University of California
Santa Barbara

(In)stability under pressure: Constraints on function, form, and diachrony

A dissertation submitted in partial satisfaction
of the requirements for the degree

Doctor of Philosophy
in
Linguistics

by

Christian D. Brendel

Committee in charge:

Professor Simon J. Todd, Chair
Professor Eric W. Campbell
Professor Stefan Th. Gries
Professor Andrew Wedel, University of Arizona

March 2025

The Dissertation of Christian D. Brendel is approved.

Eric W. Campbell

Stefan Th. Gries

Andrew Wedel

Simon J. Todd, Committee Chair

March 2025

(In)stability under pressure:

Constraints on function, form, and diachrony

by

Christian D. Brendel

Acknowledgements

While writing these acknowledgments, my truest self—a comparative–historical linguist, alas—continually interrupted me. Who gets classified *where*? Into what clade should a friend-group or mentor ‘subfamily’ be sorted, if the time depth isn’t necessarily the best indicator of their ‘relatedness’? What if the group has drifted or diversified unrecognizably? What is the most parsimonious balance between audience-specific in-joke and academic genre formality? There is a high degree of horizontality—cross-cutting—among the groupings in this section: the distinction between colleague, friend, and family is, ultimately, not discrete, nor is the ‘origin’ of an acquaintance the most accurate measurement of the nature and quantity of the burden they’ve carried for me. But I am deeply indebted to everyone below and many others, so I’ll try.

First, I want to thank my family, who are my earliest and most enduring source of support: my mother, Debbie [Wieder] Brendel, who withstood eighteen years of my education and, before all else, gave me language; my father, Dave Brendel, who taught me how to play with it (and how to do just about everything else—though I may still be *schusslig*); Stephanie Swanger, my first and best co-author; Patti [Wieder] Swanger & Jeff Swanger, whom I have always trusted and relied on; and Valeri Harteg, for her optimism, encouragement, and expertise. I owe an incredible debt to my late grandparents, who each passed away just a few months before I wrote these words: my grandfather Douglas Richard ‘Dick’ Wieder, who showed me that it’s better to be absurd than miserable; and my dear grandmother Bonnie Lou [Wagenblast] Wieder, who gave me my first Latin book and set me on this path (a long path; it’s good that she also bought me many, many socks). I’m grateful to have grown up hearing their Dutchy vowels, which made the world feel a little warmer.

To my second family, although they entered my life a little later, I owe just as much gratitude for their wholehearted support, as well as their patience and faith. Thank you to Connie Gilday, who, over many conversations about my work, has always known exactly what to say to motivate me (& what I need to eat) and has always cared deeply enough to tell me—knowing how happy she would be to see me accomplish this goal has motivated me in these past few months; and Joe D’Ignazio, whose welcoming and supportive nature makes a home for me wherever we are—I am grateful for his magnanimity, his heartfelt kindness, and his expert testimony on the inflectional morphology of *bagel*[a]. Thank you to my fellow-former-Californians: Alex D’Ignazio, who (like me) is a parallel processor and whose project & life management skills (unlike me) are so highly developed that even I become more productive just from osmosis when she’s around; and Ryan Eskridge, the Gondor to my Rohan—or vice versa?—and possibly the only person interested in my undergrad thesis I cited above. I am forever grateful to Cathy D’Ignazio, whose advice about the academy and generosity of time & space preserved my sanity; and to Sophia D’Ignazio, who, befitting a medievalist, convinced me there was another world (or two) beyond the inferno—and that it was better.

I want to thank my oldest friends who had to withstand (yet didn’t outright reject) my teenage enthusiasm about linguistics, as well as my nascent conlanging phase: Jesse Carozza; Joel

[a] In Yiddish, בייגל /bɛɪ̯gl̩/ ‘bagel’ makes no distinction between singular and plural (Weinreich, Uriel. 1977. *Modern English–Yiddish dictionary*. New York: Schocken: 26).

Goodfellow, my first firend[b]; and Jordan Quigley.

Since my time at Penn State, the friends I made as an undergraduate (or just after) have been with me through each major academic milestone and have changed me, for good. After waiting a decade, they are happier to see this dissertation completed than I am—and that's the gospel truth. Thank you to the Berks crew: Ashley Asensio, Dan Cohen, Sean Geguera, Karissa Smith, Jordan Smith, and Amy Rutter. Included in this clade are friends who, although I didn't go to school with them[c], have also become a core part of me: Ivee Baugh, Allison Ellinger, and John T. Hickman. I am grateful to all the students in the honors program/club and the writing center who created cohesive community spaces that I had never experienced before.

It's impossible to adequately convey the role of my professors at Penn State (some of whom are instead listed in the committee section below). Thank you, Cheryl L. Nicholas, for showing me not just how to speak, but what I *could* speak, for (to paraphrase another professor) unleashing my self-narrative, for having me as your teaching assistant and eventual collaborator, and for being your authentic, dynamic self. Many thanks to the late David Bender, the surest wizard of the registrar I've ever met (whose magic saved me on countless occasions); I-Ming (Alice) Shaparenko, who made me, for the first time, feel good at math; Belén Rodríguez-Mourelo, who created a graduate-level Romance historical linguistics seminar for me and another student, my first formal coursework in historical linguistics; and Yan Gu, who designed and taught an independent study in Mandarin for me. Other staff members encouraged me to grow socially, as well: Sandra Reichel, who taught me how to be a team player (and mediator); and Teri Sabatelli, who demonstrated how to *make* a team. I would not have graduated without the support of generous donors, and I owe deep thanks to: the Boscov family, in particular Shirley K. Boscov, whose genuine interest in the projects of scholarship recipients helped me realize my work was 'real'; the Ciferri family; and to the other scholarship donors who I have not met but who have affected my life forever.

As an undergraduate, I was fortunate enough to learn from other faculty outside of Penn State. Thank you to Kyle Johnson for offering an introductory linguistics course online through UMass Amherst, the first linguistics course I had ever taken. Thank you also to Leendert Plug, who introduced me to phonology and discourse during my semester at the University of Leeds and served as a letterwriter for my graduate applications so long ago.

I would also like to thank my colleagues and friends who I initially met while working at the Educational Testing Service. Many, many thanks to Shawn Steinhart, who was the first 'linguist' I ever spoke to as an expert source for a term project, a manual on 'how to be a linguistics student'. Shawn got me (and others) through countless days of work with morning cronuts and afternoon lunch pun-offs (One day we punned about yogurt, but I can't recall the jokes; I guess I lac-tose specific memories) and quickly became a dear friend and confidante, who, when I expressed doubts about finishing this program, told me that no, I could definitely finish. (I was skeptical, but he was right.) Thank you, Kristine Nugent, who has been the center of my gyroscope for a decade, who first encouraged me to think about computational linguistics rather than pure philology, who proved to me that agency exists if you can think

[b] [sic]

[c] As I mentioned, this phylogeny is imperfect!

outside the booth, and whom, after years on opposite coasts (sometimes opposite continents), I've been supremely fortunate to live near once again.

I'm especially grateful to the mentors and colleagues who took chances on me early in my undergraduate career during my experience at the Smithsonian National Museum of Natural History. Thank you to Elizabeth Cottrell, Eugene Hunt, and Virginia Power, who were instrumental to ensuring I was prepared for the National History Research Experiences; to Ives Goddard, who led me in a one-on-one seminar on comparative Indo-European linguistics; and to Neal Woodman, who helped me tie the metaphor of linguistics-as-biology to the *reality* of linguistics-as-biology (a shrewd connection indeed). During this era began the foremost and longest-lasting professional partnership in my linguistics career: thank you, Gabriela Pérez Báez, whose mentorship has touched every aspect my career, from computational lexicography in the National Museum of Natural History to practical engagement at the Folklife Festival and in Oaxaca, covering everything from career advice to impromptu language consultation (*¿lavar, o llevar, los aguacates? esa es la cuestión...*). I cannot adequately express my gratitude to Gabriela for the privilege of our continuing work together (and I will stop here before I introduce duplicates!).

In Santa Barbara, I owe thanks to many people who have helped me conduct this work, as friends, colleagues, and a combination of the two. I am unquantifiably lucky to have been part of a strong 'cohort effect'—thank you to Alexia Fawcett, Karen Tsai, Simon Peters, and Yi-Yang Cheng—and a wider 'system' of friends who offered practical or emotional support—Adam Koch, Albert Ventayol-Boada, Caroline Crouch, Chloe Willis, Dibella Wdzenczny Caminsky, Giorgia Troiani, Kendra Calhoun, Kevin Schäfer, Jared Sharp, Martin Kohlberger, Nathaniel Sims, Phillip Rogers, Sandra Auderset, Tank Willis, and Tanya Wang. I am grateful to Fermín Moscoso del Prado Martín, who was instrumental to my early conception of the ideas in Chapter 3 and to my affinity for system-theoretical approaches to language. I'm grateful to Mary Bucholtz and Matthew Gordon, each of whom worked during their respective appointments as department chair to identify engaging work opportunities which supported me during my graduate career. I also want to thank Mary for her mentorship in teaching and navigating academia early in my time at UCSB, and the many faculty and staff who made South Hall a community: in particular, Alicia Holm, Mala Poe, René Marchington, Samantha Little, and Sandy Thompson.

I am indebted to the many undergraduates at UC Santa Barbara whom I've had the privilege of seeing in class, office hours, and Linguistics Club; your enthusiasm was a continual, much-needed restoking of my own passion for linguistics. I will limit myself to acknowledging the undergraduate research assistants and readers who I've been fortunate enough to work with: Alyson Osgood, Khari Stinson, Kimberly Chiu, Leslie Su, Lu Liu, Sabrina Chaffino, and Shannon Rumsey.

I want to thank a few people in particular who all, in different ways, helped me fix problems that kept me from writing: Robby Nadler, who helped me talk; Catherine A. Picken, who helped me breathe; Emily Krause, who helped me care; and Landon Wigg, who helped me 'be done'. I also want to thank Linus Corgvalds for continually reminding me that, as important and interesting as human adaptations like words and language are, the universe of empathy and communication is much wider.

I am most grateful to my partner, Nathan Dignazio. Nate has seen me through every major milestone in my graduate program, from MA thesis to dissertation. He has selflessly assumed domestic burdens in order to free up capacity for me to write, and the weight of the final stages of my dissertation has been heavier on him than on me. He's also carried the weight of a sisyphean burden of emotional labor: me. Thank you, Nate. This dissertation is for you; it is *due* to you, in both senses.

Finally, I owe the sincerest gratitude to the members of my committees at Penn State and UC Santa Barbara, who all have shared the... interesting experience of supervising a milestone of mine (and, despite being displaced by time and institution, I think those experiences have been broadly similar—I'm sorry!).

First, many thanks my undergraduate thesis committee at Penn State: Jeanne Marie Rose, who introduced me to rhetoric and the very notion of genre conventions, who helped me navigate the practical choices in graduate school, and whose written voice I now detect in my own comments on student papers; Holly Ryan, who trained me to be a writing tutor (and stocked the office with Kit Kats), introduced me to a lifelong fascination with television criticism, and helped me project manage a thesis, twice; and thesis (& undergraduate) adviser, Sandy Feinstein, who let me write, read, teach, who brought me to the world, who refused to let my undergraduate degree drag on and taught me it's better to choose to decline than to let possibilities slide by uninterrogated. Thank you all for what you've done to shape me and for refusing to let me fold.

At UC Santa Barbara, I am eternally grateful to those who have served on my various committees over the years: thank you to Argyro Katsika, whose feedback on my MA and whose questions at my doctoral colloquium helped me think beyond the segment; to Marianne Mithun, whose enthusiasm for teaching was a model for my own graduate teaching, thank you for all that you've taught me about language and for your comments (both in many classes and on my MA and Publishable Paper committees) on the projects that became Chapters 2 and 4; to Bernard Comrie, whose edited volume[d] has been on my bookshelf since I was a teenager (my first linguistics book), thank you for sharing your knowledge and comments during your seminar on relative clauses, which was formative to the project that has culminated in Chapter 4; and to Violia Miglio in the Department of Spanish and Portuguese, whose expertise as a scholar of both translation and Icelandic made her an uncannily perfect match for my Publishable Paper committee, I thank you for devoting so much time and enthusiasm to helping me steep the study as richly as possible in early Iceland. Most relevantly, I am grateful to the members of the present doctoral committee: to Andrew Wedel, whose work has fascinated me and shaped my research interests for years, thank you for being a part of this committee; your feedback has been irreplaceable; and to Stefan Th. Gries, who introduced me to quantitative corpus research and taught me everything I know about R (though I am still tempted by a certain snake [but not by the Apple]), thank you for all you've taught me and especially for your crucial feedback on my prospectus, which significantly shaped the course of this dissertation. In particular I owe the most thanks of all to those who have served a tour (or several!) as a chair of a committee and as my primary advisers: to Eric Campbell, who changed the trajectory of my life when he first

[d] Specifically: Comrie, Bernard (ed.). 1993. *The world's major languages*. Oxford: Oxford University Press

called me in 2014 to invite me to an open house at UC Santa Barbara and who has guided me ever since (including, but not limited to, his official role my adviser for more than half a decade, in addition to serving as chair or member on every committee I've formed at UCSB), thank you for advocating for and supporting me in innumerable ways, both through your academic mentorship and through your investment in my general well-being. And finally, to Simon Todd, who has served as chair of my doctoral committee since 2021 and who, in addition to his critical role in helping me work through the framing and analyses in this dissertation, persistently developed (countless) new strategies to help me be more productive and more balanced[e]: thank you for all the time, energy, thought, and empathy you have devoted to working with me, especially during the times when I gave less than I took—frankly, I don't know how you managed it (or managed me). We first met when I drove you to your visit at UC Santa Barbara, and now you've driven me to finishing this dissertation.

My research has continually shown me the effects of environment, interaction, and community on some target—a word, a dissertation, or a person. No linguistic feature has meaning in isolation; languages create meaning through contrast and concordance. Similarly, I am nothing without the people above who have conditioned me, encouraged me, constrained me (from my worst tendencies), challenged me, and simply talked with me. Anything useful in these pages is owed to them, and any mistakes that remain are my own.

This work was partially supported by the National Science Foundation[f], the Beinecke Scholarship, and the University of California, all of whose support I am grateful for. The views expressed here are my own.

[e] Indeed, Simon, the Post-It notes are finally all down from the wall.

[f] NSF GRFP #2017223465

Curriculum Vitæ
Christian D. Brendel
March 2025

Education

2025	PhD in Linguistics, University of California, Santa Barbara
2018	MA in Linguistics, University of California, Santa Barbara
2014	BA in Communication Arts and Sciences (*summa cum laude*), Schreyer Honors College, Pennsylvania State University

Publications

Pending	Feinstein, Sandy, Nicolas Fay, Faith Iseman, **Christian D. Brendel**, Rachel Jensen, & Jennifer Muret-Bate. In press. Creative–critical coauthoring: Definition and demonstration. In *Coauthoring with undergraduates in writing studies: Revising identities & institutions*, Laurie Grobman & Jane Greer (eds.). Logan, UT: Utah State University Press
2023	**Brendel**, **Christian D.** 2023. Translation-induced interrogative relativizers and stability in Icelandic. *Transactions of the Philological Society* 121(1). 1–32. DOI: `10.1111/1467-968X.12258`
2018	**Brendel**, **Christian D.** 2018. *Cohort effects and asymmetrical word-level sound change*. Santa Barbara, CA: University of California, Santa Barbara MA thesis
2017	Feinstein, Sandy, Tiffany Wesner, **Christian Brendel**, Lily Cernak, Ashley Offenback, Rachel Jensen, & Sean Geguera. 2017. Why study literature? *Rendezvous Journal of Arts & Letters* 43(1–2). 163–174
2016	Moscoso del Prado Martín, Fermín & **Christian Brendel**. 2016. Case and cause in Icelandic: Reconstructing causal networks of cascaded language changes. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (ACL)*, Katrin Erk & Noah A. Smith (eds.), pp. 2421–2430. Berlin: ACL. DOI: `10.18653/v1/P16-1229`
2014	**Brendel**, **Christian**. 2014. *The wanderers and the analytics: A lexicostatistical study of Sindarin and Quenya*. University Park, PA: Pennsylvania State University, Schreyer Honors College honors thesis
2012	**Brendel**, **Christian**. 2012. Tutoring between language with Comparative Multilingual Tutoring. *The Writing Center Journal* 32.1(8). 78–91. DOI: `10.7771/2832-9414.1856`

Select Honors & Awards

National

2017–2022	Graduate Research Fellowship, National Science Foundation
2014–2019	The Beinecke Scholarship, Sperry Fund

University of California, Santa Barbara

2015–2019	Regents Fellowship

Pennsylvania State University

2014	Reddy Mission Award, Schreyer Honors College
	Young Investigator Award for the Humanities, Arts, and Social Sciences Division, Berks campus
	Eric A. and Josephine S. Walker Award
2011–2014	Joseph L. and Shirley K. Boscov Honors Endowment
2010–2014	The Renaissance Scholarship

Teaching & Research Appointments

Teaching Associateships, Dept. of Linguistics, University of California, Santa Barbara

2019	Linguistic Analysis
	Introduction to Language & Linguistics

Teaching Assistantships, Dept. of Linguistics, University of California, Santa Barbara

2023, 2019, 2017	Introduction to Language & Linguistics
2022, 2020, 2016	Language in Life
2021	Language, Power, & Learning
2020	Memes
	Language in Society
2018	Computational Linguistics
	Introduction to Historical–Comparative Linguistics
	Story of English
2017, 2016	Animal Communication
2017	Language, Race, & Ethnicity

Other Research

2013	Undergraduate research intern, Smithsonian National Museum of Natural History, Natural History Research Experiences (NHRE), National Science Foundation REU Site, EAR-2244445

Abstract

(In)stability under pressure:

Constraints on function, form, and diachrony

by

Christian D. Brendel

The description of the mechanisms underlying the transformation of synchronic language variation into diachronic change (Gardiner & Nagy 2017; Sankoff & Blondeau 2007; Labov 1994; Weinreich et al. 1968) and the influence on language evolution of competing synchronic pressures in language production and perception (e.g. Wedel 2012; Bybee 2007; Martinet 1955; Zipf 1949) are topics of longstanding interest in the study of language change. Constraints of this sort can be perceived through their realization on particular concrete forms or structures. In this dissertation, I present three discrete case studies organized into article-like chapters, each of which examines the interaction of form and function in synchrony and diachrony, fitting the drivers behind each scenario of language change to a particular explanatory framework.

Chapter 2 refines our understanding of the role of functional load (Surendran & Niyogi 2006, 2003; Martinet 1955; Hockett 1955) in the resistance to merger of certain phonemic contrasts (Wedel et al. 2013a,b). Building on Wedel et al. (2013a,b) by incorporating word co-occurrence vectors (Mikolov et al. 2013a), we provide evidence that phonemic contrasts in minimal pairs of high contextual similarity are less likely to undergo phonemic merger, corroborating the finding of Wedel et al. (2013a) for words that share discrete parts-of-speech without requiring the *a priori* determination and annotation of word classes.

Chapter 3 examines asymmetrical patterns of sound change within words through the lens of a theory of speech perception, the cohort model of lexical access (Marslen-Wilson & Welsh 1978), which I argue subjects different subwords to differing cognitive processes and, there-

fore, potentially-heterogenous selection pressures. By analyzing cognate pairs across six Indo-European languages and dividing them into intralexical subregions predicted by cohort theory, I corroborate earlier work by showing that, in pairs of cognates, phonemes that occur relatively further inside words are more likely to undergo change (Wedel et al. 2019; Brendel 2018). Additionally, I suggest that, when recognition points are incorporated into the analysis, material after a recognition point has a pattern of predicted change that differs from material subjected to cohort selection in addition to a general end-of-word tendency to change, suggesting that the unequal deployment of the cognitive machinery of perception shapes the long-term trajectory of sound change at a sublexical level.

In the historical case study presented in Chapter 4, formal inadequacies and system-level 'preferences' interact during a centuries-long period of variation and instability (as in Nichols 2003) in the relativization system of Early Modern Icelandic. Through a study of diachronic written corpora (Wallenberg et al. 2011), I examine the interactional, system-level, and sociocultural factors which conspired to maintain a preference for a certain strategy of relativization despite extensive contact pressure, ongoing instability in the relevant grammatical subsystem, and formal change (lexical replacement). This story ends in restored equilibrium after instability, and I suggest a sort of system elasticity or inertia, perhaps most readily observable across time, that can exert pressures on synchronic realizations.

Taken together, these chapters contribute to our understanding of how languages bear out competing pressures—variously interactional, cognitive, and, in the final chapter, sociocultural—while maintaining systemic coherence over time.

Table of Contents

Chapter 1

Introduction

The interface—and distinction—between stable variation and lasting, systemic change remains a perennial challenge to describe due to the so-called “transition problem” (Gardiner & Nagy 2017; Weinreich et al. 1968); namely, how does mutually intelligible synchronic variation translate to a diachronic tectonic shift? Even if it is assumed that the forms of synchronic variations—those which sometimes go on to spread through a speech community and across time—themselves arise randomly, these would-be innovations—in the context of the language’s other structures and for its speakers—are not necessarily arbitrary, perfectly-equal alternatives to the existing structures that they undermine or build on. Instead, these disruptors stand to affect the internal coherence of a language as a system.

If relative regularity and consistency of both the inventory and the composition of a language’s patterns is meaningful to language users—in other words, if ‘rules matter’—then a prospective innovation does not incubate in isolation, but rather must be affected by how it affects—positively, negatively, or neutrally—the ability for language users to communicate, much as the persistence of a mutation in biological evolution is tied to its effect on the fitness of the species.

As an example *partim ad absurdo* of the transition problem, say two friends who speak the same language are engaged in a perpetual ‘groundhog’s day’ event: they wake up each morning to find that their interlocutor’s idiolect has got itself new rules. Maybe it has lost its formerly-rigid fixed word order, or perhaps yesterday’s ergative–absolutive morphosyntactic alignment has become nominative–accusative. The [unfortunate] friends cannot rely on the principle of regularity, and the quick, constant disruption to internal coherence of the system shared between them—'rewriting the rules'—interferes with their ability to communicate with each other. This hypothetical scenario is, of course, extreme (at least for language users who have not suffered injury to their brains), but illustrates that, for a language to accomplish its' users communicative goals, there must be some value to having a linguistic 'rule of law'.

However, we know from observation that all elements of linguistic structure—including phonological material, syntactic patterns, and other constituents—*can and do change*, often substantially, across time. Moreover, although we don't see naturalistic examples of the extreme scenario above, people are *capable* of rapid acquisition of innovative forms, and yet language diversification proceeds at a much slower rate (Wallenberg 2019) [1]. In the aggregate, dramatic changes certainly occur over a language's history, but on the level of lived, daily experience, these extremes are not approached (absent the genesis of pidgins or the loss of language).

Such a modulating or conservative effect on what would otherwise be a sort of 'wildfire' language evolution implies the existence of a modulat*or*: some preference for continuity, which itself is realized as a pressure against formal and/or functional change. Of course, language systems are constantly changing on various dimensions, from sound to discourse conventions, on both the level of idiolect and the level of an institutionalized national variety. But my core purpose in this dissertation is to suggest that systems of a language sometimes persevere in functional characteristics even though they may change quite substantially on the surface. Just as no living adult is composed of the same cells as when they were born (and just as the Ship of Theseus no longer has any of its original planks), systems that are highly cohesive can persist through dramatic changes.

In Chapter 4, I discuss a case in which a system seemingly 'snaps back' into place (into an earlier state) after centuries of contact-induced destabilization—such systems can be thought

[1] Wallenberg (2019) terms this conundrum 'Yang's Paradox' after Charles Yang, who identified this contradiction while developing the variational grammar model proposed in Yang (2000).

of as having a sort of elasticity, like a properly kneaded bread dough.[2] Through diachrony, languages often reveal to us what structures did or did not ‘work’ in some earlier stage of development, giving us cause to examine *why* the structures were problematic in the first place, synchronically-speaking. In my first two chapters (Chapters 2 and 3), I examine some such factors which arise from common cognitive pressures affecting lexical retrieval and speech processing. Such synchronic challenges can help provide reasons why some moribund feature would have been dispreferred *in situ*. Conversely, observing diachronic outcomes can help reveal these quotidian pressures in the first place: the proof is, one may say, in the outputting.

1.1 Form and function

Like the pull of gravity, the sort of cognitive and functional pressures I work with in my dissertation—or the existence of abstract *functions* at all—are perceived through their realization on particular concrete forms or structures, such as sounds and lexemes. With this ethereal character, it might be thought that functional pressures are always the invisible hand driving language change. However, I hope to show in each chapter that the *interaction* of form and function is a crucial component to understanding each case I identify. For instance, I will argue in Chapter 2 that the heavy functional load born by some phonemes can indeed confer some protection on their forms from the ravages of (some) sound change, and in Chapter 3 I will suggest that the realities of auditory processing in lexical retrieval lead to uneven rates of sound change within individual words. In Chapter 4, the formal properties and functional preferences seem to interact to preserve a state of minimal change a centuries-long period of volatility in a syntactic system.

[2] Of course, it is possible (though unfortunate) to rip bread dough or to knead it so much that the elastic proteins are over-stretched. Similarly, language systems can (and usually probably do) break down irreparably. But a well-worked dough is difficult to ruin later, and perhaps systems that firmly occupy the functional niches of their languages are similarly resistant to replacement.

Of course, it is nothing new to ascribe 'reasons' to language change, though it is dangerous territory. Indeed, the idea that some change in a language can be motivated by a language 'wanting' to do something—that is, having teleological motivations—traces to at least Sapir, who remarked that "language has 'a slope"' (Sapir 1921); in other words, linguistic 'drift' is not really random, Brownian motion, but rather a semi-directed tumble downhill, with a particular directional vector. The concept of semi-directional change has nevertheless provided appealing explanations for changes in which an ostensible 'gap' in language is filled through an innovation: for example, chain shifts seen both in phonology and syntax (Moscoso del Prado Martín & Brendel 2016), in which one development (or instability) in a system portends (and Moscoso del Prado Martín & Brendel (2016) suggest, may be causally related to) subsequent changes throughout a paradigm. However, at face value, such goal-driven explanations ascribe agency and volition to a concept—language—that is not itself alive, but rather an incorporeal symphony produced by the vocal and manual instruments of a language community.[3]

However fraught teleological metaphors may be, there is valuable insight to be gained from understanding the impact on language change from the mechanisms of our bodies (often operating on elements of form, tied to communicative and cognitive needs, such as processing and producing sufficiently distinctive sounds or gestures) and the properties of other complex systems (some functional preference for equilibrium and stability, as in the 'wake up, new language' thought experiment above). Realized through the bodies of individuals, language change (like language usage) must be consciously-evaluated yet simultaneously-automated (Blevins & Wedel 2009). In this way, the involuntarily 'actions' of language are akin to the body's reflexes or tendencies of complex systems; water, after all, also flows without 'volition', whether conditioned by gravity or by constriction. These factors exert influence on our language usage

[3] This is not to say that teleology is never at play: even if languages do not, language *users* can have goals. In my culminating article (Chapter 4), these factors come into play in regards to the development of Icelandic, whose speakers underwent a period of conscious rejection and purging of non-native vocabulary and other perceptibly 'foreign' influences (Þráinsson 1994: 188).

constantly and it is generally agreed that the synchronic variations of today and yesterday, gathered over a lifetime, are the playfield of linguistic evolution (Labov 1994; Sankoff & Blondeau 2007). Both systemic pressures (‘water must be level’) and the accidents of form and occasion accumulate and are responsible for diversification of languages in diachrony, and both formal idiosyncrasies and functional.

1.2 Preview of chapters

I intended the chapters of this dissertation to be, more or less, ‘separate’ articles, a series of discrete lenses and inquiries into united by their shared interest in the interaction of functional and formal pressures. In many ways, the chapters are exactly that; for example, the tool of cohort theory is applied to understand the question I articulate in Chapter 3, but I do not deploy this theory in Chapter 2. However, in both interpretation and in methodology, the chapters have —as we would expect from co-occurring units—converged in some ways.

The first study refines our understanding of how functional load—the work that phonological contrasts do in distinguishing meaning—influences sound change. Building on previous work showing that minimal pairs resist merger (Wedel et al. 2013a,b), I demonstrate that this effect is modulated by syntactic context. Using distributional semantic models to capture contextual similarity, I show that minimal pairs occurring in similar syntactic environments are especially resistant to merger. This finding suggests that the pressure to maintain phonological contrasts is sensitive not just to lexical contrast but to the broader grammatical context in which words appear. The second study examines how the cognitive mechanisms of speech processing create asymmetries in sound change within words. The cohort model of lexical access (Marslen-Wilson & Welsh 1978) proposes that listeners begin identifying words as soon as they hear enough information to distinguish them from other candidates, creating a functional division within words between the material needed for early recognition and subsequent phonological content. Analysis of cognate pairs across Indo-European languages provides evidence for an end-of-word increase in sound change and a difference in pattern between regions subjected to different types of auditory processing, implying that cognitive systems of lexical access shape patterns of historical change.

The final study takes a broader view, examining how formal and functional pressures interact across centuries of language change. Nichols (2003: 283) notes that 'diversity arises when some element is relatively unstable and therefore prone to replacement in various ways'. The history of relativization in the North Germanic languages provides an illustration of this claim: in the early 2nd millennium A.D, the relative marker *er* 'that' was gradually replaced across the entire family, and novel relativization strategies increased in frequency (Brendel 2023; Sapp 2019; Wagener 2017). However, when viewed through a contemporary, synchronic lens, this dynamic period of instability has seemingly resulted not in diversity, but in convergence. Today, each of the extant languages has settled on one predominant strategy: the marking of relative clauses with a reflex of the Old Norse comparative marker *sem* 'like' (Wagener 2017), a return to the status quo in function if not in form. Through an analysis of relativization strategies in Early Modern Icelandic, I show how a grammatical system can maintain functional stability even through periods of formal change and intense language contact, revealing a kind of systemic inertia in language—a tendency to preserve functionally adequate patterns despite pressure for innovation.

Together, these studies demonstrate how languages negotiate competing demands across multiple timescales. The cognitive pressures of real-time language processing create subtle biases that, accumulated over generations, result in systematic patterns of change. At the same time, functional pressures arising from the needs of communication act as a conservative force, preserving structures that efficiently serve core linguistic functions. Understanding these dynamics requires examining both the fine-grained mechanisms of language processing and the broader systemic pressures that shape linguistic evolution.

Chapter 2

Context-predictability and functional load in the maintenance of phonological contrast

2.1 Introduction

[*] The functional load of some structure of language describes 'the extent to which contrasts among members of … a set of contrastive features contribute to the signaling of significant differences' (Meyerstein 1970: 16). Functional load has been most widely implicated (King 1967a) and quantified (Surendran & Niyogi 2006) in the study of phonemic contrasts, but the central conceit of the concept—that 'those patterns which carry a proportionally larger share of the burden of communication are more important' (Meyerstein 1970)—is a foundational idea in each chapter of this dissertation. The phenomena that are 'important' might be more resistant to change because any disruptions to them could more destabilizing to a system than change in a less crucial region, just as misplacing one of many pencils is far less disruptive to one's day than misplacing a housekey.

The maintenance of phonological contrast plays a fundamental role in human language,

[*] I gratefully acknowledge the collaborators who contributed significantly to this project: coauthor Simon Todd, whose graduate pilot study on the use of word vectors to predict phonemic mergers informed the basis of our approach and who contributed significantly to the project planning, data transformation, and core methodology, joined with me to help interpret our findings, and provided crucial feedback on several verbal and written drafts this chapter. I am also grateful for our undergraduate research assistants, Shannon Rumsey and Lu Liu, who worked for months on preprocessing corpus data, especially for the onerous work of identifying and developing language-specific normalizations for the various idiosyncrasies across the corpora. I am chiefly responsible for the writing below, for some corpus preprocessing, and for most of the programming that used the cleaned corpus data (e.g. the identification of minimal pairs, the word2vec and ByT5 model training/inference, and regression model selection). Any errors that remain in this written chapter are due to my oversight.

allowing discrete meaningful elements to be distinguished through combinations of contrastive sound categories. Much like how genetic mutations can lead to the loss of trait distinctions in biological populations, sound changes in language can result in the merger of previously distinct phoneme categories. For nearly a century, linguists have hypothesized that such mergers might be inhibited when phoneme contrasts carry a higher functional load—that is, when they are responsible for distinguishing many words in the lexicon (Martinet 1955; Gilliéron 1918).

Indeed, the extent to which language material serves a necessary, contrastive function affects its ability to resist erosion and reduction over time. Phonemic material responsible for the maintenance of some kind of contrast resists change (in the form of phonemic mergers) compared to material which bears no such functional load (Blevins & Wedel 2009). The merger of phonemes whose loss entails a loss of contrast between words is less likely than phonemes whose loss is less disruptive to other linguistic strata.

Relatively recent quantitative studies by Wedel et al. (2013a,b) provided the first large-scale empirical evidence supporting this functional load hypothesis, demonstrating that phoneme contrasts distinguishing more minimal pairs in a language's lexicon are significantly less likely to merge, and that this effect is enhanced for phonemes that distinguish a large number of contrasts between words that are the members of the same word class. In other words, when the syntactic system makes no distinction, then even more burden falls on the phonological system to carry the lion's share of the contrast. These findings suggests that the distribution of contrasts across the lexicon plays a key role in shaping the evolution of phonological systems.

This study replicates and extends a subset of the results of Wedel et al. (2013a,b). We show the effect of functional load on the avoidance of phonemic merger while making some significant methodological refinements outlined below: namely, the use of word vectors—emergent measures of contextual predictability and similarity—rather than discrete parts of speech, which perhaps better models the bottom-up inference of contextual similarity in the interactional contexts that drive language change. By measuring the contexts in which words occur, rather than deriving roles from annotated corpora, this methodology could remove the burden of requiring annotated corpora and allow the claims of Wedel et al. (2013a,b) to be tested on more diverse languages.

2.2 Background

2.2.1 Early conceptualization and formulation of functional load

Most broadly, the functional load hypothesis predicts that the probability of losing a contrast—traditionally, a phonemic contrast—is inversely related to how much work that contrast does in distinguishing meaning in a language (King 1967b: 831). The concept, at the time of writing, is over a century old and has been of perennial and varied interest in linguistics. Recognizable from at least as early as Gilliéron's (1918) notions regarding the resistance of certain forms to regular sound change, the hypothesis gained wider popularity through the contemporaneous work of Martinet (1955) and Hockett (1955).

