, 258–259
analogy 91, 100, 125, 205, 212
*andian* 14, 187, 195–197, 200–204
anglicization 60–67
ARCHER Corpus 84, 91–92, 96, 139–140, 168–170, 178

*Beowulf* 54, 111, 115–119, 122–128, 256, 259, 261, 275
Biber, Douglas 7, 10, 45, 47–48, 63, 65, 140, 170–171, 184
the Bible 61, 64, 84, 172, 253–257
– translation 255–257
borrowing 8, 10, 78, 86, 100, 226, 262, 264–265
boundary 1–15
– clausal 115
– crossing 1–2, 4, 7–9, 11, 14–15, 164
– emic 3, 5, 10–14
– etic 3, 5, 9–14
– formalist 6
– generic 45–70, 259–261
– lexical 187–204
– maintenance 3, 14
– morphological 110
– national 4–5, 8, 27, 76, 257–259, 264, 266
– orthographic 109–110, 130, 237
– perceptual 1, 4, 7, 116–117
– periodization 1, 9, 33, 65, 254–257
– phonetic 128–130
– phrasal 115, 120
– punctuation 167–184
– regional 4, 7, 73–101, 257–259
– semantic 6, 8, 87, 197–205
– sociolinguistic 4–5
– syllabic 219, 221
– syntactic 6, 174–177, 197–205
– word class 2, 3, 5–6, 8, 136–164
Braudel, Fernand 23, 27–30, 34–35
British (BrE) 13, 89–91, 96–98, 130, 136, 139–141, 148–163, 151, 170, 184, 211, 213, 218
Brown Corpus 84, 89, 91–92, 96–97

Canterbury Tales 138, 249, 253, 259
– *see also* Chaucer, Geoffrey
case marking 74, 83, 87, 100
– accusative 88, 194, 198–201, 203
– dative 81, 88, 194, 198–201, 203
– genitive/possessive 74, 83, 86, 194, 198, 201, 203, 241
– loss 83, 85, 87–91, 101
– nominative 194, 198, 200–203
Caxton, William 94, 137, 252–253, 258–259, 262
Chaos theory 75, 79–80
Chaucer, Geoffrey 93–94, 138, 249, 252–254, 258–262, 264
China English 81
Cho, Taehong 5, 116, 118, 122, 128
clause
– appositive 177–178
– binding 91, 100
– complex 194
– imperative 171, 173
– main 171–173, 175, 177, 179, 183–184, 191
– participle 180
– relative 84, 144, 146, 157, 167, 170, 178
– subordinate 171, 203
coda 125–127, 274
code-mixing, *see* code-switching
code-switching 8, 226, 231, 236, 241, 244–245, 274

https://doi.org/10.1515/9783110643282-013

codification 169
colonial lag 89, 151, 163
comma 13, 167–184, 275–276
– correlative 167–168, 171–172, 178, 275
– non-correlative 13, 167–178, 181–184
– pseudo-correlative 178, 180
community 3–4, 8–9, 12, 14, 46, 50–59, 61–68, 76, 225–231, 239–241, 245–246, 271
– cultural 51
– discourse 50–51, 64
– host 225, 230–231
– immigrant 225, 245
– of practice 52
– rhetorical 51
– religious 64
– speech 52, 76, 241
complement 143, 155–156, 160, 162–163, 178, 189, 194, 200
Complex Adaptive Systems 75
Complex Dynamic Systems (CDS) 12, 74–101, 272
– auto-organization 82, 84, 86–87, 92–93, 99, 100
– butterfly effect 76, 81, 86, 101
– complexity 75–77, 79–80, 82, 84–87, 91, 93–96, 99
– dynamics 76, 78, 83, 86, 89–91, 98–99, 101
– dynamism 78, 86–87, 93, 96, 273
– emergentism 76, 82, 84, 86, 90–91, 98–100
– network 76, 78–79, 84, 86–87, 89, 90–91, 93–94, 96, 99
– order and chaos 76, 79–80, 87, 91, 100
– systemness 76–77, 86, 99
– turbulence 79, 80, 100
Complex Systems 12, 75, 162
composition and rhetoric 7, 12, 46, 61, 67
compositionality 80, 82
consonant 109, 113, 116, 118–120, 129, 211–212, 214, 216–222, 274
Construction Grammar 6, 82, 163
cookbook 226–227, 229, 240
– para-textual dimensions 231, 233, 237, 245–246
coordination 191
– *see also* Role and Reference Grammar
cosubordination 191
– *see also* Role and Reference Grammar
Crowley, Tony 250, 260, 262, 265–266
Curzan, Anne 1, 9, 24, 28, 254

