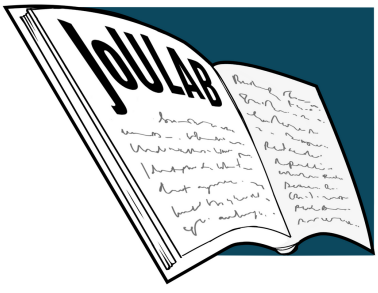
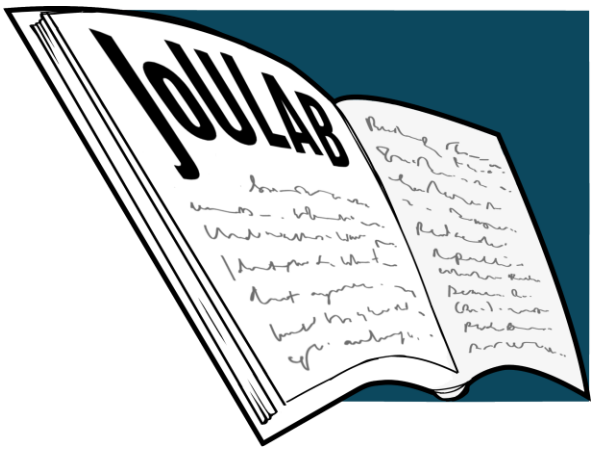
Vol.1, Issue 1
Spring 2021

ISSN 2754-0820

Journal of the Undergraduate Linguistics Association of Britain



JOURNAL OF THE UNDERGRADUATE LINGUISTICS ASSOCIATION OF BRITAIN



The *Journal of the Undergraduate Linguistics Association of Britain* (ISSN 2754-0820) was founded in July 2020 as a Subcommittee of the Undergraduate Linguistics Association of Britain. It is the only academic journal in the world taking submissions solely from undergraduates in any area of linguistics. The Journal exists to provide a forum for the publication of exceptional undergraduate research in linguistics for students across the globe and from any background. We aim to publish two issues per twelve-month period, with each volume corresponding to the Editorial Committee that oversees its production. Every manuscript is peer-reviewed over multiple rounds by two members of the Board of Reviewers, which consists of doctoral students in linguistics from a plethora of countries and institutions.

Editorial Committee of the Journal of the Undergraduate Linguistics Association of Britain

Selected Members of the Board of Reviewers

Everyone on the Board of Reviewers is currently undertaking (or was undertaking at their time of recruitment) doctoral research into linguistics or a related field. The Board consists of over forty individuals from institutions across the world, with specialisations across all areas of linguistics. The list given here represents those who consented to being included in Vol 1., Issue 1.

Reviewing Policies

All submissions to the Journal are subject to a double-blind, peer-review process. This means that the reviewers for any manuscript and that manuscript's author are unable to access identifying information about each other. Each paper is first anonymously assessed by the Editorial Board to ensure its scope meets the reviewing ability of the Journal before it is then passed on to two reviewers who provide in-depth comments through reviewing rounds.

Copyright Disclosure

Colour Blindness Declaration

All formatting within this issue has been designed specifically to be colour-blind friendly. For example, figures and tables are designed to ensure accessibility for those with monochromacy. This is part of the Journal's commitment to maximum accessibility. Please contact the Journal Editorial Committee should you have any access difficulties with this publication.

Acknowledgements

The Journal would first like to express its gratitude for the dedication and effort shown by all members of the Board of Reviewers throughout any round of review in which they were involved. Thanks must also be paid to Associate Editor Lydia Wiernik, whose impressive flair and aptitude in graphic design is chiefly responsible for the quality of our new logo, as well as the front and back covers of this publication, as well as to past Committee members Marius Henius Dreijer and Maisy Hallam for their past efforts.

Website and Social Media Links

Website: www.ulab.org.uk/journal
Twitter: @ULAB_Journal

Contents

National Committee of the Undergraduate Linguistics Association of Britain, 2021-22

NATIONAL CHAIR	*Cliodhna Hughes*
NATIONAL VICE CHAIR	*Hafren Vaughan*
NATIONAL SECRETARY; WEBMASTER	*Louis Van Steene*
NATIONAL TREASURER; ACCESSIBILITY OFFICER	*Beatrix Livesey-Stephens*
ARCHIVIST	*Lydia Wiernik*
SOCIAL MEDIA COORDINATOR; LOCAL CHAIR	*Riley Crouch*
SOCIAL MEDIA COORDINATOR	*Roma Dhasmana*
JOURNAL EDITOR	*T. R. Williamson*
MAGAZINE EDITOR-IN-CHIEF	*Stephanie Jat*
EVENTS COORDINATOR; OPPORTUNITIES COORDINATOR	*Eloise Parr*
INSTITUTIONAL REPRESENTATIVE COORDINATOR	*Caitlin Wilson*

Board of Institutional Representatives, 2021-22

UNIVERSITY OF ABERDEEN	*Roma Dhasmana*
UNIVERSITY OF ALICANTE	*Tina Wolff*
UNIVERSITY OF BIRMINGHAM	*Eloise Parr*
UNIVERSITY OF CAMBRIDGE	*James Morley*
UNIVERSITY OF EDINBURGH	*Michael Gössler*
UNIVERSITY OF OREGON	*Jaidan McLean*
UNIVERSITY OF ULSTER	*Grace Cotton*

Foreword from the Editor and the Head of the Board of Reviewers

It gives us enormous pleasure to introduce the first issue of the Journal of the Undergraduate Linguistics Association of Britain. Since the Journal's inception almost twelve months ago, we have been working tirelessly to prepare our first publication, and we are very excited to present it to you.

The phrase 'easier said than done' is one that applies to a wealth of ambitious projects. From our experience, no less applicable could the phrase be to the establishment of an undergraduate academic journal in linguistics. When we first considered founding JoULAB, it was our immediate concern to provide an opportunity for undergraduates to have their research published in light of the cancellation of the ULAB 2020 conference. We knew that opportunities like presenting at the annual ULAB conference are hard to come by and felt duty-bound to fill that gap.

Indeed, 'not easy' is a very British understatement for how founding JoULAB was. The discussions alone over how to organise submissions, reviewing, copyediting, and publishing must have taken hundreds of hours collectively of our spare time. Acting on those decisions, in whatever form that might have entailed, took even longer. This is because it has always been central to our mission to ensure the highest standards for everything we do. We have aimed to make all policies water-tight and all procedures fair. Every single reviewer met privately with at least one of the Editorial Committee to discuss the specific nature of reviewing for undergraduates and for JoULAB. All papers were anonymously reviewed by two expert PhD students and the Editorial Committee over the course of at least three reviewing rounds. Each accepted manuscript has been painstakingly copyedited at least four times by one of our Associate Copyeditors. The list goes on.

It is in this manner that we aim to continue. JoULAB has only just begun, and we aim to grow further. When we started, the Editorial Committee consisted of just four individuals. Now, over fifty make up both the Committee and the Board of Reviewers – a proudly international body with members from over twenty different institutions. There are certainly more people to thank than space to do so, and even more to be thanked than those listed on Page 4 of this issue. Special mention should be given to Cliodhna Hughes for her industrial and tireless administrative capacities – on many occasions, she has been vital in ensuring the continued efficiency of the Journal. The later-recruited Secretary Marius Henius Dreijer alongside Associate Copyeditors Maisy Hallam, Eleonora Kacl, and Rachel Shannon (the latter two remaining with us) all played vital roles in supporting the early stages of the Journal's foundation.

Inside Issue 1, you will find a fascinating array of world-class undergraduate research in linguistics. Nina Haket's paper, *Language Contact and the Phylogeny and Phonology of Early English*, brings a critical eye to a recent claim from historical linguistics. She sets out to evaluate a controversial position taken by Emonds and Faarlund that suggests Modern English is actually a North Germanic language, originating from Old Norse, rather than the traditional view that it is West Germanic, and thus that it comes from Old English. To do so, Haket takes an approach that specifically considers the phonologies of the languages implicated and finds that the only way to reconcile this controversial perspective with current theory is to posit an impossible set of phonological mappings from Old Norse to Old English. Such a conclusion rejects Emonds and Faarlund's claim, which will certainly fascinate those with an interest in historical linguistics, language change, and language contact.

The Palatalisation of the Voiceless Velar Fricative in Santiago, Chile: A Variationist Analysis, Madeleine Rees' sociophonetic paper, investigates the recent development of palatalisation for /x/ amongst Spanish speakers in Santiago, Chile. Rees first finds that palatalisation of /x/ is influenced by the vowel that follows; there was notably more palatalisation before /i/ than before /e/. More surprising, though, is the finding that palatalisation actually increased in careful speech; the more formal the context, the more this purportedly-informal phenomenon was observed. When Rees changed the context from formal to casual, it was even noted that male speakers stopped palatalising /x/ more than female speakers, but only before /e/. These findings of social and linguistic variations in producing /x/ will be interesting reading for anyone interested in sociolinguistics, phonetics, phonology, and Hispanic linguistics.

Issue 1's final paper, *The Graded Co-Salience Hypothesis for Polysemous Ambiguity*, represents a corpus pragmatics approach to theoretical psycholinguistics and polysemy. By manually annotating over two and a half thousand instances of nine polysemous words in their linguistic context from the ARCHER 3.2 corpus, T. R. Williamson yields interesting findings about the nature of polysemous words. Specifically, it is found that the significantly ambiguous polysemes share certain properties (like low frequency of discrete senses) that are reversed in the cases of the significantly unambiguous polysemes. Moreover, these properties are argued to lend themselves to an explanation in line with Giora's 'Graded Salience Hypothesis'. In so doing, Williamson proposes that the reason polysemous ambiguity sometimes arises may be due to the co-activation of co-salient senses. This work will intrigue those with an interest in corpus linguistics, pragmatics, psycholinguistics, and lexical semantics.

Soon enough, submissions for JoULAB will re-open (and on a rolling basis, too). Preparations for this, and thus for Issue 2, are already underway. Prepare your submissions, undergraduates! Fine-tune those final-year dissertations, research projects, and theses! We are very excited to receive your manuscripts and we cannot wait to see what you have all been researching.

T. R. Williamson,
***Editor**, Journal of the Undergraduate Linguistics Association of Britain*
University of Cambridge

B. Papineau,
***Head of the Board of Reviewers**, Journal of the Undergraduate Linguistics Association of Britain*
Stanford University

Foreword from the National Chair of ULAB

As ULAB's National Chair, I am delighted to have had a hand in supporting the establishment of JoULAB. As one of ULAB's primary aims is to showcase undergraduate research, the foundation of the Journal amidst the cancellation of the ULAB 2020 conference was a fine solution to the absence of the opportunities that the conference typically brings. As far as we are aware, JoULAB is the only journal solely for undergraduates that accepts submissions in all subfields of linguistics, and the fact that we received forty-four articles in our first round of submissions demonstrates that this was very much a gap in the market that needed filling.

The Journal is a huge undertaking, but one that has already proven highly rewarding. It has facilitated connections between linguistics students around the world through the recruitment of our team of ten Editorial Committee members, as well as forty-six reviewers studying at institutions across seven different countries. A number of our reviewers also kindly gave up their time to review abstracts for the ULAB 2021 conference, and some even came to talk to ULAB members about life as a PhD student at our online Postgraduate Studies Panel event in October 2020, supporting ULAB in our aim to provide information and support for undergraduates looking to pursue postgraduate studies in linguistics.

I want to thank everyone on the JoULAB Editorial Committee and the Board of Reviewers, past and present, for all of the time and effort they have devoted to this project. Given the amount of time it takes to anonymise all of the submissions and reviews, review each of the submissions, copyedit all of the final articles, compile the issue, publicise its release and perform all of the other countless tasks along the way, taking the Journal from inception to publication in less than a year is certainly something to be proud of.

Going forward, the Journal of the Undergraduate Linguistics Association of Britain will complement the ULAB conference and U-Lingua Magazine in providing a space for undergraduate linguists to showcase their work. There are many articles still being reviewed currently, and so we look forward to publishing further issues soon. In the meantime, I hope that you enjoy reading the articles here, and that perhaps they may even encourage more budding linguists to undertake their own research!

Cliodhna Hughes,
***National Chair**, Undergraduate Linguistics Association of Britain*
***Associate Editor**, Journal of the Undergraduate Linguistics Association of Britain*
University of Edinburgh

Language Contact and the Phylogeny and Phonology of Early English

Nina Haket
University of Cambridge

Abstract. This paper aims to review Emonds and Faarlund's work critically from a phonological perspective. Their work suggests that Modern English is a North Germanic language rather than a West Germanic language. After evaluating Emonds and Faarlund's usage of the literature and theories of language contact and phylogenetic relationships, it is concluded that the only way Emonds and Faarlund's theory could be reconciled with current linguistic theory, is to posit a set of mappings from the Old Norse phonology to the Old English phonology in order to allow for a simple continuation of the Old English phonology into Middle English. Using a series of etymological dictionaries, mappings are explored for gemination, Holtzman's Law, cluster simplification, and palatalization. It is concluded that mappings are an impossibility due to the intricacies of the phonology, and that the notion of mappings themselves have no place in a theory of language contact or a phylogenetic framework.

Plain English Abstract. The fact that Modern English comes from Middle English and Old English seems to be an undisputed fact. Emonds and Faarlund's book *English; The Language of the Vikings* suggests that Modern English is a North Germanic language, originating from Old Norse, rather than a West Germanic language originating from Old English. This article looks at their claim from the perspective of the phonologies, which entails their sound systems and sound changes. Looking at Emonds and Faarlund's usage of the literature and theories of language contact and relationships between languages, it is concluded that the only way Emonds and Faarlund's theory could be reconciled with current linguistic theory is to posit a set of mappings from the Old Norse phonology to the Old English phonology in order to allow for a simple continuation of the Old English phonology into Middle English. Using a series of etymological dictionaries, mappings are explored for gemination (consonant lengthening, Holtzman's Law (medial /jj/ and /ww/ become /ggj/ and /ggv/ respectively), consonant cluster simplification, and palatalization (production of segments with tongue closer to the hard palate)), it was concluded that mappings are an impossibility due to the complexity of the phonology, and that the idea of mappings themselves have no place in a theory of language contact or relationships between languages.

Keywords: North Germanic; phonology; language contact; West Germanic; mappings; phylogeny

---

1 Introduction

In Emonds and Faarlund (E&F)'s book *English: The language of the Vikings* (2014), radical claims are made about the linguistic origins of Middle English (ME), and yet phonological evidence is not sufficiently drawn upon to provide evidence for their argument. This work will focus on the phonological implications of their claim that ME is descended from North Germanic (NGmc) rather than West Germanic (WGmc), and the notion of a direct phonological continuation from Old English (OE) to ME despite its proposed origins.

1.1 A Linguistic Gap

After the Norman Conquest of 1066, exactly what happened to the Viking speakers of Old Norse (ON) and native speakers of OE is unknown. After spreading throughout England since 787, the ON speakers

had established the Danelaw, an area that covered the North and East of England, which was given up to the invading Norsemen and briefly ruled England before the Norman Conquest.

It is generally agreed that the ON speakers shifted to OE under Norman rule, bringing with them many lexical loans such as *bask*, *law*, *sky* and the pronouns *them* and *their* (Townend, 2006), and certain phonological phenomena, such as unpalatalized velars before front vowels. The extent of the influence and the exact features that can be associated with the Vikings are not agreed upon, but the idea of a shift from ON to OE is generally accepted, as is described in Townend's influential work (Townend, 2005). There are, of course, other views. For example, Bailey and Maroldt (1977) and Poussa (1982) have argued for creolisation hypotheses of the origin of ME but these have not passed into mainstream acceptance.