Martinet (1955) proposes that sounds which carry a relatively greater burden of meaning distinction in a language (*rendement fonctionnel*)—specifically, the ability to discriminate minimal pairs—would be more resistant to merger or loss. Significantly, he generalizes this idea to a more general claim about drivers of language change that are not purely internal or structural: sound change, rather than being sequestered in a terrarium of pure structuralism, is also subject to certain functional pressures, including the communicative needs of interlocutors[4] (Martinet 1955: 49). Martinet (1955: 54) specifically connects this notion of economy to Zipf's (1949) Principle of Least Effort: speakers tend toward minimizing articulatory exertion, potentially leading to phonetic reduction and loss of contrasts. Simultaneously, however, listeners benefit from the opposite: clear phonetic distinctions that facilitate accurate perception. This compe-

[4] As the literature cited on this topic primarily concerns the oral–aural modality, the terms *speaker* and *listener* will often be used when discussing the interlocutor dialectic.

tition helps explain why frequently-used forms tend to become phonetically reduced over time, while elements crucial for maintaining important contrasts often resist such reduction (see Bybee 2007).

As Martinet (1955) is responsible for the functional load hypothesis in a framework of economy, Hockett (1955: 217–218) is the first to propose an information-theoretic (Shannon 1948) formulation of functional load (refined in Hockett 1967) based on the system-level entropy of phonemic contrasts: if a contrast between a pair of phonemes is lost, then the corresponding information loss constitutes the functional load of the contrast. The application of entropy to various levels of language structure has featured in several subsequent quantifications of functional load (Surendran & Niyogi 2003, 2006; Wedel et al. 2013b)

2.2.2 Quantification of functional load and phonemic merger

Despite its intuitive appeal and century-long history, robust quantitative evidence supporting the functional load hypothesis long remained elusive, leading to challenges to the hypothesis (see, for instance, King 1967a, 1967b: 848, which regards functional load as unimportant at best as a factor in sound change, having failed to quantify it effectively given the constrains of the day). More recently, however, in a pair of landmark studies, Wedel et al. (2013a,b) formulate measures of functional load and apply them to large corpora of data, finding strong support for the role of functional load in preventing mergers. As the present study is a replication and extension of this work, this section describes them in some depth.

Wedel et al. (2013b) examine the relationship between functional load and phoneme merger across nine language varieties. The authors quantify functional load primarily through minimal pair counts, testing whether phoneme contrasts that distinguish more minimal pairs are less likely to have merged diachronically. Using a mixed-effects logistic regression model, they find strong support for this hypothesis—phoneme pairs with higher minimal pair counts showed significantly lower rates of merger. Additionally, they discover that for phoneme contrasts distinguishing few or no minimal pairs, the probability of the higher-frequency phoneme becomes a significant predictor of merger likelihood, with more frequent phonemes more prone to merging. They also find that lemma-based minimal pair counts are more predictive than surface form counts, and provide evidence that minimal pairs with similar frequencies contribute more strongly to merger resistance.

In the second study, Wedel et al. (2013a) find that the interface between phonology and syntax (and the distribution of the functional load between them) is also relevant in accounting for the likelihood of phonemic merger. They show that the likelihood of phonemic merger decreases when the phonemes are in lemmas that share a part of speech. In order words, the functional load of a phonemic contrast in maintaining a lexical distinction is heightened when other linguistic cues (such as syntactic context) cannot help disambiguate the words. In minimal lemma pairs that share a part of speech, the phonemic contrast bears the brunt of the functional load to maintain the lexical distinction; minimal lemma pairs belonging to different parts of speech will be distinguished in part by morphosyntactic context, reducing the functional load of the contrast and, accordingly, making it less resistant to merger.

The findings of these studies align with broader evidence that speakers modulate phonetic properties based on contextual predictability and information density. When linguistic units carry a greater functional load in maintaining contrasts, their phonetic cues tend to be enhanced (Aylett & Turk 2004; Cohen Priva 2012). Over time, this systematic enhancement may help preserve phoneme contrasts that play a more crucial role in the language's system of lexical distinctions. Such a possibility—that functional load not only protects contrasts from merger, but promotes the accretion of phonetic material—is one illustration of the importance of continued research on the operationalization and evaluation of functional load in language evolution.

2.2.3 Word vectors as a characterization of contextual similarity

Traditional approaches to measuring syntactic similarity (particularly, in Wedel et al. 2013a) have often relied on discrete categorizations like part of speech, which can mask more granular functional similarities in the usage of words. The functional load of a phonological contrast may depend not just on whether two words share a syntactic category, but on the degree to which they appear in similar contexts.

Word vectors offer part-of-speech-agnostic, highly granular method to capture these relationships, representing words as points in a high-dimensional space where proximity reflects distributional similarity (Mikolov et al. 2013a). Unlike categorical approaches, vector representations can encode rich information about the contexts in which words appear, in respect to both syntactic and semantic similarity. Rather than binning words into a finite set of pre-defined categories, word vectors can capture gradient similarities and complex relationships between words' usage patterns. Additionally, vectors can model these relationships from unannotated text, a benefit which potentially allows for less-resourced language varieties to be included in a wider range of studies that have traditionally focused on well-described languages. These representations have proved valuable in historical linguistics, where they can track gradual shifts in word usage and meaning (Hamilton et al. 2016).

In this study, the cosine similarity between two word vectors provides a continuous measure of their contextual similarity: words that appear in similar syntactic and semantic contexts will have vectors pointing in similar directions. We will use these vectors to test whether the functional load effects observed by Wedel et al. (2013a) reflect broader patterns of contextual similarity, attempting to replicate their finding for categorical syntactic distinctions.

2.3 Methodology

Wedel et al. (2013a) used discrete parts of speech in the construction of their dataset, which, as noted above, restricts the variety of data which can be analyzed to materials that have already been tagged for part-of-speech. Furthermore, the use of part-of-speech as a proxy for morphosyntactic context results in a dimensionality reduction of co-occurrence relationships that may be fully explained by a word class. Consequently, this study operates on text corpora and utilizes word vectors as the measure of syntactic similarity; this methodology serves both as an attempt to replicate the findings of the Wedel et al. (2013a,b) studies using an alternative operationalization of functional load, as well as an expansion of their study to investigate the extent to which a broader notion of context-similarity can be substituted for their coarser measure of word class.

The below sections detail the particular lexical resources in our analysis, as well as our preparation of candidate minimal pairs.

2.3.1 Connected corpus and lexicon selection

Wedel et al. (2013a,b) examine a set of 9 language varieties. Both studies includes English (Received Pronunciation and Standard American), German, Dutch, French, Spanish, Korean, and Cantonese. The final language differs between the two studies: Turkish (Wedel et al. 2013a) or Slovak (Wedel et al. 2013b). In this work, we examine a subset of 7 language varieties: English (Received Pronunciation and Standard American), German, Dutch, French, Spanish, and Slovak. The decision to focus on the subset of Indo-European languages in the study was motivated by the scope of the other chapters of this dissertation on language change within Indo-European; future work should certainly expand our approach to the full set of languages (and, ideally, a far broader set).

The resources for each language include connected corpora—which, in all cases, are sourced from the OpenSubstitles corpora (Lison & Tiedemann 2016) available through the OPUS project (Tiedemann 2012)—and various lexicons, which are detailed in Table 2.1.

Table 2.1: Summary of corpora used in the study

Language	Lexicon	OpenSubtitles corpus size		
		Lemmas	Tokens	Sentences
English (US)	CMUDict (Weide 1995)	2,260,434	794,963,999	102,548,133
English (RP)	CELEX2 (Baayen et al. 1995)	2,260,434	794,963,999	102,548,133
German	CELEX2 (Baayen et al. 1995)	1,226,424	148,239,595	21,227,082
Dutch	CELEX2 (Baayen et al. 1995)	1,385,333	277,640,188	36,898,831
French	Lexique (New et al. 2005)	963,045	328,956,430	42,876,478
Spanish	AnCora-ES (Taulé et al. 2008)[5]	1,864,652	557,464,819	74,593,385
Slovak	OpenSubtitles (Lison & Tiedemann 2016)[6]	569,121	49,115,061	8,003,865

Sentences were segmented based on line breaks and tokenized based on whitespace delimiters. As the OpenSubtitles corpora feature potentially multiple subtitle tracks per language, each corpus contained a high degree of sentence-level duplication. Sentences that were repeated verbatim were removed from the corpora (primarily in the interest of increasing pro-

[5] The Gigaword corpus (Mendonça et al. 2011), cited in the Wedel et al. (2013a,b) studies, does not provide lemmatized forms. It is unclear how Wedel et al. (2013a,b) prepared the list of lemmas used in those studies, so we substituted Taulé et al. (2008), a corpus which does include lemmatized forms.

[6] The Slovak National Corpus (Šimková 2006), used in the original Wedel et al. (2013a,b) studies, could only be obtained as a subset, which resulted in an unacceptably small number of minimal pair candidates. Thus the Slovak lexicon was instead inferred from OpenSubtitles. To exclude erroneous forms and unassimilated loanwords, a conservative minimum frequency threshhold for candidate forms was set at 50 tokens, based on inspection of the divergence of observed low-frequency tokens from the expected Zipfian (Zipf 1949) distribution—this threshhold also happens to approximate the one-per-million-word cutoff that Brysbaert et al. (2011: 421) references as containing the low-frequency words that lead to 'weak research'. Nevertheless, for full consistency with the other languages surveyed here, it would be preferable to use the full Slovak National Corpus.

cessing speed in the various preparation steps)[7]. Table 2.1 reflects the counts of these lemmas and de-duplicated sentences.

Initially, the research design did not depend on the existence of lexical resources. For the languages with a mapping of orthographic to phonological form that is particularly irregular (i.e. English, French) or particularly complex to model with simple substitution (i.e. German, Dutch), ByT5 grapheme-to-phoneme (g2p) models (Xue et al. 2022) were trained. For the languages whose orthographic systems are more transparently representative of phonology (i.e. Spanish, Slovak), orthographic forms from the OPUS corpora were converted to phonemic representations via ordered rewrite rules (implemented as simple regular expressions). However, as continued investigation of the OPUS corpora made clear, the removal of 'junk' forms (e.g. typographical errors introduced by OCR, high degrees of untranslated/unassimilated [often English] text in non-English corpora, encoding artefacts) was a non-trivial task. In this final analysis, the lexical resources in Table 2.1 were used as a filter for minimal pair selection in order to exclude these junk forms. As these lexicons already contain relatively-high-confidence phonemic representations of the orthographically-opaque languages, the ByT5 transformer models were omitted from the final workflow[8], and the lexical resources provided phonemic forms of the languages with particularly opaque orthographies (English, German, Dutch, & French).

[7] This strategy of de-duplication only eliminates exact duplicate sentences (e.g. from multiple identical, yet separate, subtitle tracks attached to a given piece of media). It does not eliminate sentences that differ only due to typographical errors or optical character recognition artifacts.

[8] The significant investment of resources in the generation of the ByT5 transformer models remains useful for ongoing and future research in this vein; they can be applied to connected orthography-only corpora that are already curated to remove junk forms and high degrees of non-native text. (For example, corpora from the Universal Dependencies project (Nivre et al. 2020), which are smaller in corpus size but do not share the same problems of representation as subtitle corpora.) Additionally, if retrained on a subset of phonemically-transcribed, connected language, the usage of the ByT5 models permits the possibility of generating more accurate phonemic representations for heteronyms based on context (Gorman et al. 2018)—e.g. the phonemic forms of the English heteronym *live* could be separated into the noun /laɪv/ vs. the verb /lɪv/ based on the surrounding context. Our present approach is limited by its inability to discriminate such forms, choosing the most frequent phonemic form available for an orthographic lemma regardless of heteronymy.

2.3.2 Minimal pair selection

In acknowledgment of the role of phonological similarity in the likelihood of a merger to occur, the set of candidate mergers was restricted to those phonemes which have a maximum featural distance of 1, as in Wedel et al. (2013a,b). Exceptions to this criterion were made only for contrasts which were attested but would otherwise have been excluded by the maximum featural distance criterion. A consonantal feature inventory and vocalic feature inventory was prepared for each language, informed by the cited analyses of the languages' phonological system. The set of phonemic contrasts with an attested merger is given in Table 2.2; the full set of pairwise contrasts, including those regarded as unmerged, are given in Table A.1 (Appendix A). In both tables, given pairs are categorized by 'Type', which illustrates the type of segments (consonant or vowel), as well as the any environmental conditions, that categorize the subsystem under investigation.

The feature matrices and method of comparison for complex segments (e.g. diphthongs) used by Wedel et al. (2013a,b) were unavailable at the time of writing; consequently, this study's sets of candidate mergers (given in Table A.1 in Appendix A) differ from those of Wedel et al. (2013a,b). Additionally, some inconsistencies in the reporting of attested mergers were noted across the various Wedel et al. (2013a,b) tables. These ambiguities were resolved by consulting the language-specific literature cited by Wedel et al. (2013a,b) for the attested mergers of each language. While the list of attested mergers is not exhaustive for each language, we strove to stay as close as possible to the Wedel et al. (2013a,b) studies and avoided adding additional

Table 2.2: Attested phoneme contrasts and mergers across languages

Language	Type	Contrast	Example	Source
English (US)	V∼V	ɑ∼ɔ	LOT∼THOUGHT	Labov et al. (2006)
		ɔɪ∼ɛɹ		
	V∼V/__ɹ	ɑ∼ɔ	START∼NORTH	
	V∼V/__n	ɪ∼ɛ	PIN∼PEN	
	V∼V/__l	ɪ∼i	HILL∼HEEL	
		ʊ∼u	PULL∼POOL	
		ʊ∼oʊ	BULL∼BOWL	
		ɔ∼ʌ	HALL∼HULL	
	C∼C/ V́__V	d∼t	LADDER∼LATTER	
	C∼C	ʍ∼w	WHICH∼WITCH	
English (RP)	V∼V	aɪ∼ɔɪ	PRICE∼CHOICE	Wells (1982)
		ʊə∼ɔː	CURE∼THOUGHT	
		ɪə∼ɛə	NEAR∼SQUARE	
		ɜː∼ɛə	NURSE∼SQUARE	
	C∼C	θ∼t		
		θ∼f		
	C∼C/{C,V}__	ð∼d		
		ð∼v		
French	V∼V	ɛ̃∼œ̃	VIN∼UN	Fagyal et al. (2006)
		e∼ɛ	ÉPÉE∼ÉPAIS	
		ø∼ɔ		
		ø∼œ		
German	V∼V	eː∼ɛː	GEBE∼GÄBE	Wiese (2000)
Dutch	C∼C	s∼z		Kissine et al. (2003)
		f∼v		
		x∼ɣ		
Spanish	C∼C	ʝ∼ʎ		Penny (2002)
		θ∼s		Harris (1969)
Slovak	C∼C	ʎ∼l		Krajčovič (1988)
	V∼V	æ∼a		
		æ∼e		

attested contrasts or subsystems, with one exception[9]. Future work would benefit from more comprehensive determinations of broadly attested mergers in a given language variety[10] and clearer boundaries delineating the language systems investigated[11].

Minimal contrast sets were identified by selecting distinct orthographic forms from a given lexical resource, whose phonological forms were iteratively mutated by replacing all phonemes of a given type with another phoneme. When the candidate phonological form of such a mutated word happened to exist as one or more 'real words' in the lexical resource, the original word was placed into a minimal pair with each word whose real phonological form matched the candidate. Homonyms were disambiguated by orthography. For example, take a word like American English[12] *Bob* /bɑb/. For the contrast (an unattested merger in our set) of /b/∼/m/, /bɑb/ would be matched with 'real' words like *bomb* /bɑm/, *balm* /bɑm/, and *mom* /mɑm/: homonyms, as separate lexical items, are included, and all words that distinguished by the phonological contrast, regardless of the number of segments affected, are also included.

[9] Wells (1982: 96, 328) notes TH-FRONTING in London English, the realization of interdental fricatives as labiodental or alveolar obstruents. This phenomenon is presumably the source of the attested mergers between /θ/∼{/t/, /f/, /s/} and /ð/∼{/d/, /v/, /z/} as listed in Wedel et al. (2013b: supplement p. 3), though the latter were absent from Wedel et al. (2013a: 402, 2013b: 181). Wells (1982) does not identify /θ/∼/s/ as an unconditioned merger and specifies the /ð/∼{/d/, /v/} mergers as conditioned by a non-word-initial environment. For this study, we have resolved this ambiguity by treating /θ/∼/s/ as an unattested merger candidate and recategorizing /ð/∼{/d/, /v/} into a conditioned subsystem C∼C/{C,V}__. After discovering the conditioned environment for the voiced dental fricative mergers, we attempted to proceed with only the uncontroversial set of /θ/∼{/t/, /f/} per Wedel et al. (2013b: 181). However, the large reduction of attested mergers caused singularity, forcing us to implement a full account for the conditioned /ð/ mergers. (See also fn. 13.) This solution is a fuller reflection of how the dental fricative mergers operate; we note the process in depth here since it resulted in an alteration to the dataset after the first round of model selection.

[10] Some examples of 'unattested' mergers that should likely be considered 'attested' within a given subsystem include /p͡f/∼/f/ for C∼C in German (Krech et al. 2009: 108) and /a/∼/ɑ/ for V∼V in French (Fagyal et al. 2006: 31) (though the particular latter contrast is already collapsed in the Lexique (New et al. 2005) phonemic transcriptions).

[11] The highly diversified macrolects of English are most conspicuously related to this point. For instance, the mergers of the dental fricatives (/θ/, /ð/) are not, strictly speaking, attributed to the learned, prestige Received Pronunciation, but rather to speech in London (Wells 1982: 328). Wells (1982: 297) deploys the label NEAR-RP to denote educated, regionally-sanitized English speech that is structurally different from RP—but given that one of his criteria for inclusion in NEAR-RP is the maintenance of all phonemic contrasts found in RP, the presence of these mergers suggests that the UK English system investigated in Wedel et al. (2013a,b) and here represents an even broader dialect continuum altogether.

[12] Apologies to any reader for whom the following examples do not all have the same vowel (which applies to at least one of the present authors).

Minimal pairs which were inferred from the lexicon of a language but not found in the OpenSubtitles corpus for a language were excluded from the analysis (as such pairs would have an undefined cosine similarity). Hypothetical contrasts were excluded from the analysis if the number of minimal pairs defined by the contrast was 1 or fewer. Table 2.4 shows the number of unique minimal pairs per language, categorized by whether or not a given minimal pair participates in one of the attested phonemic mergers depicted in Table 2.2.

Table 2.4: Minimal pair counts by language

Language	Minimal pairs				Total
	with merger		with no merger		
	count	%			
English (US)	202	1.77	11,191	98.23	11,393
English (RP)	317	1.68	18,552	98.32	18,869
German	38	0.65	5,793	99.35	5,831
Dutch	146	1.00	14,527	99.00	14,673
French	455	2.33	19,060	97.67	19,515
Spanish	28	1.66	1,659	98.34	1,687
Slovak	85	1.46	5,754	98.54	5,839

2.3.3 Model preparation and rationale

We next generated word embeddings, which are vector representations that encode the predictability of a target word based on its co-occurrence with some number (a window size) of neighboring context words that surround it. In linguistics research, these embeddings are often used to quantify semantic and syntactic similarity, both synchronically and diachronically (e.g.

Hamilton et al. 2016). In generating our word embeddings, use the metric of cosine similarity and follow Hamilton et al. (2016) in using a skip-gram with negative sampling approach implemented by word2vec with skip-grams, which they found to perform particularly well on historical data.

As word embeddings are, first and foremost, measures of patterns of a word's usage, the extent to which a 'usage' or a 'context' corresponds to the linguist's sense of 'semantics' and 'syntax' is non-trivial. One significant factor is the aforementioned window size, which refers to the size of the scope around a target word—essentially, it limits how many words away from the target word that a context word can be for the algorithm to consider it when calculating a vector. With these varying scopes of adjacency in mind, it might be assumed that relatively smaller windows (say a symmetrical window of size 1) may encode relatively more syntactic information than larger window sizes, as nearby words often have direct grammatical dependencies. For instance, if the window includes only immediately-adjacent neighbors, perhaps the target 'monkey' is likely to co-occur most frequently with words like 'the' and 'is' (situating it pretty firmly as a noun, in syntaxland). With a larger window, which might include, we may see words like 'zoo', 'trees', 'climbing', etc.—words which bear some degree of semantic connection with a monkey (in this example, it's a sort of scene).

In unfortunate contrast to this simple example, the relationship between window size and different types of contextual information is unclear in the literature (Levy & Goldberg 2014). While some work supports the window-monkey view of window selection—Mikolov et al. (2013b), for instance, used narrow context windows to produce embeddings that excelled at syntactic analogy tasks—other work suggests that both syntactic and semantic information are accessible at multiple window sizes (Lison & Kutuzov 2017). Mikolov et al. (2013a) found that while larger windows improved semantic task performance, they still preserved substantial syntactic information. The optimal window size may depend heavily on the specific task and type of relationship being modeled, as well as the nature of the source data, and the entwining of semantics–syntax is a significant challenge to attempts to disentangle their influences based on window sizes alone.

In order to compare the effect of window size, we generated word vectors for window sizes of 2, 3, 5, and 7. The sample of window sizes was limited to 7 at the upper end based on the average sentence length of 7.49 tokens (see the counts reported in Table 2.4).

2.4 Results

We describe here a mixed-effects logistic regression model, implemented with the `lme4` package in R (Bates et al. 2015). Considered separately, both increasing minimal pair count (Table 2.6) and increasing cosine similarity (Table 2.7) predicted the anticipated outcome of a decreased probability of merger. We initially formulated the models with a hierarchical random effect of LANGUAGE/TYPE, intended to account for the variations between different intra-language rates of merger and anticipating the possibility that sounds in a particular subsystem might feature different patterns of change relative to other subsystems (e.g. consonantal mergers vs. vowel mergers, or conditioned vs. unconditioned mergers). LANGUAGE/TYPE and LANGUAGE each resulted in singularity in the design matrix, and we were forced to reduce the random effect structure of the final models to include only TYPE.

Table 2.6: Minimal pair count model statistics

Predictor	Statistic	Value
–	AIC	270.7
Count	Effect	−0.685
	Standard Error	0.372
	z-score	−1.84
	p-value	0.065

Table 2.7: Cosine similarity model statistics by window size

Predictor	Statistic	Window Size			
		2	3	5	7
–	AIC	270.5	270.6	270.5	270.4
Cosine similarity	Effect	−0.699	−0.685	−0.693	−0.694
	Standard Error	0.379	0.374	0.374	0.373
	z-score	−1.84	−1.83	−1.85	−1.86
	p-value	0.065	0.067	0.064	0.063

2.4.1 Effect of window size

One aim this study was to investigate the extent (if any) to which the window size used for word vector calculation affected the probability of merger. As differing window sizes may capture different degrees of syntactic information, window size might be expected to affect the probability of merger if syntactic environment is a relevant factor.

Ultimately, window size did not change the predicted outcome of merged vs. unmerged, though slight differences in predictive power and significance were estimated (Table 2.7). Window size 7 is used (unless otherwise noted) in the discussion and visualization of our results, as the effect of cosine similarity had the lowest p-value of our models ($p = 0.063$).

2.4.2 Comparison of functional load metrics

True to the assessment of Wedel et al. (2013a: 404), the two measures of functional load as formulated in this study—minimal pair count and cumulative cosine similarity of the minimal pairs—are indeed highly correlated ($r = 0.8723$; Figure 2.1). To some degree, this is unsurprising—our cosine similarity metric is the simple summation of the pairwise cosine similarities of each minimal pair within a particular contrast group, meaning that contrasts borne by a large number of minimal pairs tend to have higher aggregate cosine similarity than contrasts with only a few minimal pairs, regardless of the pairwise cosine similarities in each group.

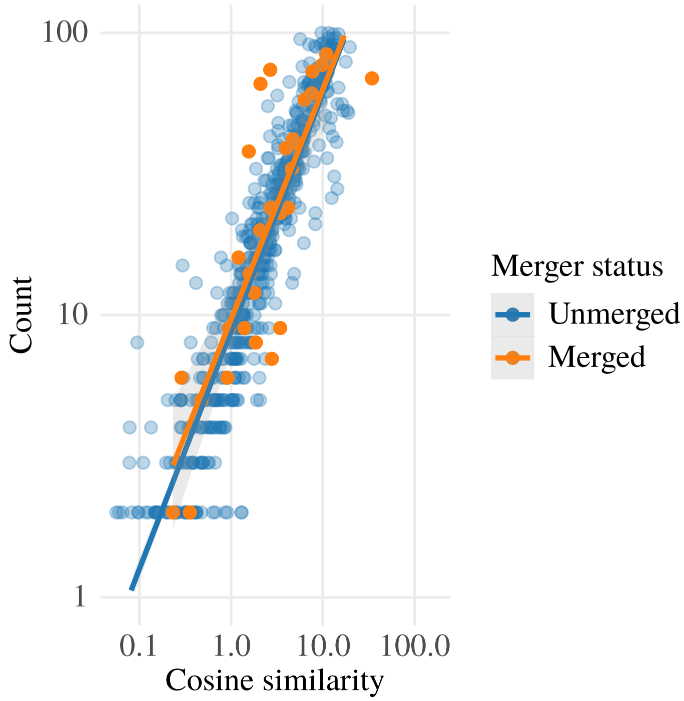


Figure 2.1: Log-scale count and cosine similarity of observed minimal pairs by attested merger status (window= 7, $r = 0.8723$)

As Wedel et al. (2013a: 404) note, the high degree of correlation between predictors such as these raises a challenge in determining which of the two predictors better predicts the probability of merger. Following their approach, we compared two models, each featuring either minimal pair count or cumulative cosine similarity as the sole fixed effect, with a superset model featuring both metrics, which were scaled to facilitate comparison and visualization. The predicted effects of this superset model are illustrated in Figure 2.2, showing the high degree of collinearity in the predicted effects.

A likelihood ratio test showed that neither predictor was significantly better than the other ($p > .5$). After model selection based on AIC (270 vs. 271), cumulative cosine similarity was the only remaining fixed effect in this maximal model ($\beta = -0.694$, $p = .063$), suggesting that of the two metrics in our study, cumulative cosine similarity shows a slightly greater absolute effect, slightly lower p-value, and explains merger probability with slightly less statis-

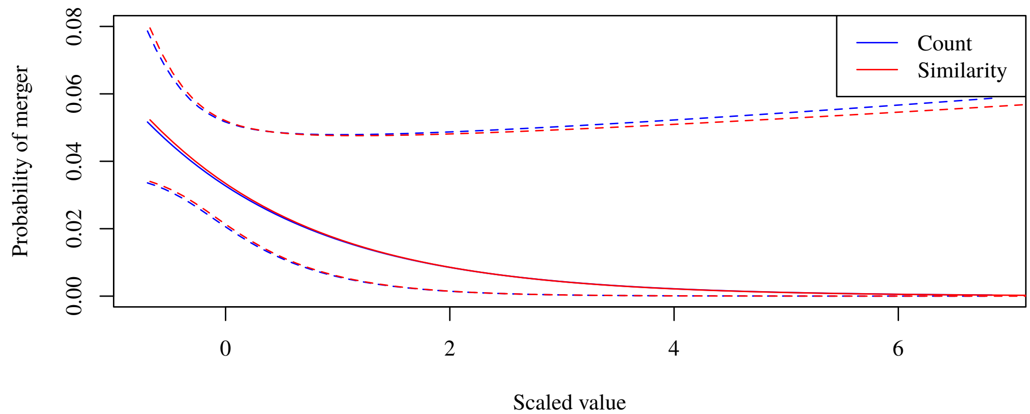


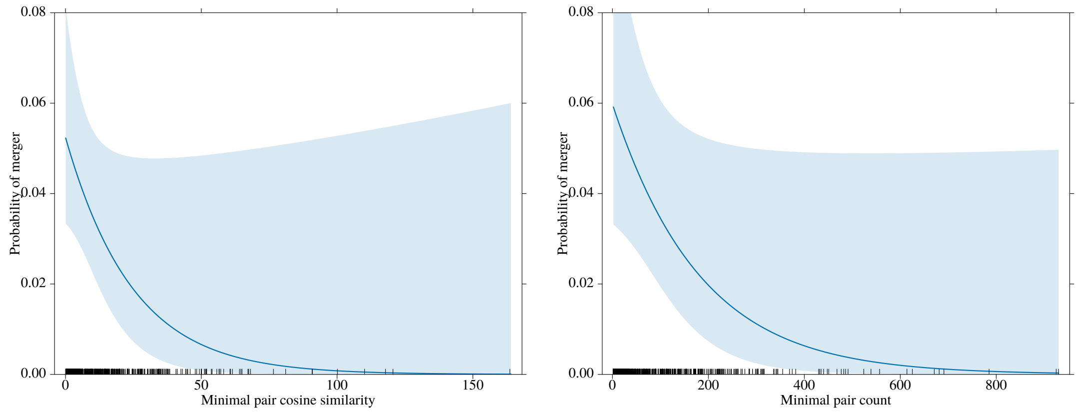
Figure 2.2: Predicted effect of count and cosine similarity on merger probability (window= 7, $r = 0.9431$)

tical complexity than minimal pair count (see Tables 2.6 and 2.7). In other words, the decrease in cumulative cosine similarity as formulated in our study is a marginally better predictor of phonemic merger over a baseline minimal pair count, but the two measures are not significantly different from each other[13].

Separate minimal models, each containing either minimal pair count or cumulative cosine similarity as the sole fixed predictor, each show a significant effect of increasing the respective functional load metric with a lower predicted probability of merger, as depicted in Figure 2.3.

As the 95% confidence band in Figure 2.3 suggests, the most confident estimate of the effect of each functional load metric occurs in a relatively narrow band corresponding to density of minimal pairs observed around a given level of the predictor. In other words, unsurprisingly, the effect of each metric is most confidently predicted for degrees of functional load which are most attested among the minimal pair data, and extrapolating the effect to very low or very high de-

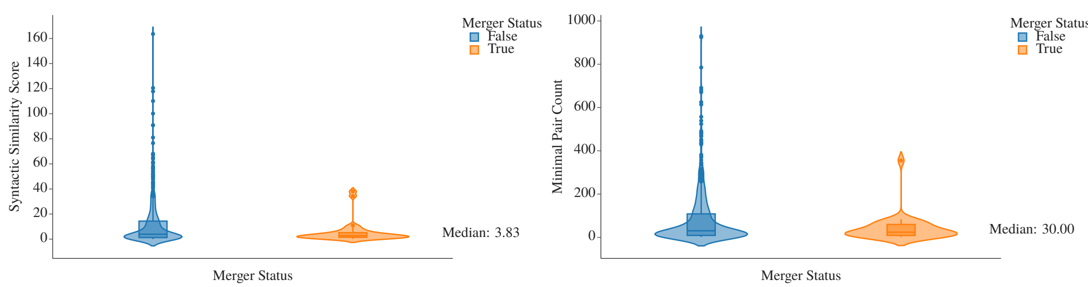
[13] It is important to emphasize that which of these two measures performs ‘better’ is highly dependent on the relatively low frequency of attested mergers in our dataset (1.63 % of minimal pairs that feature an attested merger, and across unique values for (LANGUAGE, SUBSYSTEM, CONTRAST), 4.08 % represent attested mergers). For example, during the initial model selection described in fn. 9, the change in a small number of contrast attestation decisions resulted in minimal pair count performing slightly better than cumulative cosine similarity (though, as here, the differences in their performance were not significant).



(a) Cosine similarity (window= 7) **(b)** Count

Figure 2.3: Effects of functional load metrics on merger probability **Fix slight text overlap on labels**

grees of functional load is less confident. Figure 2.4 illustrates this pattern in the observed data most straightforwardly—observed contrasts occur most frequently with a minimal pair cosine similarity score of about 2.0 (Figure 2.4a) and a minimal pair count of about 16 (Figure 2.4b). Additionally, Figure 2.4 demonstrates a distinction in the distributions of merged vs. unmerged contrasts: the observations of attested mergers are restricted to a narrower, lower range of the functional load metric compared to unmerged contrasts.



(a) Cosine similarity (window= 7) **(b)** Count

Figure 2.4: Distribution of functional load metrics by merger status

2.5 Discussion

This effect remained robust when controlling for other factors, including raw minimal pair counts. The vector-based similarity measure achieved prediction accuracy comparable to or better than the original part-of-speech classification method, particularly at window sizes 3 and 5. Additionally, we found that the effect was consistent across different language families in our dataset, suggesting this relationship is not an artifact of particular linguistic typologies but reflects a more general constraint on sound change. These findings support the functional load hypothesis while demonstrating that distributional methods can effectively quantify the relationships that affect phonological contrast maintenance.

This study provides empirical validation of the relationship between functional load and phonological merger through large-scale corpus analysis. Our results replicate the central finding of Wedel et al. (2013a,b) that higher functional load correlates with decreased probability of merger. The cosine similarity metric derived from word co-occurrence vectors performs comparably to raw minimal pair counts in predicting merger probability, though it does not appear to contribute additional predictive power beyond this baseline count measure.

The vector-based methodology employed here represents a significant advancement in operationalizing the notion of "functional space" in linguistic systems. Traditional approaches to functional load have often relied on discrete categories that may not capture the gradient nature of word relationships in actual language use. Much like how genetic systems maintain enough diversity to ensure adaptive flexibility while preserving core functional elements, phonological systems exhibit a comparable balance between variation and conservation. This approach demonstrates that computational methods drawing from distributional semantics can quantify these relationships without imposing artificial boundaries on linguistic categories. The correlation between vector similarity and resistance to merger suggests that phonological contrast maintenance is sensitive to the continuous nature of the semantic-syntactic space—a finding that bridges traditional structuralist conceptions of functional load (Hockett 1955; Martinet 1952) with contemporary usage-based approaches to language change (Bybee 2007; Wedel 2012).

Our findings suggest that bottom-up measures of distributional similarity are relevant, suggesting that the local, interactional patterns of how words are used in context may be just as important for sound change as their abstract grammatical properties, as Wedel et al. (2013a) find for part-of-speech. Such a finding aligns with usage-based theories of language change (Bybee 2007), which emphasize the role of concrete patterns of use in shaping systematic changes. Of course, like the similar studies mentioned throughout this article, our study has limited generalization to broader language production, as it relies on (naturalistic, but still artificial) written corpora. While large written datasets enable broad statistical analysis, they cannot capture the phonetic variation and gradience present in spoken language that may serve as precursors to merger (Hay et al. 2015). The lack of spoken data is particularly relevant given that mergers often begin in specific phonetic or social contexts before spreading more broadly (Labov 1994).

Our approach has both general limitations—particularly around the still-substantial work required to collect, clean, and lemmatize large corpora, which remain non-trivial tasks—and limitations specific to this particular study—for example, we did not incorporate known predictors of merger and indicators of functional load, such as entropy in the phonological system and frequency (Wedel et al. 2013a,b), the latter of which especially may have limited the performance of our cosine similarity measure over the baseline (as discussed above). Nevertheless, the ability to work with ”raw” corpus data, remaining challenges aside, enables the growth of functional load research to languages and varieties that lack extensive linguistic annotation.