D'Arcy, Alexandra 28, 163
Devitt, Amy 25, 45, 47, 51, 53–55, 58–60, 63–64, 68, 78, 273–274, 276
dialect 221, 231, 249, 257–259, 261, 263, 266, 274
– regional 4, 7, 73, 80, 85
– social 80, 85, 261
– standard 15, 261–262
– *see also* Standard English
dialectology 4–5, 219
dictionary 33–34, 36–37, 253, 260
diffusion 65, 86, 98, 100
– Standard English 60, 261–262
discourse 9–10, 24–25, 48–50, 52–53, 55, 59, 226
domain
– metrical 116
– prosodic 109, 111–113, 115, 120–121, 271
– usage 226, 229, 231, 245

edge
– Englishes 83
– genre 46, 63
– language 241
– meter 109, 111, 115
– prosody 109, 113, 116–117, 119–120, 129
– verb category 189, 204
Early Modern English (EModE) 13, 84–85, 87–89, 91–92, 94–95, 137–139, 254–256, 258–259, 263
English Language Complex 73
enhancement 117, 119–120, 126
enregisterment 227, 231, 237, 245–246, 247

Finegan, Edward 10, 61, 140, 184
foot 115–116, 123, 233
Freadman, Anne 49, 55–56
French 85–86, 258, 260–266
Fulk, R. D. 54, 111, 115, 123–128, 271

gemination 218–222
gender
– as a social category 4, 59, 85, 101, 271
– grammatical 84–85, 101
genre 7–10, 12, 13, 25–26, 30, 36, 88, 140, 168–176, 179–184, 229, 273–275
– norm 52, 58–59, 66–67
– set 46, 55–58, 61, 63–64
– social action 7, 12, 46–47, 49–54, 57–58, 61–63, 67, 273–274
– stability 46, 57–58, 65–66, 68
– typification 46, 49–51, 53, 57–58, 66
– uptake 46, 55–57, 61, 63, 64
Germanic 14, 80–81, 101, 116, 118–120, 125, 211–222, 273–274
– Main Stress Rule 81
glottal stop 119–120
Goldilocks principle 256–257
Görlach, Manfred 8, 10, 45–48, 50, 53–58, 61–64, 68, 74, 151, 254
Gothic 120, 212, 215, 218–220
grammaticalization 13, 93–94, 137–138, 160, 162, 164, 275
Grund, Peter 9, 30, 45, 55, 65, 83, 85

Helsinki Corpus 52, 84, 88, 91, 95, 138, 140
heritage language, *see* Swedish
historical linguistics 1–2, 7–12, 23–27, 39–40, 45–48, 52, 54, 56, 65, 67–68, 273, 276
historiography of the English language 23, 26, 32, 38
– annales 27–29, 33, 35, 37
– microhistory 28, 30–31, 33–34, 36, 272
– mode 24–28, 31–32, 36–39, 272
– narrative 23–26, 28, 31–32, 34–39, 272
– scale 3–4, 6, 23–28, 30–39
Hoffman, Angela 11, 14, 226, 229, 237–238, 241, 274
Hutcheson, B. R. 117, 123–127

ICE corpora 88–89, 91–92, 97–98
ictus 115–116, 120–122, 126
indeterminacy 273–274, 276
Indian English (IndE) 48, 73, 76, 78, 89, 97–98, 100, 259
inkhorn term 262, 265
isogloss 4, 212, 216, 219

Johnson, Samuel 253, 260, 263
Johnstone, Barbara 227, 231

Kalton, Michael 75–79, 82
Kaluza's law 13, 109, 123–129
Keating, Patricia 1, 5, 110, 113, 116–117, 122
Kohnen, Thomas 45, 47–48, 50
Kretzschmar, William 1, 4, 75–76, 101, 162
Kytö, Merja 5, 8, 11, 14, 170, 176, 226, 229, 237, 238, 241, 274

Labov, William 4, 8, 37–38, 80
language
– change 2, 59, 64–68, 74, 76, 86, 99–100, 169, 249–250, 273, 276
– contact 78, 81, 86, 100, 225–227, 244–246
– legitimacy 261
– shift 226–227, 229
Larsen–Freeman, Diane 75
Late Modern English (LModE) 13, 84, 167–168, 177, 182, 254–258
Latin 84, 86, 100, 110, 122, 211–212, 262–266
Lee, David 23, 32, 47–48, 50, 58
lengthening 5, 13, 113, 118, 120–122, 126–129, 211, 219–222
lenition 214–215, 218, 221
letters 30–31, 53–54, 63, 169, 181, 183–184
Levin, Beth 187–190, 193–194, 199, 202, 204–205
lexicography 26, 33–34, 36
lexicon 187, 193, 195
lexis 4, 33, 226, 230–233, 235–237, 240–243, 245
literature 33
– language 249, 252, 259, 260–261, 266, 267
– texts 8, 10
LOB corpus 84, 89, 92, 97
*la longue durée* 27–28, 30–31, 33–34, 36
Luick, Karl 217, 219, 221