1.2 Emonds and Faarlund: *English: The language of the Vikings*

In a relatively recent addition to the field of Germanic linguistics, E&F have made a bold claim regarding the origins of ME and therefore Modern English (ModE). In *English: The language of the Vikings* (2014), and in their later paper entitled *Anglicised Norse or anything goes* (2016), E&F claim that ME is descended from the NGmc branch of the linguistic family tree; the same branch that ON and modern languages such as Swedish, Norwegian, Icelandic, Faroese, and Danish are descended from. The traditional view that ME is descended from OE entails a WGmc origin, meaning E&F have changed the Stammbaum tree. What E&F claim is no minor change. The difference between the traditional Germanic tree and E&F's revised version is illustrated in Figures 1 and 2 below:

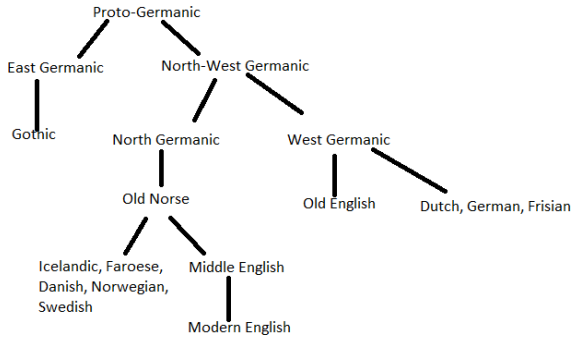


Figure 1: *The traditional Germanic family tree.*

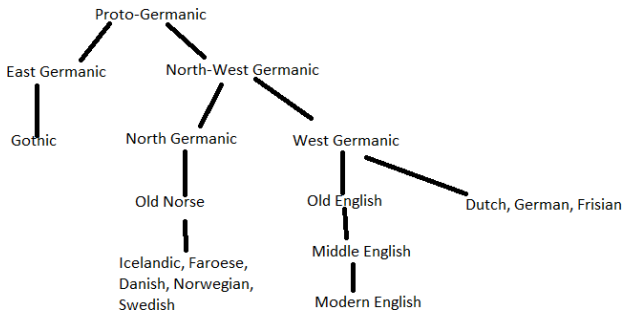


Figure 2: *E&F's modified Germanic family tree (simplified).*

E&F contend that ON did not die during the period of Norman rule, and claim that it was in fact OE that underwent language death in the Danelaw. ME is therefore descended from ON rather than OE, thereby maintaining the common Germanic ancestor, but thus originates from the NGmc branch rather

Received: 24/06/2020
Revisions: 03/02/2021
Accepted: 15/02/2021

than the WGmc. Rather than positing several changes to the OE grammar that resulted in the ME grammar, E&F point to a 'seamless continuation' (2014, p. 157) of the ON grammar, taking 20 different syntactic properties of ME, including word order, stranded prepositions, periphrastic auxiliaries, and split infinitives, and demonstrating their continuity from ON rather than OE. They conclude that ME shows no features that are not shared by ON.

In terms of the lexicon, E&F propose a lexical amalgam that formed through different origins, one of which is cultural borrowing. OE speakers would have had words for concepts that were foreign to the ON speakers, such as those that were a remainder of the Roman occupation, those related to crops and food production, and the written culture of the English. More importantly, there was a high rate of inter-ethnic marriage, with many ON-speaking men taking OE-speaking women as wives (Knooihuizen, 2009). ON was thus already heavily Anglicised before the Norman Conquest, creating what E&F name Anglicised Norse (AN). After the conquest, E&F describe a fusion of the ON and OE ethnic groups due to the social, economic, and linguistic context of the East Midlands under the rule of the Norman conquerors (2014, p. 34). Within the new fused society, E&F conclude that the AN speakers enjoyed more prestige, and thus the children in this fused society eventually settled on a single model on the basis of the higher prestige that the Scandinavian ethnicity had, despite mixed AN and OE input (2014, p. 155). The vocabulary therefore remained similar to that of the lost OE language. Referring to this as relexification, they insist that the basic structural properties of the language are indicative of a NGmc origin.

Despite the shift from OE to AN, and thus a continuation of the NGmc branch, E&F claim that the phonology of this ME is 'very much a continuation of Old English phonology' (2014, p. 157). The phonology of AN and the role of phonology in determining phylogenetic relations is only briefly discussed in E&F's appendix, where they touch on pre-vocalic velars, loss of low offglides and the reduction of vowel length contrasts, but explain these changes almost like collateral damage in the development of AN. They explain them simply as children learning both the OE and ON lexical items and change through language drift. Typically, linguists use a variety of sources of information in identifying phylogenetic relationships, such as basic vocabulary, regular sound correspondences, and in recent years, syntactic evidence (which is more controversial due to the chance similarity among binary variables and transfer from language contact, which appear to be insurmountable problems, (Campbell & Poser, 2008)). E&F consider almost exclusively syntactic evidence, and in doing so seem to take as a given that syntax cannot be, or is considerably less likely to be, transferred. This stability is taken to be a reliable marker of genealogical descent, thereby side-lining phonological evidence.

1.3 Criticism and Response

When it comes to the reactions to this work, E&F faced criticism. The majority of historical linguists reject the claim, including Trudgill (2016), Pons-Sanz (2013) and Stenbrenden (2016), citing a variety of evidence to substantiate their argument, including the falsity of certain etymological evidence presented as fact, and the circularity of their argument. Such a consensus is succinctly summarised by Munch (1962, p. 28) who believes that some authors 'feel themselves drawn to what causes a stir and seems stirring than to more painstaking endeavours that appear in a humbler guise', and E&F appear to match these criteria.

Some, however, do appreciate the line of argument put forward, such as Holmberg (2016) and Lightfoot (2016), neither of whom see a way around the conclusion that E&F draw. Lightfoot goes on to comment that those who believe the 'standard textbook view that English is a gradual and imperceptible changing object which moves through time and space' (2016, p. 475) will find this outcome difficult to accept, since it would mean accepting that change can be radical and sudden rather

than gradual. Paradoxically, Lightfoot himself has argued for a gradual drift in frequencies for linguistic forms in previous work. He attributes this to children, who cause changes in frequencies due to sensitivity to their input. It must be noted that Lightfoot appears to be somewhat individual in his ideology here, since some of those that are the most critical of E&F's work are, like Lightfoot, generative in orientation, including Van Gelderen (2016) and Walkden (2016).

A fundamental reason for the discredit of the work is the authors' own linguistic background. As Bech and Walkden (2016) point out, neither has any background in OE. Faarlund is a leading Scandinavianist who has a background in the history of Norse, and Emonds is a distinguished theoretical syntactician. Simms (2014) builds on this, and believes that the book contains major methodical and minor factual errors, such as suggesting *kid* and *child* are doublets of the same source and suggesting that *kind* and *king* are forms with non-palatalized velars (Emonds & Faarlund, 2014, p. 158). This is claimed to be due to E&F's lack of knowledge the phonological conditions for palatalization, and that the preservation of the velar when adjacent to a front rounded vowel fronted by umlaut is common.

Furthermore, despite E&F's focus on syntactic evidence, thereby implying that syntactic items cannot be borrowed, they themselves appear to promote this at several different points. At times they seem to require the transfer of a structure, especially from OE into this new AN they have created (Thomason, 2012). They therefore do not strictly follow their own framework, creating a clear contradiction. The result of basing their claim mainly on syntax is that phonology is largely ignored, even though the very nature of neogrammarian sound change is that systematic correspondences can, in fact, not be explained by anything but genetic relationship.

1.4 Phonological Claims

Phonological continuity between OE and ME is agreed upon by many scholars, including Font-Santiago and Salmons (2016). E&F do not disagree with this idea, and yet still state that ME is descended from ON. How these ideas can be reunited is not made clear in their work; they do not specify how ME can be a continuation of OE phonologically when their syntax is entirely derived from the ON system. Barnes (2016) thus concludes they have a 'cavalier' attitude to phonology, which is evidenced several times in the book, as will be shown in this work. There must therefore be a point in time where the entire OE phonology was mapped onto the AN lexicon. This seems like an unusual claim, and the theories of language contact that will be touched on later in this work agree.

How would the continuation of OE phonology to ME work? It cannot simply happen of its own accord; there must be a concrete way in which this continuation occurs. In the lexical amalgam of the AN language, if the underlying phonological inventory and rules were those of OE, then this goes against what E&F are claiming. If the underlying form is OE phonologically, for example, *drinkan* 'to drink', then ME is genealogically descended from OE, contradicting E&F. The only other option is a set of mapping rules between the ON form and the OE forms that speakers used to map the OE phonology onto the ON phonology. For example, the underlying form is ON phonologically, such as *drekka* 'to drink', to which some operation is applied to derive OE *drincan*. Since the former option contradicts E&F's argument, the latter option must be explored, which entails a pathway from ON to OE to ME phonology. While this is not directly mentioned by E&F in their book, it appears to be the only logical way in which the phonology can be a continuation of OE in their framework.

The rest of this work consists of three Sections, the first of which will aim to put forward a theory of phylogenetic relationship and contact induced change. The second Section will look at a number of sound changes in the Germanic languages and attempt to create mappings similar to the cognitive transformations the ON words must have undergone to constitute a continuation from OE phonology to ME. The final Section will conclude and summarise the findings.

Received: 24/06/2020
Revisions: 03/02/2021
Accepted: 15/02/2021

2 Theoretical Frameworks of Language Contact and Phylogeny

The claims E&F make require a solid foundation of a framework of contact induced change. Terms such as borrowing, imposition, and transfer are not defined in the book, despite forming a crucial part of their hypothesis. An attempt will be made here to clarify some of the terminology used, and point out some of the flaws in E&F's approach.

2.1 A Psychologically-Based Approach

An appropriate framework for contact induced change is a psychological approach, as described in Van Coetsem's (1988) and Winford's works (2003). Their approach centres on the idea that the traditional classification into borrowing and interference has some inconsistencies, since they are used interchangeably for both the outcome and the process, and also do not indicate the direction of the influence nor the agent. In the past, 'interference' has been used to refer to any kind of crosslinguistic influence, some have used 'transfer' in the same way, confining 'interference' to changes brought about by second language learning. Instead, they draw upon source language (SL) agentivity versus recipient language (RL) agentivity as types of crosslinguistic influence, which embody shift and borrowing respectively. In a single contact situation, both SL and RL may be involved, and while the movement is always from source to recipient, the agent may vary. 'Transfer' is a neutral term for the movement of material from one language to another.

The difference between the psychologically-based and traditional prestige-based approaches, such as that of Thomason and Kaufman (T&K) (1988), in which contact induced change is split into shift and borrowing, lies in the fact that there is a distinction in the psycholinguistic notion of speaker dominance (Winford, 2003). T&K concluded that in the ON/OE contact 'the influence is found in all parts of the language, but is not deep except in the lexicon' (1988, p. 302), and placed it at a level 2 out of 3 on their borrowing scale. A key point from T&K's work is the notion that interference cannot be compartmentalised, suggesting that even minor syntactic changes suggest the presence of minor phonological changes and vice versa. This approach did not consider the agents of the change, suggesting bilinguals and monolinguals were equally responsible.

In Van Coetsem's framework (1988), a bilingual is dominant not in their first language, but in the language that they are more proficient in, creating a dichotomy between linguistic dominance and social dominance. The latter includes the political and social position of the language, whereas the linguistic dominance is simply the language of highest proficiency and does not require the language to be dominant socially. This therefore makes bilinguals the main agents of the change. Since language is always in flux, bilinguals may be dominant in either of their languages at different times, creating differing directions of transfer.

In RL agentivity, material moves from the non-dominant language to the dominant, and in SL agentivity material moves from the dominant language into the non-dominant. Due to this, the potential of structural and radical modification is higher in SL agentivity, since the agents will be preserving the matrix structure of the language that they are dominant in and thus impose this material onto the RL. This is typical of second language acquisition, where speakers transfer aspects of their L1 to their L2 when their acquisition is incomplete. This is evidenced in arguably all linguistic domains.

Borrowing, or RL agentivity, is usually less radical, involving vocabulary transfer, which can consequently be integrated into the native vocabulary (Winford, 2003). Again, the language the agents are dominant in will provide the matrix for the change.

2.1.1 E&F's Hypothesis in this Framework

Traditionally, it is thought that the ON speakers shifted to OE, resulting in some SL agentivity in the form of shift, and some RL agentivity in the form of borrowing (Van Coetsem, 1988). The result appears to neatly fit the idea of RL agentivity in the cultural borrowings from ON, and SL agentivity in the minor phonological changes present. This is exemplified in unpalatalized ON stops contrasting the OE palatalized ones, causing doublets such as *shirt* and *skirt*, or the dialectal loss of the palatalized stops in favour of the unpalatalized stops, as is seen in Scottish and Northern *kirk* 'church'. E&F's approach does not fit so neatly into expectations. In the framework presented here, it is expected that there are some OE phonological features in ME due to SL agentivity in the shift of the population to AN. However, E&F do not suggest 'some' transfer, but complete transfer of the OE phonology.

In the creation of AN there is borrowing, and thus RL agentivity, in the cultural loans that entered ON from OE due to trade and intermarriage. The dominant language will have been ON, as the bilingual ON speakers are the agents of the change. This kind of contact can easily enable the transfer of lexical loans, and allows the creation of AN. Winford (2003) suggests that lexical borrowings often have little impact on the RL, and are usually adapted to native phonological and morphological patterns, as has been studied by Haugen for early and late borrowings from ON to OE and ME (Haugen, 1950).
This means that this initial stage of borrowing from OE would have left minimal impact on the ON phonology, especially since the phonologies of the two languages were similar.

The second stage of the creation of ME is the shift of the OE speakers to the AN, where there would be SL agentivity. It would be expected to find L1 effects on many linguistic subdomains, and thus the transfer of some OE phonological features, structural features, and semantic features, as the OE speakers impose their matrix structure on the AN language. It can result in significant changes in the phonologies of the target language, as is seen in Irish English, Welsh English, Indian English and Singapore English, where the phonologies have been radically altered due to the SL agentivity (Kortmann & Schneider, 2008), but does not support the transfer of the entire SL phonology into the RL. Yet, this is what E&F require. Any OE structures E&F find in ME, they attribute to borrowing rather than imposition, despite structural transfer being more likely in SL agentivity than RL agentivity. Their claim therefore either requires that the frameworks of Winford, Van Coetsem, and T&K need to be widened to allow this kind of relexification, or requires them to posit extra mechanisms outside of the scope of the traditional frameworks. One such mechanism is tested in this paper, which is the use of transformation rules.

2.2 A Phylogenetic Framework

A 'genetic relationship' between languages refers to descent from a common ancestor, therefore a phylogenetic relationship (Campbell & Poser, 2008), the idea of which can be traced back to the Philologer Passage (Jones, 1786). The comparative method is often used to establish such genetic relationships, and consists of a systematic comparison of different languages or dialects in order to reconstruct the language from which they developed. It relies on basic vocabulary, grammatical and morphological evidence, and regular sound correspondences. While languages may have diversified through means such as war, trade, isolation, social and economic organisation, group identity, and technological advance (Campbell & Poser, 2008) creating radically different surface level structures, underlying systematic and patterned correspondences may still be used to identify the relations.