While our findings confirm the notion that functional load (as both baseline minimal pair count and as cumulative cosine similarity) is protective against merger, our aims to demonstrate that context similarity would be better than a simple minimal pair count were not achieved in this study. This outcome is somewhat surprising given Wedel et al. (2013a)’s finding that parts of speech significantly modulate the effect of minimal pairs on merger resistance. There are several possible interpretations of this result.

One explanation is that discrete syntactic categories may have a stronger influence on contrast maintenance than gradient measures of contextual similarity. The categorical nature of syntactic roles could create more clearly delineated functional pressures compared to the potentially noisy signal captured by continuous vector similarity. The lack of explicit syntactic annotation in our dataset precluded us from investigating the performance of our distributionally-derived measure directly against a schema containing discrete categories (parts of speech); this remains a direction for future research.

Alternatively, our cosine similarity metric may be conflating multiple linguistic factors—semantic, syntactic, and distributional—in a way that obscures the specific relationship between distributional context and functional load—for instance, as operationalized in this paper, our cumulative cosine similarity metric embeds information about the number of minimal pairs within a particular contrast group. This decision made it difficult to disentangle whether merger probability is being influenced by the distributional similarity of the forms, their relative frequencies, or some interaction between these factors—initial exploratory attempts to disentangle the two (into an average cosine similarity score with an additional count predictor) yielded insignificant results, and were ultimately not pursued for the scope of the present study.

Moreover, we did not incorporate relative frequency measures that have been shown to influence merger probability (Wedel et al. 2013a). The balanced minimal pair frequency effect noted by Wedel et al. (2013a)—where minimal pairs whose members occur with similar frequencies show stronger resistance to merger—remains unexplored in our vector-based approach. Future research might explore alternative vector-space metrics that can better separate these effects, perhaps by explicitly normalizing for frequency or by focusing on specific distributional features rather than overall contextual similarity.

The substantially lower predicted probabilities of merger in our models compared to previous work likely stems from methodological differences in corpus composition. While Wedel et al. (2013a) focused on a carefully selected set of known mergers and similar unmerged contrasts, our decision to consistently and systematically include all pairs of phonemes with maximum featural distance of 1 produced a far larger set of unmerged contrasts, potentially diluting the observed effect sizes. Additionally, during model selection, we encountered convergence and singularity issues suggesting that alternative statistical approaches, such as permutation-based models, may be better suited to handling this type of heavily imbalanced data.

2.6 Conclusion

In this study, we investigated whether a bottom-up contextual measure could replicate the noted effect of shared syntactic category on the resistance of phonemes to merger (Wedel et al. 2013a). We replaced discrete part-of-speech categories with a continuous vector similarity measure derived from distributional information. Our results confirmed that phonemes which preserve contrasts between minimal pairs had lower merger probabilities across all window sizes tested. Substituting cosine similarity for the part-of-speech labels used in Wedel et al. (2013a), we were unable to replicate the Wedel et al. (2013a) finding that syntactically-similar minimal pairs confer phoneme pairs any additional resistance to merger, although our measure of contextual similarity performed about as well as the raw number of minimal pairs that a pair of phonemes distinguished, demonstrating that distributional similarity is a relevant metric to further investigate the effect of context predictability on constraining language change.

The methodology demonstrated here—using large, annotatation-agnostic corpora—opens up possibilities for investigating functional load across a broader range of languages and time periods than was previously feasible. This represents a significant advance over earlier work on functional load, which was constrained not only by computational limitations but also by the need for theoretically-annotated data. Where King (1967a) was skeptical of functional load's explanatory value due in part to the restricted data available at the time, modern corpus methods allow us to more comprehensively evaluate its role in sound change while controlling for other factors.

By demonstrating that vector-based measures of contextual similarity can replicate and extend the findings of Wedel et al. (2013a,b) without requiring a priori categorization of words, we validate not only the importance of functional load as a constraint on sound change, but also illuminate how these constraints operate in conjunction with natural distributional properties of the lexicon.

Chapter 3

Asymmetrical sound change across the word: a cohort-based quantification

3.1 Introduction

[*] In the context of language change, the extent to which language material serves a necessary, contrastive function affects its ability to resist erosion and reduction over time. Phonemic material responsible for the maintenance of some kind of contrast resists change (in the form of phonemic mergers) compared to material which bears no such functional load (Wedel et al. 2013a,b; Blevins & Wedel 2009). More generally, the effect of predictability on phonetic reduction has been a topic of recent investigation. It has been proposed that the increasing probability of a lexical item occurring in a local context is associated with the phonetic reduction of that item (Jurafsky et al. 2001), and a similar effect can be seen on sublexical units of meaning which are predictable from the context of use (Blevins 2005). There is a sort of selection pressure constraining sound change on material responsible for distinguishing meaning that does not apply to material that is predictable or redundant.

Synchronically, in lexical recognition, we see another, related type of 'reduction'—a general principle of economy to reduce cognitive effort when processing predictable material. One such proposed mechanism is proposed by the cohort model of recognition (Marslen-Wilson 1984;

[*] The study presented here is an extension of work I conducted for my MA thesis (Brendel 2018), which constitutes a pilot study for the more ambitious work here. I am grateful for useful comments on and discussions about that earlier thesis from Fermín Moscoso del Prado Martín, Argyro Katsika, Marianne Mithun, and Eric W. Campbell, as well as for the data collection work of the undergraduate research assistants Alyson Osgood and Kimberly Chiu. This study differs significantly in data, scope, and evaluation, and any errors that remain continue to be my own.

Marslen-Wilson & Tyler 1980; Marslen-Wilson & Welsh 1978), which describes a process in which lexical recognition of spoken words is driven using live processing of the auditory stream to winnow down the group of words which might correspond to the sounds being heard: an initial segment of speech activates a cohort of words which then are gradually ruled out as input is perceived. Speakers in isolated lexical decision tasks habitually make decisions about what lexical item is being perceived before the word has been fully heard. This effect occurs at what was initially labeled the uniqueness point: the point at which only one word in the lexicon begins with the sequence of sounds already heard.

Cohort competition—the state in which auditory input has not yet led to the selection of a single lexical item—among phonologically similar words is demonstrated to result in distinctions in speech processing, as discussed below. The cohort model as formulated in its earliest definitions provides a metric—the uniqueness point, recognition point, or Optimal Discrimination Point—which divides words into regions which seem to be, in some way, processed differently in the mind. Marslen-Wilson & Welsh (1978) propose that the early recognition of a spoken word allows the cognitive resources devoted to acoustic-phonetic processing to be freed up as soon as possible so that they can instead be devoted to synthesizing the wholistic meaning of a particular message; less attention "need be paid" (Marslen-Wilson & Welsh 1978: 61) to the remaining acoustic input of a lexical item. In terms of the early recognition of a spoken word, the remainder region does not contrast or distinguish one cohort competitor from another.

The cohort model implies that the region of a word after the recognition point is, in a sense, similarly predictable in that, for a given language, the remaining sequence of phonemes between the recognition point and the word-final boundary is highly predictable. Perhaps the resistance to sound change described above can also be differentially conferred by the location of a sublexical sequence in relation to the recognition point predicted by the cohort model: a region with no contrastive function in early word recognition may be freer to change. The cohort model can be useful as framework to help explain the link between cognitive effort—constantly shuttling between attentiveness and preservation of resources—and types of language material which may be particularly resistant or particularly susceptible to change.

In a previous pilot study (Brendel 2018), I found evidence for an asymmetry in sound change between the two regions delineated by a strict definition of the recognition point (the uniqueness point originally posed by Marslen-Wilson & Welsh 1978): when a word was bifurcated by the uniqueness point, the region after the uniqueness point had a higher average amount of changes in phonemes. In this study, I deepen and improve the pilot study through methodological refinements, such as the incorporation of cognate sequence alignment, statistical modeling that is more appropriate for the nature of the data, and the addition of frequency effects. In addition to finding a general increase in the probability that a slot is more likely to undergo a phonemic change the closer it is to the end of the word, I demonstrate that the recognition points predicted by the cohort theory delineate intralexical regions which are associated with different probabilities of sound change unpredicted by a baseline end-of-word bias alone.

3.2 Background

3.2.1 The cohort model

The original formulations of the cohort model (Marslen-Wilson 1984; Marslen-Wilson & Tyler 1980; Marslen-Wilson & Welsh 1978) propose a process of spoken word recognition in which listeners continually compare what they hear to possible lexical representations that they know. The 'cohort' is the range of words to which a sequence of auditory input could potentially map at any one time, and the size of the cohort is reduced as more auditory input is perceived: with more auditory input, there are increasingly fewer words which the stream of sound could correspond to (Figure 3.1). The cohort model contends that speakers make early decisions regarding the identity of a word before the entirety of auditory signal has been perceived, allowing words to be identified as soon as there are no other words in a lexicon to which the sequence of sound could possibly map. The point at which this decision can be made is, for spoken words heard in isolation, where "a particular word becomes uniquely distinguishable from any other word in the language beginning with the same sound sequence" (Marslen-Wilson 1984: 141). This point is referred to variously as the Optimal Discrimination Point (Marslen-Wilson 1984) or, in more recent formulations of the model which integrate information other than auditory input, the recognition point (Marslen-Wilson 1987).

$$tre \rightarrow \left\{\begin{matrix} treacherous \\ tread \\ trellis \\ tremor \\ trench \\ trenchant \\ trend \\ \ldots \end{matrix}\right\}$$

(a)

$$tren \rightarrow \left\{\begin{matrix} trench \\ trenchant \\ trend \end{matrix}\right\}$$

(b)

$$trench \rightarrow \left\{\begin{matrix} trench \\ trenchant \end{matrix}\right\}$$

(c)

$$trencha \rightarrow \{trenchant\}$$

(d)

Figure 3.1: A example of cohort winnowing with progressive auditory input on the left and candidates in English on the right (adapted from Marslen-Wilson 1984). In English, *trenchant* is the only word that could match the sequence beginning with *trencha*. This straightforward example presupposes the simplest sort of uniqueness point, after which the lexical entry is fully predictable. This point is immediately after *a*.

Particularly relevant for this study are the more recent formulations of the cohort model (such as Marslen-Wilson 1987), which move away from the idea of an 'all-or-nothing' identification process based on mapping of phonemes to candidates and instead implicate sources of knowledge beyond the auditory stimulus, such as the notion of word frequency. These formulations capture the intuitive notion that recognition, or comprehension in general, is not a binary state: rather, recognition happens in discourse, weighted by the the priming effect of semantic and syntactic context, as well as the likelihood for a word to be rare or common in general. This study is particularly interested in the latter effect. Accordingly, I will usually adopt the term 'recognition point': the end boundary of an initial sequence after which a given match is the most *frequent* in usage, not necessarily (and, in fact, almost never) unique.

A challenge for the original 'all-or-nothing' formulation of the cohort model concerns the retroactive identification of an item as a non-word based on data which follows the uniqueness point. In a lexical decision task, if no processing were to occur after the uniqueness point of a word, then listeners would be unable to distinguish non-words like THOUSAIDING /θaʊzaidɪŋ/ from lexical items like THOUSAND /θaʊzənd/, which becomes unique in English at /z/ (Taft & Hambly 1986). Some processing must therefore occur after the resolution of the initial stage of processing involved in selecting an increasingly smaller set of candidates in the cohort model, but the processing of input after the identification of a uniqueness point is not necessarily similar to the method of processing which occurs in advance of the uniqueness point (Cole & Jakimik 1980; Marslen-Wilson & Welsh 1978).

An additional complication is the ability of speakers to identify a non-word input such as GROCODILE /grɒkədaɪəl/ as being a clear misproduction of the valid word CROCODILE /kɹɒkədaɪəl/. If the cohort model were able to fully explain all lexical identification, words such as GROCODILE would never produce a cohort that includes the intended CROCODILE since the first phoneme /g/ can only produce a cohort of words beginning with /g/. And yet, speakers are able to identify such non-words as likely being misproductions of valid words which share a high degree of similarity after the initial phoneme. Taft & Hambly (1986) conduct several experiments to investigate the degree to which the cohort model can account for the processing of such words. The finding most relevant to this discussion is the delay in lexical decision for non-words like MEP /mɛp/ and MEPSIG /mɛpsɪg/. Both are non-words which diverge from all real English words at the phoneme /p/, and the cohort model would thus predict an equal delay in identification of non-wordhood on average. However, the results of a decision task show that participants take more time to identify the longer MEPSIG as a non-word. One reason for this finding is that the inclusion of additional phonetic information in the stimulus for the lexical decision task delays the full acceptance of the result of cohort model selection (Marslen-Wilson & Tyler 1980). Additional sounds after a uniqueness point, as discussed above, must be processed in some way to rule out non-words like THOUSAIDING, and perhaps even in non-words this effect can be seen: as long as there is phonetic input, a speaker processes it in some way. Marslen-Wilson (1987) refines his earlier conception of lexical recognition, proposing that bottom-up effects (as seen in the original formulations of the cohort model) are indeed used for early identification of a word in isolation, but that this process competes with a separate but parallel process that evaluates the likelihood of the candidate word to be an appropriate fit in terms of the context.

Even if the beginnings of words are most important for lexical recognition, what comes after the recognition point is not irrelevant: the ends of words are more important for lexical access than material in the middle of words (a point which my results also seem to endorse) and disruptions to the ends of words accordingly affect lexical access more acutely than corruptions of word-medial segments (Hawkins & Cutler 1988). The cohort model has previously been implicated as an explanatory factor for the cross-linguistic typological preference for suffixation over prefixation[14]: if the model is accurate, neural machinery races to select a semantic representation from auditory input as quickly as possible and therefore works most efficiently when content words are initial and grammatical elements are postposed (Hall 1988). The bifurcation of words suggested by the cohort model does not imply that one region is more important than the other, but that beginnings and ends of words are best suited to semantic and morphological

[14] Languages whose primary affixation strategy is suffixing outnumber those which are primarily prefixing by a ratio of 7:1 (Dryer 2013).

functions, respectively.

The discussion above primarily concerns the access of isolated words in contrived experimental settings. This context is, of course, not the natural environment of connected utterance. Even in isolation, the idea of a 'word' is nebulous: Haspelmath (2011) ultimately concludes that no satisfactory criteria for a cross-linguistic definition of the word exist, a notion more recently reinforced by Tallman (2021), who found no empirical evidence for an emergent constituent of wordhood in Chácobo. While descriptive linguists reject the idea that natural discourse is contingent on institutional prescriptions of grammar and pronunciation, the very idea of a 'word' as the basic unit of syntactic analysis in naturally occurring speech is a perception shaped by this same artificial force. However, some evidence suggests that the early recognition of a word applies in, and may even facilitate, the processing of real-time discourse. Marslen-Wilson & Welsh (1978) create a task in which participants listen to prose passages which feature mispronunciations of particular phonemes, and the participants repeat these passages aloud nearly simultaneously. When participants repeat the passages, they fluently fix highly noticeable mispronunciations that occur in the third syllable of a word which is primed by context[15]. The authors conclude that discourse context and the early recognition of individual words result in a system of lexical recognition optimized "to devote the minimum feasible processing capacity to the detailed interpretation of the incoming acoustic-phonetic input" (p. 61); the earlier that a word can be identified from a stream of phonemes, the sooner that this processing power can be devoted to the continuous input still being received (such as markers of grammatical relation or, indeed, new lexemes). The context of an utterance is not incorporated into the methodology of this thesis, but these findings suggest the cohort effect applies in naturalistic speech, the domain of sound change (an unsurprising notion, given Chapter 2).

Both my pilot study and the present chapter are motivated by the notion that the distinc-

[15] A sentence in which the mispronunciation is primed by context is something like *He wanted to smoke a *cigaredge (cigarette)*, as opposed to something like *He wanted to purchase a *cigaredge (cigarette)*.

tions in functional parallelism in the word recognition process (Cole & Jakimik 1980; Marslen-Wilson 1987), centered around the recognition point in the word, lead to differences in sound change in each region. This point is used as the basis for dividing words into regions which seem to be processed differently by speakers, and this study examines whether these distinctions in process result in distinctions in sound change. However, my current approach differs from my prior research in that this study does not suppose there is *no* selection pressure at word-ends; prior research and my own findings below instead seem to suggest two semi-independent selection pressures, springing from founts of different cognitive machinery, on either end of the word, and that these pressures both can act to constrain sound change.

3.3 Methodology

This analysis applies the cohort model to the dissemination of sound change within a particular language family. I investigate a set of cognate words—selected on the basis of the likely domain of the cohort candidate (lemmas and derived forms, as described above)—and divide them into regions delineated by the recognition point (whose existence has been asserted by formulations of the cohort theory). The quantification of difference between these subword regions is taken to be a diagnostic for sound change undergone between etymon and reflex—a difference now implies a historical diversification—and could reveal patterns of asymmetries in the distribution of sublexical sound change.

3.3.1 Data selection and preprocessing

Six languages from the Indo-European family were selected for this analysis. The set here differs from the eight languages used in Brendel (2018), which included some of the present set—UK English, German, Dutch, French, and Spanish—but added US English, Russian, and Serbian. The alteration to the pilot study set was motivated by the fact that corpus preprocessing, lemmatization, and grapheme-to-phoneme (g2p) conversion were already performed on OpenSubtitle corpora (Lison & Tiedemann 2016) for the current set of languages in the preparation of the analysis in Chapter 2 (see §2.3.1). Given the noted issues with the age and calculation of the frequency counts in lexical sources like Baayen et al. (1995) and the evidenced sufficiency of large subtitle corpora (Brysbaert et al. 2011), fresh lemma frequencies were calculated on the OpenSubtitle corpora (after removing exact duplicate sentences, an artifact of some media having multiple subtitle tracks per language for a given piece of multimedia). As lemma frequency was used only to identify the most frequent full word for a given initial substring, the extensive attempts to exclude low-frequency junk data (typographical errors, optical-character-recognition errors, and encoding artifacts) conducted in §2.3.1 were not performed here, under the assumption that the target words extracted from the manually-reviewed IE-CoR database (Heggarty et al. 2024) were unlikely to feature a high rate of data corruption.

The grapheme-to-phoneme conversion strategy varied; for languages where an authoritative lexical resource with phonemic transcriptions was available, a simple lookup was performed for words that were present in the lexicon. For undocumented words of orthographically opaque languages, a series of monolingual ByT5 models (Xue et al. 2022) was trained on the authoritative lexicons using the `transformers` Python library (Wolf et al. 2020) with a CUDA backend. Given the computational effort of training and inference, the ByT5 models were omitted for languages that are generally regarded as orthographically transparent, for which simple linear substitution rules were instead used to map the orthographic forms to phonemic forms. The languages along with the lexicographic resources and grapheme-to-phoneme conversion strategy utilized are listed in Table 3.1.

Table 3.1: Summary of corpora used in the study

Language	g2p strategy	Source of phon. form	OpenSubtitles corpus size		
			Lemmas	Tokens	Sentences
English (UK)	lookup + ByT5	CELEX2 (Baayen et al. 1995)	2,260,434	794,963,999	102,548,133
German	lookup + ByT5	CELEX2 (Baayen et al. 1995)	1,226,424	148,239,595	21,227,082
Dutch	lookup + ByT5	CELEX2 (Baayen et al. 1995)	1,385,333	277,640,188	36,898,831
French	lookup + ByT5	Lexique (New et al. 2005)	963,045	328,956,430	42,876,478
Spanish	substitution		1,864,652	557,464,819	74,593,385
Slovak	substitution		569,121	49,115,061	8,003,865

3.3.2 Cognate identification

None of the six dictionaries provide etymological information or glosses in any of the other languages represented. Consequently, these relationships were determined through the recently-completed Indo-European Cognate Relationships database (Heggarty et al. 2024). This database subsumes, corrects, and supplements the traditional 200-word list of core vocabulary (Swadesh 1952) compiled by Isidor Dyen (Dyen et al. 1992) on punch cards in 1970s, which I used for the pilot analysis (Brendel 2018). The availability of this resource is a significant advancement from the pilot study to the present chapter: although fewer cognates were selected from IE-CoR (995) than from the Dyen list (2,204) in Brendel (2018), the cognacy decisions are more principled and the data is richer (including, for example, the most recent common ancestor of a cognate set). Table 3.2 shows (1) the raw numbers of cognate set members for a language in IE-CoR, (2) the number which had at least one other cognate set member among the languages in the survey set, and (3) how many of those cognates were determined to have a recognition point.

Table 3.2: Cognates in survey set

	in IE-CoR	in pair/set	with RP
English	170	101	86
Dutch	141	129	108
German	172	144	131
French	170	126	96
Spanish	171	126	100
Slovak	172	46	41
Total	995	672	562

Inflectional suffixes, such as German and Dutch infinitive marker *-en* (e.g. German. *sitz-en* (sit-INF, 'to sit') and the Romance infinitive *-r*[16] (e.g. Spanish *caza-r* 'to hunt') were manually removed from the data.

Notes on the choice of Indo-European

Aside from the synergy with the datasets in Chapter 2, this project was conducted using Indo-European data for four primary reasons:

1. The availability of accessible and digitized phonological data
2. The availability of accessible and digitized resources establishing etymological relationships among specific words
3. The lack of lexical tone in the languages identified in Table 3.1
4. The (relative) clarity in segmentation of words[17]

[16] The thematic vowel, when present, was retained, as they are lexically-specified and unpredictable, seemingly the status quo since at least Proto-Indo-European (Sihler 1995: 482).

[17] Of course the definition of wordhood is a fraught topic, as I discussed above, but since this analysis (and as many of the analyses involving lexical access described above) relies on prescriptive sources (ie, dictionaries) for information, I refer here to the relative unambiguity of headwords in dictionaries of major European languages.

Firstly, while the application of the cohort model is not limited to Indo-European languages, the ready availability of lexical databases featuring phonological transcriptions of all items is necessary to identify at which phoneme a given word becomes most frequent in a language's vocabulary—in other words, the more complete the lexicon, the more accurately the recognition points of words can be identified, so languages which feature relatively complete lexicographic resources provide a higher number of datapoints on which the analysis can be performed. While ideally this analysis would feature a much larger number of languages which sampled all branches of the Indo-European language family, these six were selected as a starting point due to the accessibility and completeness of the data.

Secondly, the present study relies heavily on the establishment of cognate sets featuring unique subwords across languages, and the longstanding attention of historical linguists to Indo-European, coupled with more recent efforts to digitize these resources, facilitates the mass comparison of cognates across the family. For the purpose of this study, the most accessible lexicographic resources were selected, but optimally the majority of the languages included in the IE-CoR database could be included in the survey, especially since automated grapheme-to-phoneme models removes the criterion of needing digital lexical sources with phonemic transcriptions; even if a language is orthographically opaque, ByT5 models trained on a smaller, more attainable subset of transcribed data could obviate the need for comprehensive lexicons.

In addition to the Indo-European data in IE-CoR (Heggarty et al. 2024) and the earlier IELex (Dunn 2012), similar databases of high quality indeed already exist for other language families, such as the Sino-Tibetan Etymological Dictionary and Thesaurus (STEDT) project (Bruhn et al. 2015). The existence of these sources provides an avenue for replication of this study, including the ability to investigate the effect of lexical tone, an additional character relevant for some languages which is not examined in this paper due to its absence (in general) from the Indo-European dataset. This concern is the third reason for the selection of Indo-European data: the parsing of lexical tone adds an additional layer of complexity to lexical access. For instance, in

Cantonese, lexical tone has a high functional load but has been seen to induce greater error in lexical decision tasks between tonal minimal pairs than in segmental minimal pairs (Cutler & Chen 1997), and in Mandarin tone has been shown both to prime segmentally-identical pairs in certain contexts (Lee 2007) as well as to induce competition (that is, increased decision time) among segmentally-dissimilar candidates which share lexical tone (Poss et al. 2008). Additionally, the locus of tone processing in lexical tasks may lie in different cognitive machinery[18].

The final consideration, as discussed above, is that cross-linguistically and even within a single language, the definition of *word* varies greatly. The assumption that a word written in a dictionary is a unit which visually represents an identical unit (of a spoken word) stored in the mental lexicon provides some separation from the issue of defining the morphosyntactic or phonological word in natural language. The cost of this abstraction is that we study an artificial conception of 'the word' which is defined by a written standard, and moreover a chiefly Western standard. Accordingly, the decisions in data selection in the present study were made with the intention of remaining as close as possible to the linguistic sources of the empirical evidence which underpins the cohort model. The challenge of working with data from languages whose literary traditions diverge more greatly from the European languages on which the cohort model is based increasingly necessitates further questions of wordhood which will not be addressed in this study.

The present study focuses exclusively on Indo-European data, but the question of the generalizability of this model to non-Indo-European languages remains a topic of significant importance. While studies on isolating Sino-Tibetan languages are not uncommon, the applicability

[18] In an event-related potential (ERP) study, Malins & Joanisse (2012) find that Mandarin speakers display more activity in the left hemisphere when processing tonal minimal pairs (such as *huá* 'flower' and *huâ* 'painting') than when processing words that comprise a cohort based on early segmental similarity (such as *huá* 'flower' and *huí* 'gray'). It is uncertain whether these findings are limited to languages like Chinese with relatively small segment inventories and whose lexical tones bear high functional load in distinguishing near-homophones. The selection of Indo-European data enables these concerns regarding suprasegmentals to be reserved for future study and allows this initial examination of cohort-induced effects in diachrony to be tested on data which raises comparatively fewer questions regarding the synchronic processing of the candidates themselves.

of the cohort model to highly synthetic languages remains an open question. If the cohort model holds for these languages (as supported through research on these speakers' processing of lexical access), the present diachronic methodology could be applied to non-Indo-European languages in future projects.

3.3.3 Selection and conversion of lexical resources

Once cognate identification was complete, the phonemic representations of the relevant words in the six dictionaries were generated using one of the strategies noted in Table 2.1. Phonemic representations were chosen as they provide a layer of abstraction that should reduce noise in the data and allow this study to focus on the most meaningful distinctions.

Two other important factors, syllabification and lexical stress, were also excluded from this study, even though word-level stress is important in the segmentation of words in continuous speech (Cutler 1989). Stress is transcribed discretely only in half of the resources (CELEX2), and syllabification only in four (CELEX2 and Lexique 3). There is immense potential explanatory value in including these features in this analysis: syllable prominence could contribute to later-than-expected recognition points. However, even though stress is not treated in this study as an independent character, some effects of stress are implicitly present in the data. For example, when stripped of stress marking, CELEX2 English encodes *advertise* as /adv\u0259ta\u026az/ but *advertisement* as /\u0259dv\u025cti\u026azm\u0259nt/. While stress *per se* is not depicted, the difference between initial and post-initial stress is realized in a difference in vowel quality. Although this study does not enable the examination of stress as a discrete predictor, the effects of stress nonetheless are present in the data.

The full number of cognate pairs selected for analysis is listed in Table 3.3). Unsurprisingly, the languages that belong to more recent clades (West Germanic, Romance) have more shared cognates. The level of common subgrouping between pairs of these related languages is included as a random effect in the regression analysis described below.

Table 3.3: All cognate pairs identified

	German	English	Dutch	French	Spanish
English	95				
Dutch	122	85			
French	42	34	36		
Spanish	41	34	35	121	
Slovak	33	28	31	35	36

3.3.4 Alignment and comparison of cognates

The comparison of homologous sequences first required sequence alignment for the cognates. Cognate sets were aligned using the Multiple SCA algorithm in the Python processing library `lingpy` (List & Forkel 2024), which converts phonological sequences into sound classes using a predefined sound class model. Initial multiple alignment rather than pairwise alignment was chosen in order to give the alignment algorithm access to the most cognate data possible, leading to more accurate alignments of homologous material. The sound class model groups similar phonemes together based on their phonological features. Compared to a naive sequence alignment model, the approximation of natural classes increases the likelihood that sequences are aligned with cognate material that has diverged yet is nonetheless share a common proto-sound. An illustration of one particular aligned pair is in Table 3.4.

Table 3.4: Example alignment for cognate set TWO

German	*zwei*	t͡s	v	a	i	-
Slovak	*dva*	d	v	a	-	-
Spanish	*dos*	d	-	o	-	s

After alignment, I implemented a library in Rust (a compiled programming language) to efficiently calculate the recognition points of each word (in phonemic form) within the lexicon of its language[19], which I defined as the minimal initial substring required at which the full target word becomes the most frequent cohort candidate in the lexicon. Words which did not feature a recognition point of this sort were still included in this analysis: without a recognition point, the cohort theory would suggest the entire word participates in the initial discrimination process, and so should be subject to the same selective pressures (if any). The calculated recognition points were then converted from character-based indices to the slot-based indices necessary for applying the recognition point to a word in an aligned cognate set.

Next, words in a cognate set were compared in a pairwise fashion. The alignment of a cognate set across multiple languages often produced superfluous gaps (i.e., 'nothing' aligning with 'nothing') in a pair. For example, for the set in Table 3.4, German and Slovak both feature an empty slot in word-final position. This gap is necessary in order to account for the Spanish data, but would not be relevant if only treating German and Slovak. In the pairwise comparison, such shared gaps were skipped when comparing slots. This transformation retains the accuracy benefit of multiple alignment while maintaining parsimonious representations for pairs of cognates.

Finally, sublexical regions of words in the cognate pair were defined based on the indices of the recognition point of each previously-aligned word in the pair.[20] In most (but not all) cases, the positions of recognition points in each pair were not the same. The BEFORE REGION represents aligned lexical material before both recogntion points, and the AFTER REGION represents material after both recognition points. In cases where the recognition points were not aligned, an additional BETWEEN REGION was formulated between the two recognition points. This speci-

[19] This difference suggests the utility of native libraries of compiled code for mass language analysis; whereas the Python library I created in Brendel (2018) took over a minute to generate uniqueness points for all words included in a cognate pair, the native Rust library performs these calculations in under a second and was easily integrated with a typical data exploration workflow in a Jupyter notebook, using the `polars` data framework. On a practical level, optimization of this sort enables painless exploration and regeneration of the data, encouraging algorithmic refinements. (Appropriately enough for this chapter, removing [time] pressure encourages freer variation.)

[20] This is unlike Brendel (2018), where I aligned words at the shared uniqueness point and analyzed two regions.

fication is outlined in Equation (3.2), where r is the index of the slot after which the recognition point has been crossed, and a schematic representation of a fully aligned pair is given in Figure 3.3. Words without recognition points were treated as if they featured only one region, a BEFORE REGION[21].

$$
\begin{aligned}
R_{\text{before}} &= (0, \min(r_1, r_2)) && \text{(3.2a)}\\
R_{\text{between}} &= (\min(r_1, r_2) + 1, \max(r_1, r_2) + 1) && \text{(3.2b)}\\
R_{\text{after}} &= (\max(r_1, r_2) + 1, |S|] && \text{(3.2c)}\\
r_1, r_2 &\in \{0, \ldots, |S| - 1\}
\end{aligned}
$$

Equation 3.2: Region specifications as positions in word (0-indexed string indices). $R =$ region, $r =$ recognition point, $S =$ aligned slots

It is possible (and uninterrogated here) that the recognition point is crossed during the realization of the phoneme and not after production; however, due to the treatment of discrete phonemes and the operationalization of recognition point chosen, the recognition point is regarded as occurring immediately after a given phoneme, since this phoneme is the last element in the sequence required to select the target word as most frequent[22].

Finally, per region, slots in the same position were counted as undergoing change if they were non-identical. I implemented an algorithm that is essentially a slot-level implementation of Levenshtein-Damerau distance (Damerau 1964; Levenshtein 1965), which measures the minimum number of single-character edit operations (insertions, deletions, substitutions, and adjacent transpositions[23]) needed to transform one string into another. In linguistics, this distance measure has been variously used, for example, in the quantification of genetic relation-

[21] Including the presence vs. absence of a recognition point as a fixed effect in the analyses below was not significant ($p > 0.1$).

[22] The notion that the recognition point could be 'inside' the phoneme could be investigated in this study by allowing BETWEEN REGION to include the final phoneme slots that are required for the recognition point (overlapping 1 character each with BEFORE REGION and AFTER REGION). However, this option was not investigated in this study, as such overlap would likely introduce difficult-to-disentangle collinearities among the regions.

[23] The inclusion of adjacent transpositions allows metathesis (e.g. *aks*>*ask* in English) to be treated as 1 'change' rather than 2 changes, placing it on the same level as other historically-known common sound changes captured by a plain Levenshtein distance, which would treat the *aks*>*ask* example as two processes (syncope and epenthesis).

ship (Wichmann et al. 2010), pronunciation differences (Wieling et al. 2014), and phonological distance (Sanders & Chin 2009). Levenshtein distance applied to phonemic representations of cognate material can function as a diagnostic for sound change. If the phonemic representations of homologous regions in two cognates are dissimilar, historical change must have applied.

The full alignment, region specification, and summed changes per region are given in Figure 3.3. To prepare for Poisson regression, the final dataframe contains a row per cognate pair per region with non-zero length in that pair.

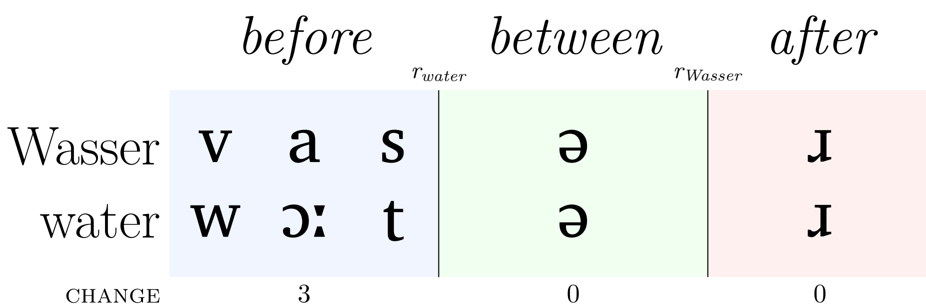


Figure 3.3: Example of region alignment

3.4 Results

Initial exploration of the assembled dataset, as shown in Figure 3.4, reveals a similar impressionistic finding as Brendel (2018): namely, the distribution of the proportion of phonemic change varies in the region after a recognition point. The BEFORE REGION shows a wider range of change than the other two regions: while BEFORE REGION can certainly undergo full phonemic replacement ($y = 1$), its mean and median change is lower than the other regions, which feature very high mean and median proportions of change.