majority language 225, 231, 244
matrix language 226, 230–231, 245
meter 13, 109, 111, 112, 115–117, 127, 129
Middle English (ME) 7, 80, 84–87, 252, 254, 254–261, 263
Miller, Carolyn 7, 10, 30, 46, 49–51, 53, 59
Milroy, James 79, 260–261, 267
Milroy, Lesley 79, 261
Milton, John 252, 259–260
Minkova, Donka 5, 11, 37, 112–113, 116–119, 122, 271, 274–275
Mobus, George 75–79, 82
Modern English 80, 167–169, 174, 176–177, 182–184, 213
Moessner, Lilo 45, 47–48
morpheme 74, 77

Nevalainen, Terttu 45, 63, 78, 84, 138, 261
New Englishes 74, 97–98
– *see also* World Englishes
*niðan* 187, 196–197, 202–205
Norse 27, 81, 120, 212, 214–219
noun 2–3, 93–94, 98, 116–117, 138, 146, 161, 241–242
– phrase (NP) 78, 138, 143, 146, 160, 172, 177, 189–190, 199–200
numeral
– antecedent 93
– *one* 135–137, 142–143, 151–152, 154–155

object 85, 88–89, 155–156, 160, 162, 189–190, 202
*ofunnan* 14, 187, 195–197, 202–205
Old English (OE) 9–10, 13–14, 54, 76, 79–81, 84–85, 109–110, 112–122, 124–125, 128–130, 136–138, 187–188, 204–205, 214–215, 217–221, 254–257, 273–274, 276
Old Irish 118–119
orality 55, 112, 114, 116, 127, 130, 170, 184, 275

palatalization 216–221, 274
Parkes, Malcolm 110–111
Philippines English (PhilE) 84, 89–91, 97–98, 100
phoneme 14, 119, 129, 211–218, 273
poetry 112, 124–125, 127, 129–130, 260, 274
Polychronicon 253, 262
Postcolonial Englishes, *see* World Englishes
postmodifier 145, 169, 177–180
postvernacularity 226, 245–246
pragmatics 9, 28, 31, 49
predicate 177, 187, 189–190, 192, 194, 201, 203, 205, 275
prescriptivism 3, 85, 261–262, 275
Present-day English (PDE) 1, 87, 109–110, 168, 172, 184, 254–258
pronominalization 85, 137, 152
pronoun 6, 12–13, 78, 82–85, 87–88, 91, 94, 99–101, 121, 129, 135–139, 141, 147–155, 162, 172, 182, 272–273
– demonstrative 79, 84–85,
– indefinite 137–138, 142, 152–153,
– relative 60, 94, 100–101,
– *see also wh*-pronouns
propword
– *one* 138–139, 142, 152, 157, 162–163
prosody 109–110, 112, 115–117, 128–130, 170, 178–184, 271
– break 176–177, 179, 182–184
public records 54, 60–64, 66–67
punctuation 14, 110–112, 129, 167–172, 174–177, 179–184
– grammatical 168, 176, 181–182
– rhetorical 168, 176–177, 179–181, 183–184
Puttenham, George 252–253, 258, 262–264,

Quirk, Randolph 5, 167–168, 171, 176, 180

re-allocation 135
recipe 14, 48–50, 53, 57–58, 226–227, 230–246, 274
records of the Privy Council 65–66
reference 94–95, 98–100, 137–138, 161
– animacy 85–86, 92–93, 99–100, 145, 160, 198, 201, 203
referent 85–86, 92–94, 98–99, 136–141, 145, 148, 153, 160–164
register 47, 50, 61, 63, 274
relativizer, *see wh*-pronoun
religious treatise 60–62, 64–65, 67

Remlinger, Kathryn 231, 233, 237
resolution 122–130
– suspension of resolution 123–125, 127, 129–130
rhetorical genre studies 45–49, 51, 55, 56, 58
rhotacism 211–212, 214–216, 218, 221
Rissanen, Matti 45, 48, 52, 62, 135–139, 141, 146, 152, 155, 157, 164
Role and Reference Grammar (RRG) 187, 188, 190, 204
– argument 187–205
– core 188–191, 194, 197–205, 249
– direct core argument 197–203, 205
– juncture 187, 190–191, 194, 197, 204
– M-transitivity 192
– nexus 187, 190–191, 194, 197, 204
– oblique core argument 187, 194, 197–205
– prepositional government 187, 193–194, 197–198, 204
– S-transitivity 192, 200
Romaine, Suzanne 45, 83–85
Rule of Precedence 117
Russom, Geoffrey 115, 119, 123, 126–127