It is generally agreed that using the lexicon alone to determine phylogenetic relationships is dangerous, since as presented in our theory of language contact, vocabulary is commonly transferred. While core vocabulary is less susceptible, even this can be transferred. Furthermore, instances of

Received: 24/06/2020
Revisions: 03/02/2021
Accepted: 15/02/2021

nursery forms, onomatopoeia, and sound symbolism make it unreliable for deterministic classification. Linguists have therefore mainly focussed on phonological and grammatical correspondences.

The need for regular sound correspondences in the comparative method was identified by Hübschmann (1875) long before the neogrammarian doctrine detailing the exceptionlessness of sound laws. It is now recognised that sound correspondences are necessary in determining phylogenetic relationships, and almost all consider them strong evidence (Campbell & Poser, 2008). While there are other sources for phonological similarity, if regular correspondences can be identified, then irregularities can be excluded. Similarities from other sources include analogy, borrowing through RL agentivity, and L1 effects through SL agentivity. Furthermore, false correspondences can be created if there are enough instances of accidental similarities involving the same sounds, and also if linguists are lax about the semantic meanings of the cognates discussed, since it is easier to find correspondences if meanings are relegated. As a result, there can be systematic correspondences between items that look different on the surface; resemblance simply is not enough, and regular correspondences are crucial. While there are certain aspects of using sound correspondences to determine genetic relationship that are problematic, they are still accepted as an integral part of the comparative method.

There is a longstanding ideology that structural features are less likely to have been transferred in a language contact situation, and this has led many to believe that structural evidence is superior in determining genetic relationships, including Sapir (1921), Meillet (1925), and E&F. They refer to the fact that syntax has been thought of as the deepest structure with the most cohesion, and thus hard to influence. They believed that borrowing of structural elements can only occur when the languages are typologically similar. This is most clearly seen in instances of dialect borrowing. An example would be the case of the Hvar dialect (spoken on the Croatian island of Hvar) which had a distinction of genitive/locative vs. dative/instrumental cases which was replaced with the Serbo-Croatian pattern of genitive vs. dative/instrumental/locative (Thomason, 2006).

This has since been argued against by linguists such as Thomason (1980), who showed that morphology is not as stable as thought; cognates vanishing but the morphology remaining unchanged is unlikely. Utmost care must therefore be taken to exclude borrowings from evidence. In the framework of language contact discussed, structural properties do transfer. Phonological and syntactic transfer are typically among the last to be transferred in RL agentivity, but they are still borrowed. In terms of SL agentivity, syntactic transfer is common and usually among the first features to be transferred in a contact scenario.

Winford (2003) discusses the difference between direct and indirect structural transfer. Indirect structural transfer refers to structural changes mediated by lexical borrowing, in which large quantities of lexical borrowing could introduce new structural and phonological features into a language such as derivational affixes in ModE from French or Latin, or phonological fricative voicing due to French influence. In the rare cases where there is direct structural borrowing, it usually involves free morphemes such as prepositions or conjunctions where a simple replacement can occur with RL morphemes of a similar form or function; this naturally requires a high degree of bilingualism, thus being more associated with SL agentivity than with RL agentivity. It is clear that structural borrowing and transfer is subject to some strong constraints, but to say that structural transfer does not occur is therefore untrue, meaning that they must be mitigated against in our application of the comparative method.

2.2.1 E&F's Treatment of the Comparative Method

The crux of E&F's argument lies in their treatment of the comparative method; they work in a Labovian framework in which structural patterns are not subject to the same kind of conditioning as phonemes

and morphology (Labov, 2001, pp. 28–9). E&F do not use phonological evidence in the traditional way in which it is required to evidence phylogeny, focussing solely on a misinformed view that syntax does not change in contact scenarios. This misinformed view is well exemplified by Elderkin (1976, p. 296) who said, 'classification of languages rests on the selection of one part of a language to typify that language, and this selection is arbitrary'. E&F have selected syntax, and their method is unconvincing. To create a stronger claim, the phonology and syntax need to be unified and a way to explain the transfer of the entire phonology needs to be found. If it is proven that a language can have a phonology that is entirely transferred from another language, leaving little or no trace of its own original phonology, this discovery would be catastrophic for the comparative method.

2.3 Learnability

The learnability of the rules proposed will form the central argument as to whether they are plausible in regard to E&F's hypothesis. It has been pointed out that it is not possible to provide definitive proof that something is not learnable (Lai, 2015), since any experiments only prove that a pattern was not learnt by its set of participants. As a result, there is not a widely accepted definition of learnability (Lewis & Elman, 2001). This has not stopped learnability theories emerging, such as the possibility of data compression, or generalisation, as is explored in Gerken, Balcomb, and Minton (2011). This theory suggests that rules which are not generalisable, have too many exceptions, or are context dependent are unlikely to be learnable.

However, this is contested by the fact that speakers are able to learn non-regular patterns, as is shown by irregular verb morphology. Ringe and Yang (2015) looked at this phenomenon and created a tolerance principle consisting of a threshold which, if not exceeded, allows the rule to be extended and lead to the loss of exceptions. This is evidenced by the historical tendency of exceptions to become regular.

While creating mappings from an underlying ON form to a surface OE form may be radically different to other examples of learnability in the literature, it is comparable with generative phonology, or more specifically derivational phonology (Vaux, 2008), in the sense that generative phonology also involves underlying forms subject to a series of rewrite rules or transformations that create the surface forms (Dresher, 2015). Rules may take several forms, but one is in the style of *The Sound Pattern of English* (Chomsky & Halle, 1968) and takes the form A→B/C_D. This is the basic formula adopted in this study.

Kaplan and Kay (1994) suggested that all phonological patterns belong to the regular class within the Chomsky hierarchy, and as a result can be defined by finite state automata (Lai, 2015). This means that the transformations between underlying forms to surface forms can be considered regular, and that such rule-based grammars can describe all phonological patterns (Hyman & Plank, 2018), thus being universal.

Since there is no definitive way to define learnability in this context, it is important to generalise the ideas put forward from these different approaches. These phonological operations have been found to be regular, so this work hypothesises that the mappings created between OE and ON should be regular too. It is evident that these are the types of patterns which speakers are able to learn, though it does not disprove that speakers cannot learn others. Regardless, this view shall be taken into the creation of the mappings.

2.4 Intelligibility and Dialect Congruity

Received: 24/06/2020
Revisions: 03/02/2021
Accepted: 15/02/2021

The idea of creating mapping rules between two closely related varieties has been utilised in intelligibility and second dialect acquisition studies. Hockett (1987) created a model for intelligibility involving two modes, which are listening to word identity and listening to word shape. Word identity depends on implicit motor matching, gestalt perception and, most importantly, the switching-code. This code is a series of transformations from the foreign dialect to the hearer's own dialect, and once this code is established, words can be automatically coded and recognised through these transformations. Milliken and Milliken (1993) take a similar approach to dialect congruity and posit that intelligibility correlates with phonological correspondences between two varieties, thereby making intelligibility a property of phonological correspondence rather than phonetic similarity. Key to their theory is dialect congruity, which describes whether correspondences can be generalised and exceptionless. While Milliken and Milliken offer no quantifiable dimension to their proposal, the idea behind their theory resonates with the mappings posited in this dissertation. Townend (2005) used the above theories to explore intelligibility in Viking age England, looking in particular at place names and the systematic swaps made by ON speakers in their references to these places or people. Townend (2005) concluded that the level of intelligibility suggested that ON and OE should be treated as different dialects rather than different languages, and looked briefly at some phonological correspondences identified in these place names, including palatalization and assibilation.

The approach taken in this study will be broader and look at the language in general rather than at place names only, and also consider the two varieties as different languages rather than different dialects, making this a question of shift and contact induced change rather than intelligibility.

Furthermore, the mappings posited here will be in the opposite direction, from ON to OE rather than from OE to ON, to test E&F's theory of a NGmc origin for ME. Regardless, it is useful to see the same methods applied in other domains.

3 Hypothesis

Having now looked at why mappings are necessary to entertain E&F's hypothesis from a phonological perspective, it must be determined what these mappings look like and whether they are realistic or achievable. The hypothesis of this work is that the mappings required to transfer the OE phonology onto the ON phonology will be arbitrary, and therefore cannot be learnt by neither the children using it as primary linguistic data, nor adults who are in the process of shift.

3.1 Methodology

This work aims to create a set of transformation rules to map the ON phonology to the OE phonology using gemination, Holtzman's Law, and palatalization, which are traditionally used as a diagnostic for NGmc and WGmc branch membership. Examples will be found in dictionaries of OE, ON and PGmc. The aim will be to create plausible, logical, and cognitively simple mappings, and most importantly, these rules must be learnable for L2 learners and native speakers alike. If they become too numerous, complicated, sporadic, or a word-for-word enterprise, then these mappings cannot be learnt and therefore cannot account for the continuation of the OE phonology.

Mappings will take the form of one segment or cluster becoming another in a collection of transformation rules. They do not need to be plausible sound changes, since this is not an instance of sound change, but an instance of cognitive mapping by the L2 learners of AN transferring their OE phonologies to the ON/AN words.

Take initial /v/ in ON. There is a rule between Proto-Norse and ON that all initial /w/ become /v/, whereas this rule did not apply in the WGmc branch, where it remained /w/. ON therefore has *verpa* 'to warp' where OE has *weorpan*. The mapping here is thus as simple as /v/ to /w/ word-initially. Other examples include ON *vika* 'week' compared to OE *wicu*, ON *vara* 'to warn' compared to OE *war(e)nian*, ON *vapn* 'weapon' compared to OE *wǣpen*, and ON *vif* 'wife' compared to OE *wīf* (Ross, 2002). This is a straightforward mapping and appears to pass the learnability test, and if the entire ON phonology could be mapped to the OE phonology in this way, E&F's hypothesis may be plausible.

3.2 Resources and Data

The resources that will be used for the present study are the *Bosworth-Toller Anglo-Saxon Dictionary* (Bosworth, 2010), the *Oxford English Dictionary* (The Oxford English Dictionary Online), *the English-Old Norse Dictionary* (Ross, 2002) and the *Etymological Dictionary of Proto-Germanic* edited by Guus Kroonen, which is part of the Leiden Indo-European Etymological Dictionary Series edited by Alexander Lubotsky (Kroonen, 2009). This work will be taking examples of words exemplifying gemination, Holtzman's Law, cluster simplification, and palatalization to demonstrate that the mappings required by E&F are not only difficult, but impossible.

4 Gemination, Holtzman's Law and Cluster Simplification

The way that NGmc and WGmc treat Gemination, Holtzman's Law, and cluster simplification highlight the impossibility of such mappings. PIE had no length contrasts in consonants, only in vowels, meaning that the development of gemination occurred in the Gmc branch separately.

4.1 Divergences between West Germanic and North Germanic

In OE, all consonants except /r/ double between a vowel and a /j/ (Townend, 2005). This resulted in forms such as **bedjan* (Kroonen 2009, s.v. *bedjan*-) 'to request' in PGmc to OE *biddan* (OED Online, s.v. *bid* v1-). This can be compared to Gothic (Go) *bidjan*, whereas ON *biðja* underwent different processes to gemination. In a second round of gemination, voiced stops geminated before /r/, as in *blæddre* 'bladder'. In terms of the treatment of Holtzman's Law, also known as sharpening, medial /jj/ and /ww/ become /ggj/ and /ggv/ respectively (Townend, 2005). Intervocalic glides in strong position geminate, then the first semivowel is vocalised, thus becoming a diphthong. The vowels are then separated, creating for instance OE *trēowe* from PGmc **trewwu* 'true' (Kroonen 2009, s.v. *trewwu*-). When not in these contexts, consonants remain single.

In ON, there is a different outcome. Widespread syncope caused consonant clusters to simplify and assimilate, involving mainly /r/, /h/ and /n/, creating large scale gemination that is not present in WGmc. These lost vowels can be considered to be due to 'the strong stress accent on the first syllable', which 'caused in Germanic a progressive weakening of unaccented syllable, which is particularly marked in the case of final syllables' (Prokosch, 1939, p. 133). Phonemes bearing secondary stress were still weakened, causing medial loss of short vowels which has been considered as one of the most distinctive ON sound changes. Holtzman's Law in the NGmc family geminates intervocalic glides in strong position, which then sharpen and form occlusive onsets, causing PGmc**trewwu* 'true' to become ON *tryggr*. This ON sharpening is regularly attested in ON and Gothic, despite only being visible in a small number of words (Dance, 2019). While Dance notes that the operation remains obscure, it is a definitive marker separating OE from ON and Gothic (despite ON and OE forming a NWGmc family

Received: 24/06/2020
Revisions: 03/02/2021
Accepted: 15/02/2021

excluding Go). As a result, there are instances such as ON *tveggja* and Go *twaddje* compared to OHG *zweiio* and OE *twēġa*, meaning 'two' (Dance, 2019).

4.2 Consonant Clusters

The geminated consonants caused by syncope in ON must be mapped onto the consonant clusters that are maintained in OE. It could be hypothesised that this occurs through dissimilation of the two consonants to restore an arbitrary consonant that was there before, but without there being a particular conditioning environment for it. It would therefore be necessary to posit spontaneous arbitrary dissimilation. However, there were numerous consonants involved in the ON syncope gemination resulting from both progressive and regressive assimilation, such as /mp/ → /pp/, /nk/ → /kk/, /nt/ →/tt/, /ʀd/ → /dd/, /ʀn/ → /nn/, /ht/ → /tt/, /ðl/ → /ll/ /dt/ → /tt/, /lð/ → /ll/, and /nð → /nn/ (Townend, 2005). Many of these processes have precisely the same outcome despite involving different consonants. Without some teleological knowledge of the PGmc form, how would learners know which mapping to opt for? There will be consonant clusters in OE in cognate words, which would aid in the mapping, but it would still remain arbitrary and spontaneous, and on a case-by-case basis.

For example, a /tt/ in ON could be traced back to /ht/ /nt/ or /dt/ in PGmc and WGmc, but since the OE word shows a /nt/, a /tt/ → /nt/ mapping will be chosen. Another word with /tt/ could be cognate to another OE word with /ht/, and so a /tt/ → /ht/ mapping would be chosen. There is no regularity here, no rule that can be applied; it is a case-by-case basis. Taking the /tt/ cluster in ON, words such as *brattr* 'sheer' equate to the OE cognate *brant* (Kroonen 2009, s.v. *branta-*) (Townend, 2005). The /tt/ cluster in *brattr* would need to map to /nt/ to create the correct correspondence, meaning this can be posited as a mapping rule. However, this does not hold when looking at other words. ON *réttr* 'right' would by this logic map to OE **rent*, but in reality, it corresponds to OE *riht* (Kroonen 2009, s.v. *rehta*).

There would also need to be a mapping from /tt/ to /ht/. This is also found in other lexical items, such as OE *drihten* 'lord' compared to ON *dróttinn* (Kroonen 2009, s.v. *druhti-*), as well as OE *neaht* 'night' compared to ON *nǫ́tt* (Kroonen 2009, s.v. *nahti-*). Whether /tt/ maps to a /ht/ or /nt/ is not evident from the surface realisation, again thereby invoking either comparison with the OE form or teleology to explain how the correct mapping occurred. Since the contexts presented here are close phonologically, both involving word final clusters in OE following vowels, it must be more than a game of chance as to whether the correct mapping is chosen.

Another way of making a mapping would be to posit that the geminate falls into a singleton, but then spontaneous and unconditioned epenthesis of a consonant adjacent to it would have to occur. ON *brattr* 'sheer' would become **bratr* followed by epenthesis of the /n/ (along with loss of the /r/ suffix) to form OE brant. ON *dróttinn* would become **drotinn* followed by ephenthesis of /h/ (along with a change in the vowel) to form OE *drihten*. Similarly, this leads to the problem of how speakers would know which consonant to insert, since there are once again have examples of an /n/ and /h/ being inserted without a specific conditioning environment. Languages do not have memories, so inserting a consonant to match a historical form that arose via simplification is a counterintuitive explanation.