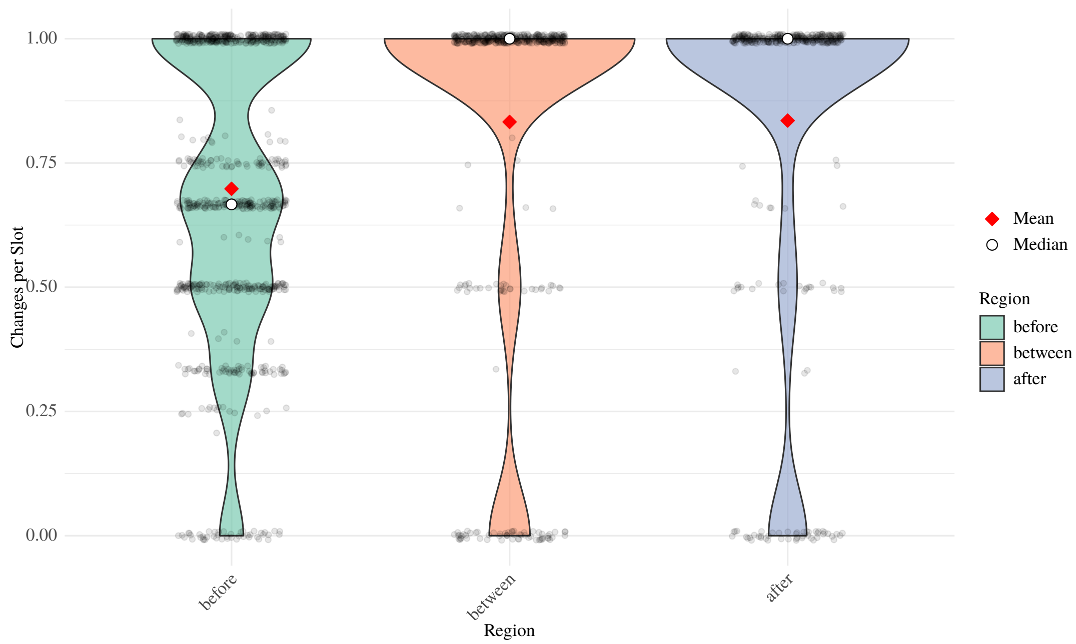


Figure 3.4: Observed proportion of phoneme slots changed per region, with density and scatter of observed datapoints (region per cognate pair)

3.4.1 Region model

To model the relationship between phonemic changes and word region in the Indo-European cognates, I constructed a mixed-effects Poisson regression using the `lme4` package in R (Bates et al. 2015). The model examines how regional position within a word influences the probability of phonemic changes in that region (Change) For the fixed effects, I included the categorical variable Region which divides each word into three segments (*before*, *between*, *after*) as well as an offset term for the number of slots (Slots) available for potential change in each region, which accounts for the varying opportunities for change across different word lengths (and intralexical region lengths). Region was treatment-coded, with Region=*before* set as the reference level.

To control for phylogenetic relationships in the data, I added two ordered categorical variables as random effects , since older cognates would be expected to feature more changes than recent cognates. Recency captures the genetic relationship between a language pair and, therefore, the time depth since the etymon for that cognate pair began diversifying in this subfamily. The time depth is simplified into two levels (*recent* and *old*) and denotes the number of levels of subgrouping that separate the language pair from their most recent common ancestor (e.g., West Germanic, Romance, Proto-Indo-European). Some intra-family cognates which were borrowings from Latin but proceeded to diversify in the recipient subfamily (e.g. German *Frucht*, English *fruit* < Latin *frūctus*) were present in the data in small numbers; these were treated as *recent* borrowings, as the borrowings were introduced well after the diversification of Indo-European. Consequently, to further distinguish the age of etymon regardless of borrowing, I added a more granular Ancestor which represents the number of levels of subgrouping that separate the extant language pair from the proto-language/language which innovated the etymon (levels *recent*, *middle*, *old*, corresponding to the 3 levels of subgrouping represented in IE-CoR). As Slovak was the only Slavic language, it has only *old* values for Recency and Ancestor, aside from Latin borrowings, which are Recency=*recent*. This information is summarized in Table 3.5.

Table 3.5: Categorical structure of language relationship functions

Factor	Language grouping	Simplified level	Ordering
Ancestor	Proto-Indo-European	Old	Recent < Middle < Old
	Proto-Germanic/Proto-Italic	Middle	
	Latin/West Germanic	Recent	
Recency	West Germanic	Recent	Recent < Old
	Romance	Recent	
	Latin borrowings	Recent	
	Other cross-family	Old	

Finally, LanguagePair, the sorted pair of languages for the given cognate pair, was added as a random effect, accounting for levels of subgrouping that were not noted in IE-CoR and for other language-pair-specific variations. An additional nested random effect of cognate_set/word was initially added to the model, but I was forced to remove the nested and independent effects as they caused singularity issues and explained no variance. The other three random effects described above were retained, leading to the final formulation of the Poisson regression, which I'll call the Region model (3.5):

$$\text{Change} \sim \text{Region} + \text{offset}(\log(\text{Slots})) + (1|\text{Recency}/\text{Ancestor}/\text{LanguagePair}) \quad (3.5)$$

The model (Table 3.6) shows a significantly lower predicted rate of phoneme change in Region *before* and significantly higher predicted rates of change in the other two regions. When transformed to rate ratios, the Region=*between* effect ($exp(0.172) = 1.188$) indicates that cognate changes occur at a rate approximately 18.8% higher between recognition points compared to the region before both. Similarly, the Region=*after* effect ($exp(0.160) = 1.174$) suggests a 17.4% higher rate of changes. The intercept ($exp(-0.413) = 0.662$) represents the baseline rate when all other variables are at their reference levels, indicating that, on average, 66.2% of the slots in Region=*before* undergo change (as compared to 78.7% of the slots in *between* and 77.6% of slots in *after*).

Table 3.6: Region model mixed-effects Poisson regression results

Effects	**Estimate**	**std. error**	**z-value**	**p-value**
Intercept (Region=*before*)	-0.413	0.140	-2.961	< 0.01
Region=*between*	0.172	0.048	3.614	< 0.001
Region=*after*	0.160	0.051	3.171	< 0.01
Random Effects	*(Std. Dev.)*			
LanguagePair:Ancestor:Recency	0.068			
Ancestor:Recency	0.135			
Recency	0.121			

Although REGION=*between* shows higher predicted change than REGION=*after* (an intriguing finding), pairwise comparison of the regions with `emmeans` shows that while BETWEEN REGION and AFTER REGION are significantly different from the BEFORE REGION ($p < .001$, $p < .01$ respectively), they are not significantly different from each other ($p > .5$). This finding suggests that the theoretical distinction in this study between a BETWEEN REGION and an AFTER REGION could be conflated into one region. The comparison of these regions is visualized in Figure 3.6, which also shows the difference in predicted means for cognates of differing time depth.

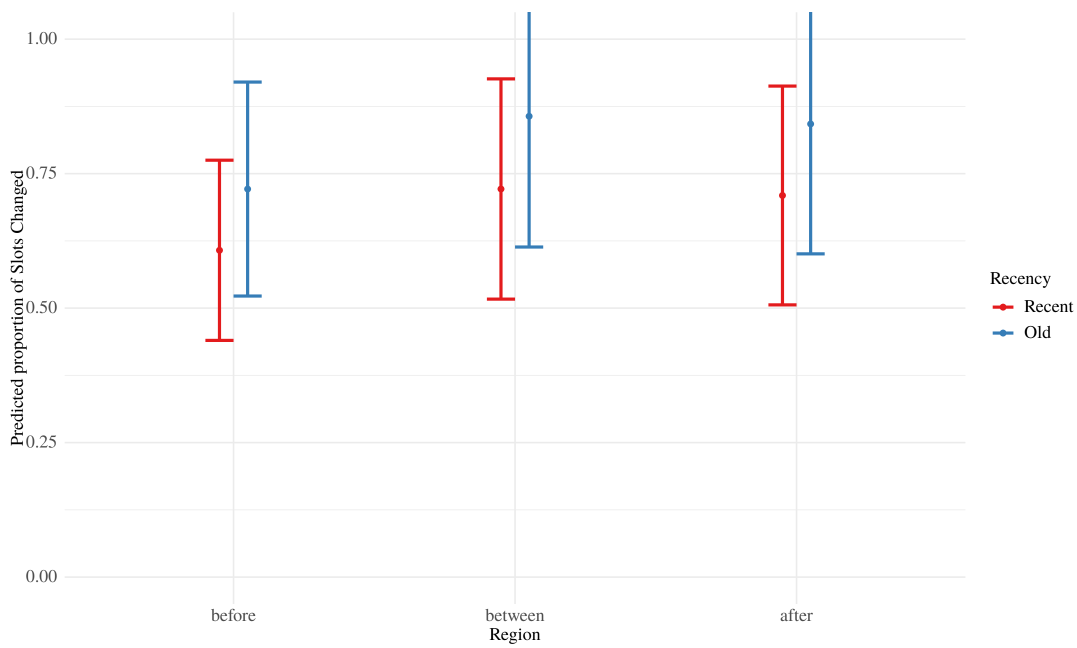


Figure 3.6: Marginal predicted probabilities of phoneme change as portion of region length (95% confidence interval), separated by region and cognate recency (time depth)

3.4.2 Position model

In addition to the REGION MODEL Poisson regression conducted above, a second model, the POSITION MODEL, was created and run on a transformation of the dataset in order to examine the extent to which the results of the first model could be explained by a general end-of-word tendency to undergo greater amounts of change, motivated by the slight (but statistically insignificant) increase in predicted probability of change in BETWEEN REGION compared to AFTER REGION.

The dataset assembled previously was ungrouped for region. The indices of aligned slots were transformed into a relative index of its position within the aligned pair ($slot_i/len(slot)$), termed POSITION. Each slot was compared for change as outlined in §3.3.4 and was stored as a separate row in the recreated dataframe. The tripartite REGION factor was re-created *post hoc* by comparing the POSITION to the recognition points calculated for that minimal pair. If either word lacked a recognition point, rather than excluding the data, the choice was made to treat the word as if the word had its recognition point at its end-of-word boundary; this decision is a simplification, but captures the idea that a word that lacks a recognition point may be subject to the cohort competition process across its entire length, similar to the BEFORE REGION of a word with a recognition point. When this assumption was made, the existence of a null was noted in a (NULLRP) examine whether this decision had a meaningful impact on the results of the subsequent logistic regression.

A logistic regression was fit, incorporating both the linear and nonlinear fixed effect of REGION in order to allow for investigation of the unexpected, slight increase in probability of change in REGION=*before* seen in the REGION MODEL; a cubic polynomial was chosen to allow for the possibility of an inflection at each of the two REGION boundaries. REGION and NULLRP were also included in the maximal model. All fixed effects were allowed to interact. For this model, the same random effects structure was adopted as in the REGION MODEL, with the addition of a random slope for POSITION. During model selection, convergence issues forced the simplification of the random-effect structure to estimate a random slope only for the highest-order random effect and only random intercepts for the others. Higher-order REGION polynomials had insignificant ($p > 0.1$) effects and were not selected. AIC comparison and likelihood ratio tests revealed that model fit was improved only by removing NULLRP and its interactions, suggesting that the treatment of words lacking a recognition point as having an end-of-word recognition point was an acceptable assumption for this analysis. k-fold cross-validation showed that the cubic model had the least mean error rate of the linear, quadratic, and cubic models ($.245$, still relatively large). The final formulation of the model after selection is shown in Equation (3.7), and the results are given in Table 3.7.

$$\begin{aligned}\textsc{Change} \sim \textsc{poly}(\textsc{Position}, 3) \times \textsc{Region} \\ + (1 + \textsc{Position} \mid \textsc{LanguagePair:Ancestor:Recency}) \\ + (1 \mid \textsc{Ancestor:Recency}) + (1 \mid \textsc{Recency}) \quad (3.7)\end{aligned}$$

Table 3.7: POSITION MODEL mixed-effects logistic regression results

Effects	**Estimate**	**std. error**	z**-value**	p**-value**
Intercept	0.709	0.511	1.386	0.166
POSITION (orthogonal polynomials)				
Linear	73.687	8.624	8.544	< 0.001
Quadratic	−44.749	6.062	−7.381	< 0.001
Cubic	24.895	4.063	6.127	< 0.001
REGION (treatment-coded)				
Linear	−1.236	0.188	−6.588	< 0.001
Quadratic	−0.895	0.151	−5.920	< 0.001
POSITION × REGION interactions				
Linear × REGION (linear)	97.225	13.592	7.153	< 0.001
Quadratic × REGION (linear)	−77.336	11.497	−6.727	< 0.001
Cubic × REGION (linear)	34.256	6.548	5.232	< 0.001
Linear × REGION (quadratic)	51.819	10.996	4.712	< 0.001
Quadratic × REGION (quadratic)	−39.052	9.352	−4.176	< 0.001
Cubic × REGION (quadratic)	19.753	6.565	3.009	0.003
Random Effects		**Parameter**		***Std. Dev.***
LANGUAGEPAIR:ANCESTOR:RECENCY		Intercept		0.321
		POSITION		0.982
ANCESTOR:RECENCY		Intercept		0.218
RECENCY		Intercept		0.668

The Position model (Table 3.7) shows that all fixed effects are highly significant predictors of change (aside from the interaction of Position^3×Region, which could not be dropped), supporting the notion that change is more likely the further into the word a phoneme is found. Any attempts to remove the effects of Region and Position resulted in a worse model: AIC comparison revealed that the model including the effect of Region ($AIC = 3603$) performed better than a version where Region was excluded ($AIC = 3610$).

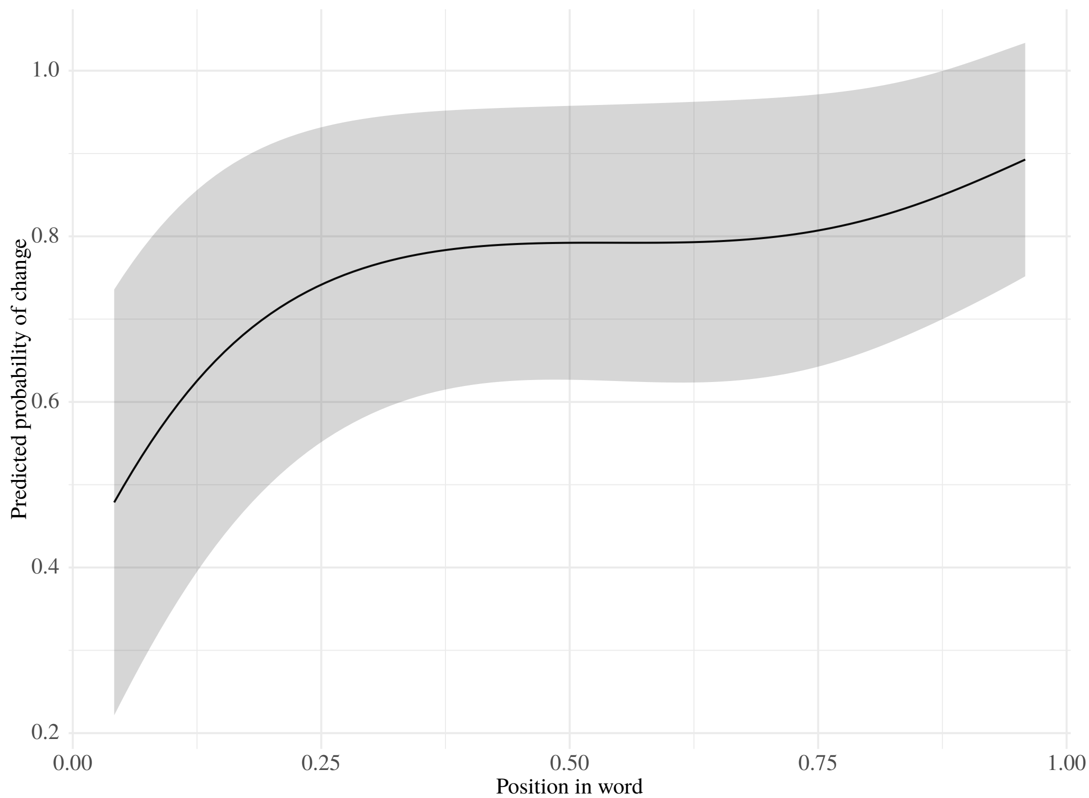
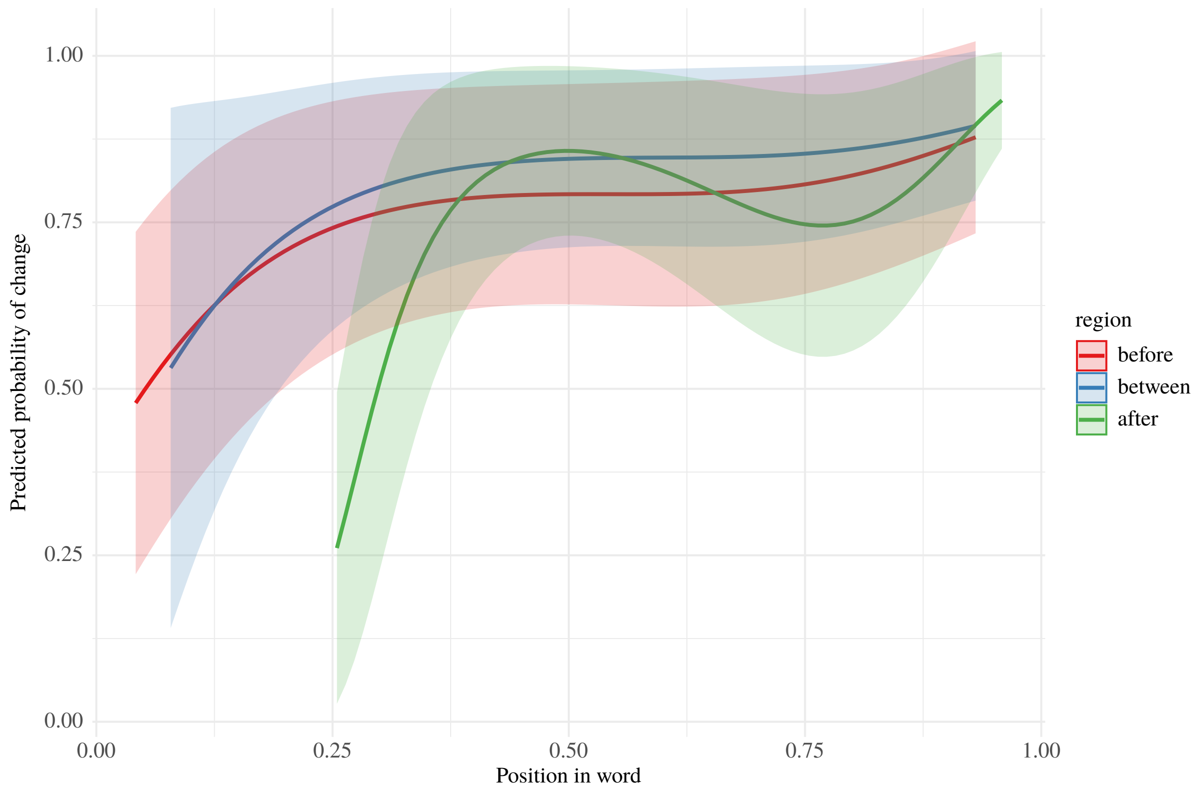
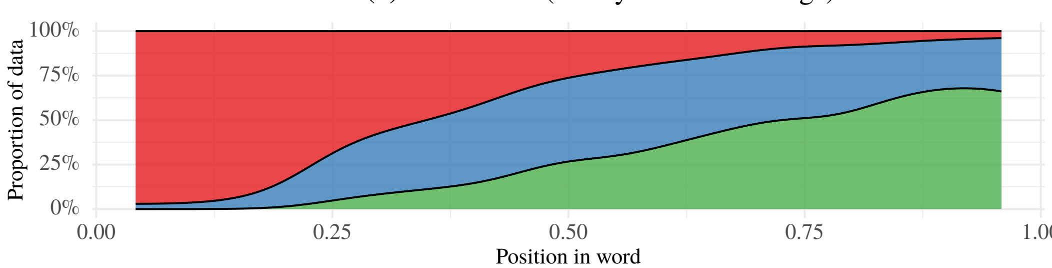


Figure 3.8: Marginal predicted probability of change of Position model (95% confidence interval)

Figure 3.8 visualizes the marginal predicted probability of change from the position model. Note that, impressionistically, Figure 3.8 seems structured into three regions: an increase in change probability after the start of the word (about first third of the word); a plateau in the rate of increase in change around the middle of the word (around 0.33–0.73), where change probability is still elevated, but does not grow more likely depending on position within that zone; and then a more gradual increase in rate of change towards the end of the word (after around 0.70). The nonlinearity aside (discussed in §3.5), Figure 3.8 shows that the POSITION MODEL suggests a tendency for change to become more likely toward word-ends than beginnings.



(a) Predictions (x only in attested range)



(b) Relative frequency of data

Figure 3.9: Predicted probabilities of change in slot at position, separated by region (95% confidence interval)

The interaction of REGION and POSITION is visualized in Figure 3.9, where the marginal prediction in Figure 3.8 is split into probabilities by REGION. By separating the predicted probabilities of change by region, Figure 3.9 suggests that there is a deeper pattern of asymmetry across the word than Figure 3.8 implies: whereas the shape of predicted change impressionistically follows the marginal prediction of Figure 3.8 in REGION=*before* and *between*, in the *after* region, the pattern is markedly divergent, featuring a relatively low probability of change when the recognition point is early in the word[24], but is followed by an immediate sharp increase toward the middle of the word. The probability of change then follows a shallow U-shape, momentarily *decreasing* in probability before increasing towards the edge of the word. This second dip is not predicted by the design of this study, but will be discussed in §3.5.

Table 3.8: Effects of position score within each region

Region	F-ratio	p-value
Before	4.992	.025
Between	2.145	.143
After	53.351	$< .001$

To quantify the extent to which the effect of POSITION varied across REGION, analysis with `emmeans::joint_tests` (Table 3.8) was conducted, which showed significant effect of position score in the REGION=*before* ($F = 4.99$, $p = .025$) and *after* ($F = 53.35$, $p < .001$), but no significant effect in the Between region ($F = 2.15$, $p = .143$), confirming the pattern of interaction indicated in the model (Table 3.7) and confirming that the strongest relationship between the two is in REGION=*after*.

For parallelism with the discussion of the REGION MODEL, the main effect of REGION underwent pairwise comparison with `emmeans` (Figure 3.10 and Table 3.9).

[24] It should be noted that REGION=*after* tends to become more common later in the word, so there are fewer examples at this early range; see §3.4.2.

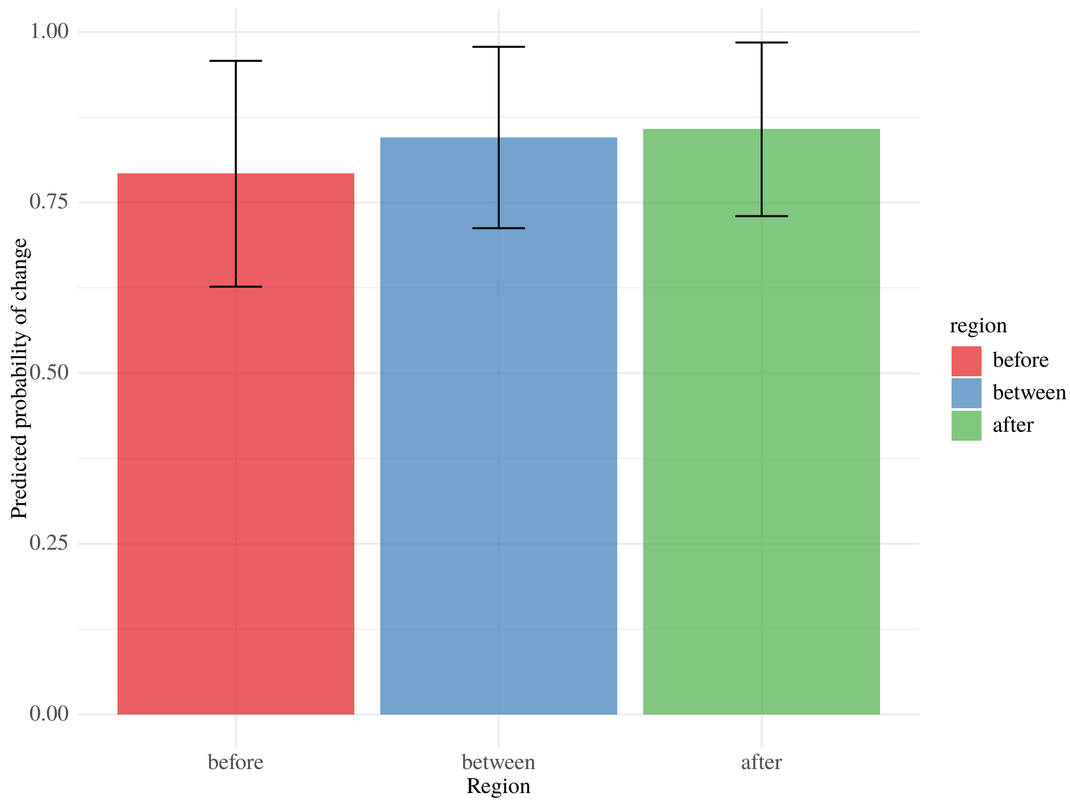


Figure 3.10: Estimated marginal means of region (averaged over all other effects)

Table 3.9: Pairwise comparisons between REGION main effect (on response scale)

REGION contrast	Estimate	std. err.	z-ratio	p-value (Tukey)
before ∼ between	−0.0531	0.0305	−1.740	> 0.1
before ∼ after	−0.0651	0.0352	−1.851	> 0.1
between ∼ after	−0.0120	0.0280	−0.427	> 0.1

The Position model does not display the same significant differences in predicted change between regions (as main effects) that were seen in the Region model. However, it should be noted that the Region model lacked a significant interaction between Position:Region, which the Position model does have, so this incongruence is not surprising. This finding suggests that position alone might be enough to account for differing rates of change before a recognition point (or imperfect gray area between two recognition points in a cognate pair), but the region that is unambiguously post-recognition point (as this study treats it) undergoes a pattern of change that is better explained by the interaction of both word position and post-recognition status.

In summary, then, like the Region model, the Position model, shows a slight increase in predicted change in Region=*after*, though (unlike in the Region model) it is not a statistically significant difference from any of the other predicted marginal means of Region (Table 3.9). However, more intriguingly, Region=*after* shows a markedly different pattern of change compared to the regions before a recognition point[25], matching the expectation that the region after the recognition point would change, if not *more*, then at least *differently* than the material before, regardless of the end-of-word bias that seems to affect all material irrespective of its relationship to a recognition point (if any).

3.5 Discussion

The results of this study provide evidence for asymmetrical patterns of sound change within words that align with predictions from both cognitive processing models and historical linguistic research. The Region model shows a significant increase in phonemic change after a recog-

[25] The shared pattern between the before region and the between region suggests that these regions could perhaps be conflated in future extensions of this study; if there is a differential effect of cohort candidate selection, the findings in this study imply that that effect is observed in what I've defined as the after region, not in what I've termed the between region (see Table 3.5 and the surrounding discussion).

nition point, which aligns with previous research showing that material with lower functional load is more susceptible to change. The Position model partially disentangles the effect of these recognition-point-defined regions from other asymmetries within the word: namely, the end-of-word bias for phonological neutralization, noted by Wedel et al. (2019), is confirmed and observed for sound change in general, not just neutralization. Moreover, the Position model shows that the effect of position on sound change varies significantly on either side of a recognition point, which supports the hypothesis that regions of words that serve different functions in lexical access may be subject to different types or degrees of change. Just as Marslen-Wilson & Welsh (1978) demonstrated that cognitive resources could be redirected once a word becomes uniquely identifiable, these findings suggest that this functional distinction may influence patterns of historical sound change. While it is not clear from this study whether the region after cohort selection undergoes *more* change (more than would be expected just from being nearer the end of a word), the difference in pattern of predicted change could, perhaps, be explained by the different cognitive machinery that attend to either region.

While the exceptional pattern of the after region is not surprising under this study's assumptions, another matter that is more perplexing is the clear, middle-of-word elevated plateau in the marginal predicted change of the Position model, shown in Figure 3.8. This pattern is more challenging to explain within the frameworks invoked in this chapter. In fact, word middles should be, perhaps, *least* constrained by the selection pressures of auditory recognition, as they are, by definition, less likely than beginnings to participate in cohort competition and also further away from both the next word in connected speech (which would, once again, participate in cohort selection) and from, perhaps, suffixal material that should be accompanied by an increase in attention[26]. Of course, middles of words are particularly important in languages with stem-internal changes or infixation. Modern Indo-European languages themselves have varying

[26] See §3.2.1 for a discussion about word middles (Hawkins & Cutler 1988) and suffixation (Dryer 2013; Hall 1988).

degrees of ablaut, with varying degrees of extant systematicity and predictability. If ablaut has high functional load in distinguishing inflectional variants of high-frequency words, and, moreover, if ablaut is an increasingly irregular system, then perhaps attending more closely to the irregular-but-important middles of words could protect middles of words from the exponential effect of position on the beginnings and ends of words[27]. This study is agnostic to the status of middles: contrary to the marginal predicted probability of the Position model, the per-region probabilities (Figure 3.9) show that the middle of a word is a local *maximum* of change (and inflection point) if a phoneme is in the after region. This line of thought won't be investigated here, but would make a fascinating and informative future study.

3.5.1 Limitations

The recognition points calculated in this study do delineate regions that seem to be under different selection pressures. However, several limitations must be considered. First, while the Position model suggests a different pattern of change in the after region, the use of relatively short words means that the between region and after region of words tend to be short compared to the before region. This shortcoming could perhaps be ameliorated by extending this analysis to include derived forms in order to increase the length of these regions and to increase the number of cognate words which possess regions other than the before region.

Second, while the results of this study show clear positional and regional effects, these patterns cannot definitively be attributed exclusively to the cohort model's predicted recognition points or the particular way that this study operationalized the concept of 'recognition point'. A bifurcation—perhaps a distinction between not-after and after a recognition point—appears to be justified by this analysis, but the reality of the location of the particular recognition points

[27] This notion may be partly undermined, however, by languages that show the replacement of fossilized stem-internal inflectional morphology with productive, regular suffixation strategies, as in English *brethren* > *brothers*, etc.

calculated here is unclear: is this study's definition of 'between' or 'after' the best such definition? For example, the interaction of Position and Region in the after region bin in this study could be so seemingly varied because the after region includes sounds that could just as well fit into the between region bin, if the recognition point were defined slightly differently. One potential avenue for future work is a ground-up approach that infers the most-well-supported delimiters for points of inflection or discontinuity in sound change patterns inside words, which then could be compared with the boundaries predicted by theories like the cohort model (or alternative formulations of recognition points).

Additionally, the treatment of phonemes as discrete units, without consideration of featural relationships or prosodic information, may oversimplify the complex patterns of sound change. As noted by Marslen-Wilson (1984), phonetic and featural properties play important roles in cohort selection and lexical access: for example, consonant clusters delay the production of a subsequent vowel and therefore the transitional cues between consonants and vowels, and this phonetic effect has been found to delay the identification of a word (Marslen-Wilson & Tyler 1980). The work set out in this chapter could be improved by taking a sub-phonemic characterization of aligned slots in order to quantify the types, not just quantities, of sound change which may occur in any of the three region, or across the word generally.

The focus on Indo-European languages, while methodologically necessary due to data availability and shared characteristics, limits the generalizability of our findings. As discussed by Haspelmath (2011), even the concept of wordhood itself varies significantly across languages, and our reliance on dictionary-based word definitions may not capture the full complexity of word recognition in natural speech. The exclusion of tonal languages and highly synthetic languages also leaves open questions about how these patterns might manifest in typologically different linguistic systems.

Furthermore, while this study controlled for genetic relationships through random effects, the interaction between time depth and sound change (Figure 3.6) patterns merits deeper investigation. The visible separation in predicted probabilities between recent and old cognates suggests that the relationship between recognition points and sound change may not be uniform across different time scales. While this finding is not surprising, it has a greater effect than was assumed prior to model selection in this study. Additionally, it must be noted that this study is conducted on extant languages, using incongruencies in phonemes as a proxy for historical change. While, to some degree, this is a tautology—units that we know were once the same and are now no longer the same *must* have 'changed' at some point—it is also a simplification that obscures the extent of the historical journey of a sound change.

3.6 Conclusion

This study provides solid evidence of the existence of highly significant asymmetries in sound change within the word. First, it provides evidence that the end-of-word tendency for phonological neutralization (Wedel et al. 2019) extends to rates of sound change more generally. Additionally, the analysis here suggests that the incorporation of regions defined by a recognition point informed by the cohort model (here, there were three regions; but perhaps merely two, BEFORE and AFTER, are sufficient; see §3.4) reveals a pattern of sound change in the region after a recognition point that is not explained solely by a general end-of-word increase in probability of sound change. Such an asymmetry in intralexical sound change could be linked to differences in the deployment of cognitive resources to each region, and future research could examine whether there are, accordingly, differences in the *types* of sound change on either side of a recognition point.

More generally, the positive effect of the position of a phoneme within a word on the chance of sound change, for which this study finds strong evidence, itself illustrates the significance of cognitive and interactional pressures in sound change. Asymmetries of the type explored here, after all, are anathema to traditional, Neogrammarian assumptions about the regularity of sound change. This study does not investigate the extent to which the sound changes that occur in the surveyed cognates are regular or irregular, but future work in this vein could encode attested sound changes in the histories of these languages in order to determine if exceptions to irregular sound change cluster in the AFTER REGION or, at least, at the end of words. Such a finding would further support the notion that, regardless of whether a discrete recognition point can be satisfactorily operationalized in a study like this one, the cohort effect may constrain the mutability of the sounds most necessary for lexical recognition. Not all change is alike: random mutation may well be less destabilizing to language users than the systematic, *regular* change that historical linguistics has traditionally documented.

Chapter 4

Diversity and (in)stability in the relativization system of Icelandic

4.1 Introduction

[*] Modern Icelandic, like the other members of the North Germanic subgroup of the Indo-European language family, marks relative clauses (RCs) with a relativizer. Most frequent in Modern Icelandic is the relativizer *sem*, which is RC-initial and invariant in form regardless of the semantic or syntactic role of the noun in the embedded clause or in the matrix clause—for example, the same relativizer marks both object relativization in (1a) and subject relativization in (1b). Other invariant relative particles, such as *er*, are used in similar syntactic constructions and are described in §4.2.1[28].

[*] This chapter is based on earlier work I conducted and published in Brendel 2023. Many thanks to Eric W. Campbell, Viola Miglio, Marianne Mithun, Bernard Comrie, Hólmfríður Hannesdóttir, Khari Stinson & Sabrina Chaffino, and the two anonymous reviewers of the publication version of this project. Any errors which remain are my own.

[28] Square brackets in these examples indicate relative clause boundaries. Underlined text indicates the head of the noun phrase (NP) in the matrix clause which is coreferential with the relativized NP in the relative clause. Boldface text indicates the relativizer(s). The blank slots represent the gap in which the relativized NP would typically occur in a matrix clause; I mark these gaps when relevant to the discussion. I gloss the most important relativizers in this data with their lemmas since some relativizers feature a suppletive paradigm (e.g. the Latin relative pronoun *cuius* is glossed QUI.GEN.SG). Below, I often vary the level of glossing, choosing narrow glosses when the morphological characteristics of a given token are relevant to the discussion (e.g. most relative pronouns and demonstratives) and otherwise prioritizing broader, more readable glosses.