Sapir, Edward 82, 88
Scandinavian 214–216, 218–219
Schneider, Edgar 8–9, 11, 13, 75–81, 83, 85, 88–89, 91–94, 98, 135, 155, 162, 225, 245, 272, 273, 275–276
science 13, 75, 167–168, 170–178, 180–184
Scots 12, 48, 60–64, 66–67, 273
S-curve 60, 65, 78
selection
– Standard English 261–263
semantic role 86, 92, 194
– actor 192–194
– experiencer 192, 193, 197, 201–203
– macroroles 192–194
– patient 10, 192, 197, 199, 201, 203
– target 193, 197, 199–203
– undergoer 192, 194
semantic valence, *see* semantic role, Role and Reference Grammar
sermon 13, 50, 64, 167–168, 170–177, 179–184
Sevenker, Justin 250, 267
Shakespeare, William 33, 89, 252–253, 259, 261
Sievers, Eduard 115–117, 123, 219, 220
simplification 88–89, 91
*Sir Gawain and the Green Knight* 257, 260
Smirnitskij, A. I. 215, 217–218
Smitterberg, Erik 8, 11, 13, 167–169, 174, 177, 180, 181–183, 275
sociolinguistics 4–5, 25, 32–33, 37–38, 45
Standard English 73, 221, 241, 259, 261, 264–265, 267
standardization 58, 60–62, 180, 258, 262–264
strengthening 5, 81, 116–118, 129–130, 164, 274–275
stress 5–6, 79, 81, 101, 113, 115–117, 119–123, 125, 128–130, 275
style 24, 36, 47, 50, 87–91, 99, 128, 263
subject 13, 84–85, 91, 99, 138, 155–157, 160, 162–163, 167–169, 171–184, 275
– non-subject 87–89, 91, 100
– *see also* object
subjectification 137
subordination 191, 197, 202, 204
– *see also* Role and Reference Grammar
Swales, John 24, 49, 50–54
Swedish 14, 225–231, 233, 235–246, 274
– Heritage Swedish 14, 225–226, 231, 237–238, 240–242, 274
– Swedish–American 225, 227, 229–231, 233, 237, 245, 246, 274
– Standard 226, 230–231, 238, 240, 242–243
Swift, Jonathan 252–253, 263
syllable 120–130
– final 120, 125, 129
– first 81, 101
– length 173
– monosyllable 110, 120–121, 129
– penultimate 125–126
– stressed 116–117, 119, 122
syntactic valence, *see* Role and Reference Grammar

Taavitsainen, Irma 45, 48, 57–58, 62
Tagliamonte, Sali 37–38, 83, 147, 173
tax accounting 53, 55–56

text type 45–48, 62, 68, 88–89
textbook 9, 14–15, 217, 249–264, 266–267, 274, 275
– canon of examples 14, 249, 251–252, 254, 257, 260, 274
thematic role, *see* semantic role
Trevisa, John 258, 262, 264 (*252*, *253*)

Van Valin, Robert 187, 190–193
verb 2–3, 14, 167–169, 171–172, 187–196, 276
– alternation 187–190, 193–195, 198–199, 201, 204–205
– classes 80, 187–188, 190, 193, 195, 204–205
– ditransitive 192, 201
– intransitive 192, 194, 201
– irregular/strong 79–80
– phrase 113, 177, 179, 183
– regular/weak 80
– transitive 189, 192, 194, 200, 202, 204
Verner's law 212, 215–216
verse 109–112, 114–130, 275
vowel 119–122, 124–130, 213–214, 217, 220
– alliteration 117, 119–120, 129
– duration 118, 122, 129–130
– lengthening 120, 122, 127, 129

*whereof* 92, 95–96, 99–100
*wh*-pronouns 12, 74, 83–85, 91, 99–100, 272–273
– animacy 85–86, 92–93, 99–100
– antecedent 85, 93–94, 98, 273
– definiteness 93–94, 100
– interrogative 83–84, 87–88, 92, 100, 273
– possessive 91–99,
– post-nominal position 94–96, 100
– pre–nominal position 93–96, 98
– pied piping 88
– relative 84–85, 87–99,
– zero 36, 84–85, 99, 192
word class, *see* boundary
World Englishes 8, 12, 15, 73–77, 79, 81, 83, 89, 97–98, 100–101, 135, 164, 256, 258–259, 266