To exemplify this, the development of PGmc **drinkana* 'to drink' shows the difficulty of this mapping. **drinkana* comes from PIE **dhrénge* (Kroonen 2009, s.v. *drinkan-*) The ON form is *drekka*, therefore showing the consonant cluster simplification from /nk/ to /kk/, as well as the loss of final /n/ which is common and advanced in ON compared to OE (Townend, 2005). The vowel is lowered from/i/ to /e/ through a-mutation of the following /a/. On the other hand, there is OE *drincan* (OED Online, s.v. *drink* v1-), where the <c> represents a /k/, meaning that the form has undergone little or no change from the PGmc form. To map from *drekka* to *drincan*, the onset cluster would have to be maintained, and the vowel would have to raise back up to /i/. However, the consonants pose a larger problem. /kk/ needs to

dissimilate back to /nk/, and thus reinsert a segment after an assimilatory sound change got rid of it. As discussed, an ON cluster could stem from multiple origins, meaning that only teleology or the OE version contributing to arbitrary decisions could cause the correct reinsertion of the /nk/ cluster.

4.3 Holtzman's Law

Moving on to the treatment of Holtzman's law in the two languages, it becomes evident this is equally complicated. Both involve semivowels geminating in strong prosodic position, but those in ON would have to lenite the onsets back to glides. The problem occurs when considering the fact that these then have to diphthongize and turn into separate vowels. What is needed is a direct mapping from /ggj/ to the diphthongs or vowels created by Holtzman's law in OE. The same problem as with the assimilations thus arises; what segments would /ggj/ need to map to in certain situations?

Looking at some instances of Holtzman's law in NGmc and WGmc, it appears as if there is not a simple correspondence between the words in question. Holtzman's law is often not as uniform as we would like to believe, and the words displaying the NGmc outcome of Holtzman's law in ON have several different consonant/vowel clusters, rather than the simple /ggj/ and /ggv/ expected.

To start, some words do show the typical /ggj/, such as ON *tveggja* 'two', which corresponds to OE *twēġa* (genitive singular of *twēġen*) (Smith, 1998). Due to OE orthographical practice, *twēġa* is pronounced /twe:.ja/, meaning that the correspondence here is /ggj/ to /e:j/. If this rule could be applied to all instances of /ggj/ then mappings may be possible There are also several instances where the outcome in NGmc is /ggr/, as is the case in ON *tryggr* 'true' and ON *glǫggr* 'sharp minded'. The OE cognate for *tryggr* is *trēowe*, while the OE cognate for *glǫggr* is *glēaw* (Smith, 1998). This would mean that positing the mappings /ggr/ to /ēow/ as well as /ggr/ to /ēaw/. While these are close, multiple mappings are still required.

There may also be a vowel following the geminated /g/ in NGmc, and these again need different mappings onto OE words. ON *skuggi* 'shadow' corresponds to OE *sċua* (Smith, 1998). This would require a /ggi/ mapping to a complete diphthong /ua/. ON *bryggia* 'bridge', corresponding to OE *brycg*, is different from any of the others discussed (Smith, 1998), since the outcome is not a diphthong or semi-vowel, but a voiced affricate /yd͡ʒ/. Finally, there is ON *tryggua* to OE *trūwian/trugian* (Smith, 1998), where there would need to be a mapping of /yggua/ to /ūwia/ or /ujia/. This serves to highlight that there is not only the problem of many OE outcomes for a single ON feature, but also many OE outcomes for many ON features. There does not seem to be a single cohesive unit that would make the comparison straightforward, neither to linguists nor to speakers.

PGmc **trewwu* (Kroonen 2009, s.v. *trewwu-*) 'true' derived from from PIE **dreuHu* developed to **triwwjaz* in the proto-ON branch. Then *i*-umlaut changed /e/ to /i/ and, most importantly, there was gemination of /w/ in strong prosodic position to /ww/ which then hardened to /ggw/. Finally, the ON form gained an /r/ ending from /iz/ *i*-stem nominative singular ending in PGmc. This then formed ON *tryggr*. To get to the OE version, **trewwu* becomes **triuwjaz*, which already shows the vocalisation of the first semi-vowel which is characteristic of the WGmc branch. The diphthong formed then separates the segments, so there is OE *trēowe* (OED Online, s.v. *true*, adj., n., adv., and int), and later ME *treu* (Middle English Compendium Online, s.v. *treu*, adj.). As previously mentioned, going from /ygg/ to /yww/ to /iuw/ to /ēow/ is plausible since it involved lenition and assimilation processes, but going from /ygg/ to /ēow/ will be difficult. Not every /ygg/ will map onto /ēow/ as is exemplified by the above discussion, and thus we also need to consider /e:j/, /ēow/, /ēaw/, /ua/, /yd͡ʒ/, /ūwia/, and /ujia/.

4.4 Old English Gemination

Received: 24/06/2020
Revisions: 03/02/2021
Accepted: 15/02/2021

Moving on to mappings involving consonants that geminate in OE but not in ON, even more mappings must be posited. ON gemination only occurred in Holtzman's Law due to the syncope and resulting assimilation, whereas OE had many more contexts, such as all consonants except /r/ doubling between a short vowel and /j/, and voiced stops geminating before /r/. The remaining OE geminate consonants would have to be derived from those that remain singletons in ON.

Concerning the consonants doubled between a short vowel and /j/, a mapping would consist of all single consonants except /r/ geminating between a short vowel and a /j/. It is not as simple a case as this; instead, there is a change from ON /ð/ to OE /dd/ in some instances. An example is PGmc **medja* 'middle' which becomes ON *miðr* compared to OE *middel* (Kroonen 2009, s.v. *medja*-1). Another is PGmc **bedjan* which became ON *biðja* compared to OE *biddan* (Kroonen 2009, s.v. *bedjan*-). This therefore suggests that a ON /ð/ to OE /dd/ mapping is required rather than only a rule positing doubling between a short vowel and a /j/. This requirement is mirrored in the outcome of second gemination, which involves voiced stops doubling before /r/. Rather than simply geminating voiced stops before /r/, ON /ð/ to OE /dd/ is evidenced. This is seen in ON *bláðra* 'bladder', which would thus need to map to OE *blaeddre* (OED Online, s.v. *bladder*, n), and the same is seen with ON *naðra* 'snake' compared to OE *næddre* (OED Online, s.v. adder, n). Mappings are therefore needed from ON /ð/ to OE /dd/.

This rule is very deterministic, since not all ON /ð/ can go to OE /dd/. Indeed, there are several words where ON /ð/ is not equivalent to OE /dd/, such as in ON *garðr* 'garden' for OE *geard* (Kroonen 2009, s.v. *garda*-) or in ON *maðr* 'man' for OE *mann* (Kroonen 2009, s.v. *mannan*-). The change must therefore only happen in words and contexts where the PGmc form had either a consonant between a short vowel and a /j/, or had a voiced stop followed by an /r/. As has been previously discussed, the speakers would not have had any knowledge of the proto forms, again creating a mapping that is idiosyncratic. On surface level, there is no reason why ON *miðr* 'middle' becomes OE *middel* but ON *garðr* becomes OE *geard* and ON *maðr* becomes OE *mann*. The environments are almost identical in the ON words, and differences can only be seen in the proto-forms and the changes they have undergone to become the ON words. Speakers do not have this knowledge, therefore mappings become unlearnable.

4.5 Summary of Gemination, Holtzman's Law and Cluster Simplification

This goes to show that mappings which are created by the processes of gemination in English appear to be significantly easier to map onto their ON counterparts than the outcome of cluster simplification and Holtzman's Law, where the variable outcomes make mappings impossible due to the sheer number of exceptions. The learnability is thus compromised, making them cognitively unlikely in the language learners; if they become individualistic then they cannot be learned, neither by L2 learners nor L1 speakers.

5 Palatalization

One of the most problematic areas for mappings is palatalized consonants, but not due to the complexity of the mappings. The Gmc velar consonants /k/, /g/, and /ɣ/ remained as such in ON in all positions (Dance, 2019). However, in OE these have palatalized allophones /ʧ/, /d͡ʒ/, and /ʝ/ before front vowels, and the unpalatalized variants in other positions, which led to the development of these as independent phonemes, still represented orthographically by /k/ and /g/. In terms of clusters, ON retains the Gmc /sk/ cluster, but the same cluster was palatalized to /ʃ/ in nearly all contexts in OE, except where it was preserved across boundaries and with /r/. Another condition can be placed on the palatalization process in OE, which is that preservation of the velar stop when adjacent to a front rounded vowel fronted by

umlaut is common, meaning this will have to be accounted for in the mappings posited for the palatalization process (Simms, 2014).

One important thing to note is the fate of these consonants and clusters in ME and ModE. In ME, according to the traditional viewpoint, there are loans from ON with the unpalatalized velars and clusters rather than the palatalized ones. These replace the palatalized English words in words such as *give*, exist as doublets with English words such as *shirt* and *skirt*, or remain as dialectal variants for palatalized words, such as *kist* for 'chest' (OED Online, s.v. *kist* n.1-) in some Northern dialects, specifically in Scottish English.

5.1 Palatalized Versus Unpalatalized

Within the palatalized versus unpalatalized simpleton consonants, there should be a mapping which mirrors the process that happens in OE. /k/, /g/, /sk/, and /ɣ/ would be mapped onto their palatalized counterparts when adjacent to front vowels. This would be simple, and also mirrors processes that are later seen in daughter languages of ON, for example the palatalized velar in Icelandic *gefa* [gj] 'to give'. Norwegian and Swedish /g/ did develop into a glide, and Faroese /g/ affricated with sibilant release (Harbert, 2007). While the mappings do not have to mirror realistic phonological changes, it is interesting to see the same pattern in NGmc language later in their development.

Considering that preservation of the velar stop when adjacent to a front rounded vowel fronted by umlaut is common in OE, the mappings become more complicated, since the source of these front vowels must be considered. For example, OE *cyssan* 'to kiss' has an unpalatalized /k/ in initial position despite being followed by front vowel /y/. It derives from PGmnc **kussjan*-, where the /j/ in the next syllable caused umlaut of /u/ to /y/ (Kroonen 2009, s.v. *kussjan*). In this case, the ON *kyssa* 'kiss' would need to maintain the velar stop rather than palatalising it after the front vowel. The rule must therefore be refined to state that /k/, /g/, and /ɣ/ palatalise before all front vowels except /y/, since this phoneme was not present in the proto-language.

So, at surface level, this would be an instance that would provide evidence to E&F's proposal, which would require mappings between the ON and OE versions. However, this is not the case.

5.2 Loanwords

According to the traditional approach, in which OE survived and was simply influenced by ON, palatalized variants of words are native and unpalatalized variants are loans, leading to doublets like *shirt* and *skirt*, *ditch* and *dyke*, *witch* and *wicca*, and *church* and *kirk*. In this approach, this would be a simple case of both SL agentivity and RL agentivity; RL agentivity in lexical borrowings from ON that were not adapted fully to the OE phonology, and SL agentivity in phonological transfer from the ON speakers shifting to OE.

Posing mappings to account for these palatalized and unpalatalized variants in E&F's framework becomes problematic. If a conditioned rule is posited such as 'palatalize /k/ before front vowels', then ON *skyrta* 'shirt' would indeed become OE *scyrte*, in which the first sound is palatalized (OED Online, s.v. *shirt*, n; OED Online, s.v. *skirt*, n). However, if this happens then the problem arises of how to account for the lexical doublet in ME *skirte*. The phonological environment is the same, since they both stem from the same cognate, and yet one undergoes palatalization whereas the other does not. The same can be shown for every palatalized-unpalatalized doublet. In cases such as this, the mappings appear incomplete, leaving both the palatalized and unpalatalized variants in the amalgam lexicon.

Even more problematic are cases where only the ON unpalatalized consonant survives, such as in *give*, since that would mean that the transformation rules simply did not apply to this word, even

Received: 24/06/2020
Revisions: 03/02/2021
Accepted: 15/02/2021

though there was the correct phonological environment for them. Not only do the mappings seem sporadic, but they also seem word specific. This would heavily impact the learnability of the mappings proposed. Another option would be that the rules were indeed regular, and these unpalatalized variants were borrowed into ME by RL agentivity due to later contact with NGmc speakers, whether that be through remnants of Norse in England, or through relations across the North Sea.

5.3 How to Account for Variation

Suggesting that the mappings are sporadic or word specific would go against the nature of the mappings in being regular and learnable. This means that to account for the variation found in ME and ModE, extra mechanisms would need to be posited. However, these mechanisms through which they were introduced are unknown. RL agentivity as the source for these stops is a similar proposal to what is traditionally thought to have happened through ON influence on OE. ON speakers shifting to OE brought with them some unpalatalized stops, and loanwords containing unpalatalized stops were welcomed into the OE language through both SL and RL agentivity. Two ways in which ME could have gained unpalatalized stops in E&F's proposal would be through either SL agentivity or through RL agentivity.

SL agentivity has a higher likelihood of impacting the phonology of a language, and yet when this transfer should have occurred, the OE population would have already switched to AN. The OE speakers could not have brought with them these unpalatalized stops since their own phonology did not have them. Since ME under this proposal is a form of relexified Norse, it could be posited that the Norse speakers imposed their L1 habits onto the newly formed ME as they learnt the new phonology transferred by the OE speakers. This would not go against the idea that ME phonology is a continuation of OE, since languages can be genetically related and still incorporate borrowed features (Campbell & Poser, 2008). However, this seems like a specific part of the language to impose L1 habits onto. Phonological transfer is most common in SL agentivity, but one would expect an influence over more aspects of the phonology. What is seen is phonological transfer limited to unpalatalized stops in some contexts but not in other contexts, leaving doublets, some unpalatalized stops, and some palatalized stops. While transfer is not be particularly regular in its outcome, to restrict it to such a domain appears to postulate some modularity in transfer that is unlikely.

As a result of these problems, the most likely proposal is that the unpalatalized forms were borrowed through RL agentivity from NGmc after the mappings were complete, rather than in the creation of AN. This would explain the apparent sporadic nature of the remaining words, in the sense that there are both doublets and single surviving variants without a particular pattern under which to accumulate them. The question is therefore when, and why.

ON did not survive in England, as is evidenced by the lack of written texts in ON after the written tradition restarted in 1150. The chance of borrowing from remaining ON in England after ME had been created though the relexification of ON is therefore small. The other option is that this variant was transferred through language contact across the Norse Sea. This is possible, but it is unknown when the borrowing happened, which groups initiated it, and how the contact occurred. Trade would have put the new ME language in contact with NGmc, but was this contact frequent enough and did it involve enough bilinguals to propose RL agentivity?

Regardless, E&F cannot simply invoke borrowing as an explanation whenever their mappings do not produce the intended outcome, thereby weakening their theory. Borrowing is not a magical process that solves any anomalies. Their proposal appears to violate Occam's Razor (unnecessary multiplication of rules or entities is to be avoided) due to the operation of the mappings, and then

sporadic borrowing via some unspecified mechanism rather than only borrowing and transfer through SL and RL agentivity.