(1) Canonical *sem* relativization in Modern Icelandic (Þráinsson 2007: 407)

a. *þetta er* <u>*maður-inn*</u> ***[sem** María hitti __ í gær]*
this is man.NOM-DEF SEM Mary.NOM met yesterday
'This is the man [that Mary met yesterday]'

b. *þetta er* <u>*maður-inn*</u> *[__ **sem** hitti Maríu í gær]*
this is man.NOM-DEF SEM met Mary.ACC yesterday
'This is the man [that met Mary yesterday]'

In contrast to this typical pattern involving invariant relativizers, another, much rarer pattern—relativization marked by relative pronouns sourced from interrogative pronouns (*hv*-words, such as *hver* 'who', *hvað* 'what', etc.)—is attested in both modern grammars (Þráinsson 2007; Einarsson 1945) and the historical record (illustrated by an extensive corpus analysis of Wallenberg et al. 2011 throughout this paper). Unlike the invariant *sem*, the marking of such interrogative–relative pronouns (IRPs) varies according to the semantics of the head NP in the matrix clause (NP_S) and its relativized function (NP_{REL}) in the interrogative–relative clause (IRC). These relative pronouns fully share both the inflectional paradigm and phonological material of their corresponding interrogative pronouns—in other words, they show interrogative–relative polysemy. This inflectional exponence is seen in the dative case marking in (2a) in contrast to the accusative case marking in (2b)[29], corresponding to the function of the relativized NP in each example.

[29] Citations for examples of the format used in (2b) are passage identifiers taken from IcePaHC (Wallenberg et al. 2011). Unless otherwise noted, I have retrieved, glossed, and transcribed such examples myself.

(2) Interrogative (*hv*-word) relativization

a. Modern Icelandic (Þráinsson 2007: 407)

þetta	*er*	*<u>maður-inn</u>*	***[með***	***hverjum***	*María*	*fór*	*í gær]*
this	is	man-SG.NOM.DEF	with	HVER.M.SG.DAT	Mary	went	yesterday

'This is the man with whom Mary went yesterday'

b. Early Modern Icelandic (1790.FIMMBRAEDRA.NAR-SAG,.525)

kotungar	*þar*	*urðu*	*upphafsfinnarar*	*þess*	*<u>athæfis</u>*	***[hvað***
peasants	there	became	instigators	DEM.GEN	behavior.GEN(N)	HVER.N.SG.ACC

þá	*prestur*	*fornam]*
then	priest	noticed

'The peasants became instigators of the behavior that the priest then noticed'

Though the interrogative–relativization strategy is well-described in Icelandic grammatical treatments, interrogative–relativization is exceedingly rare in Modern Icelandic. It is generally regarded as archaic, poetic, and rather un-Icelandic, a pattern which exists in the language 'presumably because of foreign influence' (Þráinsson 2007: 407). As Proto-Germanic is not thought to have featured an interrogative–relative pronoun (Harbert 2006: 437–438; König & Auwera 1994: 27, 36), the development of the interrogative–relativization strategy must result from an independent innovation in Icelandic (or earlier in Old Norse) or through language contact. Interrogative–relative polysemy is cross-linguistically rare outside European languages and even within Indo-European languages is thought to spread primarily through contact with other Indo-European languages (Heine & Kuteva 2003: 541; see §4.4.2 for discussion). Of the three options above, we can rule out the first: although Old Norse featured several invariant relative particles which grammaticalized to varying extents in the Scandinavian languages (such as *sem* in Icelandic; see §4.2.1), it did not feature an IRP (Wagener 2017), and extant Scandinavian languages either lack true relative pronouns[30] in non-formal speech altogether or, particularly in Danish, derive them from demonstrative forms (Wagener 2017; Platzack 2002; Smits 1989), which suggests that the strategy is unlikely to originate from the closest siblings of Icelandic in the North Germanic family. Since the pattern is not inherited from Old Norse and is likely not borrowed from contact with other Scandinavian languages, the strategy must originate in one (or both) of the following ways:

- through an independent innovation in the attested history of Icelandic, or
- from another language which was in close contact with Icelandic when the strategy emerged.

In this paper, I review the ever-drifting relative clause formation strategies of Old–to–Early Modern Icelandic and examine both internal grammaticalization processes and contact effects on these systems by comparing the translation of the morphosyntatic structures of relative clauses among the languages involved in this contact scenario: Icelandic, German, and Latin. Using the Icelandic Parsed Historical Corpus (IcePaHC, Wallenberg et al. 2011), a resource spanning the entire written history of Icelandic from 1150 A.D. to the present, I conduct a diachronic corpus study that quantifies the development of this relativization strategy. In particular, I focus on a subset of two Biblical texts which had strong effects on the national standard of Icelandic (see §4.3.3). I argue that the close parity between the Latin model and the Icelandic translation, as attested through parallel Bible passages, shows that the proliferation of the Icelandic interrogative–relativization strategy is a contact phenomenon stimulated by a borrowing or replication of grammatical patterns of the sort described by, e.g., Gardani et al. (2015), Matras (2009), & Sakel (2007). In Icelandic relativization, this borrowing is made primarily from Latin, but was contextualized and made possible by a state of systemic variability and instability

[30] In this article, a 'true relative pronoun' (i) fulfills Comrie (2006)'s three criteria of a relative pronoun—it marks the RC-internal semantic or syntactic role of the head noun and is preposed to the beginning of its relative clause, which is head-external—and (ii) moreover, following Sapp (2019), it is specifically non-case-attracting—the relative pronoun must index the syntactic qualities of the head's relativized role within the RC (NP_{REL}) as opposed to reflecting the syntactic role of the antecedent (NP_{S}).

in the relative paradigm in Icelandic during the 16th-18th centuries.

In regards to the role of independent innovation, grammaticalization of interrogatives into some relative functions—a common pathway in European languages (Heine & Kuteva 2006)—had been underway since the 12th century in the most archaic Icelandic texts (§4.4). Additionally, another pronoun, the demonstrative *sá* (§4.4.1), has been argued to have grammaticalized into a full relative pronoun, like *hver*, before falling out of use (Sapp 2019). As I discuss in §4.2, *hver* emerged as a relativization strategy in an environment of myriad relative paradigms (including *sá*) present in Early Modern Icelandic, and I propose that the interaction between *hver* and *sá* influenced the grammaticalization and de-grammaticalization of each pattern.

As for the second hypothesis—external origin—the grammaticalization of *hver* (and even *sá* before it) is heavily influenced by contact. The use of interrogatives as markers of headed relative clauses is extremely rare in Icelandic prior to the 16th century (§4.2), late into the attested literary history. The most extensive usage of this strategy coincides with the beginning of the Protestant Reformation (16th century[31]), a time of intense scholastic activity in Iceland involving engagement with German and Latin sacred texts (§4.3.3). Both of these languages possess relative pronoun strategies and could serve as a model for the Icelandic usage. The examples in (3) show three translations of the same Bible passage. The Icelandic (3a) shows the usage of a *hver* pronoun in a relative sense. In both Latin and German, a relative pronoun strategy is used in these verses: Latin (3b) features a relativizer from an interrogative *qui*, cognate with *hver*, and German (3c) shows a relative pronoun grammaticalized from a demonstrative *der*.

[31] In Iceland, the effects of the Reformation surfaced in the late 1520s and in 1541 the country was nominally converted to Lutheranism by its colonial liege, the king of Denmark. However, in practice Iceland remained politically divided among Catholic and Lutheran factions until the execution of the last Catholic bishop in Iceland, Jón Arason, in 1551 (Kalinke 1996: 25).

(3) Acts 7:52

a. Gottskálksson (Icelandic, 1540)

… og þá í hel slegið sem fyrir fram boðuðu til komu
… and 3PL.ACC P death beat.3PL SEM P advance foretold P coming

*ins réttláta [**hvers** falsarar og morðingjar þér*
DEM.SG.GEN just HVER.M.GEN.SG betrayers and murderers 2PL.NOM

eruð nú orðnir]
be.2PL.PRS now become.PERF

b. Vulgate (Latin)

… et occiderunt eos qui praenuntiabunt de adventu
… and killed.3PL 3PL.ACC QUI.NOM.PL foretold.3PL about arrival.ABL

*iusti [**cuius** vos nunc proditores et homicidae fuistis]*
justice.GEN, QUI.GEN.SG 2PL.NOM now betrayers and murderers be.PERF.2PL

c. Luther (German, 1522)

…und sie haben getötet, die da zuvor verkündigten die
…and they PERF killed DER.PL.ACC there before foretold DET.ACC

*Zukunft dieses Gerechten, [**dessen** Verräter und Mörder ihr*
future(ACC) DET.GEN just DER.GEN.SG betrayers and murderers 2PL.NOM

nun geworden seid]
now become.PERF be.2PL

'And they have killed those who foretold the coming of the Just [One], of whom you have now become traitors and murderers'

Although IRPs never become the dominant strategy in Icelandic[32], the use of Icelandic *hver* in these passages is correlated with Latin and German structures (§4.3.4). More notably, although the forms of the root and inflectional marking of each pronoun is native to its respective language, the resemblance of the syntactic and morphological patterns across these verses is remarkable. All three texts show similar syntactic structures for the relativizer: each encodes the pronoun in the genitive case. Additionally, Icelandic (3a) and Latin (3b) each feature a relative pronoun polysemous with an interrogative pronoun. The close similarity in these constructions suggests that the strategy was transmitted through the borrowing not of forms or phonological material, but rather of the grammatical patterns (i.e., calqued) (Gardani et al. 2015; Sakel 2007) or, more specifically, the import of a stage of grammaticalization that a related language had already undergone (Mithun 2013; Heine & Kuteva 2003) (§4.3.1).

Before advancing to these claims, I first review the predominant relativization strategies which have been used in Icelandic, discuss the functional and typological differences between *hver* and canonical relativizers, and chart the diachronic frequency in usage of these strategies (§4.2). In §4.3, I describe the grammaticalization of the *hver* relative pronoun through the lens of contact and grammatical pattern borrowing, quantified by the structural similarity between Icelandic Bible passages and German and Latin models. In this section, I also review the cultural context of this contact, examining Bible translation in Iceland during the Reformation. In §4.4, I analyze the language-internal factors which shaped the grammaticalization of *hver*, including the evolution of the relative pronoun *sá* as a predecessor to the widespread usage of *hver* as a relativizer. Following an account of the grammaticalization of *hver* through the stages of the interrogative–relative pathway (Heine & Kuteva 2006), I propose that the acceptance of *hver* as a relative pronoun is facilitated through the instability of the Icelandic relativization system, concluding in §4.5.

[32] For example, in (3a), two occurrences of *qui* in the same sentence are variously translated: first with the canonical relative particle *sem* and next the interrogative *hver*.

4.2 Icelandic relativization strategies

Over the literary period of Icelandic, several relativizers have come into and fallen out of favor. Consequently, before addressing the roles of contact and innovation on the grammaticalization of the *hver* IRP, I will first briefly orient the reader to the most common relativizers and the typological characteristics of the most attested relativization strategies. Most of these relativizers will then be discussed in greater detail below.

4.2.1 The relativizers

Canonical relativizers

In Modern Icelandic, relative clauses use the gap strategy, are head-external, and are obligatorily marked with an invariant preposed relativizer: *er*, *sem*, or *þar*. These relativizers are generally regarded by Icelandic scholars as relative particles or complementizers rather than relative pronouns (Þráinsson 2007: 5), primarily because they do not display the morphological case system of Icelandic pronouns (Jónsson 2017; Þráinsson 1980). From a typological standpoint, since these relativizers do not index the semantic or syntactic qualities of the head's role in the RC, they do not fulfill a key criterion of Comrie (2006)'s classification of relative pronouns[33].

The most archaic of these strategies, relativization with *er* (4a), is the most common pattern until the 15th century. Of the three non-controversial relative particles, *er* is the most frequent during the period of the highest usage of *hver*, and is discussed at length in §4.4.2.

The primary relativizer in Modern Icelandic, *sem* (4b), grammaticalized from a comparative marker 'like' and spread across Scandinavian languages to become the default complementizer in all modern varieties (Wagener 2017: 263).

A third marker, *þar* (4c), was semantically constrained to relativization of location in West

[33] See fn. 30.

Nordic languages. It is scarcely-attested in IcePaHC, remaining restricted to locative domains, but grammaticalized to some degree into a general relativizer in East Nordic, surviving as a marked relativizer of subjects, *der*, in Modern Danish (Wagener 2017: 270).

(4) a. *er* (1150.HOMILIUBOK.REL-SER,.3)

*menn þeir [**er** bera kennimanna-nöfn og vígslur yfir ólærðum mönnum]*

men SÁ.M.PL.NOM C bear priest-name and ordination over uneducated men

'Those men that have priestly titles and preside over uneducated men'

b. *sem* (1540.NTJOHN.REL-BIB,210.905)[34]

*Lasarus [**sem** sjúkur lá]*

Lazarus SEM sick lay

'Lazarus, who was sick'

c. *þar* (1540.NTACTS.REL-BIB,237.40)

*fyllti upp allt hús-ið [**þar** þeir sátu í]*

filled up whole house-ACC.DEF where they sat at

'[They] filled up the whole house that they sat in'

Demonstrative pronoun *sá*

I will very briefly introduce here the demonstrative *sá*, an inflectionally-variant pronoun which I discuss at length in §4.4.1 as its evolution in relative function is intertwined with that

[34] This RC, unlike the majority of the RCs analyzed in this paper, seems to be non-restrictive: it does not delimit one of many possible Lazaruses, but rather adds additional information about a known entity. Wagener (2017: 24) contends that there is 'no formal way of distinguishing between restrictive and non-restrictive RCs in […] Early Nordic', and Þráinsson (2007: 407) indicates that *sem* can be used with both non-restrictive and restrictive RCs.

of *hver*. *Sá*, which generally agrees in gender, number, and case with the head NP in the matrix clause (rather than its role in the RC), often co-occurs with the canonical relativizer *er* in early texts and sometimes lacks strong demonstrative function (5), leading to divergent views that it is purely a demonstrative pronoun (Wagener 2017; Faarlund 2004) or that it functions as relative pronoun (Sapp 2019; Åfarli 1995). As *sá* indexes the syntactic features of NP_{S}, I represent it here as external to the RC inside of the head NP in the matrix clause.

(5) (1150.FIRSTGRAMMAR.SCI-LIN,.5, from Sapp 2019)

hver	*tunga*	*hefir*	<u>*hljóð*</u>	*þau*	*[**er***	*eigi*	*finna-st*	*í*	*annarri]*
each	tongue	has	sound(N).PL.NOM	SÁ.N.PL.NOM	C	not	find-PASS	in	another

‘Every language has (*those) sounds that are not found in others’

Interrogative pronouns

Icelandic features several interrogative pronouns, and those most commonly seen as IRPs (interrogative–relative pronouns) are listed in Table 4.1. I variously refer to these as IRPs or *hv*-words (on analogy with the use of the term *wh*-word in English, as most of these pronouns possess an English cognate).

Table 4.1: Most attested *hv*- words as relativizers

Pronoun	gloss	occurrences in IcePaHC
hver	‘who’/‘what’	585
hvar	‘where’	92
hvor/hvörri	‘which (of two)’	15
hvaðan	‘whence’	5
hvað	‘what’	5
others		9

In Modern Icelandic, these pronouns are uncontroversially used in word questions (6a) and indirect questions (6b), which lack a head in the matrix clause. Their ability to mark true headed

RCs, which was productive in literature in the 16th–18th centuries (§4.4), is now perceived as an archaism (Þráinsson 2007).

(6) a. Word question (Þráinsson 2007: 157)

__hver__ er þetta?

hver.sg.nom is there

'Who is there?'

b. Indirect question (headless RC) (Þráinsson 2007: 408)

ég veit ekki bróðir __hver-s__ hann er

I know not brother hver-m.sg.gen he is

'I don't know whose brother he is'

c. Headed relative clause [=2a]

þetta er <u>maður-inn</u> __[með hverjum__ María fór í gær]

this is man-sg.nom.def with hver.m.sg.dat Mary went yesterday

'This is the man with whom Mary went yesterday'

As the overwhelmingly most common example of an interrogative pronoun used in a relative context, *hver* is the primary focus of my discussion in this paper.

4.2.2 Typological features

Pied-piping

Relativizers in Icelandic, with the exception of the interrogative–relative pronouns, obligatorily strand their prepositions. In (7a), *með* 'with' remains in the slot that it would normally occupy if NP_{REL} were realized in the typical order of an independent clause (cf. *María fór* ***með hann*** *í gær* 'Mary went **with him** yesterday'). The fronting of *með* in (7b), by contrast, is ungrammatical with a canonical relativizer such as *sem* or *er*.

(7) Pied-piping in Modern Icelandic (Þráinsson 2007)

a. *þetta er maður-inn [**sem** María fór **með** í gær]*
this is man-SG.NOM.DEF SEM Mary went with yesterday
'This is the man that Mary went with yesterday.'

b. **þetta er maður-inn [**með sem** María fór í gær]*
this is man-SG.NOM.DEF with SEM Mary went yesterday

The interrogative–relative pronouns, however, pied-pipe these elements, preposing them at the beginning of the clause (8).[35]

(8) *þetta er maður-inn [**með hverjum** María fór í gær]*
this is man.NOM-DEF with HVER.M.SG.DAT Mary went yesterday
'This is the man with whom Mary went yesterday' [=2a]

Accessibility

The canonical relativizers in Modern Icelandic—*sem* and much less frequently *er*—are capable of relativization of every function outlined by Keenan & Comrie (1977) in their NP accessibility hierarchy except for relativization on possessors:

(9) **þetta er maður-inn [systur **sem** ég hitti í gær]*
this is man.NOM-DEF [sister SEM I met yesterday]
'This is the man whose sister I met yesterday' (Þráinsson 2007)

Some sources claim that Old West Nordic was able to relativize on all nominal functions, including possession (Wagener 2017: 55; Faarlund 2004: 259). However, the only such example

[35] This situation can be compared with English. Consider *The man that I went with* and **The man with that I went*, the latter of which is ungrammatical with the complementizer *that*, but is acceptably pied-piped when a *wh*-word is used as a relativizer, as in *The man with whom I went*.

that Faarlund (2004: 260) provides (10) is ambiguous due to a seeming conflation of genitive case-marking with semantic possession[36]. The verb *þurftu* 'needed' may take its object ($\textsc{np}_{\textsc{rel}}$) in the genitive case[37], but (10) does not show a semantic relationship of possession. It is uncertain, then, if possessor NPs were ever accessible to canonical relativization over the history of Icelandic.

(10) Genitive case marking without relativization on possessor (Faarlund 2004: 260)[38]

þeir *hǫfðu* *þat* *með* *sér* *[**sem** þeir þurftu*
they.M had.3PL that.ACC with themselves.DAT SEM they.M needed.3PL(>GEN?)

__]
($\textsc{np}_{\textsc{rel}}$)

'They had with them what they needed'

By contrast, the IRP is historically able to relativize on possessors unambiguously, as the examples in (11) show. Typically, these possessors are supported with the use of the complementizer *að*, as in (11a) and (11b), though a few examples are seen without this additional complementizer (11c).

[36] Many Icelandic prepositions and verbs are lexically-specified to take arguments of a certain case with no synchronically-recognizable function prototypically associated with that case (e.g. a genitive noun does not always signify some relationship of possession, as in the Modern Icelandic *Ég fór til Ísland-s* 'I went to Iceland-GEN', where the preposition *til* 'to' requires a genitive object). If prepositions and verbs can take objects in multiple cases, there is often a change in meaning associated with the different case.

[37] Note that Cleasby & Vigfússon (1957: 749–750) offer examples in which *þurfta* 'to need' takes either accusative or genitive objects. I am unsure why Faarlund (2004) treats the instance in (10) as requiring an object in the genitive rather than the accusative.

[38] The gap and parenthetical glosses are mine.

(11) Relativization on possessor with *hver*

a. Acts 10:6 (1540.NTACTS.REL-BIB,257.613)

*hann herbergjar hjá Símoni sútara einum [**hvers** hús að*
he lodges with Simon tanner one HVER.SG.GEN house(NOM) C

leggur við sjó-inn]
lies by sea.ACC-DEF

'He lives with a certain tanner Simon whose house is by the sea'

b. John 4:16 (1540.NTJOHN.REL-BIB,192.295)

*Og þar var þá einn smákonungur [**hvers** sonur að lá sjúkur til*
And there was then a nobleman HVER.GEN.SG son C lay sick at

Kaparnaum]
Capernaum

'And there was a nobleman whose son was sick at Capernaum'

c. Acts 13:25 (1540.NTACTS.REL-BIB,265.847)

*heldur sjáið að sá kemur eftir mig [**hvers** skóklæði eg em eigi*
but see C he comes after me HVER.GEN.SG shoes I am not

verðugur að leysa]
worthy to loosen

'But see that he comes after me, whose shoes I am not fit to loosen'

4.2.3 Frequency of strategies

Figure 4.1 summarizes the frequency of the relativizers described in §4.2.1 as a percentage of overall marking of headed relative clauses for each century from the 12th century to the 21st. The data used to create the graph is also presented in Table 4.2, which provides both raw counts

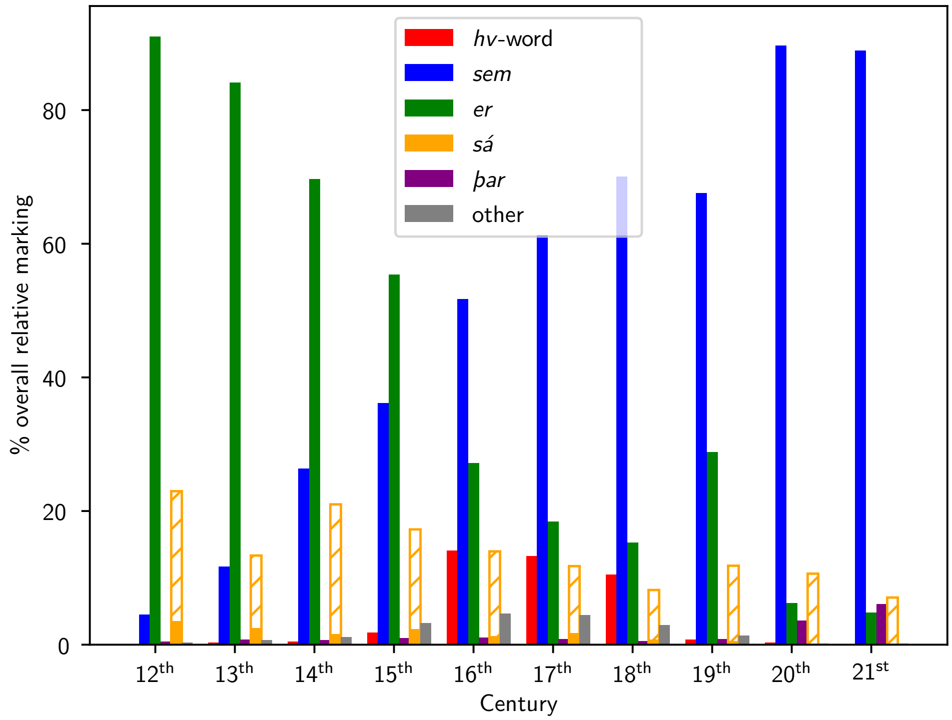


Figure 4.1: Markers of headed relative clauses by frequency of all relativization per century

(Table 4.2a) and frequencies of overall relative clause occurrence (Table 4.2b) by century and across the corpus as a whole. Table 4.2 and Figure 4.1 are based on my parsing of the corpus (see §4.2.3). The invariant relativizers *er* and *sem* dominate RC marking across the corpus; as discussed below, these relative particles are used in similar syntactic environments, suggesting that the predominant pattern used to form RCs in Icelandic has remained stable through its history; though the relativizer *er* has been slowly replaced by *sem*, it occupies the same slot in the same construction.

In Figure 4.1, the special case of *sá* occurring in an NP which contains a RC is represented by the orange bar and dashed orange box. The solid color represents unambiguous cases of relative *sá* in which *sá* occurs immediately before a RC and after—but not *immediately* after—its head noun, suggesting a non-demonstrative usage per Sapp (2019). The hashed bar represents *sá* occurring in other positions in the NP which contains the RC, including pre-nominally in the typical demonstrative position and immediately adjacent to both head noun and RC. The ambiguity between the demonstrative and relative functions of *sá* is discussed in §4.4.1.

This graph confirms that IRPs in headed IRCs was scantly attested before the 16th century, at which time it becomes much more common. This drastic increase occurs during and after the Reformation, the time in which I contend the strategy was reinforced through translation of Latin and German texts that feature these structures. I discuss this origin in §4.3.

This increase in frequency of IRPs also coincides with a decrease in the usage of *sá* in relative contexts, and comes also at a time when the previously dominant pattern of *er* was being replaced by *sem*. Indeed, the 16th century is the period in which the frequencies of each RC strategy are the most similar to each other; *sem* is clearly the most frequent strategy, but in this century the gap between the most predominant strategy and the less popular strategies is the smallest in the entire corpus (compare with the 21st century, which shows *sem* as the most prevalent strategy by the widest margin in the entire corpus).

The wealth of relativization strategies and their commonness during the period between the 16th and 18th centuries suggests that this period was a time of great instability in the relativization system of Icelandic. The longstanding *er* was falling out of use to be replaced by *sem*, and moreover, the grammaticalization of two pronouns which were inflectionally variant, a very different pattern from *er* and *sem*, was underway. I discuss the language-internal contexts and motivation for change in §4.4, revisiting the idea of systemic instability in the conclusion §4.5.

[39] One anonymous reviewer noted the unusual recent increase in the frequency of *þar* 'there' as a relativizer, a strategy which has never been frequent in Icelandic (see §4.2.1). I do not take up this issue in depth, but I find that all the 20th and 21st century *þar* instances indicated here are part of the collocation *þar sem*, which generally corresponds with English 'where' (Þráinsson 1994: 361), e.g. *…í fjöru* ***þar sem*** *hann æpir upp …* '…at the shore **where** he called out …' (2008.MAMMA.NAR-FIC,.1445). These occurrences are captured here because *þar* is the initial lexeme in the relative clause as coded in IcePaHC, even though *sem* performs the relative function. This data does not suggest that *þar* is becoming a domain-general relativizer in Modern Icelandic.

Table 4.2: Relative clause occurrences by relativizer per century (visualized in Figure 4.1)

(a) raw count

	12th	13th	14th	15th	16th	17th	18th	19th	20th	21st	total
sem	47	173	522	476	822	1220	1138	818	1139	441	*6796*
er	943	1248	1380	729	432	367	248	349	79	24	*5799*
sá (ambig.)	238	198	416	228	223	234	134	143	135	35	*1984*
sá (unambig.)	37	37	32	31	21	34	12	7	0	0	*211*
hv-word	1	5	10	24	224	264	170	9	4	0	*711*
other	3	10	23	43	74	88	48	17	2	1	*309*
þar	5	11	14	13	17	17	9	10	46[39]	30	*172*
total	*1274*	*1682*	*2397*	*1544*	*1813*	*2224*	*1759*	*1353*	*1405*	*531*	*15982*

(b) percent frequency of all relative clauses in period

	12th	13th	14th	15th	16th	17th	18th	19th	20th	21st	over all texts
sem	(%) 3.7	10.3	21.8	30.8	45.3	54.9	64.7	60.5	81.1	83.1	*42.5*
er	74.0	74.2	57.6	47.2	23.8	16.5	14.1	25.8	5.6	4.5	*36.3*
sá (ambig.)	18.7	11.8	17.4	14.8	12.3	10.5	7.6	10.6	9.6	6.6	*12.4*
sá (unambig.)	2.9	2.2	1.3	2.0	1.2	1.5	0.7	0.5	0.0	0.0	*1.3*
hv-word	0.1	0.3	0.4	1.6	12.4	11.9	9.7	0.7	0.3	0.0	*4.4*
other	0.2	0.6	1.0	2.8	4.1	4.0	2.7	1.3	0.1	0.2	*1.9*
þar	0.4	0.7	0.6	0.8	0.9	0.8	0.5	0.7	3.3	5.6	*1.1*

Corpus processing

I pause here to explain the methods I used to arrive at the above data. The Icelandic Parsed Historical Corpus (IcePaHC, Wallenberg et al. 2011) provides the Icelandic data I examine in this paper. The resource comprises 61 syntactically-parsed and part-of-speech tagged texts dating from the 12th to 21st centuries. The parsed syntax trees in IcePaHC are bracket-delimited plain text, generally the schema of the Penn Historical Corpora (Marcus et al. 1993), and are mostly compatible with the framework of the Natural Language Toolkit (Bird et al. 2009) treebank processing library for Python. Over the course of this project, I built a framework on top of these generic functions to enable efficient parallelized processing of syntactic structure, supporting more accurate tabulation of patterns and analysis of subordination as well as facilitating the manual inspection of the data through a human-readable visualization of syntactic structure. The development of a tool which is aware, to some degree, of the particular syntactic structure of Icelandic also allowed for flagging of inconsistencies in the tagged IcePaHC data, which I manually corrected when necessary.

The status of the non-canonical relativizers examined in this paper is a non-trivial issue in automated parsing (or manual assay). At times, the syntactic parse is insufficient: IcePaHC labels relative clauses, but does not explicitly mark the connective element as a relativizer, nor does it index shared semantic or syntactic features with the referent. It does, however, distinguish between interrogative forms used in word questions (e.g. ***Hver*** *fór?* 'Who went?') from those used in any subordinate clause (e.g. the indirect question *Ég veit* ***hver*** *fór* 'I know who went'), though its role within the subordinate clause is not directly coded. Additionally, the demonstrative pronoun *sá* which has arguably relative function (Sapp 2019; see §4.4.1) is not coded as a connective. Identifying relative pronouns that function as connectives in relative clauses thus requires additional classification of clauses coded as relative in the corpus.

To address these issues, I systematically identified the first word in each relative clause. If this word was a preposition, the clause was next analyzed for pied-piping (in which case a preposition precedes a relativizer, as in <u>*Maður*</u> *[**með hverjum** ég fór]* '[A] <u>man</u> **with whom** I went'). In order to identify such constructions as *á meðal* ***hverra*** 'in the middle **of which**', chains of prepositions in these potential pied-piped constructions were parsed word-by-word until an interrogative pronoun was found. These instances were counted as an occurrence of the relative pronoun which pied-pipes them.

For the identification of *sá*-demonstratives in relative contexts, I selected all morphological variants of *sá* which occurred in an NP which contained a RC. I then divided these instances into those which occur as an unambiguous relativizer per Sapp (2019): those occurring immediately before a RC but *not* immediately after the head noun in its NP.

I performed this analysis for each year represented in the corpus, calculating the number of clauses connected by a particular relativizer, with inflectional variants for *hv*-words and *sá* grouped under their respective lemmas, yielding the output presented in Figure 4.1.

4.3 The role of contact

As König & Auwera (1994: 27, 36) note, the diversity of relative markers in extant Germanic languages suggests that Proto-Germanic had no single relative marker. Proto-Germanic is not thought to have used interrogative pronouns as relativizers, instead employing relativization strategies such as nominalization in the form of participial relative constructions (Harbert 2006: 421) or the use of demonstrative pronouns (Harbert 2006: 436). König & Auwera (1994: 68) note that Old Norse, the most direct ancestor to Icelandic, lacks relative pronouns, instead using *er*, and that *sem*—which stems from a comparative particle similar in function to English *as* (Harbert 2006: 426)—replaces *er* by the 14th century. Other Scandinavian languages primarily use a cognate of *sem* for relativization[40]. This wealth of evidence suggests that Icelandic could only have acquired relative pronouns through contact or internal innovation[41], not through inheritance.

This section details the linguistic and cultural circumstances surrounding a contact explanation for the use of *hver* as an IRP and presents an analysis of this scenario. After reviewing the borrowing of grammatical patterns (§4.3.1), I overview the typological proximity and functional similarities in the patterns and languages involved in the Icelandic interrogative–relativization

[40] With the notable exception of Danish, which today uses a cognate of *þar* ‘there’, *der*, as a relativizer on subjects (Wagener 2017: 270)

[41] As one anonymous reviewer suggested, it is also possible that the frequency of *sem*-relativization in other Scandinavian languages, such as Danish and Norwegian, could have influenced the increasing usage of *sem* in Icelandic. Indeed, *er* seems to have almost entirely fallen out of usage in East Nordic by the 15th century, (Wagener 2017: 263), while it is still attested as the predominant relativizer in Icelandic texts until the 16th century (see §4.2.3). Such intra-family contact effects are likely to be at play, but family-internal borrowing alone does not explain why Icelandic did not acquire other idiosyncrasies of relativization in its contemporaries. As one example, *þar* relativization, not *sem*-relativization, was the most common strategy in 13th century East Nordic (Wagener 2017: 261), but Old Icelandic shows no corresponding increase in *þar*-relativization.

scenario (§4.3.2) and the social environment in which these languages, their texts, and their translator existed (§4.3.3). After this discussion, I present a comparative analysis of Bible passages from Icelandic, Latin, and German (§4.3.4).

4.3.1 Grammatical pattern borrowing

Traditionally, the borrowing of lexical material—loanwords—was the primary locus of language contact studies at the expense of morphological change, a situation exacerbated by the 'apparent relative infrequency of morphological borrowing (Gardani et al. 2015: 1). More recently, a distinction has been made between the borrowing of phonological material (MAT-borrowing) and the borrowing of the patterns of a language (PAT-borrowing) (Gardani et al. 2015; Sakel 2007). PAT-borrowing obligatorily involves a process of grammaticalization, which frequently produces different patterns in the recipient language compared to the donor language (Sakel 2007: 17).

The Icelandic patterns detailed in this paper feature neither borrowed phonological material nor borrowed inflectional morphemes; the inflectional paradigm of the *hver* interrogative pronoun, identical to the IRP paradigm, dates to Old Norse (cf. Wagener 2017; Faarlund 2004), centuries before the usage of *hver* to mark headed RCs was regularly attested. Instead, the change observed is in function and in stage of grammaticalization. Mithun (2013: 262–267) suggests that in addition to the types of MAT and PAT borrowing outlined above, contact effects among related languages can also include the transfer of grammatical function and of stages of grammaticalization. Heine & Kuteva (2003: 539), too, propose 'replica grammaticalization', the transference of processes of grammaticalization, rather than the borrowing of foreign patterns of derivation or inflection to be filled in by native phonological material. In this view, changes in function of some shared structure between genetically-related languages can incubate in one variety and be transmitted to the other. Such transference could serve as an accelerant for already-ongoing process of grammaticalization. For 16th century Icelandic, the progress down the pathway of interrogative–relative grammaticalization (outlined in §4.4.2) reveals a language which had already reached an advanced stage of interrogative–relative polysemy, teetering on the cusp of the final stage (headed IRCs). Contact with two related languages which had already reached the final stage—Latin and German—pushed Icelandic to the end, as well.