An example of palatalization is from the *ditch/dyke* alternation in modern English. Both words come from PIE **dheighno*, which gave rise to PGmc **dika* (Kroonen 2009, s.v. *dika*-). This **dika* then becomes ON *diki*, where the velar stop is maintained as expected. In OE it becomes *dic*, with a palatalized consonant and loss the final vowel, whereas the rest remains unchanged. To map between the ON and OE forms is simple; /k/ becomes /ʧ/ due to the front vowel. This leaves the question of how to account for the variance in the palatalized versus non-palatalized consonants in the modern doublets *ditch* and *dyke*. The difference is typically associated with loans, where *dyke* is a loan from ON, in order to explain the maintained velar. In E&F's hypothesis, the only options are to posit a rule that applied the mappings sporadically, leaving some words changed and some untouched, but also some words changed and unchanged simultaneously in order to cause the lexical split, or to say that these words are borrowed. To say that all words with an unpalatalized /k/ are borrowed is plausible, since this is the approach taken in the traditional framework. The difference is that the traditional framework offers a scenario in which this would happen and offers the agentivity and directionality needed to establish the doublets. E&F offer no such explanation. All words with a /k/ before a front vowel in ModE cannot be borrowed in their framework. E&F are unable to account for who it was borrowed from, when it was borrowed, why it was borrowed, or where it was borrowed.

5.4 Summary of Palatalization

To summarise, the mappings viewed in isolation are deceptively simple, but when faced with the complicated reality of the ME and ModE phonologies, mappings do not work without positing some extra mechanisms through which the unpalatalized variants were restored or transferred in some contexts but not others.

6 Conclusion

Now that the complexity of potential mappings has been made clear, it is necessary to evaluate the results of the transformation rules and return to the notion of mappings within our framework of language contact and phylogenetic relation. Then, the notion of mappings generally will be looked at, and finally the problems associated with implementing the mappings in a second language acquisition context will be evaluated.

Looking at the mappings from the study of gemination, Holtzman's Law and cluster simplification, the idea of spontaneous unconditioned epenthesis or dissimilation, complicates E&F's proposal. Singletons do not become geminates in a single definable context, and the idea of teleology and forms returning to their historical states without a conditioning environment is unlikely at best. At first sight, the mappings between the palatalized OE variants and the unpalatalized ON variants was straightforward, but when considering lexical doublets and the large amount of ModE words which have the unpalatalized variants, mappings do not seem sufficient. Extra transfer must have occurred and yet the time, location, and reason of these borrowings are not explained. Mappings cannot be executed and then undone selectively, also causing this mapping to fall apart. Even the small selection of words and sound changes investigated provide evidence that mappings are not a plausible reparation strategy to make up for E&F's shortcomings in not abiding to a theory of language contact or phylogenetic relationship.

The exploration of the theory of language contact of Van Coetsem (1988) and Winford (2003) concluded that a shift from OE to AN should have resulted in features transferring in different domains,

Received: 24/06/2020
Revisions: 03/02/2021
Accepted: 15/02/2021

since contact induced change is unlikely to be modular, meaning that there should be phonological, lexical, and syntactic transfer of OE into AN to create ME. However, what E&F propose is an entirely maintained ON syntax, and an entirely maintained OE phonology, leading them to conclude that ME is of NGmc origin due to the superior diagnostic powers of syntax, and also seemingly creating modular aspects of the ME language. How this entirely OE phonology was transferred to AN is unclear, since even under SL agentivity, where the most radical transfer of syntax and phonology occur, the theories do not predict transfer of the entire phonology. No clues can be found in Van Coetsem's theory of language contact to explain how these mappings should function. This led to the development of the mappings as a possible solution to this problem which, as has been demonstrated, is implausible and cognitively demanding at best, impossible at worst.

The traditional approach, in stark contrast, does fit well with knowledge of language contact scenarios. With OE as the RL and ON as the SL, there is no requirement of such mapping, and fits with what is known in terms of SL agentivity in lexical loans and RL agentivity in shift. There is influence in the lexicon, the phonology, and the syntax, all to a reasonably minor level, and OE is maintained while ME is recognisable as descended from it. The phonological evidence fits the traditional approach more closely.

Similarly, in failing to use regular sound correspondences and phonology in their application of the comparative method, E&F reject the accepted way of creating the most likely genetic relationships. It is important to their argument that they do this, because using regular sound laws would yield only one conclusion; ME is definitely descended from OE. By introducing the idea of sound correspondences into their hypothesis, the only way that it could still be considered is through mappings. These mappings would devastate the assumptions of the comparative method, since an entirely borrowed phonology could result in altogether different phylogenetic relationships.

The mappings were necessary to make E&F's claim line up with a theoretical foundation, but their lack of success only goes to show that E&F have no phonological basis for what they are claiming. The problems extend not only to the failure of the mappings themselves, but to the failure of mappings altogether. There is no indication as to when these rules developed. This leaves the question of whether they formed when the first loanwords from OE transferred to ON, or later, when the OE speaking population shifted to the new AN. The latter seems more likely due to the greater power of SL agentivity in contact induced change.

Finally, the question of what kind of cognitive mechanisms would account for these rules cannot be answered. Whether they have any kind of psychological reality, or form a more abstract knowledge, the sheer quantity of rules and their exceptions must not be simple cognitively, and so would be difficult to account for via psycholinguistic mechanisms. Theories of second language phonological acquisition, or second dialect acquisition, may be able to shed some more light on the topic in terms of how second language learners map their L1 phonemes onto their new L2, but even mappings like this are unlikely to account for the replacement of an entire phonological system. It would also be interesting to see a large-scale corpora study of the application of the mappings proposed in this dissertation, to quantitatively determine the proportion of the lexicon correctly rendered by the transformation rules.

E&F's solution brings up a lot more questions than it answers, and their hypothesis is weakened, or disproven, by deeper consideration of the phonological evidence available. While it is always interesting to see linguists challenging the traditional assumptions of historical linguistics, in this case significantly more evidence and research would be needed to justify diverging from the norms established.

7 References

Arbor, A., & McSparran, F. (2018). *Middle English Dictionary*. Retrieved from Online edition in Middle English Compendium <http://quod.lib.umich.edu/m/middle-english-dictionary/>.

Bailey, C.J., & Maroldt, K. (1977). The French lineage of English. In J. M. Meisel (Ed.), *Langues en contact: Pidgins-Creoles-Languages in Contact* (pp. 21–53). Tübingen: Narr.

Barnes, M. P. (2016). Joseph Embley Emonds and Jan Terje Faarlund: English: The language of the Vikings. *Maal og Minne,* 108(1), 173–179.

Bech, K., & Walkden, G. (2016). English is (still) a West Germanic language. *Nordic Journal of Linguistics,* 39(1), 65–100.

Bosworth, J. (2010). An Anglo-Saxon Dictionary Online. (T. N. Toller, Ed.) Prague: Faculty of Arts, Charles University. Retrieved from <http://www.bosworthtoller.com/>.

Campbell, L., & Poser, W. J. (2008). *Language Classification: History and Method.* Cambridge: Cambridge University Press.

Chomsky, N., & Halle, M. (1968). *The Sound Pattern of English.* New York: Harper & Row.

Dance, R. (2019). Words derived from Old Norse in Sir Gawain and the Green Knight: an etymological survey. *Volume 1. Transactions of the Philological Society,* 116(2), 1– 238.

Dresher, E. (2015). Rule-based Generative Historical Phonology. In P. Honeybone, & J. Salmons (Eds.), *Oxford handbook of Historical Phonology* (pp. 501–522). Oxford: Oxford University Press.

Elderkin, E. D. (1976). Southern Cushitic. In L. Bender (Ed.), *The non-Semitic languages of Ethiopia* (pp. 278–297). Michigan: Michigan State University.

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

Emonds, J. E., & Faarlund, J. T. (2016). Anglicised Norse or anything goes. *Language Dynamics and Change,* 6(1), 49–56.

Gerken, L., Balcomb, F., & Minton, J. (2011). Infants avoid 'labouring in vain' by attending more to learnable than unlearnable linguistic patterns. *Developmental Science,* 14(5), 972–979.

Harbert, W. (2007). *The Germanic Languages.* Cambridge: Cambridge University Press.

Haugen, E. (1950). The Analysis of Linguistic Borrowing. *Language,* 26(2), 210–31.

Hockett, C. (1987). *Refurbishing our foundations: elementary linguistics from an advanced point of view*. Amsterdam: Benjamins.

Holmberg, A. (2016). Norse against Old English: 20–0. *Language Dynamics and Change,* 6(1), 21–23.

Hübschmann, H. (1875). Über die Stellung des Armenischen im Kreise der Indogermanischen Sprachen. *Zeitschrift für vergleichende sprachforschung auf dem gebiete der Indogermanischen sprachen,* 23(1), 5–49.

Hyman, L. M., & Plank, F. (2018). *Phonological Typology: Phonology and Phonetics 23*. Berlin: Walter de Gruyter.

Jones, W. (1786, February 2). The Philologer Passage. *Third Anniversary Discourse to the Asiatic Society*.

Kaplan, R., & Kay, M. (1994). Regular models of phonological rule systems. *Computational Linguistics,* 20(3), 331–378.

Knooihuizen, R. (2009). Shetland Scots as a new dialect: phonetic and phonological considerations. *English Language and Linguistics,* 13(3), 483–501.

Kortmann, B., & Schneider, E. W. (Eds.). (2008). *A Handbook of Varieties of English. A Multimedia Reference Tool. Volume 1: Phonology*. Berlin: Mouton De Gruyter.

Received: 24/06/2020
Revisions: 03/02/2021
Accepted: 15/02/2021

Kroonen, G. (2009). Etymological Dictionary of Proto-Germanic. In A. Lubotsky (Ed.), *Indo-European Etymological Dictionaries Online*. Retrieved from <https://dictionaries.brillonline.com/search#dictionary=proto_germanic&id=pg0001>.

Labov, W. (2001). *Principles of Linguistic Change: Volume 2 Social Factors*. Oxford: Blackwell.

Lai, R. (2015). Learnable vs. Unlearnable Harmony Patterns. *Linguistic Inquiry*, 46(3), 425– 451.

Lewis, J., & Elman, J. (2001). A Connectionist Investigtion of Linguistic Arguments from the Poverty of the Stimulus: Learning the Unlearnable. In J. D. Moore, & K. Stenning (Eds.), *Proceedings of the Twenty-Third Annual Conference of the Cognitive Science Society* (pp. 552–557). Mahwah: Erlbaum.

Lightfoot, D. (2016). English: The language of the Vikings by Joseph Embley Emonds and Jan Terje Faarlund (review). *Language*, 92(2), 474–477.

Meillet, A. (1925). *La méthode comparative en linguistique historique*. Paris: H. Aschehoug & Company.

Milliken, M. E., & Milliken, S. R. (1993). System Relationships in Dialect Intellegibility. *International Language Assessment Conference*, (preprint). Horsley's Green.

Munch, P. A. (1962). Til Udgiveren af "Skandinavisk Gazette". *Illustreret Nyhedsblad* 6, 27– 28.

Pons-Sanz, S. (2013). *The lexical effects of Anglo-Scandinavian linguistic contact on Old English.* Turnhout: Brepols.

Poussa, P. (1982). The evolution of early standard English: the creolization hypothesis. *Studia Anglica Posnaniensia,* 14, 69–85.

Prokosch, E. (1939). *A Comparative Germanic Grammar.* Philadelphia: Linguistic Society of America.

Ringe, D., & Yang, C. (2015). *The threshold of productivity and the "irregularization" of verbs in Early Modern English.* Retrieved from University of Pennsylvania: <https://www.ling.upenn.edu/~ycharles/digdug.pdf>.

Ross, A. (2002). *English-Old Norse Dictionary*. Ontario: In Parentheses Publishing: Linguistic Series.

Salmons, J., & Font-Santiago, C. (2016). The Descent of English: West Germanic any way you slice it. *Language Dynamics and Change*, 6(1), 37–41.

Sapir, E. (1921). *Language: An introduction to the study of speech*. New York: Harcourt.

Simms, D. P. (2014). Reviews: English: The Language of the Vikings by Joseph Embley Emonds and Jan Terje Faarlund. *Journal of Germanic Linguistics,* 28(2), 171–178.

Smith, L. C. (1998). What's all the fuss about 16 words? A new approach to Holtzman's Law. *Göttinger Beiträge zur Sprachwissenschaft,* 1, 75–100.

Stenbrenden, G. F. (2016). Why English is not dead: A rejoinder to Emonds and Faarlund. *Folia Linguistica,* 37(1), 239–279.

The Oxford English Dictionary Online. Retrieved from <www.oed.com>.

Thomason, S. (1980). Morphological instability, with and without language contact. In J. Fisiak (Ed.), *Historical Morphology* (pp. 359–72). The Hague: Mouton.

Thomason, S. (2006). Language Change and Language Contact In Brown, K. (Ed.), *Encyclopaedia of Language & Linguistics* (pp. 339-46), Amsterdam and Boston: Elsevier.

Thomason, S. (2012, December 4). English or Englsk? Retrieved from *Language Log*: <https://languagelog.ldc.upenn.edu/nll/?p=4351>.

Thomason, S., & Kaufman, T. (1988). *Language contact, creolization, and genetic linguistics*. California: University of California Press.

Townend, M. (2005*). Language and History in Viking Age England; Linguistics Relations between Speakers of Old Norse and Old English*. Turnhout: Brepols N.V.

Townend, M. (2006). Contacts and conflicts: Latin, Norse, and French. In L. Mugglestone (Ed.), *The Oxford History of English* (pp. 61–85). Oxford: Oxford University Press.

Trudgill, P. (2016). Norsified English or anglicised Norse? *Language Dynamics and Change,* 6(1), 46–48.

Van Coetsem, F. (1988). *Loan phonology and the Two Transfer Types in Language Contact*. Doordrecht: Foris.

Van Gelderen, E. (2016). Split Infinitives in Early Middle English. *Language Dynamics and Change,* 6(1), 18–20.

Vaux, B. (2008). Why the Phonological Component must be Serial and Rule-Based. In B. Vaux, & A. Nevins (Eds.), *Rules, Constraints and Phonological Phenomena* (pp. 20–60). Oxford: Oxford University Press.

Winford, D. (2003). *An Introduction to Contact Linguistics*. Oxford: Blackwell.

Winford, D. (2005). Contact-induced changes: Classification and processes. *Diachronica,* 22(2), 373–427.

About the Author

Nina Haket graduated from the University of Cambridge in 2020. She is currently studying at Universiteit Leiden for her Research Master's in Linguistics. In addition to this, she is participating in the project *When Philosophers Meet Linguists: The Conceptual Engineering Project*, working on collaborating with philosophy to bring a linguistic insight into the issues surrounding conceptual engineering. Her research interests continue to lie in historical phonology, especially in Germanic languages, as well as contact linguistics, and hopes to continue in these areas in the future.

Received: 24/06/2020
Revisions: 03/02/2021
Accepted: 15/02/2021

The Palatalisation of the Voiceless Velar Fricative in Santiago, Chile: A Variationist Analysis

Madeleine Rees
University of Cambridge

Abstract. Allophonic palatalisation of velar consonants in Chilean Spanish has been hastily attributed to the inevitable anticipatory assimilation that occurs between a back consonant and front vowel. In this study, the extent of palatalisation of the voiceless velar fricative /x/ was measured before /e/ and /i/, in male and female speakers, and in controlled and casual speech, to ascertain the relationship between palatalisation, the gender of the speaker and the speech style, from a variationist standpoint. Using reading and interview tasks, a total of 1586 /x/ tokens were taken from 14 participants from Santiago. Results show a significant drop in frequency (indicative of lesser palatalisation) in casual speech from males, but not in casual speech from female speakers. Concurrently, the following vowel also influenced the degree of palatalisation. It can be proposed that palatalisation is dictated to an extent by articulatory effort and caution, linked to differences in sociolinguistic behaviours of both genders. [h] has also been registered as a more infrequent allophone of /x/, serving as a replacement for [x] and [ç].

Plain English Abstract. Palatalisation is the production of soft palate-originating (velar) consonants further forward in the mouth than typically expected. In Chilean Spanish, this process has been hastily attributed to the physical limitations on articulation when a speaker articulates a sound originating in the back of the mouth followed by one originating in the front. However, gradient changes like palatalisation may also be controlled by social pressures exerted on speech: the variationist analysis here follows the reasoning that no socially-driven language variation of this type is random. In this study, palatalisation of the voiceless velar fricative /x/ (as in the first sound in 'gente' – /ˈxen.te/) was measured before front vowels /e/ and /i/, in male and female speakers, and in controlled and casual speech, to determine the relationship between palatalisation, speaker gender and speech style. Using controlled reading tasks and informal interviews, 1586 instances of /x/ were obtained from 14 participants from Santiago. Results indicate that the following vowel mediates palatalisation to an extent. Concurrently, male speakers show significantly less palatalisation before /e/ in casual speech, while female speakers maintain relatively constant levels of palatalisation in both speech styles. It is proposed that palatalisation is mediated by both physical articulation pressures, caused by the following front vowel, and by social pressures, particularly gendered differences in speech styles.

Keywords: velar palatalisation; Chilean Spanish; variationism; coarticulation; sociophonetics; language variation

1 Introduction

1.1 Overview

The allophonic palatalisation of the voiceless velar fricative /x/ into [ç] before [e], [i] or [j] is especially marked in Chilean Spanish (González, 2014) (see Figure 1). Currently, the most prevalent opinion surrounding this allophone suggests that it is produced from assimilation between a velar fricative and a following front vowel. The front location of the vowels /e/ and /i/ in the oral acoustic space causes the tongue dorsum to move forward anticipatorily, in turn causing the fronting of the previous consonant (Hualde, 2014). To a certain extent, the production of slightly palatalised velar fricatives is inevitable in all dialects of Spanish, due to the subconscious nature of coarticulation. However, in Chile, this

process seems to intensify: as such, the voiceless palatal fricative [ç] has been added as an allophone to its phonetic alphabet (Sadowsky & Salamanca, 2011).

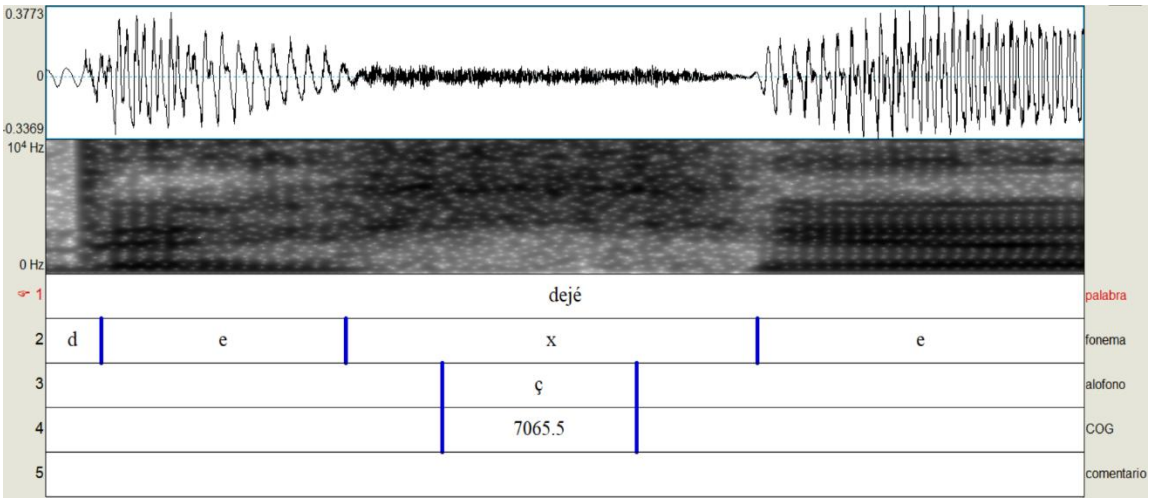


Figure 1: *Voiceless palatal fricative (tagged), female speaker.*

This topic has, so far, not been explored in depth. This investigation seeks firstly to fill a gap in the literature, which has been communicated explicitly. Sadowsky (2015, p. 82) notes that "it cannot be discounted that sociolinguistic variation may have gone undetected in other phonetic or phonological phenomena, such as the degree of palatalisation of the allophones [c], [ɟ, ɣ̟, ɣ̠] and [ç] of /k/, /g/, and /x/, respectively"[1].

Furthermore, there is another gap in previous research: the key feature of the studies on this variation, carried out by Tapia Ladino and Valdivieso (1997) and Flores (2016), is that they used speech samples obtained from television and radio, from people whose speech tends to be judged by the standards of a professional environment (Ávila, 2003). As such, the most probable result of this is that these people will change their speech to align with a more neutral professional standard. Up until now, there has been no investigation using more naturalistic speech samples taken from reading tests and linguistic interviews. Similarly, previous literature only mentions one study (Flores, 2016) in which this phenomenon has been measured using spectrograms and spectra. Therefore, this experiment seeks to analyse, instrumentally, the extent of allophonic palatalisation in natural Chilean speech.

Specifically, this paper will analyse the extent of palatalisation before two front vowels (/e/ and /i/), in male and female speakers and in two speech styles (careful and casual) following a variationist paradigm, which will be discussed more in Section 1.4. Variationism focuses on the links between the speech, social characteristics and social motivations of speakers. As such, characteristics like socioeconomic strata (Labov, 1966), gender (Eckert, 1989), perceived ethnic identity (Barrett, 1999), level of schooling (Silva Corvalán, 1987 [2001]), degree of social integration into the immediate community (Labov, 1972), and style in different contexts (Spolsky, 2003) can be influential factors in the choice of certain linguistic variants. Variationism is an interesting point of view from which to study this phenomenon, given that it is not known with certainty why palatalisation occurs, aside from the proposed theory of assimilation.

[1] "No se puede descartar que exista variación social aún no detectada en otros fenómenos fonético-fonológicos, tales como el grado de palatalización de los alófonos [c], [ɟ, ɣ̟, ɣ̠] y [ç] de /k/, /g/ y /x/, respectivamente" (Sadowsky, 2015, p. 82)

Received: 31/07/2020
Revisions: 29/10/2020
Accepted: 08/03/2021

1.2 Palatalisation in Spanish

Many Spanish dialects do not see such a strong palatalisation as that of Chilean Spanish. In Mexico, /x/ tends to be realised as voiceless velar fricative [x] regardless of the following vowel, although in some cases the place of articulation changes to be pharyngeal or even laryngeal (Butragueño, 2014). Similarly, in southern Spain and the Caribbean, the velar fricative tends to be realised as the voiceless glottal fricative [h] (Coloma, 2011), again regardless of the following vowel.

Some dialects show different changes in place of articulation of /x/ before a front vowel. In Spain, especially in central and northern areas, /x/ tends to be realised as the uvular fricative [χ] with certain inevitable coarticulatory velar advancement to a postvelar place of articulation (Hualde, 2014). Other dialects also show a minor advancement before front vowels: a certain degree of palatalisation has been seen in Uruguay and the high Andean territories of Peru (Lipski, 1996). In Bolivian Spanish, /x/ encompasses many allophones: before front vowels, its place of articulation can extend to post-palatal (Quilis & Sanz, 2003, cited in Aleza Izquierdo, 2010). Meanwhile, in Paraguayan Spanish, /x/ is realised as [x] except before a before a front vowel, whereupon it becomes [h] (Krivoshein & Corvalán, 1987). It can be concluded that although the frontness of the following vowel habitually affects the allophone of /x/ produced, the process in Chile comprises a degree of fronting quite apart from that of other Spanish dialects.

1.3 Palatalisation in Chilean Spanish

It is common to hear high frication juxtaposed with a sound similar to a palatal approximant, before a front vowel, in Chilean Spanish. The presence of the palatal fricative [ç] was noted for the first time by Rodolfo Lenz around 1890 (Lenz et al., 1940). As has been mentioned, this process differs from lesser coarticulatory fronting in other Spanish dialects, which implies that Chilean Spanish palatalisation may have social or stylistic connotations.

Studies dedicated to this phenomenon have so far been few and far between. In some cases, it only receives a brief mention: Gladys Cepeda (1991) in her book 'Las Consonantes de Valdivia' [Consonants of Valdivia] mentions a stronger degree of palatalisation in the speech of female speakers, while Cartagena (2002, p. 347) proposes that palatalisation "is found in all social strata and registers of language" [2].

With regards to more in-depth investigations, Tapia Ladino and Valdivieso (1997) analysed speech samples taken from television, finding no correlation between the gender of the speaker and the degree of palatalisation in their speech. Flores (2016), using speech samples from Chilean radio programmes, also did not find any correlation between the prevalence of any [x] allophone and the gender of the speaker. However, the results of this study bring to light various interesting results: in the 'Talk' genre, consisting of formal conversations, the palatal fricative was more frequent, but in sports broadcasts dominated by male speakers, the popularity of the palatal variant dropped.

Finally, Huskey (2010) shows that realisations of /x/ before both /e/ and /i/ have equally high centre of gravity frequencies, adding that her results showed no significant difference between the places of articulation of /x/ before both /e/ and /i/. This proposition would suggest that allophones of /x/ before /e/ were produced with the same degree of palatalisation as the /x/ allophones before /i/, which may imply that the extent of palatalisation is not dictated entirely by the extent of coarticulation, and that palatalisation is a separate process less inevitable than first thought. Importantly, from a

[2] "se encuentra en todos los niveles sociales y registros de lengua" (Cartagena, 2002, p. 347)

sociophonetic point of view, it can be proposed that in this instance, if palatalisation is separate, it is not avoided on purpose by speakers, nor is it stigmatised. This result raises the question of whether a certain degree of palatalisation can possess social or stylistic connotations.

1.4 Variationism

Fundamentally, the variationist paradigm supports the view that linguistic variation does not exist in a vacuum, without being subject to social and sociolinguistic pressures (Labov, 1972). Here, it cannot be assumed that linguistic variation is free nor accidental, but instead that it is dictated systematically (Milroy & Gordon, 2003). It is also not assumed that trajectories of synchronic and diachronic linguistic change are separated (Weinreich et al., 1968, p. 188): patterns of variation in specific speech communities help to understand the trajectory of a given linguistic change over time. These notions may be useful with regards to palatalisation, with its scarcity in different Spanish dialects worldwide raising the questions of the linguistic factors involved, as well as the extent and permanence of such a change.

The intensification of this palatalisation process uniquely in one country suggests that a certain quantity of social variation, as well as the extent of assimilation, may dictate its presence in Chilean Spanish. Previous experiments by multiple linguists support the notion that speech can be influenced by the gender of the speaker and the formality of the conversation. Women tend to favour the standard speech variety, especially during a formal conversation which demands a more controlled speech style, whereas male speech tends to focus on higher usage of nonstandard variants (Trudgill, 2000; Tagliamonte, 2011; Kiesling, 2011). Having mentioned the concept of normativism and standard versus nonstandard variants, it is important to add to this the concepts of overt and covert prestige (Trudgill, 1972). The prestige of a variant has the power to dictate the behaviour of speakers and their choice of preferred allophone, according to the characteristics of the speaker and the social environment in which the speech act takes place. Female speakers tend to choose standard variants, which enjoy overt prestige within the speech community and on a broader societal level, whereas male speakers tend to favour variants with covert prestige, which have a certain prestige amongst the members of the immediate speech community, but not necessarily in larger society.

López Morales (2004, pp. 56–57), adds that choice of variants can be determined jointly by linguistic and social factors. Therefore, it would be useful to investigate the interface of coarticulation and sociolinguistics factors, to see the results of the union of physical factors (which cause ‘inevitable’ allophonic realisations of a given phoneme, out of the control of the speaker) and social factors (which cause productions of a given phoneme to be chosen on purpose). Having said this, if it becomes evident that no social variation controls the palatalised allophone, it can be proposed that the variation is now a concrete change, and that the degree of palatalisation has social connotations which are at least neutral, to avoid its eradication at the hands of speakers.

2 Aims of the investigation

2.1 Objectives

This investigation will focus on three facets of this phenomenon. Firstly, linguistic variation can be dictated by purely linguistic factors (Moreno Fernández, 2009). Therefore, it is necessary to compare degree of palatalisation before both /e/ and /i/: given that /e/ is articulated slightly further back in the oral space, it would be logical to suggest that palatalisation should be lesser before /e/.

Received: 31/07/2020
Revisions: 29/10/2020
Accepted: 08/03/2021

Secondly, this investigation will focus on the effect of two speech styles on palatalisation. The first style is careful speech, measured from a reading task, while the second is casual speech, obtained during a free conversation interview with no fixed questions. It is assumed that the speech style obtained from the reading task will be more formal, with higher usage of standard variants, than during the more informal interview.

Thirdly, this paper will investigate if the gender of the speaker affects their propensity to palatalise more, which can be an indicator of the prestige of a variant. The combination of gender and style has the possibility to further illuminate the prestige of palatalisation, as a combination of female gender and formal conversation has been shown previously to lead to the use of the most prestigious variants.

2.2 Hypotheses

(1) $H_{0(1)}$: The following vowel exerts no effect on the degree of palatalisation.
H_1: The following vowel influences palatalisation, with /i/ triggering a more fronted realisation of /x/ than /e/.

(2) $H_{0(2)}$: Palatalisation is not influenced by any environment or speech style.
H_2: Coarticulation increases in informal speech (Browman & Goldstein, 1987). If the allophone is produced by assimilation, we would expect a greater degree of palatalisation in casual speech elicited by informal interview.

(3) $H_{0(3)}$: Palatalisation is not influenced by gender of speaker.
H_3: Differences in the degree of palatalisation produced by male and female speakers: previous research (Cepeda, 1991; Flores, 2016) suggests that female speakers will favour a more palatal realisation of /x/.

3 Investigation Design and Methodology

3.1 Speaker Sample

3.1.1 Speaker Characteristics

The speaker pool was composed of 7 men and 7 women, all monolingual Chilean Spanish speakers from Santiago Metropolitan Region and all between 25 and 40 years old. No participant had lived outside Santiago for more than 2 years. No participant reported any speech or hearing difficulties. To avoid unexpected sociophonetic variation dictated by membership of a socioeconomic stratum or group of strata (see Haska, 2016) all participants belonged to socioeconomic strata ABC1 (pre-October 2018) according to their answers to a sociodemographic questionnaire (based on UK NRS (Ipsos MediaCT, 2009)). Participants were recruited through social media: 2 pairs of the 14 participants knew each other previously and habitually interacted with each other; the other 10 did not know each other. The researcher was functionally bilingual in Spanish, but not a native speaker of the Chilean variety.

3.1.2 Ethical Protocol

Participants were reminded of their right to retire from the investigation at any point, without repercussion or the need to provide an explanation. To maintain anonymity, all participants received a

code to replace their name. Sensitive data was password protected. Interview locations were well connected and structurally protected in case of earthquake.

3.2 Elicitation Tools

Two individual tests were administered in one session. Firstly, participants carried out a reading task consisting of a presentation of phrases and short texts, designed to elicit productions of /x/ read out loud (see Appendix). The test contained 30 instances of /x/ before /e/ and 30 instances of /x/ before /i/. Slides were changed by the participant using a keyboard, which let each speaker read the slide at their own pace. During this section, the investigator left the room, so as not to influence the speech style of the participant. This test also contained various texts without /x/, so as not to reveal the motives of the experiment. The reading task was followed immediately by a casual conversational interview without pre-prepared questions, with the aim of generating a more casual speech style in participants.