Thomason (2015, 2014) asserts that the likelihood of a morphological structure being transmitted is correlated with the typological similarity between the two structures, arguing that contact-induced morphological changes 'occur much more frequently at typologically congruent structure points than at typologically disparate structure points' (Thomason 2015: 29). In the Icelandic contact scenario, typological similarity in paradigm seems to facilitate functional change in the use of that paradigm, as well. The declension systems of the interrogative pronouns in Latin and Icelandic comprise one such typologically-similar juncture whose morphological exponents are realized across similar dimensions: each distinguishes three nominal genders (masculine, feminine, and neuter), number (singular and plural), and case (nominative, genitive, dative, and accusative, while Latin additionally has ablative case). The full paradigms of each of the two pronouns used to mark relative clauses in the texts analyzed in the present study—*qui* in Latin and *hver* in Icelandic—are listed in Table 4.3 and Table 4.4.

Table 4.3: Icelandic *hver* interrogative

	MASC	FEM	NEUT
	singular		
NOM	*hver*[1]	*hver*	*hvert*[2]
ACC	*hvern*	*hverja*	*hvert*
DAT	*hverjum*	*hverri*	*hverju*
GEN	*hvers*	*hverrar*	*hvers*
	plural		
NOM	*hverir*	*hverjar*	*hver*
ACC	*hverja*	*hverjar*	*hver*
DAT	*hverjum*		
GEN	*hverra*		

[1] *hverr* in older Icelandic; standardized to *hver* in IcePaHC

[2] Modern Icelandic licenses a suppletive paradigm for the neuter nominative and accusative in which *hvað* 'what' may instead be used

(Faarlund 2004; Einarsson 1945)

Table 4.4: Latin *qui* relative pronoun

	MASC	FEM	NEUT
	singular		
NOM	*qui*	*quae*	*quod*
ACC	*quem*	*quam*	*quod*
DAT	*cui*		
GEN	*cuius*		
ABL	*quo*	*qua*	*quo*
	plural		
NOM	*qui*	*quae*	
ACC	*quos*	*quas*	*quae*
DAT	*quibus*		
GEN	*quorum*	*quarum*	*quorum*
ABL	*quibus*		

(Pompei 2011)

4.3.2 Relativization in potential donor languages

I now briefly describe the relativization strategies in Latin and German, both of which could serve as potential donors for the development of IRCs in Icelandic, and the similarity of these strategies to *hver*.

Latin

In Latin, aside from the usage of participles with meanings sometimes interpreted as relative in translation[42], the only relativizing particle is an inflectionally-polymorphic relative pronoun, *qui*, which is also derived from an interrogative form (Pompei 2011). This paradigmatic similarity and history of contact between Icelandic and Latin suggests that Latin is the most likely donor for the interrogative–relative strategy in Icelandic.

(12) Structural parallelism in two analogous Bible passages (John 17:11)

a. Vulgate (Latin)

*pater sanct-e, serva eos in nomine tuo, [**quos** dedi-sti*
father holy-VOC save.IMP them in name your QUI.PL.ACC give.PERF-2SG

mihi]: ut sint unum, si-cut et nos
1SG.DAT COMP be.SBJV.3PL one just-as like us

'Holy father, keep in your name those whom you have given me that they may be one, as we are.'

b. Gottskálksson (Icelandic)

*heilagur faðir geym þá í þínu nafni [**hver-ja** þu gaft mér]*
holy father keep.IMP them in your name HVER-ACC.PL you give I.DAT]

að þeir sé eitt svo sem við
COMP they be.SBJV one so COMP we

'Holy father, keep in your name those whom you have given me that they may be one, as we are.'

[42] e.g. in sentences such as Latin *homō [**correns** ab lupo]*, 'The man [**running** from the wolf]', the participle phrase restricts the domain of the head NP in a manner functionally-similar (but syntactically-dissimilar) to a relative clause (cf. 'The man that runs from the wolf'). Some participle phrases with this restrictive function are translated as syntactic RCs in the texts of the Icelandic corpus.

In (12a) and (12b), two homologous passages are shown—the same Bible passage with a Latin transferendum and a Icelandic translatum. As the glossing illustrates, the Latin relative pronoun is translated with the Icelandic interrogative pronoun rather than a canonical relativizer like *sem* (which would be grammatical in this instance). Moreover, *hverja* and *quos* are both marked as accusative plurals in their respective languages, though the phonological realizations of these morphemes are unconnected. This perfect parity in morphosyntactic structure suggests the existence of a constriction for faithfulness to the Latin original in the Icelandic text: the Latin relative clause structure—in particular, the inflection of the interrogative–relative pronoun—is realized not through direct borrowing of structural material (matter such as phones), but rather through the stimulation of the native morphological processes of the Icelandic interrogative pronoun *hver*, which captures the number and case of the Latin original.

German

While Latin has one relativizer used to form relative clauses, German has historically deployed several strategies, many of which are present in the Lutherbibel.

The variant relative pronoun, *der*, is derived from a demonstrative pronoun (13a & 13b). This strategy—though not the relativizer itself—was developed presumably under the influence of Latin during the Old High German period circa 8th–12th century (Wal & Quak 1994: 106). This relative pronoun usage is attested in Martin Luther's German Bible, which was available to Icelandic scholars during this period (see §4.3.3).[43]

[43] Additionally, German speakers and Icelandic speakers were in direct contact: German-speaking merchants had established permanent houses in Hafnarfjördur, Iceland by the mid-16th century (Karlsson 2000: 95). It is uncertain if this literary relative clause strategy was yet used in vernacular German at the time, or if it was ever used in spoken Icelandic, so I have not investigated this explanation.

(13) Modern German demonstrative–relativizer

a. Nominative function

da	*ist*	*der*	*Mann,*	*[**den***	*ich*	*kenn-e]*
there	is	DET.M.SG.NOM	man(M)	DER.M.SG.ACC	1SG.NOM	know-1SG.PRS

'There is the man who I know.'

b. Accusative function

da	*ist*	*der*	*Mann,*	*[**der***	*mich*	*kenn-t]*
there	is	DET.M.SG.NOM	man(M)	DER.M.SG.NOM	1SG.ACC	know-3SG.PRS

'There is the man who knows me.'

In addition to this strategy, a more archaic pattern based on the interrogative pronoun *welcher* 'which' is attested in the Lutherbibel. This strategy is regarded to be the result of Latin influence itself (Dal & Eroms 2014: 23).

(14) (Formal) German interrogative–relativizer (Smits 1989)

a. Dative function

es gab	*kein*	*Ereignis,*	*[**an***	***welchem***	*der*	*Kaiser*	*nicht*	*persönlich*
EXPL.PST	no	event	in	which.DAT	DET	emperor	not	personally

teilgenommen	*hätte]*
taken.part	had

'There was no event in which the emperor had not personally taken part.'

b. Accusative function

er	*sagte*	*"Guten*	*Tag",*	*[**welchen***	*Gruß*	*sie*	*freundlich*	*erwiderte]*
he	said	"good	day"	which.ACC	greeting	she	amicably	returned

'He said "good day", which greeting she friendlily [sic] returned.'

Additional relativizers show up less frequently in the Lutherbibel, including the interrogative pronoun *was* 'what' as markers of headless RCs and *da* 'there' as a semantically-constrained locational relativizer, similar to the archaic Scandinavian usage of the cognate *þar*. Both of these relativizers are invariant, although *da* can take a postposed pied-piped element.

(15) *da* relativizer (Acts 1:25, Lutherbibel)

...	*apostelamt,*	*[**da=von***	*Judas*	*abgewichen*	*ist]*
...	apostleship	there=from	Judas	fallen	is

'[He may take this] apostleship, from which Judas strayed.'

The *der* and *welcher* strategies commonly appear in homologous Bible passages which use *hver* in Icelandic (§4.3.4). Other strategies such as *da* and *was*–relativization appear in homologous passages as well, but much less frequently.

4.3.3 Cultural context of contact

Icelandic is a language which has traditionally been isolated from the degree of contact which influenced its continental siblings, leading to its generally-regarded slow rate of historical change in several dimensions. For example, most Old Icelandic vocabulary survives into Modern Icelandic (Kvaran 1996: 46–47), and formal syntactic change across this period is minimal, informally estimated to be comparable with change undergone between Early Modern English and Modern English (Barðdal & Eyþórsson 2003: 464). When diffusion does occur, nativization strategies such as calquing subverbal elements of compound words or significant phonological assimilation—or, on the societal dimension, prescriptive revisions to perceived outside linguistic influences—are preferred to the transparent borrowing of material (Þráinsson 1994: 188). This long-established inertia stems from the emergence of Icelandic as national language distinct from the notion of a monolithic pan-Nordic *dönsk tunga* ('Danish tongue') that prevailed across the Scandinavian world until the 12–13th centuries (Leonard 2012), even though mutual intelligiblity among Scandinavian dialects (the precursors of modern Icelandic, Faroese, Norwegian, Swedish, and Danish) had likely ceased by the 11th century (Haugen & Markey 1972).

Although Icelandic shows scant evidence of grammatical change on the surface, shift in the usage of existing structures is attested. For instance, the functions of the four nominal cases have undergone drift despite the lack of formal change in the nominal inflectional paradigm (Moscoso del Prado Martín & Brendel 2016; Barðdal 2011). Contact has not been implicated as catalyst for these particular developments, but such areas of covert change could be particularly susceptible to contact stimulus: the effects of contact can manifest through alteration of the function and domain of existing structures rather than through more overt formal/material changes to which languages such as Icelandic have demonstrated greater resistance. I show in this paper that this scenario describes the development of the interrogative–relativizer in Icelandic.

Texts

The claim that the interrogative–relative pronoun is stimulated by contact is supported by an analysis of the New Testament books of John and Acts, which are included in IcePaHC from the first printed New Testament Bible translation made during the 16th century, *Nýja testamenti Odds Gottskálkssonar*. The translator Oddur Gottskálksson, a statesman and early proponent of Protestantism, had begun the Icelandic translation somewhat clandestinely in thenpredominantly-Catholic Iceland[44]. The edition was later printed during the subsequent period of forced religious conversion to Protestantism mandated throughout the Kingdom of Denmark (which included Iceland) by Christian III. This New Testament is the oldest surviving printed literature in Icelandic, but was not published until 1540, in Denmark. This work later served as the basis for the New Testament in the first Icelandic translation of both the Old and the New Testament (the *Guðbrandsbiblía*, compiled by Guðbrandur Þorláksson) in 1584 (Eggertsdóttir 2006).

Gottskálksson's New Testament, along with the *Guðbrandsbiblía*, is a significant anchor for the Icelandic literary tradition and the legitimacy of Icelandic language as distinct from other Scandinavian tongues. With the publication of Gottskálksson's New Testament came the recognition within the Danish church of the independence of Icelandic from Danish, a particularly important political development that would cement the identity of Iceland during its status as a possession of the Danish crown until the 20th century. Additionally, the lack of purported influence from foreign languages on these seminal translations (particularly *Guðbrandsbiblía*, perceived to be even more purist than Gottskálksson's[45]) empowered the resilience of Icelandic during the subsequent centuries of Danish rule and Danish linguistic imperialism (Eggertsdóttir 2006).

There is some uncertainty regarding which languages and which textual sources primarily informed Gottskálksson in his Icelandic translation. Gottskálksson was proficient in multiple languages: the son of the Norwegian bishop of the Icelandic bishopric of Hólar and an Icelandic

[44] Some of Gottskálksson's translation was based on earlier (unprinted) work by Jón Arason (Eggertsdóttir 2006: 177), the last Catholic Icelandic bishop who was exiled and later beheaded for his resistance to the Reformation. While the idea of a Catholic bishop vernacularizing the Bible may contradict the position of the Catholic Church of this era, religious sectarianism is not the only explanation for these translations. Icelandic clergy in this period strove to disseminate the sagas and elevate Icelandic as a language of culture in order to dispel European conceptions of Iceland as a barbarian land (Loftsdóttir 2019: 20–21). The production of Icelandic texts—religious or secular—can be understood, then, as an assertion of ethnic and national identity rather than merely defiance of the Pope.

[45] Astås (2005: 76) comments, however, that the extent to which Guðbrandur Þorláksson modified Gottskálksson's text is 'uncertain'. More recently, Kvaran (2015: 25) attests that Þorláksson adopted Gottskálksson's New Testament almost wholesale, making only minor changes to perceived translation errors.

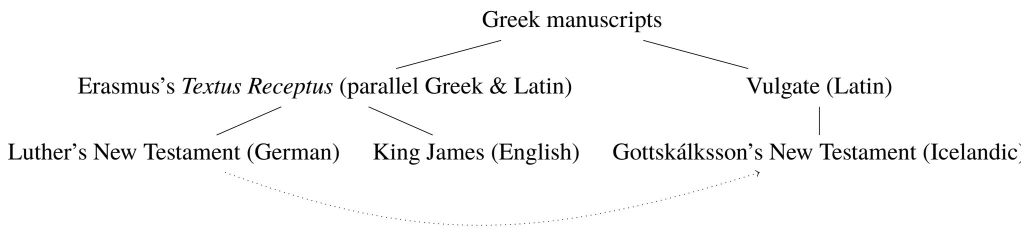


Figure 4.2: Relevant New Testament translation lineage

mother, Gottskálksson spent his childhood in Bergen with his Norwegian relatives after his father's death and went on to study in Denmark and Germany (Kvaran 2015: 20). Accordingly, we might expect artifacts of these languages in his translation. Additionally, after returning to Iceland and beginning to conduct his translation, Gottskálksson had access to not only the Latin Vulgate and, purportedly, even Erasmus' Greek-to-Latin translation (Astås 2005), but also the German translation of the New Testament completed by Martin Luther in 1522, as evidenced by the fact that the forewords included in Gottskálksson's New Testament are translations of Luther's own prologues (Kvaran 2015: 25)[46]. This literary genealogy is pictured in Figure 4.2. Astås (2005: 1198) comments that Gottskálksson worked primarily from Luther's German translation, using the Vulgate as a 'control text'. Accordingly, Astås (2005) claims that Gottskálksson's translation style was simultaneously rote and creative—although he introduced several German loanwords into the text and was prone to verbatim translation, he nevertheless was as likely to produce simple, purist Icelandic as he was to write stylized, high-register prose divorced from the vernacular. The artificiality of the language which Gottskálksson employed intentionally elevated the sacred text above idiomatic language.

Other evidence suggests, however, that faithfulness to the Vulgate, rather than Luther's German translation, was prioritized (justifying, I trust, my perhaps-controversial choice for dotted line versus solid line in Figure 4.2). Gottskálksson's printing included a letter from King Christian III of Denmark asserting that he had subjected Gottskálksson's text to scrutiny by scholars

[46] Gottskálksson might have made use of another Low German translation from around 1530 (Karlsson 1989), but the inclusion of Luther's forewords suggests he at least had access to these High German materials.

who verified the text's fidelity to the Latin Vulgate before approving its publication (Kvaran 2015: 21). This faithfulness is seen in some verses where parenthetical content present in the Vulgate is omitted from Luther's version, but is preserved in the Icelandic, as in (16)[47].

(16) John 1:38

a. Gottskálksson (Icelandic, 1540)

*þeir sögðu honum Rabbí [**hvað** að þýði-st meistari] hvar*
they said to.him Rabbi(N) HVER.N.SG.NOM C translates-REFL master where

ert þú heima
are 2SG at.home

'They said to him, "Rabbi (which is translated as master), where do you live?"'

b. Vulgate (Latin)

*qui dixerunt ei: Rabbi [**quod** dicit-ur interpretat-um Magister], ubi*
they said to.him Rabbi QUI.NOM said-PASS translate-PART master where

habita-s?
live-2SG.PRS

'They said to him, "Rabbi (which is translated as master), where do you live?"'

c. Luther (German, 1522)

sie aber sprachen zu ihm: Meister, wo bist du zur Herberge?
they however spoke to him master where are you at home

'But they said to him, "Master, where do you live?"'

In regards to Gottskálksson's translation of subordinate clauses in particular, Helgason (1929: 270) remarks that 'Oddur often follows his originals more than popular language prac-

[47] *hvað* here may also be analyzed as part of the paradigm for the distinct *hv*-word HVAÐ 'what'. However, *hvað* is polysemous with the neuter nominative and accusative of *hver*, and IcePaHC codes such instances as belonging to the lemma HVER.

tice', particularly compared to Luther. Gottskálksson in at least some cases modeled relative clause constructs on the relative pronoun strategy used by Latin where the form of the relative pronoun originates from the interrogative pronoun. Such a case is shown below in (17):

(17) John 11:2

a. Gottskálksson (Icelandic, 1540)

*en María ... [**hver-rar** bróðir var sá Lasarus sem sjúkur lá]*

but Mary ... HVER-F.SG.GEN brother was DEM Lazarus SEM sick lay]

'But Mary, ... whose brother Lazarus was sick'

b. Vulgate (Latin)

*Maria autem ... [**cuius** frater Lazarus infirma-ba-t-ur]*

Maria however ... [QUI.SG.GEN brother Lazarus weaken-IMPF-3SG-PASS]

'But Mary, ... whose brother Lazarus was sick'

c. Luther (German, 1522)

*Maria aber ... [**der-en** Bruder Lazarus war krank]*

Mary but ... DER-F.SG.GEN brother Lazarus was sick

'But Mary, ... whose brother Lazarus was sick'

d. 21st Century Bible (Icelandic, 2007)

en María ... bróðir hennar Lasarus var sjúkur

but Mary ... brother she.GEN Lasarus was sick

'But Mary, ... Her brother, Lazarus, was sick.'

In both the homologous Icelandic (17a)[48] and Latin passages (17b), the possessor is rel-

[48] This passage additionally shows subsequent canonical relativization with *sem*.

ativized with a *wh*-word. While Luther's translation (17c) shows relativization as well[49], the relative pronoun sources the German demonstrative *der*. Modern Icelandic Bible translations, such as (17d) typically show a paratactic strategy in these cases, avoiding relativization altogether, consistent with the inability for a headed relative clause to relativize a possessor in Modern Icelandic (Þráinsson 2007: 408), a constraint which the interrogative–relativizer lacks. Per the work of Sapp (2019), Icelandic also featured a demonstrative, *sá*, which, when accompanied by the canonical complementizer *er*, was capable of relativizing a possessor[50]; despite the existence of the demonstrative relative pronoun strategy, the interrogative material was chosen, suggesting that the Latin text, with its *wh*-word relative pronoun, was the source for Gottskálksson in this case.

Passages sometimes feature hybridized structures which can be traced to syntactic patterns in both German and Latin texts simultaneously. In particular, complement clauses marked by a relativizer which are stylistically fronted in Latin passages, as in (19b), typically correspond to postnominal headed RCs in German, as in (19c). The Icelandic translation (19a) shows features of the German model—the same clause order, a headed RC, and the usage of a 3SG pronoun to serve as the head for this RC—and the Latin model—namely, the usage of an interrogative–relative pronoun lacking the pied-piped element in the German.

[49] Notably, unlike Gottskálksson, Luther referred to Greek sources, not only the Vulgate, when composing the New Testament (Burger 2014), and the German translation here is not a calque of the Vulgate strategy. I have not attempted in this study to include Greek texts, but compare the Luther passage with the 1550 Stephanus New Testament (a printing of Erasmus's *Textus Receptus*), which features a relative pronoun in the genitive:

(18)	ἧς	ὁ	ἀδελφὸς	Λάζαρος	ἠσθένει
	hēs	ho	adelphos	Lazaros	ēsthenei
	REL.F.SG.GEN	DEM.M.SG.NOM	brother	Lazarus	was.sick

I will not attempt to trace the development of the Koine Greek relative pronoun here, but the marking of the relative clause in Luther's translation bears similarity to the inflectional marking of the relativizer in the Greek.

[50] If this strategy were employed, we might expect a result resembling *en María ... [**sú er** bróðir]* 'But Mary ... DEM.FEM.NOM that brother...', where *sá* would be predicted to display case attraction to the domain nominal.

(19) Acts 13:25

a. Gottskálksson (Icelandic, 1540)

*sagði hann eg em eigi hann [**hvern** þér mein-ið mig vera]*

said he I am not he HVER.M.SG.ACC you think-2PL.PRS 1SG.ACC be

'... he said, "I am not the one whom you think me to be" '

b. Vulgate (Latin)

*dice-bat [**quem** me arbitra-mini esse], non sum ego*

say-3SG.IMPF QUI.M.SG.ACC me think-2PL.PRS be not am I

' ... he said, "Whom you think me to be, I am not" '

c. Luther (German, 1522)

*sprach er ich bin nicht der, [**für den** ihr mich*

said he I am not DEM.M.SG.NOM for DER.MASC.SG.ACC 2PL.NOM me

halt-et]

regard-2PL.PRS

'... he said, "I am not the one for whom you take me [to be]" '

4.3.4 Analysis of parallel passages

I now turn to the quantification of the link between Latin and German relative clause structures and Icelandic IRCs in homologous Bible passages.

For the systematic comparison of Icelandic relativizer structure with Latin and German relativizer structure, I selected a subset of IcePaHC—the Gospel of John and the Book of Acts, the only texts included in this corpus from Gottskálksson's New Testament, the first printed Icelandic translation of the New Testament (see §4.3.3). The passages featuring a RC with a *hver* relativizer ($n = 146$) were manually aligned with parallel passages from the Latin Vulgate and the German Lutherbibel, which were then coded for inflectional marking of case and pied-piping. The results of this comparison are given in the tables below. The chief takeaway of the tables in this section is that there is a tendency for German and Latin passages to feature relative clauses with IRPs inflected in the same manner as the Icelandic IRP, but that the correspondence between German and Icelandic is looser than that between Icelandic and Latin.

Of all 146 IRPs investigated, 11 lacked an explicit Latin model (each representing an innovative usage of an IRP by Gottskálksson), 15 IRCs could not be linked to a German relative clause at all (Table 4.5), and only 5 lacked both models.

Table 4.5: *hver* passages ($n = 146$) with matching RC in model

Language	variant rel.	invariant rel.	*total*
Latin	135	0	*135 (92.5%)*
German	105	26	*131 (90.0%)*

An example of a *hver* IRC with a corresponding Latin IRC, but no German RC at all, is shown in (20). Here, both Icelandic (20a) and Latin (20b) translations feature a relative clause, whereas the German (20c) features only a prepositional phrase.

(20) Acts 2:10

a. Gottskálksson (Icelandic, 1540)

...	*in*	*álfum*	*Líbýe*	*[**hverjar***	*að*	*eru*	*nærri*	*Kýrenia]*
...	in	region.F.DAT.PL	Libya.GEN.SG	HVER.F.PL.NOM	C	are	near	Cyrene

'in regions of Libya, which are near Cyrene'

b. Vulgate (Latin)

...	*in*	*partes*	*Libyae*	*[**quae***	*est*	*circa*	*Cyrenen]*
...	in	part.F.NOM.PL	Libya.F.GEN.SG	QUI.F.SG.NOM	is	near	Cyrene

'in parts of Libya, which is near Cyrene'

c. Luther (German, 1522)

... *an*	*den*	*Enden*	*von*	*Lybien*	*bei*	*Kyrene*
... in	DET.ACC.PL	parts	of	Libya	by	Cyrene

'in the parts of Libya by Cyrene'

Table 4.6 shows instances in which case marking corresponds between translations featuring RCs with variant relativizers. Gender marking was not analyzed because the lexically-specified gender of nouns in Icelandic would override any Latin or German gender marking; for the same reason, pied-piped forms were excluded, as prepositions in each language require their objects to be in certain cases. Both Latin and German primarily agree with Icelandic case in applicable RCs, but this case agreement obscures other incongruence in translation.

Table 4.6: Non-pied-piped *hver* ($n = 116$) matching case of model

Language	relativizer	occurrences
Latin	*qui*	97
German	*der*	48
	welcher	31

[1] Latin ablative here is conflated with dative, as the only non-pied-piped Latin ablative in this data was adopted as dative in Icelandic. German *der* and *welcher* are the only pronouns included since they are the only relative pronouns which are variant.

Example (20) features such an irregularity, though: even though an IRC is present in both Icelandic and Latin, there is a mismatch in referent. The Icelandic IRP (20a) agrees in number and gender with *álfum* 'regions', but the antecedent of the Latin IRP (20b) must be the noun *Libya*, as the verb *est* is singular. The German (20c) is ambiguous, as well. This difference could be explained by the inflectional paradigm of *qui* (refer to Table 4.4): *quae* can be both F.GEN.SG and N.NOM/ACC.PL. If Gottskálksson was paying attention to the inflection of *qui* more than the form of *est* (perhaps, in this case, in order to resolve the ambiguity of referent in the

German), it is possible the *qui* here was mistakenly identified as a plural form.

Ultimately, the correspondence between forms is high, but there are not enough data from these two books to make claims of formal statistical significance regarding which source more accurately predicts the resultant Icelandic construction. In most cases, a particular IRC in Icelandic could have been taken from either Latin or German simply due to the fact that, in the majority of cases, all three languages feature an IRC with matching case marking. However, the data here still show that Latin, in these two books, corresponds more often with the Icelandic form. For instance, as seen in Table 4.7, a pied-piped construction in a Latin RC always has a corresponding pied-piped construction in Latin, whereas there are a few verses featuring German pied-piped constructions with no corresponding pied-piping in Icelandic.

Table 4.7: *hver* passages with pied-piped model

Language	PP in Ice.	no PP in Ice.
Latin	20	0
German	22	7

Another consideration which suggests that Latin RC structure, rather than German, is more closely tied with Icelandic translation is the fact that not all of the RCs analyzed above feature an IRP in German (Table 4.8). While Latin exposes only one model to Icelandic—the IRP *qui*—most homologous RCs in German are encoded with the demonstrative pronoun *der*. As I discuss below in §4.4, Icelandic had access to a grammaticalizing demonstrative–relative pronoun, yet the translator chose the interrogative–relative *hver* in these passages, suggesting either that Gottskálksson sourced most of these constructions from the Latin text or that he chose *hver* over *sá* since *hver* was, in some way, more similar semantically or syntactically to *qui*. One possibility is the phonological similarity between *hv-* words and *qu*-words, which are ultimately cognate: in Icelandic, *hv-*, now usually pronounced [kv], was likely pronounced [x^w] during this period until perhaps as late as the 18th century (Karlsson 2004: 17; 21), and resembles the pronunciation of the Latin *qu-* [k^w]. In terms of pure syntax, *hver* does resemble *qui*

Table 4.8: *hver* passages ($n = 146$) with matching RC in model, by relativizer

Language	relativizer	occurrences
Latin	*qui*	135
German	*der*	54
	welcher	51
	was	15
	da	9
	other	2

in that, unlike *sá*, it marks the syntactic relationship of the NP_{REL}, the role of the head within the RC. Before commenting more on this issue in §4.5, we must first talk about the syntactic peculiarities of this other relative pronoun, *sá*.

4.4 Contact and diachrony of relatives

Having reviewed the evidence for the effect of contact on the grammaticalization of IRCs, I now put these insights into dialogue with language-internal changes that Icelandic underwent in the relative paradigm. In §4.4.1, I discuss the history of *sá*, another variant pronoun which underwent grammaticalization towards a full relative pronoun, and then in §4.4.2 I discuss the language-internal factors that, together with the contact effects described above, induced shift of the relativization paradigm.

4.4.1 The case(-attraction) of *sá*

Though *sá* originates, and in most cases remains, as a demonstrative pronoun, it frequently co-occurs with relative clauses. Sapp (2019) reports that 63–81% of RCs in the 12$^{\text{th}}$–15$^{\text{th}}$ centuries were preceded by a form of *sá*[51], much more commonly than other demonstrative pronouns like *þessi* and *inn*, and that this frequent collocation is evidenced in even earlier (circa 9th century) poetry, becoming more common up until the prose documents in IcePaHC (beginning at 1150 AD). The sentence in (21) shows a prototypical example of such a construction. Here, the dative form of *sá* agrees with the dative head of the matrix NP, whereas there is a subject gap in the RC, a function usually marked by the nominative[52] (I have marked this gap for clarity).

(21) *sá* before RC (1150.FIRSTGRAMMAR.SCI-LIN,.7)[53]

en	*þó*	*rita*	*enskir*	*menn*	*ensku-na*	*latínu-stöfum*	*öllum*
But	still	write	English	men	English-DET.ACC	Latin-letters.PL.DAT	all.PL.DAT

þeim	*[___*	***er***	*réttræðir*	*verða]*	...
SÁ.PL.DAT	(SBJ)	C	rightly.pronounced	be.FUT	

'But English men still write English in Latin letters that will be rightly pronounced…'

This agreement in case and plurality is, of course, not surprising for a dependent of an NP. Partly due to this reason, Wagener (2017: 132) (among others, cf. Faarlund 2004) argues against the idea of *sá* as a relative pronoun. Wagener (2017) cites further evidence against this view: although *sá* is common in relative contexts, it is not obligatory in the way that a canonical complementizer (such as *er*) is required in these constructions. Furthermore, *sá* also occurs pre-nominally, which is the typical slot for other Icelandic demonstratives. Such a function is shown in (22).

[51] Though I do not dispute the ubiquity of this construction, I am unsure of how Sapp identified this particular frequency and of the particular query he used to generate this value—we both use the same corpus, but I calculate that during these years, 38–49% of relative clauses were preceded by DP (a determiner phrase, featuring an N and a D) with *sá* in it. This discrepancy may stem from differences in parsing between my library and Corpus Search or in different query constructions.

[52] Many Icelandic constructions can take non-nominative subjects (cf. Barðdal 2011; Eyþórsson 2000), but this is not one of them.

[53] The glosses and free translations of examples (21) and (27) draw from Benediktsson's (1972) translation of this text, the *First Grammatical Treatise*.

(22) Pre-nominal demonstrative *sá* (1260.JOMSVIKINGAR.NAR-SAG,.505)

það *gull* *hafði* *hann* *flutt* *af* *suðurlöndum*

SA.N.SG.ACC gold(N).ACC had he.NOM moved to southlands

'He moved that gold to the south.'

Responding to these claims, Sapp (2019: 25) argues that in Old Icelandic, post-nominal *sá* in particular was reanalyzed into a relative pronoun which shows case attraction to the antecedent, based on semantic and syntactic grounds. Semantically, *sá* typically functions as a marker of anaphora, referring to some unique referent in the linguistic context, rather than as a marker of physical deixis (Wagener 2017: 124). But in relative contexts in particular, this function can be redundant—as in (23a), where it marks a proper noun known already to the reader—or occur with altogether indefinite referents—such as (23b), where *jarl* is unidentifiable to the reader. In these instances, *sá* seems to function more like a marker anticipating a RC rather than as a demonstrative indexing features of the referent NP, which Sapp (2019: 3) terms a 'cataphoric demonstrative'.

(23) a. 1150.HOMILIUBOK.REL-SER,.237 (from Sapp 2019)

synir *<u>Herodi-s</u>* ***þess,*** *[**er** börnum lét fara]*

sons Herod(M)-GEN SÁ.M.SG.GEN C children let kill

'… sons of (that) Herod, who had the children killed.'

b. *Heimskringla 98* (from Sapp 2019)

hann *setti* *<u>jarl</u>* *í* *hverju* *fylki,* ***þann*** *[**er** dœma skyldi lǫg]*

he set earl(M).ACC in each district SÁ.M.SG.ACC C judge should law

'He placed an earl in each district, who should judge the law.'

I have not examined my data to track the presumed increase of semantic bleaching of pre-RC *sá*; Sapp (2019: 11) points out that identifying these indefinite usages in IcePaHC is difficult unless *sá* occurs with another clearly indefinite marker, such as *einn* 'a certain'[54] or *nokkur* 'some'. The syntactic evidence I have found in my data, however, demonstrates the most likely locus of grammaticalization of *sá* into a relative pronoun. In an example such as (24), although *sá* still agrees with the case of NP_{s} (in the matrix clause), it also coincidentally agrees with the NP_{rel} (in the RC), which would be marked with the accusative case if it were realized with an explicit NP. Consequently, the pronoun *sá* could be seen to index the syntactic features of either NP.

(24) *sá*—marking NP_{s} or NP_{rel}? (1650.ILLUGI.NAR-SAG,.7611)

hann	*tók*	*þá*	*sverð*	***það***	***[er***	___	*Ketill*	*hafði*	*haft]*
he	took	then	sword(N).ACC	SÁ.M.SG.ACC	C	(OBJ)	Ketill(NOM)	had	had

'Then he took the sword that Ketill had had.'

Furthermore, this ambiguity seems to have been the dominant state of affairs. In my database, among sentences in IcePaHC which have been indexed for syntactic relation and reference of NPs across clause boundaries[55], the syntactic affinity of *sá* in relative contexts is consistently and primarily ambiguous in each century (Table 4.9).

Table 4.9: Affinity of *sá* to NP_{S} & NP_{REL}—instances by century

	12th	13th	14th	15th	16th	17th	18th	19th	20th	21st	*total*
Ambiguous	69	87	140	101	102	103	88	64	37	16	807
Non-ambiguous	40	26	86	37	38	40	21	19	30	5	342
% ambiguous	63.3	77.0	61.9	73.2	72.9	72.0	80.7	77.1	55.2	76.2	70.2

In other words, most sentences look like example (24) above, in which *sá* could be interpreted to index the syntactic relation of either NP in either clause. Due to the relatively low frequency of constructions reinforcing the formal paradigm which explicitly links *sá* to the matrix NP, the syntactic referent of *sá* is ripe for reanalysis.

[54] Though even this diagnostic is trepidatious, as *einn* can also refer to the numeral 'one'.
[55] Not all of the sentences in IcePaHC are tagged in such a way, however.

There is some evidence for this reanalysis having once been in the process of actualization in Icelandic. Firstly, it must be said that early examples of *sá* which demonstrate agreement with NP$_{REL}$ are generally regarded as examples of the so-called 'learned style', a formal register which is presumably in imitation of Latin RC structure (Wagener 2017; Nygaard 1905). Sapp (2019: 25–26) shows that Old Icelandic examples of this pattern are constrained overwhelmingly to religious texts as an artifact of translation. This pattern carries over to some degree in the Gottskálksson vernacular translation of the New Testament in the 16th century. For example, in (25a), *sá* is marked in the dative case to match the role of *asna* 'donkey' in the RC, not the matrix clause[56]. Unlike the ambiguous example (24) above, *sá* in this passage can be analyzed as marking the syntactic properties of NP$_{REL}$, indexing only the semantic properties of NP$_{S}$. I provide here also the equivalent passage in the Vulgate (25b), which shows an instance of *qui* in the same case.

(25) Numbers 22:30

a. Icelandic (Þráinsson 1980)

er	*eg*	*ecke*	*þijn*	*Asna,*	***[þeirre***	*þu*	*hefur*	*riðeð]*
am	I	not	your	donkey(F).NOM	SÁ.F.SG.DAT	you	have	ridden

'Am I not your donkey, whom you have ridden?'

b. Vulgate (Latin)

non=ne	*animal*	*tuum*	*sum,*	***[cui***	*semper*	*sedere*	*consue-vi-sti]*
NEG=Q	animal	your	be.1SG.PRS	QUI.SG.DAT	always	sit	tend-PRET-2SG

'Am I not your animal (donkey), whom you have grown accustomed to ride at all times?'

[56] Furthermore, it also lacks the typical accompanying *er*; this is a very rare pattern in the corpus and is likely the result of clerical error.

While this pattern never overtakes the case-attracting *sá*, the affinity for NP$_{\text{REL}}$ is observed in later 17th century texts which are original works with no Latin translation model, such as (26).

(26) *sá* unambiguously case marked for NP$_{\text{REL}}$ (1661.INDIAFARI.BIO-TRA,36.280 from Sapp 2019, brackets mine)

einn	*...*	*átti*	*sér*	*unga*	*og*	*dægilega*	<u>*kvinn-u,*</u>	***[sú***	***er***	*Anna*
one	...	had	REFL	young	and	pretty	wife-ACC	SÁ.F.SG.NOM	C	Anna(NOM)

hét]
was.called

'One (tailor ...) had a young and pretty wife, who was called Anna.'

Sapp (2019) identifies only 16 instances of this pattern (out of 700 ambiguous or clearly case-attracting instances) in the 17th century, but concludes that *sá* was on the path to grammaticalize to a true, non-case-attracting relative pronoun, just before the ubiquity of the *sá er* co-occurrence was replaced entirely by *sem*. The stage of grammaticalization which *sá* seems to have achieved at the end of its usage as a relative pronoun—indexing the syntactic features of NP$_{\text{REL}}$—is one which had already been achieved by *hver* IRPs, which were in widespread use in headed IRCs by the 16th century (refer back to Figure 4.1). Perhaps the contact-aided grammaticalization of the *hver* IRP played a role in the arrest of the usage of *sá* as a true relative pronoun. I will return to this thought in §4.5 after reviewing the stages of development of *hver* in §4.4.2.

4.4.2 The grammaticalization of *hver*

We have already seen evidence in §4.3.4 that the period of most frequent usage of *hver* as an IRP in writing can be traced to the influence of Bible translation, most likely primarily from a Latin model. In this section, I now discuss the development of *hver* prior to this period of contact, which is also necessary in understanding the grammaticalization of *hver* to fill the functions of an IRP. Finally, since we have now reviewed the evolution of *sá*, I will discuss the interaction of the life cycles of these two relative pronouns.

The interrogative–relative grammaticalization pathway

The use of interrogative forms as markers for headed relative clauses is cross-linguistically unusual, but has spread throughout European languages, likely radiating outward from two centers of innovation—Slavic and Latin/Romance—with the strategy being adopted by some languages in contact with these families (Heine & Kuteva 2006; Thomason & Kaufman 1988). The progression of grammaticalization has been theorized to proceed along the following path (adapted from Heine & Kuteva 2006):

Table 4.10: Stages of interrogative–relative grammaticalization

	Able to mark	English example
1	word questions	***Who*** *came?*
2	indefinite adv./comp. clauses	*I don't know* ***who*** *came*
3	definite adv./comp clauses	*You also know* ***who*** *came*
4	headed relative clauses	*You know the woman* ***who*** *came?*

From the beginning of the literary period, Icelandic is at least at Stage 3 of the above pathway. In (27), an interrogative serves in a word question, and the examples in (28) show usages of an interrogative as a marker of complement clauses.

(27) Word questions (Stage 1)—1150 A.D. (1150.FIRSTGRAMMAR.SCI-LIN,.162)

hvað *þá skal að hafa?" kvað ég.*

HVER.N.SG.NOM/ACC then shall to do said I

' "What then is to be done?" said I.'

(28) Subordinate clauses with interrogative connectives

a. Indefinite (Stage 2)—1250 A.D. (1250.STURLUNGA.NAR-SAG,446.2037)

*Gissur spurði nú biskup og frændur sína [**hvað** nú skal*
Gissur asked now bishop and relatives his [HVER.N.SG.NOM/ACC now must

til ráða taka]
in advice take]

'Gissur now asked the bishop and his associates what now must be done'

b. Definite (Stage 3)—1260 A.D. (1260.JOMSVIKINGAR.NAR-SAG,.535)

*og Braut segja konungi svo búið [**hvað** **er** þá hefir*
and Braut says king.DAT at.present [HVER.N.SG.NOM/ACC C there had

í görst]
happened]

'And presently Braut is telling the king what had happened there'

That Icelandic advanced to Stage 3 of the interrogative–relative pathway is not necessarily surprising. Firstly, Icelandic has been in contact with Latin literature, one of the centers of innovation and spread of the IRC strategy, since at least the earliest IcePaHC texts (1150 A.D.–). Several of these texts are isolated sermons and Bible stories which predate the full vernacularization of the Bible during the Reformation. These translators were familiar with Latin and were heavily influenced by its established literary conventions, which Old Norse lacked (Wagener 2017: 128). Secondly, even though the (headed) IRP is associated with Indo-European languages, Lehmann (1984: 325–329) notes that many languages outside of the Indo-European family use interrogatives to mark indirect questions or other headless clauses (which includes Stage 2 and Stage 3 of the Heine & Kuteva (2006) interrogative–relative grammaticalization pathway). Furthermore, Auderset (2020) suggests that even within Indo-European, there may be more instances of independent innovation of Stage 4—namely, in Anatolian—rather than through contact solely with Latin or Slavic, as theorized by Heine & Kuteva (2006).

The grammaticalization of *hver* to pronoun

To some extent, then, having achieved the above stages does not *per se* necessitate a contact explanation; however, in the case of Icelandic, the overwhelming evidence from parallel corpora (presented in §4.3.4 and elsewhere) leads me to conclude that grammaticalization to Stage 4—IRPs marking headed relative clauses—is chiefly driven by contact through translation. Unambiguous instances of true headed IRCs in the Icelandic corpus are rare prior to the 16th century translation period. IRPs before this time are often supported by a subsequent canonical complementizer (29a)[57] or function as indefinite relatives (a characteristic of Stage 3) despite the collocation with an NP in the matrix clause (29a).

(29) a. Stage 4 (but supported by *er*)—1325 A.D. (1325.ARNI.NAR-SAG,.789)

þetta sumar kom og til Árna biskups bréf Jón
that summer came also to(>GEN) Árni.GEN bishop-GEN letter(N).NOM Jón.GEN

*erkibiskup-s [í **hverju** **er** hann bauð ...]*
archbishop-GEN in HVER.N.SG.DAT C he asked

'That summer, a letter from Archbishop Jón also came to Bishop Árni, in which he asked ...' '

b. Indefinite relative (only Stage 3)—1310 A.D. (1310.GRETTIR.NAR-SAG,.1232)

*skulum vér hafa líf Grettis [**hvað** **er** kostar ...]*
shall we have life Grettir HVER.N.SG.NOM/ACC C costs

'We shall take Grettir's life, whatever it costs'

[57] Namely, *eð*, *að*, and the relativizer *er*.

The usage of these canonical complementizers in double-marked IRCs grows until the 18th century. This pattern of relative pronoun paired with canonical complementizer bears similarity to the demonstrative *sá*, which typically precedes relativizer *er* and arguably grammaticalized to a case-attracting relative pronoun (see §4.4.1). Such redundant double complementation strategies have been proposed to help prop up 'phonologically light' and functionally varied words like *er*, which serves several other complementizer functions and moreover is homophonous with another high-frequency function word, *er* 'to be' (Sapp 2019: 36). Indeed, the frequency of *er* in relative constructions has gradually decreased nearly every century (as pointed out prior in Figure 4.1); such double-marking could have served to save an ailing relative particle by adding phonological and semantic bulk.

Conversely, this symbiosis may also benefit the newcomer relativizer *hver*, perhaps more than the ailing canonical relativizer *er*, by aiding in the early normalization of a foreign strategy. Although in terms of absolute frequency the double complementation of IRCs increases during the Reformation, relative to other strategies for marking IRCs the canonical complementizer strategy actually wanes, as shown in Figure 4.3.

In contrast to this decline, interrogative–relativization with 'bare' IRPs (30a) grows more common from the 16th century onward. At the same time, the pied-piping of prepositions dependent on IRPs (30b) becomes much more frequent, as well.

(30) a. Bare interrogative— 1593 A.D. (1593.EINTAL.REL-OTH,.882)

en *þér* *eruð* *þeir* ***[hvörjir*** *hjá* *mér* *voru* *í* *mínum*
but 2PL.NOM be.2PL they HVER.M.PL.NOM with 1SG.DAT be.3PL.PST in my

freistingum]
temptations

'But you were the ones who were with me in my temptation.'

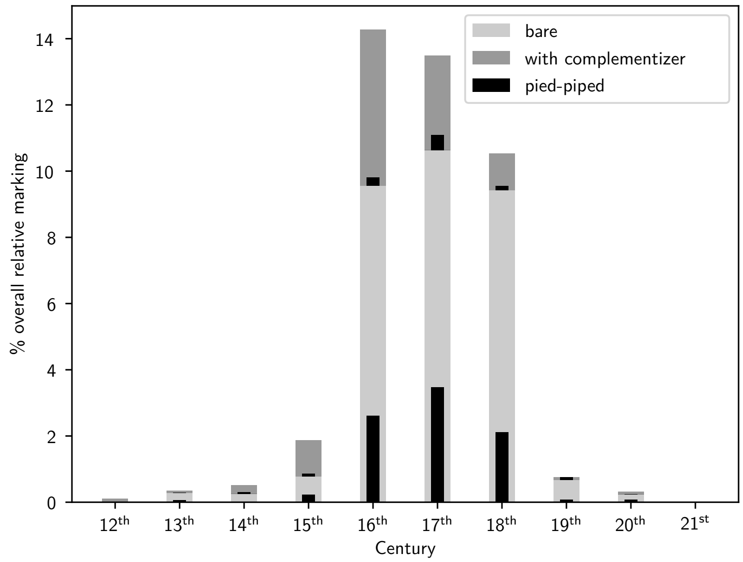


Figure 4.3: Frequency of interrogative–relative pronouns by co-occurring elements

b. Pied-piped interrogative—1630 A.D. (1630.GERHARD.REL-OTH,.144)

sjá	*ðu*	*þá*	*nagla*	*[með*	***hvörjum***	*mitt*	*hold*	*verður*
See	you	SÁ.M.SG.ACC	nail(M).ACC	[with	HVER.PL.DAT	my	flesh	becomes

í gegnum	*grafið]*
through all	carved]

'You see this nail which will pierce all my flesh'

Most notably, Figure 4.3 shows a difference in prevalence of pied-piping for bare IRCs versus IRCs double-marked with a canonical complementizer. Occurrence of pied-piping is shown with the two split black bars, each marking the frequency of pied-piping in bare and in double-marked *hv-* constructions. IRCs with a canonical complementizer can still feature pied-piping, but pied-piping is mostly observed in bare IRCs beginning in the 15th century. If double complementation of an IRC served the function of normalizing an unfamiliar structure via a familiar frame, this function was gradually replaced by pied-piped prepositions, which similarly increase the markedness of IRC formation while conveying clause-internal syntactic relations more explicitly. While IRP double-complementation becomes quite rare by the 18th century, pied-piping of bare IRPs remains well-attested. Perhaps the syntactic precedent of double complementation with *sá* and *hver* provide a template for the pied-piping of IRPs, a strategy which, like double-complementation, increases the bulk and markedness of the relativizer.

4.4.3 The race to true relative pronoun

With the discussion above, a question remains: why was the final stage of grammaticalization of the interrogative-to-relative pathway adopted from a Latin model in lieu of the nascent (and native) *sá* RC? This *sá*, we have seen, seems to have internally undergone a process of grammaticalization separate from the earlier 'learned style' construction and was developing into a true relative pronoun at the same time that *hver* IRCs became frequent. One potential anthropological explanation is, perhaps, rooted in the genre of the texts which normalized the usage of *hver* as an IRP: the concern for faithfulness, as closely as possible, to the Latin texts, a key consideration of the translators and commissioners of the vernacular versions of these sacred texts, who were pressured to balance vulgar readability with accuracy to a presumably authoritative text (§4.3.3). Such a concern could explain why the Latin source texts, which exclusively use an interrogative–relative pronoun in RCs, might have been prioritized over the equivalent German constructions, which used a variety of relativization strategies including a demonstrative–relative pronoun and whose referential qualities would perhaps lead one to suspect their translation as *sá*.

On a more linguistic level, as well, the properties of *hver* more closely mirrored Latin structure. As typological similarity catalyzes pattern borrowing (Thomason 2015), in this case it also promotes the selection of similar structures in translation. Like Latin *qui*, Icelandic *hver* unambiguously indexes the syntactic properties of NP_{REL} and the semantic properties of NP_{S}. Although *sá* in a few instances also shows evidence of developing these properties in a few late examples (§4.4.1), at best this strategy was concurrent with overwhelming usage of *sá* as a case-attracting relative pronoun indexing both semantic and syntactic properties of NP_S. With the preference for *hver* for the translation of relative constructions in the most seminal text establishing the national and literary identity of Icelandic (§4.3.3), perhaps the full grammaticalization of *sá* was interrupted as the functional space for a true relative pronoun was fulfilled through mass borrowing.

4.5 Conclusion

We have seen above that the development of *hver* as an interrogative–relative pronoun is stimulated by contact through translation, and that the period in which this stage of grammaticalization was induced (or transferred, per Mithun 2013; Heine & Kuteva 2003) is the era of the Protestant Reformation in the 16th century (§4.2.3). This finding confirms the long-standing intuitions of Icelandic linguists and grammarians that the strategy is foreign, but asserts that the contact scenario was during the Reformation rather than the earliest of Latin texts. The corpus study presented in §4.3.4 suggests that almost all cases of *hver* as an IRP in the Bible translations surveyed correspond to interrogative–relative clauses in Latin and German, and that in most cases, these constructions correspond in case-marking, as well. The monolithic nature of IRCs in Latin—formed solely with the use of the IRP *qui*—suggests that the Icelandic interrogative pronoun *hver*, which had already achieved some degree of relative function (§4.4.2), was chosen to translate Latin relative constructions due to the social context of translation and structural similarity (§4.4.3).

This strategy is not merely restricted to translation, however. The IRC strategy survived for a time after the Reformation and was productively used in texts which were not modeled on a Latin or German translation, such as example (31), an autobiographical text which dates to

1628[58].

(31) Innovative use of IRC (1628.OLAFUREGILS.BIO-TRA,.571)

og	*af*	*mínum*	*peningum*	*átti*	*eg*	*eftir*	*4*	*dali*	***[hverja***	*eg*	*geymdi*	*á*	
and	of	my	money	had	I	still	4	dollars(M)	HVER.M.PL.ACC	I	kept	on	

mér	*heimuglega]*
me	secretly

'And of my money, I still had 4 dollars which I concealed.'

However, after the 18th century, relative pronouns become much less frequent, and the central strategy of using an invariant canonical relativizer is restored—*sem* rather than *er* is the relativizer of choice after this century, but the syntactic structure is much the same as that of *er* before.

Though this paper focuses on the factors behind the evolution of relative pronouns in Icelandic and their interplay, I will briefly discuss some potential explanations for the ultimate failure of both *sá* and *hver* to persist in the language after the 18th century. One explanation is that these strategies may not have been used in the spoken language; there is no evidence presented here to determine either way. It seems likely that the phonological lightness of *er* conditioned the shift to *sem* in contemporary spoken Icelandic and that this process was reflected in the written language. However, it is uncertain if the other strategy to reinforce *er*—double complementation with *sá*—also existed in the spoken language, just as there is no evidence to suggest that IRPs like *hver* were ever used in true relative clauses in colloquial spoken Icelandic. It is possible, then, that in the spoken language *er* rather straightforwardly underwent lexical replacement by *sem*, preserving the dominance of the invariant relative particle paradigm in spoken Icelandic. Additionally, even if these usages entered the spoken language, the con-

[58] The author here, Ólafur Egilsson, was a pastor himself and would have been exposed to the instances of *hver* IRPs from translation.

scious rejection of perceived foreign elements (as described in §4.3.3) could have contributed to the ultimate refusal to accept of *hver* IRPs in spoken Icelandic usage. Similarly, *sá* might have also been a casualty of this process: even if it briefly became a true relative pronoun, it always seems to have been associated with the foreign 'learned style' of old Latin translation. Regardless, either *sem* out-competed the grammaticalizing relative pronouns or there was never any competition in the spoken language.

In terms of the written language, at least, the entire period leading up to the 20th century could be viewed as a destabilization of the equilibrium of the invariant relative particles as markers of relative clauses, triggered by the erosion of solitary *er* as a relativizer and the need for its replacement (§4.4.2). The almost-full grammaticalization of *hver* and *sá* to true, productive relative pronouns gained traction when the traditional system of relativization was least stable. In regards to the notion of stability, Nichols (2003: 283) notes that 'diversity arises when some element is relatively unstable and therefore prone to replacement in various ways'. Indeed, the mid-16th century system of relativization in Icelandic features the most diverse assortment of relativization strategies in the history of the language: the replacement of *er* with *sem*, the ambiguity of the syntactic relationship between $\textsc{np}_{\textsc{s}}$ and $\textsc{np}_{\textsc{rel}}$ of *sá*, and the proliferation of *hver* relative pronouns. As the once-monolithic *er* is replaced over centuries by *sem*, no longer is *er* always an obligatory component of a relative clause; in this state, there is ambiguity in which markers can occupy the slot of relativizer, and as *sem* overtakes *er* in frequency (between the 15th–16th centuries), case-attracting *sá* and *hver* gain popularity as IRPs. Conversely, when stability in the system is regained (as *sem* approaches its peak frequency per century), the alternative relativization strategies become very infrequent. With these considerations in mind, the origin (and later disappearance) of the *hver* IRC is not solely a contact phenomenon, but rather also the product of ongoing changes in the relative space of Icelandic, a system in flux.

Despite contact with Latin texts earlier in the history of Icelandic, *hver* is not widely used as marker of true RCs until the Reformation, the period of maximum instability in the relativization system. By this time *sá* had already proceeded down a demonstrative–relative grammaticalization pathway: as a case-attracting relative pronoun, *sá* varied according to the case of its head (§4.4.1) and it functioned in a double complementation strategy with the phonologically weak and semantically bleached canonical relativizer *er* (§4.4.2). The reanalysis of *sá* normalized the occurrence of relative markers that were variant in form. Additionally, the double complementation strategy itself is mirrored by early examples of interrogative–relativization, which are often supported by either double complementation (e.g. *hver er* 'who that') or another compound structure, pied-piping (*með hverjum* 'with whom'). In other words, the grammaticalization of *sá* into a variant relative marker created a point of greater typological congruence (in the terms of Thomason 2015; see §4.3.1) with the Latin true relative pronoun *qui*, facilitating the pattern borrowing of the interrogative–relative paradigm onto *hver*. In terms of the interrogative–relative pathway (Heine & Kuteva 2006), these developments constitute additional factors (of the sort described by Mithun 2012) which accelerate progress down the pathway: in the case of Icelandic, the contact-induced grammaticalization of interrogative to true relative pronoun was contingent on the parallel, language-internal development of the demonstrative pronoun. On a broader level, the weakening of *er* triggered language-internal changes which allowed for the pattern borrowing of structures that were incongruent to *er*. If an unstable system promotes diversity in paradigms, perhaps it also facilitates the adoption of foreign strategies which would otherwise have to compete with a clearly dominant, monolithic paradigm.

Chapter 5

Conclusion

This dissertation has revealed how pressures at multiple scales—from millisecond-level cognitive processing to centuries of cultural contact—shape the evolution of linguistic systems. Through three studies examining different aspects of language change, we see that the interplay between processing constraints and functional needs creates both stability and innovation in language structure. The way speakers access and process words in real time leaves an imprint on language that persists across generations. The cohort model of lexical access shows us that not all parts of a word bear equal weight in recognition—speakers can identify words before hearing them in their entirety. This cognitive reality is reflected in patterns of historical sound change: the parts of words crucial for early recognition resist change more strongly than material that comes after the uniqueness point. Similarly, when surrounding syntactic context helps disambiguate minimal pairs, the pressure to maintain phonological contrasts is reduced. These findings confirm and further illuminate the notion that the moment-by-moment demands of language processing gradually sculpt linguistic systems over time.

The first two studies focused on how the functional load of phonological material impacts its resistance to sound change. In Chapter 2, I demonstrated that syntactic predictability modulates the likelihood of phonemic merger—when surrounding grammatical context helps disambiguate minimal pairs, the phonological contrast between them is more likely to be lost. This finding refines our understanding of how functional load operates in natural language by showing that it is not just the raw number of minimal pairs that matters, but their distribution across syntactic contexts.

Chapter 3 revealed that the cohort effect in lexical access creates an asymmetry in how sound change operates within words. Sounds within a word are more likely to change the further into the word they occur, and sounds that occur after a cohort-selection recognition point change differently than material before such a point, even if these sounds occur in the same relative position within words. This finding suggests that the cognitive machinery of speech processing shapes the long-term trajectory of sound change.

Yet these cognitive pressures do not operate in isolation. The case of relativization in Icelandic, shown in Chapter 4, demonstrates the interaction of social and cultural factors with system-level preferences to shape language change. I showed that despite centuries of contact pressure and the grammaticalization of novel relativization strategies, the language ultimately retained a functionally-preferred pattern of relativization. The ultimate retention of functionally-preferred patterns despite contact pressure and internal motivations for lexical replacement suggests that languages prioritize communicative needs even through periods of upheaval, both cultural and system-internal. In other words, even instability can be, when viewed from a sufficiently circumspect vantage, surprisingly stable. This illustrates how functional adequacy can act as a conservative force, causing linguistic systems to “snap back” to stable configurations even after periods of instability.

Taken together, these studies demonstrate that language change operates under competing pressures—the drive toward ease of articulation and processing must be balanced against the need to maintain crucial linguistic contrasts. The interaction between these forces helps explain why languages exhibit both stability and change over time. Just as biological evolution results from the interplay between mutation and selection, linguistic evolution reflects a balance between forces promoting and inhibiting change. Understanding these dynamics requires examining both the cognitive mechanisms underlying language processing and the functional constraints that shape linguistic systems.

Future work should expand this framework to examine how cognitive and functional pressures interact in other domains of language change. Additionally, while this dissertation has focused primarily on Indo-European languages, investigating these effects in typologically diverse languages would help establish their cross-linguistic generality. The methodological advances presented here—particularly the refined measures of functional load and the automated analysis of historical corpora—provide tools for pursuing these broader questions about the nature of language change.

What emerges is a picture of language as a complex adaptive system shaped by forces operating at multiple timescales. The millisecond-level processes of speech perception and production create subtle biases that, accumulated over generations, result in systematic patterns of change. At the same time, broader functional pressures arising from the needs of communication act as a conservative force, preserving structures that efficiently serve core linguistic functions. The interaction between these factors helps explain both the remarkable stability of some linguistic features and the inevitability of change in others. These dynamics mirror patterns seen in other complex systems, from biological evolution to cultural transmission. Just as genetic drift can gradually alter populations while selection maintains adaptive traits, the small processing biases of individual speakers can accumulate to create language change while functional pressures preserve crucial contrasts. Understanding language change requires examining both these subtle cognitive influences and the larger functional constraints that shape linguistic systems. The findings presented here suggest that language change is neither purely mechanistic nor purely functional. Rather, it emerges from the complex interplay between the cognitive machinery humans use to process language and the communicative needs that language serves. This optimization problem helps explain why languages show both persistent structural patterns and ongoing change—they must balance the competing demands of processing efficiency and communicative effectiveness. In studying these dynamics, we gain insight not just into how languages change, but into the fundamental nature of language as a human cognitive and social system.

Appendix A

Supplemental data

Table A.1: All contrasts coded by attested or unattested mergers (Chapter 2)

Language	Type	Merged Pairs	Unmerged Pairs
English (US)	C∼C	ʍ∼w	b∼d b∼g b∼m b∼p b∼v b∼w d∼g d∼l d∼n d∼t d∼z d∼ɹ dʒ∼d dʒ∼l dʒ∼n dʒ∼z dʒ∼ɹ f∼h f∼p f∼s f∼v g∼k g∼w h∼s j∼w j∼ɹ k∼p k∼t l∼n l∼z l∼ɹ m∼n m∼v m∼w n∼z n∼ɹ p∼t s∼t s∼z tʃ∼dʒ tʃ∼s tʃ∼t tʃ∼ʃ tʃ∼θ v∼w v∼z ŋ∼g ŋ∼m ŋ∼n ɹ∼w ɹ∼z ʃ∼f ʃ∼h ʃ∼s ʃ∼t ʃ∼z ʃ∼θ ʒ∼d ʒ∼dʒ ʒ∼l ʒ∼n ʒ∼s ʒ∼v ʒ∼z ð∼d ð∼dʒ ð∼f ð∼l ð∼n ð∼s ð∼v ð∼z ð∼ɹ ð∼ʃ ð∼ʒ ð∼θ θ∼f θ∼h θ∼s θ∼t θ∼v θ∼z
	C∼C/V'_V	d∼t	b∼d b∼g b∼m b∼p b∼v d∼g d∼l d∼n d∼z d∼ɹ dʒ∼d dʒ∼l dʒ∼n dʒ∼z dʒ∼ɹ f∼p f∼s f∼v g∼k k∼p k∼t l∼n l∼z l∼ɹ m∼n m∼v n∼z n∼ɹ p∼t s∼t s∼z tʃ∼dʒ tʃ∼s tʃ∼t tʃ∼ʃ v∼z ŋ∼g ŋ∼m ŋ∼n ɹ∼z ʃ∼f ʃ∼s ʃ∼t ʃ∼z ʒ∼d ʒ∼dʒ ʒ∼l ʒ∼s ʒ∼v ʒ∼z ð∼d ð∼dʒ ð∼f ð∼l ð∼n ð∼s ð∼v ð∼ɹ ð∼ʒ θ∼t θ∼v
	V∼V	ɑ∼ɔ əɹ∼oɪ	aɪ∼aʊ aɪ∼i aɪ∼oɪ aɪ∼æ aɪ∼ɛ aɪ∼ɪ aɪ∼ʊ aʊ∼u aʊ∼ɪ aʊ∼ʊ aʊ∼ʌ eɪ∼i eɪ∼oɪ eɪ∼oʊ eɪ∼ɛ i∼u oɪ∼aʊ oɪ∼i oɪ∼oʊ oɪ∼u oɪ∼æ oɪ∼ɛ oɪ∼ɪ oɪ∼ʊ oɪ∼ʌ oʊ∼u oʊ∼ʊ oʊ∼ʌ ɑ∼aɪ ɑ∼aʊ ɔ∼aɪ ɔ∼aʊ ɔ∼oɪ ɔ∼oʊ ɔ∼u ɔ∼ʌ ə∼ɛ ə∼ʌ əɹ∼ə əɹ∼ɛ əɹ∼ʌ ɛ∼æ ɛ∼ɪ ɛ∼ʌ ɪ∼i ɪ∼æ ɪ∼ʊ ʊ∼u ʊ∼ʌ
	V∼V/_l	oʊ∼ʊ ɔ∼ʌ ɪ∼i ʊ∼u	aɪ∼aʊ aɪ∼i aɪ∼oɪ aɪ∼æ aɪ∼ɛ aɪ∼ɪ aɪ∼ʊ aʊ∼u aʊ∼ɪ aʊ∼ʊ aʊ∼ʌ eɪ∼i eɪ∼oɪ eɪ∼oʊ eɪ∼ɛ i∼u oɪ∼aʊ oɪ∼i oɪ∼oʊ oɪ∼u oɪ∼æ oɪ∼ɛ oɪ∼ɪ oɪ∼ʊ oɪ∼ʌ oʊ∼u oʊ∼ʌ ɑ∼aɪ ɑ∼aʊ ɑ∼ɔ ɔ∼aɪ ɔ∼aʊ ɔ∼oɪ ɔ∼oʊ ɔ∼u ə∼ɛ əɹ∼oɪ əɹ∼ə əɹ∼ɛ əɹ∼ʌ ɛ∼æ ɛ∼ɪ ɛ∼ʌ ɪ∼æ ɪ∼ʊ

	V∼V/_n	ɛ∼ɪ	aɪ∼aʊ aɪ∼i aɪ∼oɪ aɪ∼æ aɪ∼ɛ aɪ∼ɪ aʊ∼u aʊ∼ɪ aʊ∼ʌ eɪ∼i eɪ∼oɪ eɪ∼oʊ eɪ∼ɛ i∼u oɪ∼i oɪ∼oʊ oɪ∼u oɪ∼æ oɪ∼ɛ oɪ∼ɪ oɪ∼ʌ oʊ∼u oʊ∼ʌ ɑ∼aɪ ɑ∼aʊ ɑ∼ɔ ɔ∼aɪ ɔ∼aʊ ɔ∼oɪ ɔ∼oʊ ɔ∼u ɔ∼ʌ ə∼ɛ ə∼ʌ əɹ∼oɪ əɹ∼ə əɹ∼ɛ əɹ∼ʌ ɛ∼æ ɛ∼ʌ ɪ∼i ɪ∼æ
	V∼V/_r	ɑ∼ɔ	aɪ∼aʊ aɪ∼i aɪ∼ɛ aɪ∼ɪ aɪ∼ʊ aʊ∼u aʊ∼ɪ aʊ∼ʊ ɑ∼aɪ ɔ∼aɪ ɔ∼aʊ ɔ∼u ɔ∼ʌ ɛ∼æ ɛ∼ɪ ɪ∼i ɪ∼æ ɪ∼ʊ
English (RP)	C∼C	θ∼f θ∼t	b∼d b∼g b∼m b∼p b∼v b∼w d∼g d∼l d∼n d∼t d∼z d∼ɹ dʒ∼d dʒ∼l dʒ∼n dʒ∼z dʒ∼ɹ f∼h f∼p f∼s f∼v g∼k g∼w h∼s j∼w j∼ɹ k∼p k∼t l∼n l∼z l∼ɹ m∼n m∼v m∼w n∼z n∼ɹ p∼t s∼t s∼x s∼z tʃ∼dʒ tʃ∼s tʃ∼t tʃ∼ʃ tʃ∼θ v∼w v∼z ŋ∼g ŋ∼m ŋ∼n ɹ∼w ɹ∼z ʃ∼f ʃ∼h ʃ∼s ʃ∼t ʃ∼z ʃ∼ʒ ʃ∼θ ʒ∼d ʒ∼dʒ ʒ∼l ʒ∼n ʒ∼s ʒ∼v ʒ∼z ð∼d ð∼dʒ ð∼f ð∼l ð∼n ð∼s ð∼v ð∼z ð∼ɹ ð∼ʃ ð∼ʒ ð∼θ θ∼h θ∼s θ∼v θ∼z θ∼ʒ
	C∼C/{C,V}_	ð∼d ð∼v	b∼d b∼g b∼m b∼p b∼v b∼w d∼g d∼l d∼n d∼t d∼z d∼ɹ dʒ∼d dʒ∼l dʒ∼n dʒ∼z dʒ∼ɹ f∼p f∼s f∼v g∼k h∼s k∼p k∼t l∼n l∼z l∼ɹ m∼n m∼v n∼z n∼ɹ p∼t s∼t s∼z tʃ∼dʒ tʃ∼s tʃ∼t tʃ∼ʃ v∼z ŋ∼g ŋ∼m ŋ∼n ɹ∼w ɹ∼z ʃ∼f ʃ∼s ʃ∼t ʃ∼z ʒ∼d ʒ∼dʒ ʒ∼l ʒ∼s ʒ∼v ʒ∼z ð∼dʒ ð∼f ð∼l ð∼n ð∼s ð∼ɹ ð∼ʒ θ∼f θ∼t θ∼v θ∼z θ∼ʒ
	V∼V	aɪ∼ɔɪ ɔː∼ʊə ɜː∼ɛə ɪə∼ɛə	aɪ∼aʊ aɪ∼iː aɪ∼æ aɪ∼əʊ aɪ∼ɛ aɪ∼ɛə aɪ∼ɪ aɪ∼ɪə aɪ∼ʊ aʊ∼iː aʊ∼uː aʊ∼æ aʊ∼ɛ aʊ∼ɛə aʊ∼ɪ aʊ∼ɪə aʊ∼ʊ aʊ∼ʌ eɪ∼aɪ eɪ∼aʊ eɪ∼iː eɪ∼æ eɪ∼ɔɪ eɪ∼ə eɪ∼əʊ eɪ∼ɛ eɪ∼ɛə eɪ∼ɜː eɪ∼ɪ eɪ∼ɪə eɪ∼ʊ eɪ∼ʌ ɑː∼aɪ ɑː∼aʊ ɑː∼əʊ ɑː∼ʊə ɔɪ∼aʊ ɔɪ∼iː ɔɪ∼uː ɔɪ∼æ ɔɪ∼ɒ ɔɪ∼əʊ ɔɪ∼ɛ ɔɪ∼ɪ ɔɪ∼ʊ ɔɪ∼ʌ ɔː∼uː ɔː∼ɔɪ ə∼ɛ ə∼ʌ əʊ∼aʊ əʊ∼iː əʊ∼uː əʊ∼ə əʊ∼ɛ əʊ∼ɛə əʊ∼ɪ əʊ∼ʊ əʊ∼ʌ ɛ∼æ ɛ∼ɪ ɛ∼ʌ ɛə∼iː ɛə∼æ ɛə∼ə ɛə∼ɛ ɛə∼ɪ ɛə∼ʊə ɛə∼ʌ ɜː∼ə ɜː∼əʊ ɜː∼ɪə ɜː∼ʊə ɪ∼iː ɪ∼æ ɪ∼ʊ ɪə∼iː ɪə∼æ ɪə∼ə ɪə∼ɛ ɪə∼ɪ ɪə∼ʊ ɪə∼ʊə ɪə∼ʌ ʊ∼ʌ ʊə∼iː ʊə∼uː ʊə∼ə ʊə∼ɛ ʊə∼ɪ ʊə∼ʊ ʊə∼ʌ

German	V∼V	ɛː∼eː	ai∼aː ai∼iː ai∼ɔy au∼ai au∼aː au∼ɪ au∼ʊ a∼ai a∼au a∼aː eː∼iː eː∼øː iː∼yː oː∼uː oː∼øː uː∼yː yː∼øː œ∼ɔ œ∼ɔy œ∼ɛ ɔy∼oː ɔy∼yː ɔy∼ʏ ɔ∼ɔy ə∼ai ə∼au ə∼ɛ ɛ∼ai ɛ∼ɪ ɛː∼ai ɛː∼ɛ ɪ∼ai ɪ∼ɔy ɪ∼ʊ ʊ∼ai ʏ∼yː
Dutch	C∼C	f∼v s∼z ɣ∼x	b∼d b∼g b∼m b∼p b∼v b∼w d∼g d∼l d∼n d∼t d∼z d∼ɹ dʒ∼d dʒ∼l dʒ∼n dʒ∼z dʒ∼ɹ f∼h f∼p f∼s f∼x g∼k g∼w h∼s h∼x j∼w j∼ɹ k∼p k∼t k∼x l∼n l∼z l∼ɹ m∼n m∼v m∼w n∼z n∼ɹ p∼t s∼t s∼x v∼w v∼z ŋ∼m ŋ∼n ɣ∼g ɣ∼v ɣ∼w ɣ∼z ɣ∼ŋ ɣ∼ʒ ɹ∼w ɹ∼z ʃ∼f ʃ∼h ʃ∼s ʃ∼t ʃ∼x ʃ∼z ʃ∼ʒ ʒ∼d ʒ∼l ʒ∼n ʒ∼s ʒ∼v ʒ∼z ʒ∼ɹ
French	V∼V	ø∼œ ø∼ɔ œ∼ɔ œ̃∼ɛ̃ ɛ∼e	a∼e a∼i e∼i e∼ə i∼y o∼u o∼ɔ̃ u∼y u∼ɔ̃ ø∼e ø∼o ø∼y ø∼ɔ̃ ø∼ɛ ø∼ɛ̃ œ∼e œ∼o œ∼y œ∼ɔ̃ œ∼ɛ œ̃∼e œ̃∼o œ̃∼y œ̃∼ø œ̃∼ɔ œ̃∼ɔ̃ œ̃∼ɛ ɑ̃∼a ɔ∼o ɔ∼u ɔ∼ɔ̃ ɛ∼a ɛ∼i ɛ∼ə ɛ̃∼a ɛ̃∼e ɛ̃∼i ɛ̃∼œ ɛ̃∼ə ɛ̃∼ɛ
Spanish	C∼C	ʝ∼ʎ θ∼s	b∼d b∼g b∼m b∼p b∼w d∼g d∼l d∼n d∼r f∼p f∼s f∼t g∼k g∼w j∼w k∼p k∼t l∼n l∼r m∼n m∼w m∼ɲ n∼r n∼ɲ p∼t ɾ∼d ɾ∼l ɾ∼n ɾ∼r ɾ∼ʎ ʎ∼d ʎ∼l ʎ∼n ʎ∼r ʝ∼w ʝ∼ɲ θ∼f θ∼p θ∼t
Slovak	C∼C	ʎ∼l	b∼g b∼m b∼n b∼p b∼v c∼d c∼k c∼p c∼t dz∼dʐ dz∼r dz∼z dz∼ʐ d∼k d∼p d∼s d∼t dʐ∼g dʐ∼l f∼p f∼s f∼v g∼k g∼l k∼p k∼t l:∼l m∼n m∼v n∼v p∼t r:∼r r∼z s∼t s∼z ts∼d ts∼dz ts∼k ts∼s ts∼t ts∼ts ts∼ʂ v∼z ɟ∼b ɟ∼g ɟ∼j ɟ∼ɲ ɦ∼v ɦ∼z ɲ∼j ɲ∼m ɲ∼n ʂ∼d ʂ∼f ʂ∼s ʂ∼t ʂ∼z ʂ∼ʐ ʎ∼j ʎ∼l: ʎ∼ɲ ʐ∼r ʐ∼s ʐ∼v ʐ∼z ʐ∼ɦ
	V∼V	æ∼a æ∼e	a:∼a a:∼o: a:∼u: a:∼ʊɔ a∼o a∼u e:∼e e:∼i: e:∼o: e∼i e∼o i:∼i i:∼u: i∼u o:∼o o:∼u: o:∼ʊɔ o∼u u:∼u u:∼ʊɔ æ∼i ʊɔ∼a ʊɔ∼e ʊɔ∼i ʊɔ∼o ʊɔ∼u

References

Åfarli, Tor A. 1995. A unified analysis of relative clause formation in Old Norse. In *Papers from the XVth Scandinavian conference*, Inger Moen, Hanne Gram Simonsen, & Helge Lødrup (eds.), pp. 533–543. Oslo: Department of Linguistics, University of Oslo.