3.3 Recording Techniques

Acoustic data was recorded with a Fostex FR-2LE recorder and Audix HT5 microphone with a sampling rate of 24 bits/48KHz. The microphone used was head-mounted, which helped to minimise the obvious and potentially intrusive presence of a recording implement. All interviews were carried out in offices or apartments with some degree of soundproofing: due to the nature of the experiment, it was necessary to record acoustic data without interference from external sources.

3.4 Corpus

The reading task lasted around 20 minutes, while the conversation lasted around 40 minutes. Ultimately, the total corpus collected consisted of approximately 4.6 hours of careful speech from the reading task and approximately 9.3 hours of casual speech from interviews, shared equally between male and female participants.

3.5 Data Analysis and Processing

It was decided to directly measure centre of gravity, instead of assigning one production to allophone [x] or [ç]. The place of articulation of a palatal and velar fricative is not a strict binary difference, but rather more of a continuum, which does not entirely correspond with the assignation of one allophone or another to a recording. Centre of gravity is the average of the frequency components of a sound in a given time period, which shows the highest concentrations of energy in the signal, therefore indicating which frequencies have the highest amplitude (Maniwa et al., 2009). Although raw centre of gravity figures in themselves cannot be directly mapped to a place of articulation, it is widely accepted (e.g., Jongman et al., 2000) that the higher the centre of gravity frequency, the further forward the constriction in the oral cavity that produces the fricative. Therefore, it is possible to measure the difference between centre of gravity frequencies and correlate it with the distance between places of articulation of an /x/ allophone.

Spectrograms of each interview were generated using Praat 6.0.37 (Boersma & Weenink, 2018; latest version Boersma & Weenink, 2020). Tokens of /x/ before /e/ and /i/ were manually found and annotated. Each centre of gravity frequency was taken from a spectrum. Spectra were created from the middle third (selected manually at 33% and 66%) of the frication segment, where frication was most

Received: 31/07/2020
Revisions: 29/10/2020
Accepted: 08/03/2021

stable (see Figures 2–3). The program took multiple centre of gravity readings from the selected segment, generating a final average for the spectrum. Centre of gravity values were transferred to a spreadsheet, along with the word in which the segment appeared, the speech style, and participant code.

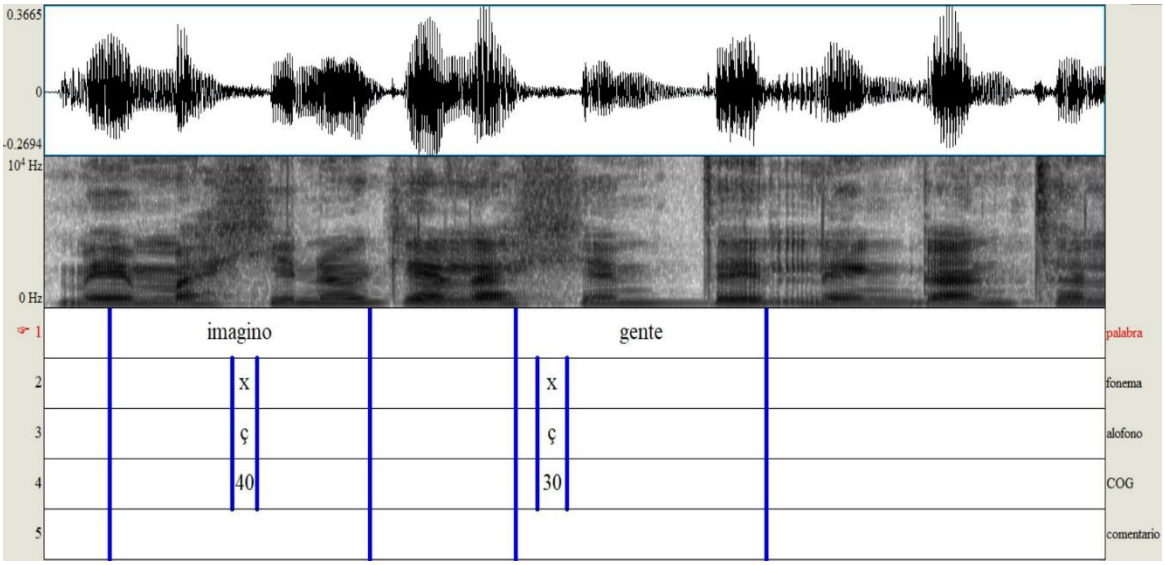


Figure 2: *Exemplar spectrogram, female informant.*

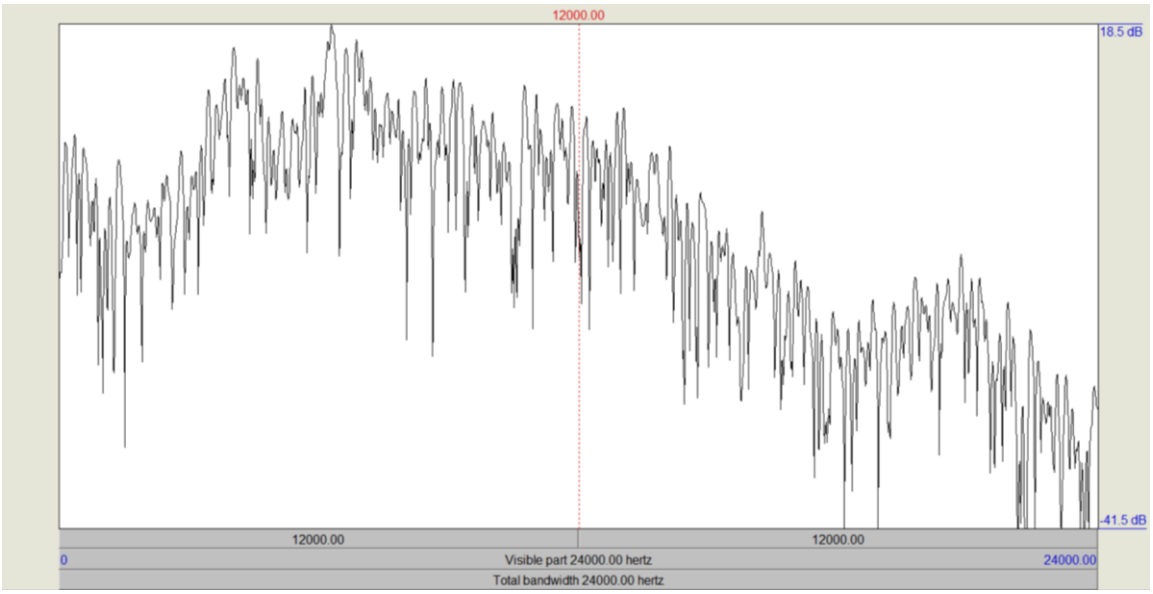


Figure 3: *Spectrum of palatal fricative, female informant, centre of gravity 7128.8 Hz.*

To ensure a high quality of speech sample, tokens with acoustic interference were deleted. Repetitions of the same token caused by speech errors during the test were also deleted. The final database consisted of 856 careful tokens (of which 437 were from male and 419 from female speakers) and 700 casual tokens (337 from male speakers, 363 from female speakers), creating a database of a total of 1556 tokens. To obtain paired data for each participant in the different conditions, means were created from each list of COG frequencies. Paired t-tests were used on variable means to determine statistical significance. Graphical output was generated in Python (Van Rossum & Drake, 2009) using the package MatPlotLib (Hunter, 2007) and the Seaborn wrapper (Waskom, 2017).

4 Results

4.1 General Results

Table 1: *Average centre of gravity frequencies (Hz) sorted by style and following vowel, male participants*

	Careful, before /e/	**Careful, before /i/**	**Casual, before /e/**	**Casual, before /i/**
M2-STG	4079.4	4376.2	3301.1	3576.4
M3-STG	3246.7	3614.2	3160.5	3793.3
M4-STG	3442.7	4179.9	3411.5	4478.1
M5-STG	3331.6	3595.3	2369.6	3579.3
M6-STG	2758.5	3099.9	2655.8	3552
M7-STG	3758.9	3762.5	3168.8	3545.2
M8-STG	2061.4	2704.3	2057.4	2754.6
AVG	3239.9	3618.9	2874.9	3611.3

Table 2: *Average centre of gravity frequencies (Hz) sorted by style and following vowel, female participants*

	Careful, before /e/	**Careful, before /i/**	**Careful, before /e/**	**Casual, before /i/**
F1-STG	2755.5	2914.3	2521	2644.6
F2-STG	4561.7	5284.9	4411.1	5066.1
F4-STG	3609.2	4242.8	2633.4	4002.5
F5-STG	3733.8	4528.2	3537.9	5088.3
F6-STG	5784.5	6826.1	4037.6	5911.3
F7-STG	5602.1	5767.8	5672.2	5756.2
F8-STG	2541.1	2580.1	2943.4	2948.7
AVG	4084	4592	3679.5	4488.2

Received: 31/07/2020
Revisions: 29/10/2020
Accepted: 08/03/2021

4.2 Results: Following Vowel

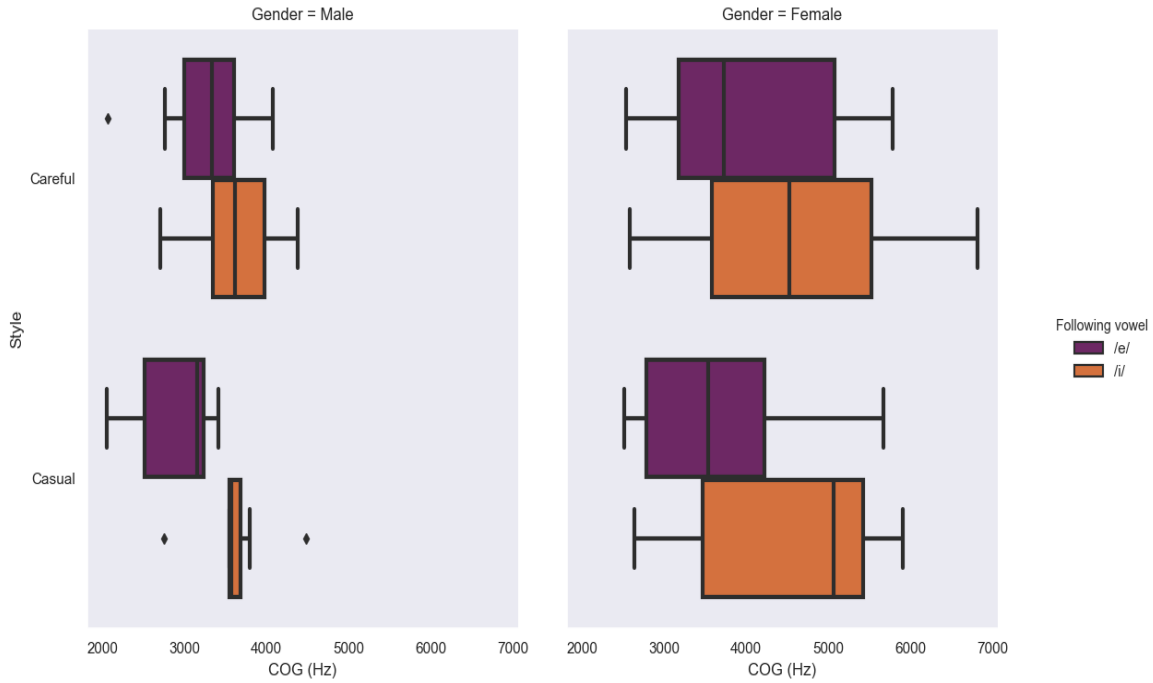


Figure 6: *Average centre of gravity according to following vowel, in all 8 categories.*

The results in Figure 6 demonstrate a pattern in the general extent of palatalisation exhibited by both genders. Centre of gravity tends to decrease before /e/ and to stay consistently higher before /i/. Significant differences were found in the centre of gravity frequencies before /e/ and /i/, for both genders in both speech styles: Male/careful: $t(6) = -4.0911$, $p = 0.006$; Male/casual: $t(6) = -5.651$, $p = 0.001$; Female/careful: $t(6) = -3.494$, $p=0.01$; Female/casual: $t(6) = -2.739$, $p = 0.03$.

It can be suggested that this pattern is the logical result of coarticulation before two vowels with different places of articulation, with /i/ being further forward than /e/: stronger coarticulatory behaviour emerges to facilitate production of the furthest forward vowel after a velar consonant.

4.3 Results: Style

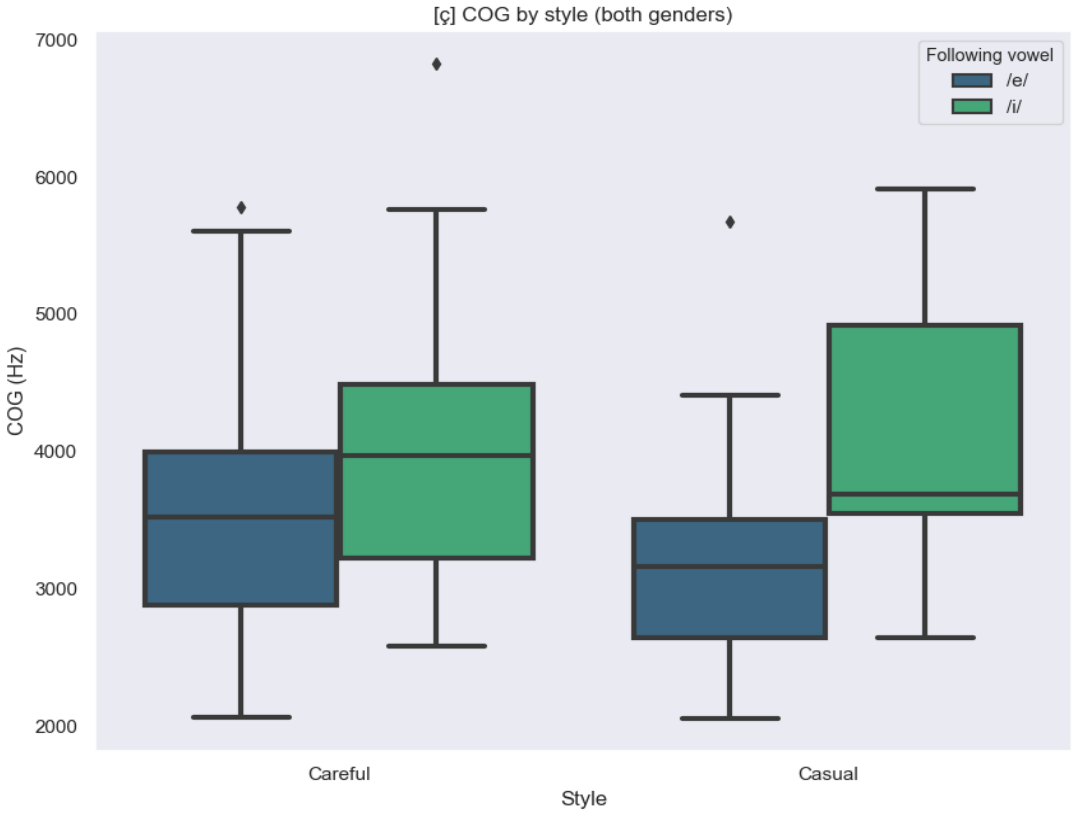


Figure 7: *Average centre of gravity frequencies according to style, taken from results of both genders.*

As shown in Figure 7, careful speech is associated with an increase in palatalisation for both men and women, especially before /e/. The combined frequencies indicate that this increase is significant ($t(13) = 2.559013$; $p = 0.02$). On the other hand, casual speech appears to cause a drop in frequencies, implying that palatalisation takes place on a lesser scale when speakers pay less attention to their speech. In comparison to /e/, the frequencies before /i/ remained relatively constant for all participants regardless of speech style ($t(13) = 0.481457$; $p = 0.6$).