Astås, Reidar. 2005. From Old Nordic to Early Modern Nordic: The language of the translations I: Icelandic and Norwegian translations. In *The Nordic languages*, Oskar Bandle, Kurt Braunmüller, Ernst Håkon Jahr, Allan Karker, Hans-Peter Naumann, & Ulf Teleman (eds.), ch. 132. Berlin: de Gruyter. DOI: `10.1515/9783110197068-014`.

Auderset, Sandra. 2020. Interrogatives as relativization markers in Indo-European. *Diachronica* 37(4). 474–513. DOI: `10.1075/dia.19030.aud`.

Aylett, Matthew & Alice Turk. 2004. The smooth signal redundancy hypothesis: A functional explanation for relationships between redundancy, prosodic prominence and duration in spontaneous speech. *Language and Speech* 47. 31–56. DOI: `10.1177/00238309040470010201`.

Baayen, R. Harald, Richard Piepenbrock, & Léon Gulikers (eds.). 1995. *CELEX2 (LDC96L14)*. Philadelphia: Linguistic Data Consortium.

Barðdal, Jóhanna. 2011. The rise of dative substitution in the history of Icelandic: A diachronic construction grammar account. *Lingua* 121(1). 60–79. DOI: `10.1016/j.lingua.2010.07.007`.

Barðdal, Jóhanna & Þórhallur Eyþórsson. 2003. The change that never happened: The story of oblique subjects. *Journal of Linguistics* 39(3). 439–472. DOI: `10.1017/s002222670300207x`.

Bates, Douglas, Martin Mächler, Benjamin Bolker, & Steve Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67(1). 1–48. DOI: `10.18637/jss.v067.i01`.

Benediktsson, Hreinn. 1972. *The First Grammatical Treatise*. Reykjavík: Institute of Nordic Linguistics.

Bird, Steven, Edward Loper, & Ewan Klein. 2009. *Natural Language Processing with Python*. Sebastopol, CA: O'Reilly. ISBN: 0596516495.

Blevins, Juliette. 2005. The role of phonological predictability in sound change: Privileged reduction in Oceanic reduplicated substrings. *Oceanic Linguistics* 44(2). 517–526. DOI: `10.1353/ol.2005.0028`.

Blevins, Juliette & Andrew Wedel. 2009. Inhibited sound change: An evolutionary approach to lexical competition. *Diachronica* 26(2). 143–183. DOI: `10.1075/dia.26.2.01ble`.

Brendel, Christian D. 2018. *Cohort effects and asymmetrical word-level sound change*. Santa Barbara, CA: University of California, Santa Barbara MA thesis.

Brendel, Christian D. 2023. Translation-induced interrogative relativizers and stability in Icelandic. *Transactions of the Philological Society* 121(1). 1–32. DOI: `10.1111/1467-968X.12258`.

Bruhn, Daniel, John Lowe, David Mortensen, & Dominic Yu. 2015. *Sino-Tibetan etymological dictionary and thesaurus database software (STEDT)*. DOI: 10.6078/d1159q.

Brysbaert, Marc, Matthias Buchmeier, Markus Conrad, Arthur M. Jacobs, Jens Bölte, & Andrea Böhl. 2011. The word frequency effect: A review of recent developments and implications for the choice of frequency estimates in German. *Experimental Psychology* 58(5). 412–424. DOI: 10.1027/1618-3169/a000123.

Burger, Christoph. 2014. Luther's thought took shape in translations of scripture and hymns. In *The Oxford handbook of Martin Luther's theology*, Robert Kolb, Irene Dingel, & L'ubomír Batka (eds.), ch. 35, pp. 481–488. Oxford: Oxford University Press. DOI: 10.1093/oxfordhb/9780199604708.001.0001.

Bybee, Joan L. 2007. *Frequency of use and the organization of language*. Oxford: Oxford University Press. DOI: 10.1093/acprof:oso/9780195301571.001.0001.

Cleasby, Richard & Guðbrandur Vigfússon. 1957. *An Icelandic–English dictionary*. Ed. by William A. Craigie. 2nd edn. Oxford: Clarendon Press.

Cohen Priva, Usama. 2012. *Sign and signal: Deriving linguistic generalizations from information utility*. Stanford: Stanford University Doctoral dissertation.

Cole, Ronald A. & Jola Jakimik. 1980. A model of speech perception. In *Perception and production of fluent speech*, Ronald A. Cole (ed.), pp. 133–164. Hillsdale, NJ: Lawrence Erlbaum.

Comrie, Bernard. 2006. Syntactic typology: Just how exotic ARE European-type relative clauses? In *Linguistic universals*, Ricardo Mairal & Juana Gil (eds.), ch. 6, pp. 130–154. Cambridge: Cambridge University Press.

Cutler, Anne. 1989. Auditory lexical access: Where do we start? In *Lexical representation and process*, William D. Marslen-Wilson (ed.), ch. 11, pp. 342–356. Cambridge, MA: MIT Press. DOI: 10.7551/mitpress/4213.003.0016.

Cutler, Anne & Hsuan-Chih Chen. 1997. Lexical tone in Cantonese spoken word processing. *Perception and Psychophysics* 59. 165–179. DOI: 10.3758/BF03211886.

Dal, Ingerid & Hans-Werner Eroms. 2014. *Kurze deutsche Syntax auf historischer Grundlage* [Short German syntax on a historical basis]. Berlin: DeGruyter.

Damerau, Fred J. 1964. A technique for computer detection and correction of spelling errors. *Communications of the Association for Computing Machinery (ACM)* 7(3). 171–176. DOI: 10.1145/363958.363994.

Dryer, Matthew S. 2013. Prefixing vs. suffixing in inflectional morphology. In *The world atlas of language structures online (WALS)*, Matthew S. Dryer & Martin Haspelmath (eds.). Leipzig: Max Planck Institute for Evolutionary Anthropology. URL: https://wals.info/chapter/26.

Dunn, Michael. 2012. *Indo-European lexical cognacy database (IELex)*. Max Planck Institute for Psycholinguistics. URL: https://ielex.mpi.nl.

Dyen, Isidore, Joseph Kruskal, & Paul Black. 1992. An Indoeuropean classification: A lexicostatistical experiment. *Transactions of the American Philosophical Society* 82(5). 1–132. DOI: 10.2307/1006517.

Eggertsdóttir, Margrét. 2006. From Reformation to Enlightenment. In *A history of Icelandic literature*, Daisy Neijmann (ed.). Lincoln, NE: University of Nebraska Press.

Einarsson, Stefán. 1945. *Icelandic: Grammar, texts, and glossary*. Baltimore: Johns Hopkins.

Eyþórsson, Þórhallur. 2000. Dative vs. nominative: Changes in quirky subjects in Icelandic. *Leeds Working Papers in Linguistics and Phonetics* 8. 27–44.

Faarlund, Jan Terje. 2004. *The syntax of Old Norse*. Oxford: Oxford University Press. DOI: 10.1093/acprof:oso/9780199235599.001.0001.

Fagyal, Zsuzsanna, Douglas Kibbee, & Fred Jenkins. 2006. *French: A linguistic introduction*. Cambridge: Cambridge University Press. DOI: 10.1017/CBO9780511791185.

Gardani, Francesco, Peter Arkadiev, & Nino Amiridze (eds.). 2015. *Borrowed morphology*. (Language Contact and Bilingualism 8). Berlin: de Gruyter. DOI: 10.1515/9781614513209.

Gardiner, Shayna & Naomi Nagy. 2017. Stable variation vs. language change and the factors that constrain them. *University of Pennsylvania Working Papers in Linguistics* 23.2(10).

Gilliéron, Jules. 1918. *Genéalogie des mots qui designent l'abeille* [Genealogy of words that refer to the bee]. Paris: Champion.

Gorman, Kyle, Gleb Mazovetskiy, & Vitaly Nikolaev. 2018. Improving homograph disambiguation with supervised machine learning. In *Proceedings of the 11th International Conference on Language Resources and Evaluation (LREC)*. European Language Resources Association (ELRA). ISBN: 979-10-95546-00-9.

Hall, Christopher. 1988. Integrating diachronic and processing principles in explaining the suffixing preference. In *Explaining language universals*, John A. Hawkins (ed.), pp. 321–349. Oxford: Blackwell.

Hamilton, William L., Jure Leskovec, & Dan Jurafsky. 2016. Diachronic word embeddings reveal statistical laws of semantic change. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (ACL)*, pp. 1489–1501. Berlin: ACL. DOI: 10.18653/v1/P16-1141.

Harbert, Wayne. 2006. *The Germanic languages*. Cambridge: Cambridge University Press. DOI: 10.1017/CBO9780511755071.

Harris, James. 1969. *Spanish phonology*. Cambridge, MA: MIT Press.

Haspelmath, Martin. 2011. The indeterminacy of word segmentation and the nature of morphology and syntax. *Folia Linguistica* 45(1). 31–80. DOI: 10.1515/flin.2011.002.

Haugen, Einar & Thomas L. Markey. 1972. *The Scandinavian languages: Fifty years of linguistic research (1918–1968)*. The Hague: Mouton. DOI: 10.1515/9783110870664.

Hawkins, John A. & Anne Cutler. 1988. Psycholinguistic factors in morphological asymmetry. In *Explaining language universals*, John A. Hawkins (ed.), pp. 280–317. Oxford: Blackwell.

Hay, Jennifer B., Janet B. Pierrehumbert, Abby J. Walker, & Patrick LaShell. 2015. Tracking word frequency effects through 130 years of sound change. *Cognition* 139. 83–91. DOI: 10.1016/j.cognition.2015.02.012.

Heggarty, Paul, Cormac Anderson, & Matthew Scarborough (eds.). 2024. *Indo-European cognate relationships database (IE-CoR version 1.2)*. Leipzig: Max Planck Institute for Evolutionary Anthropology. URL: `https://iecor.clld.org/`.

Heine, Bernd & Tania Kuteva. 2003. On contact-induced grammaticalization. *Studies in Language* 27(3). 529–572. DOI: `10.1075/sl.27.3.04hei`.

Heine, Bernd & Tania Kuteva. 2006. *The changing languages of Europe*. Oxford: Oxford University Press. DOI: `10.1093/acprof:oso/9780199297337.001.0001`.

Helgason, Jón. 1929. *Málið á nýja testamenti Odds Gottskálkssonar* [The language in the New Testament of Oddur Gottskálksson]. Copenhagen: S. L. Møller.

Hockett, Charles F. 1955. *A manual of phonology*. (International Journal of American Linguistics 21.4). Bloomington: Indiana University Publications in Anthropology & Linguistics.

Hockett, Charles F. 1967. The quantification of functional load. *Word* 23(1-3). 320–339. DOI: `10.1080/00437956.1967.11435484`.

Jónsson, Jóhannes Gísli. 2017. Why *sem* is (still) a complementizer and not a relative pronoun. In *A Schrift to fest Kyle Johnson*, Nicholas LaCara, Keir Moulton, & Anne-Michelle Tessier (eds.), pp. 169–175. (Linguistics Open Access Publications 1). Amherst: University of Massachusetts Amherst. DOI: `10.7275/R57D2S95`.

Jurafsky, Daniel, Alan Bell, Michelle Gregory, & William D. Raymond. 2001. Probabilistic relations between words: Evidence from reduction in lexical production. In *Frequency and the emergence of linguistic structure*, Joan Bybee & Paul Hopper (eds.), pp. 229–254. (Typological Studies in Language vol. 45). Amsterdam: John Benjamins. DOI: `10.1075/tsl.45.13jur`.

Kalinke, Marianne E. 1996. *The Book of Reykjahólar: The last of the great medieval legendaries*. Toronto: University of Toronto Press. ISBN: 0802078141.

Karlsson, Gunnar. 2000. *The history of Iceland*. Minneapolis: University of Minneapolis Press. ISBN: 0816635897.

Karlsson, Stefán. 1989. Brot úr barnaprédikunum í þýðingu Odds Gottskálkssonar [Fragment of the children's sermons translated by Oddur Gottskálksson]. *Árbók Landsbókasafns Íslands: Nýr flokkur* 15. 43–72. URL: `https://timarit.is/page/3138857#page/n145/mode/2up`. rev. 1991.

Karlsson, Stefán. 2004. *The Icelandic language*. Trans. by Rory McTurk. Exeter: Short Run Press Ltd. ISBN: 090352161X.

Keenan, Edward L. & Bernard Comrie. 1977. Noun Phrase accessibility and Universal Grammar. *Linguistic Inquiry* 8(1). 63–99.

King, Robert. 1967a. A measure for functional load. *Studia Linguistica* 21(1). 1–14. DOI: `10.1111/j.1467-9582.1967.tb00545.x`.

King, Robert. 1967b. Functional load and sound change. *Language* 21(43). 831–852. DOI: `10.2307/411969`.

Kissine, Mikhail, Hans van de Velde, & Roeland van Hout. 2003. An acoustic study of standard Dutch /v/, /f/, /s/, and /z/. In *Linguistics in the Netherlands*, Leonie Cornips & Paula Fikkert (eds.), pp. 93–104. Amsterdam: John Benjamins. ISBN: 1588114627.

König, Ekkehard & Johan van der Auwera (eds.). 1994. *The Germanic languages*. New York: Routledge. DOI: 10.4324/9781315812786.
Krajčovič, Rudolf. 1988. *Vývin Slovenského jazyka a dialektológia* [Development of Slovak language and dialectology]. Bratislava: Slovenské Pedagogické Naklatel'stvo.
Krech, Eva-Maria, Eberhard Stock, Ursula Hirschfeld, & Lutz-Christian Anders. 2009. *Deutsches Aussprachewörterbuch* [German pronunciation dictionary]. Berlin: de Gruyter Mouton. DOI: doi:10.1515/9783110215564.
Kvaran, Guðrún. 1996. Þættir úr sögu orðaforðans [On the history of the Icelandic vocabulary]. Í *Erindi um íslenskt mál*, bls. 35–48. Reykjavík: Íslenska málfraeðifélagið.
Kvaran, Guðrún. 2015. Brautryðjendur í biblíuþýðingum á Íslandi: Um þátt þeirra Odds Gottskálkssonar og Gissurar Einarssonar [Pioneers of Bible translation in Iceland: On the role of Oddur Gottskálksson and Gissur Einarsson]. *Ritröð guðfraeðistofnunnar* 41.
Labov, William. 1994. *Principles of language change: Internal factors*. Oxford: Blackwell.
Labov, William, Sharon Ash, & Charles Boberg. 2006. *The atlas of North American English: Phonetics, phonology and sound change*. Berlin: Mouton de Gruyter. DOI: 10.1515/9783110167467.
Lee, Chao-Yang. 2007. Does Horse activate Mother? Processing lexical tone in form priming. *Language and Speech* 50(1). 101–123. DOI: 10.1177/00238309070500010501.
Lehmann, Christian. 1984. *Der Relativsatz* [The relative clause]. Tübingen: Narr.
Leonard, Stephen Pax. 2012. *Language, society and identity in early Iceland*. (Publications of the Philological Society 45). Oxford: Wiley-Blackwell.
Levenshtein, Vladimir I. 1965. Двоичные коды с исправлением выпадений, вставок и замещений символов [Binary codes capable of correcting deletions, insertions and reversals]. Доклады Академий Наук СССР [*Proceedings of the USSR Academy of Sciences*]*. 163(4). 845–848.
Levy, Omer & Yoav Goldberg. 2014. Dependency-based word embeddings. In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (ACL)*, Kristina Toutanova & Hua Wu (eds.), pp. 302–308. Baltimore: ACL. DOI: 10.3115/v1/P14-2050.
Lison, Pierre & Andrey Kutuzov. 2017. Redefining context windows for word embedding models: An experimental study. In *Proceedings of the 21st Nordic Conference on Computational Linguistics*, pp. 284–288.
Lison, Pierre & Jörg Tiedemann. 2016. OpenSubtitles2016: Extracting large parallel corpora from movie and TV subtitles. In *Proceedings of the 10th International Conference on Language Resources and Evaluation (LREC)*, pp. 923–929. European Language Resources Association (ELRA).
List, Johann-Mattis & Robert Forkel. 2024. *LingPy. A Python library for historical linguistics. Version 2.6.13*. DOI: 10.5281/zenodo.5144474. URL: https://lingpy.org.
Loftsdóttir, Kristín. 2019. *Crisis and coloniality at Europe's margins: Creating exotic Iceland*. Oxon: Routledge.
Malins, Jeffrey & Marc Joanisse. 2012. Setting the tone: An ERP investigation of the influences of phonological similarity on spoken word recognition in Mandarin Chinese.

Neuropsychologia 50(8). 2032–2043.
DOI: 10.1016/j.neuropsychologia.2012.05.002.

Marcus, Mitcell, Beatrice Santorini, & Mary Ann Marcinkiewicz. 1993. Building a large annotated corpus of English: The Penn Treebank. *Computational Linguistics* 19. 313–330. URL: https://www.ling.upenn.edu/~beatrice/corpus-ling/annotation/.

Marslen-Wilson, William D. 1984. Function and process in spoken word recognition. In *Attention & performance X*, H. Bouma & D. Bouwhuis (eds.), pp. 125–150. Hillsdale, NJ: Lawrence Erlbaum.

Marslen-Wilson, William D. 1987. Functional parallelism in spoken word-recognition. *Cognition* 25(1). 71–102. DOI: 10.1016/0010-0277(87)90005-9.

Marslen-Wilson, William D. & Lorraine K. Tyler. 1980. The temporal structure of spoken language understanding. *Cognition* 1. 1–71. DOI: 10.1016/0010-0277(80)90015-3.

Marslen-Wilson, William D. & Alan Welsh. 1978. Processing interactions and lexical access during word-recognition in continuous speech. *Cognitive Psychology* 10. 29–63. DOI: 10.1016/0010-0285(78)90018-X.

Martinet, André. 1955. *Économie des changements phonétiques: Traité de phonologie diachronique* [Economies of phonetic changes: Treatise on diachronic phonology]. Bern: A. Francke.

Martinet, Andre. 1952. Function, structure, and sound change. *Word* 8 (1). 1–32. DOI: 10.1080/00437956.1952.11659416.

Matras, Yaron. 2009. Converging structures: Pattern replication. In *Language contact*, pp. 234–274. (Cambridge Textbooks in Linguistics). Cambridge: Cambridge University Press. DOI: 10.1017/CBO9780511809873.010.

Mendonça, Ângelo, David Graff, & Denise DiPersio. 2011. *Spanish Gigaword Third Edition LDC2011T12*. Web Download. Philadelphia. DOI: 10.35111/v6z1-cb91.

Meyerstein, R. S. 1970. Functional load: Descriptive limitations, alternatives of assessment and extensions of application. *Janua Linguarum* 99. DOI: 10.1515/9783111354163.

Mikolov, Tomas, Kai Chen, Greg Corrado, & Jeffrey Dean. 2013a. Efficient estimation of word representations in vector space. *arXiv preprint*. DOI: 10.48550/arXiv.1301.3781.

Mikolov, Tomas, Wen-tau Yih, & Geoffrey Zweig. 2013b. Linguistic regularities in continuous space word representations. In *Proceedings of the 2013 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL-HLT)*, pp. 746–751. Atlanta: Association for Computational Linguistics (ACL).

Mithun, Marianne. 2012. Questionable relatives. In *Typological aspects of relative clauses*, Bernard Comrie & Zarina Fernandez (eds.). Amsterdam: John Benjamins. DOI: 10.1075/tsl.102.13mit.

Mithun, Marianne. 2013. Challenges and benefits of contact among relatives: Morphological copying. *Journal of Language Contact* 6(2). 243–270. DOI: 10.1163/19552629-00602003.

Moscoso del Prado Martín, Fermín & Christian Brendel. 2016. Case and cause in Icelandic: Reconstructing causal networks of cascaded language changes. In *Proceedings of the 54th*

Annual Meeting of the Association for Computational Linguistics (ACL), Katrin Erk & Noah A. Smith (eds.), pp. 2421–2430. Berlin: ACL. DOI: 10.18653/v1/P16-1229.

New, Boris, Christophe Pallier, & L. Ferrand (eds.). 2005. *Lexique 3*. URL: http://www.lexique.org/.

Nichols, Johanna. 2003. Diversity and stability in language. In *The handbook of historical linguistics*, Brian D. Joseph & Richard D. Janda (eds.), ch. 5, pp. 283–310. DOI: 10.1002/9780470756393.ch5.

Universal Dependencies v2: An evergrowing multilingual treebank collection (2020). In *12th International Conference on Language Resources and Evaluation (LREC)*, Joakim Nivre et al. (eds.), pp. 4034–4043. Marseille: European Language Resources Association (ELRA). ISBN: 979-10-95546-34-4.

Nygaard, Marius. 1905. *Norrøn syntax* [Norse syntax]. Christiania: Aschehoug.

Penny, Ralph J. 2002. *A history of the Spanish language*. 2nd edn. Cambridge: Cambridge University Press. DOI: 10.1017/CBO9780511992827.

Platzack, Christer. 2002. Relativization in the Germanic languages, with particular emphasis on Scandinavian. In *Relativisation on the North Sea littoral*, Patricia Poussa (ed.), pp. 77–96. (LINCOM Studies in Language Typology (LSLT) 7). Munich: LINCOM Europa. ISBN: 9783895863660.

Pompei, Anna. 2011. Relative clauses. In *New perspectives on historical Latin syntax: Complex sentences, grammaticalization, typology*, Philip Baldi & Pierluigi Cuzzolin (eds.), pp. 427–548. Berlin: de Gruyter Mouton. DOI: 10.1515/9783110253412.427.

Poss, Nick, Tsun-Hui Hung, & Udo Will. 2008. The effects of tonal information on lexical activation in Mandarin. In *Proceedings of the 20th North American Conference on Chinese Linguistics (NACCL)*, Marjorie K. M. Chan & Hana Kang (eds.), pp. 205–211. Columbus: Ohio State University. ISBN: 978-0-9824715-0-0.

Sakel, Jeanette. 2007. Types of loan: Matter and pattern. In *Grammatical borrowing in a cross-linguistic perspective*, Yaron Matras & Jeanette Sakel (eds.). (Empirical Approaches to Language Typology 38). Berlin: de Gruyter. DOI: 10.1515/9783110199192.15.

Sanders, Nathan C. & Steven B. Chin. 2009. Phonological distance measures. *Journal of Quantitative Linguistics* 16(1). 96–114. DOI: 10.1080/09296170802514138.

Sankoff, Gillian & Hélène Blondeau. 2007. Language change across the lifespan: /r/ in Montreal French. *Language* 83. 560–588. DOI: 10.1353/lan.2007.0106.

Sapir, Edward. 1921. *Language*. New York: Harcourt Brace & Company. DOI: 10.1017/CBO9781139629430.

Sapp, Christopher D. 2019. Arrested development: Case attraction as a transitional stage from Old Icelandic demonstrative to relative *sá*. *Language* 95(1). e1–e40. DOI: 10.1353/lan.2019.0007.

Shannon, Claude E. 1948. The mathematical theory of communication. *The Bell system technical journal* 27(3). 379–423. DOI: 10.1063/1.3067010.

Sihler, Andrew L. 1995. *New comparative grammar of Greek and Latin*. Oxford: Oxford University Press. DOI: 10.1093/oso/9780195083453.001.0001.

Šimková, Mária. 2006. Slovak National Corpus—History and current situation. In *Insight into the Slovak and Czech corpus linguistics*, Mária Šimková (ed.), pp. 151–159. Bratislava: Veda. ISBN: 80-224-0880-8.

Smits, R. J. C. 1989. *Eurogrammar: The relative and cleft constructions of the Germanic and Romance languages*. Dordrecht: Foris. DOI: 10.1515/9783110882704.

Surendran, Dinoj & Partha Niyogi. 2003. Measuring the functional load of phonological contrasts. Technical report TR-2003-12. Chicago: University of Chicago. DOI: 10.48550/arXiv.cs/0311036.

Surendran, Dinoj & Partha Niyogi. 2006. Quantifying the functional load of phonemic oppositions, distinctive features, and suprasegmentals. In *Competing models of linguistic change: Evolution and beyond*, Ole Nedergaard Thomsen (ed.), pp. 43–58. (Current Issues in Linguistic Theory 279). Amsterdam: John Benjamins. DOI: 10.1075/cilt.279.05sur.

Swadesh, Morris. 1952. Lexicostatistic dating of prehistoric ethnic contacts. In *Proceedings of the American Philosophical Society*, pp. 452–463. Philadelphia: University of Pennsylvania Press.

Taft, Marcus & Gail Hambly. 1986. Exploring the cohort model of spoken word recognition. *Cognition* 22(3). 259–282. DOI: 10.1016/0010-0277(86)90017-X.

Tallman, Adam J. R. 2021. Constituency and coincidence in Chácobo (Pano). *Studies in Language* 45(2). 321–383. DOI: 10.1075/sl.19025.tal.

Taulé, Mariona, M. Antònia Martí, & Marta Recasens. 2008. Ancora: Multilevel annotated corpora for Catalan and Spanish. In *Proceedings of the 6th International Conference on Language Resources and Evaluation (LREC)*, Nicoletta Calzolari, Khalid Choukri, Bente Maegaard, Joseph Mariani, Jan Odijk, Stelios Piperidis, & Daniel Tapias (eds.). Marrakesh: European Language Resources Association (ELRA).

Thomason, Sarah Grey. 2014. Contact-induced language change and typological congruence. In *Congruence in contact-induced language change: Language families, typological resemblance, and perceived similarity*, Juliane Besters-Dilger, Cynthia Dermarkar, Stefan Pfänder, & Achim Rabus (eds.), pp. 201–218. Berlin: de Gruyter. DOI: 10.1515/9783110338454.201.

Thomason, Sarah Grey. 2015. When is the diffusion of inflectional morphology not dispreferred? In *Borrowed morphology*, Francesco Gardani, Peter Arkadiev, & Nino Amiridze (eds.), pp. 27–46. (Language Contact and Bilingualism 8). Berlin: de Gruyter. DOI: 10.1515/9781614513209.27.

Thomason, Sarah Grey & Terrence Kaufman. 1988. *Language contact, creolization, and genetic linguistics*. Berkeley: University of California Press. DOI: 10.1525/9780520912793.

Tiedemann, Jörg. 2012. Parallel data, tools and interfaces in OPUS. In *Proceedings of the 8th International Conference on Language Resources and Evaluation (LREC)*, Nicoletta Calzolari et al. (eds.). Istanbul: European Language Resources Association (ELRA). ISBN: 978-2-9517408-7-7.

Wagener, Terje T. 2017. *The history of Nordic relative clauses*. Berlin: de Gruyter Mouton. DOI: 10.1515/9783110496536.

Wal, Marijke J. van der & Aad Quak. 1994. Old and Middle Continental West Germanic. In *The Germanic languages*, Ekkehard König & Johan van der Auwera (eds.), pp. 72–109. New York: Routledge. DOI: 10.4324/9781315812786.

Wallenberg, Joel C. 2019. A variational theory of specialization in acquisition and diachrony. In *The determinants of diachronic stability*, Anne Breitbarth, Miriam Bouzouita, Lieven Danckaert, & Melissa Farasyn (eds.), pp. 245–262. (Linguistics Today 254). Amsterdam: John Benjamins. DOI: 10.1075/la.254.10wal.

Wallenberg, Joel C., Anton K. Ingason, Einar F. Sigurðsson, & Eiríkur Rögnvaldsson (eds.). 2011. *Icelandic parsed historical corpus (IcePaHC) 0.9*. URL: https://linguist.is/wiki.

Wedel, Andrew. 2012. Lexical contrast maintenance and the development of sublexical contrast systems. *Language and Cognition* 4(4). 319–355.

Wedel, Andrew, Scott Jackson, & Abby Kaplan. 2013a. Functional load and the lexicon: Evidence that syntactic category and frequency relationships in minimal lemma pairs predict the loss of phoneme contrasts in language change. *Language and Speech* 56(3). 395–417. DOI: 10.1177/0023830913489096.

Wedel, Andrew, Abby Kaplan, & Scott Jackson. 2013b. High functional load inhibits phonological contrast loss: A corpus study. *Cognition* 128(2). 179–186. DOI: 10.1016/j.cognition.2013.03.002.

Wedel, Andrew, Adam Ussishkin, & Adam King. 2019. Crosslinguistic evidence for a strong statistical universal: Phonological neutralization targets word-ends over beginnings. *Language* 95(4). e428–e446. DOI: 10.1353/lan.2019.0086.

Weide, Robert L. 1995. *The CMU pronouncing dictionary*. URL: http://www.speech.cs.cmu.edu/cgi-bin/cmudict.

Weinreich, Uriel, William Labov, & Marvin. I Herzog. 1968. *Empirical foundations for a theory of language change*. In *Directions for historical linguistics*, Winfred P. Lehmann & Yakov Malkiel (eds.), pp. 95–195. Austin: University of Texas Press.

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

Wichmann, Søren, Eric W. Holman, Dik Bakker, & Cecil H. Brown. 2010. Evaluating linguistic distance measures. *Physica A: Statistical Mechanics and its Applications* 389(17). 3632–3639. DOI: 10.1016/j.physa.2010.05.011.

Wieling, Martijn, John Nerbonne, Jelke Bloem, Charlotte Gooskens, Wilbert Heeringa, & R. Harald Baayen. 2014. A cognitively grounded measure of pronunciation distance. *PLoS ONE* 9(1). e75734. DOI: 10.1371/journal.pone.0075734.

Wiese, Richard. 2000. *The phonology of German*. Oxford: Oxford University Press. ISBN: 9780198299509.

Wolf, Thomas et al. 2020. Transformers: State-of-the-art Natural Language Processing. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language*

Processing: System Demonstrations, pp. 38–45. Online: Association for Computational Linguistics (ACL). DOI: 10.18653/v1/2020.emnlp-demos.6.
Xue, Linting et al. 2022. ByT5: Towards a token-free future with pre-trained byte-to-byte models. *Transactions of the Association for Computational Linguistics (TACL)* 10. 291–306. DOI: 10.1162/tacl_a_00461.
Yang, Charles. 2000. Internal and external forces in language change. *Language Variation and Change* 3(12). 231–250. DOI: 10.1017/S0954394500123014.
Zipf, George Kingsley. 1949. *Human behavior and the principle of least effort*. Cambridge, MA: Addison-Wesley Press.
Þráinsson, Höskuldur. 1980. Tilvísunarfornöfn? [Relative pronouns?]. *Íslenskt mál og almenn málfraeði* 2. 53–96. URL: https://timarit.is/view_page_init.jsp?pubId=832.
Þráinsson, Höskuldur. 1994. Icelandic. In *The Germanic languages*, Ekkehard König & Johan van der Auwera (eds.). New York: Routledge. DOI: 10.4324/9781315812786.
Þráinsson, Höskuldur. 2007. *The syntax of Icelandic*. Cambridge: Cambridge University Press. DOI: 10.1017/CBO9780511619441.