Therefore, it can be concluded that participants have a greater tendency to palatalise realisations of /x/ in environments where they pay more attention to their speech, and that this trend is more frequent before /e/. However, it is apparent that before /i/, place of articulation remains relatively fixed despite the change in style. In the next Section, the results according to gender will be analysed, in order to illuminate any potential intersection between gender and style.

Received: 31/07/2020
Revisions: 29/10/2020
Accepted: 08/03/2021

4.4 Results: Gender

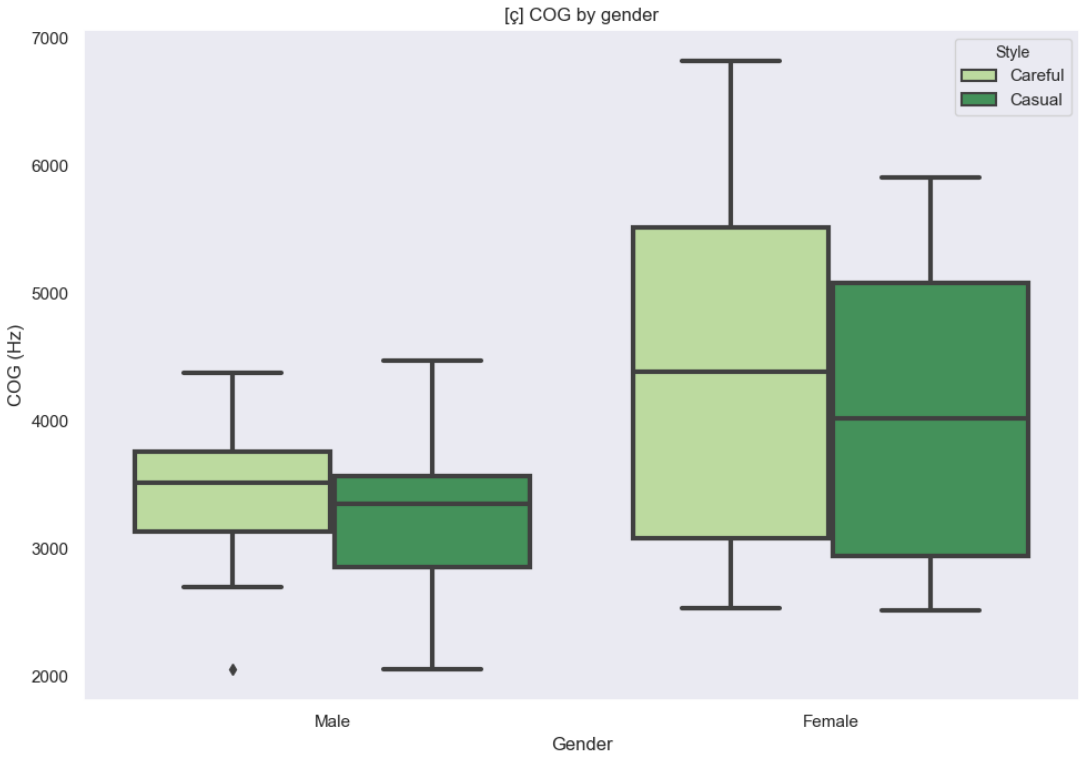


Figure 8: *Average centre of gravity frequencies from both genders. It is expected that the women will have higher voices and thus typically higher COG values due to the different physical characteristics of their vocal tracts.*

4.4.1 Results: Male Gender

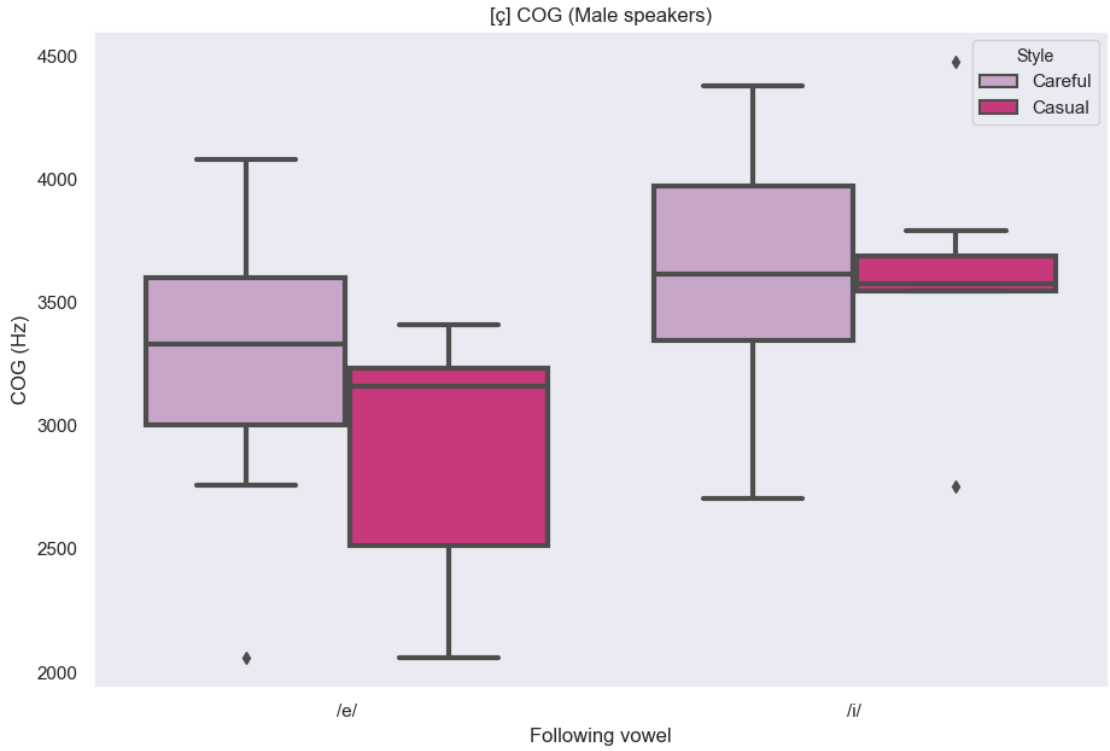


Figure 9: *Average centre of gravity, careful and casual speech, male participants.*

Bearing in mind the centre of gravity frequencies registered in careful speech from male speakers (see Figures 8–9) these frequencies dropped significantly in casual speech before /e/ ($t(6) = 2.406$; $p = 0.05$). The difference in frequencies before /e/ suggests that the realisation of /x/ in this phonological environment tends to be significantly further back in casual masculine speech in comparison to their careful speech. As the difference in frequencies before different vowels is greater in casual speech, it appears that the male participants follow the pattern of consonant-vowel coarticulation more closely in this style. It could even be proposed that they are more susceptible to the influence of the further-back vowel /e/ when they pay less attention to their speech.

Contrary to their results before /e/, the male participants did not present any significant difference in centre of gravity frequencies before /i/ ($t(6) = 0.049$, $p = 0.96$). This result suggests that place of articulation remains constant and that there is no sociolinguistic factor that influences their speech in this phonological environment: in other words, centre of gravity is only influenced by the following vowel.

Received: 31/07/2020
Revisions: 29/10/2020
Accepted: 08/03/2021

4.4.2 Results: Female Gender

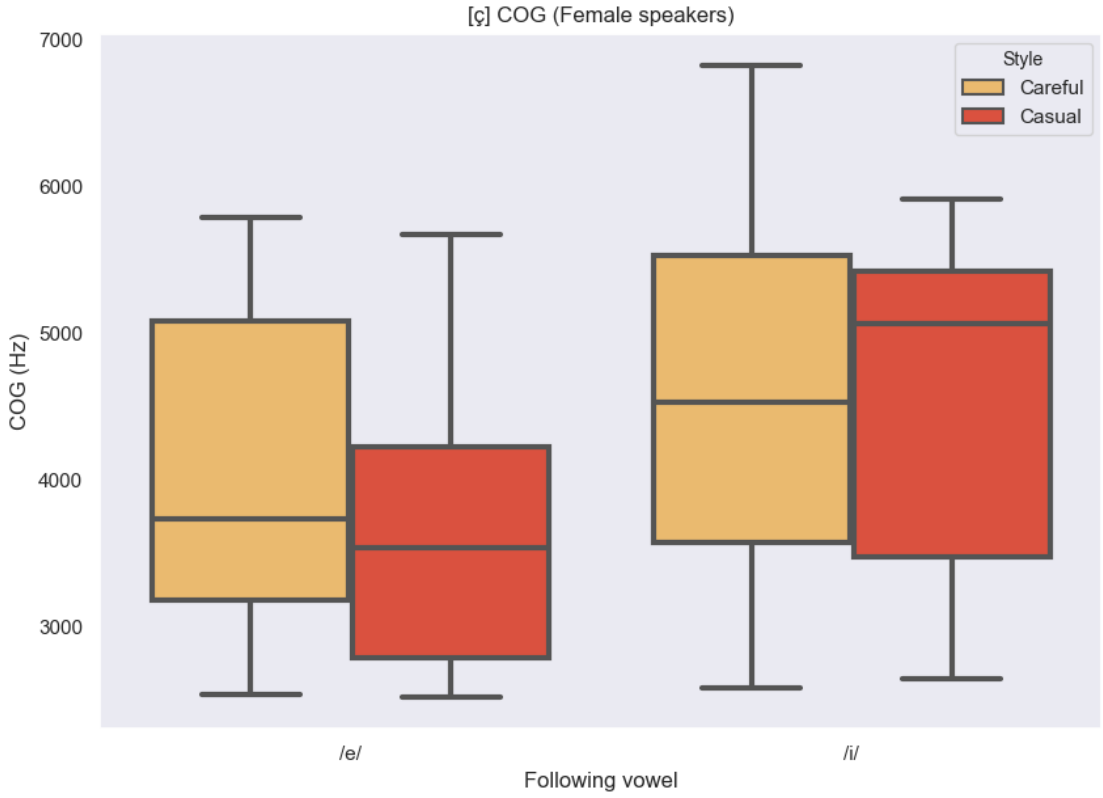


Figure 10: *Average centre of gravity, careful and casual speech, female participants.*

Firstly, as opposed to the male participants, female speakers did not present any significant frequency differences in careful and casual speech before /e/ ($t(6) = 1.479$; $p = 0.19$). Although Figures 8 and 10 show a decrease in centre of gravity, this change was not found to be significant: while there were indeed some female participants whose average frequencies dropped during the casual interview, there were also some participants whose frequencies increased during the informal conversation. Therefore, although their place of articulation of /x/ oscillates, this fluctuation does not equal a significant change in the degree of palatalisation, implying that place of articulation in female speech remains more constant in the two speech styles than it does in male speech.

Secondly, in a similar fashion to the male participants, the female cohort did not demonstrate significant changes in centre of gravity before /i/ in both speech styles ($t(6) = 0.571$; $p = 0.59$), although the averages dropped slightly in casual speech. Consequently, it can be concluded again that neither gender nor style influence productions of /x/ before a very front vowel.

5 Discussion

5.1 Discussion of Results by Style

Reiterating the second hypothesis previously mentioned in Section 2, it was proposed that there would be a difference in the degree of palatalisation in careful and casual speech styles. Normally coarticulation and the resulting assimilation is assumed to be a sign of a more casual speech style, which implies that it should be more prevalent during the interview. However, both masculine and feminine

centre of gravity averages (especially before /e/) dropped during the interview, implying that there was a lesser degree of palatalisation during this activity. Therefore, the decrease in frequencies appears to be linked to a decrease in articulatory caution: correspondingly, an increase in frequency (indicative of a more palatal realisation) appears to be correlated with an increase in speech control.

According to previous literature, an allophone caused by assimilation is unlikely to be more frequent in careful speech. However, this unexpected result could have arisen because the voiceless palatal fricative is articulated with the greatest intensity of all voiceless fricatives, implying that air pressure is high during articulation (Strevens, 1960). As a result, it is possible that the production of such a sound requires a special articulatory effort and therefore that its presence is indicative of a more careful articulation. Similarly, velarisation of pre-palatal fricatives (the most similar to [ç]) in old Spanish is cited as an example of lenition (Ariza, 2004, p. 13), implying that a palatal place of articulation requires articulatory effort to sustain. This effort to produce a more fronted palatal variant could become more evident before /e/, whose location further back in the oral acoustic space reduces the degree of inevitable coarticulation that helps to generate a palatal fricative.

At times, speakers used alternative allophones. In several instances they produced cases of the voiceless glottal fricative [h], which has a markedly lower centre of gravity than [x] or [ç], as shown in Figure 11.

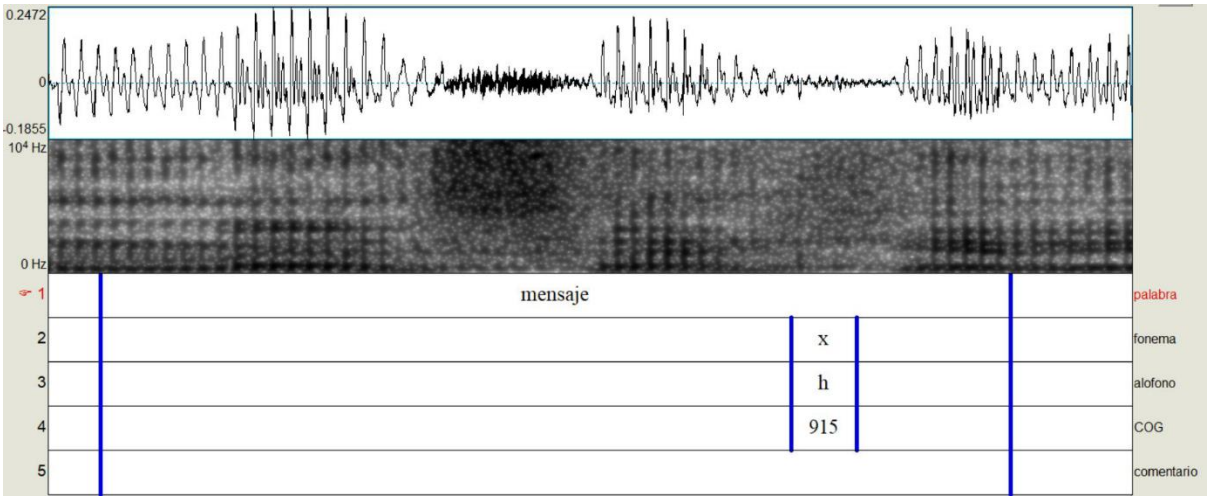


Figure 11: *Example of [h] as an allophone of /x/ (COG 915Hz), male participant.*

O'Brien (2012) proposes that this process happens due to debuccalisation, where the consonant loses its place of articulation and becomes a laryngeal allophone (it is assumed that the appearance of [h] is due to lenition and not to other processes such as exonorms or dialect contact, given the perseverance of [x] and [ç] before other vowels in the recordings). The frequency of lenition seems to be positively correlated with the informality of the situation. It appears that the generation of a palatal fricative (especially before /e/) demands more articulatory effort and a careful articulation from the speaker. In return, to let /x/ be realised as [h] permits a fast articulation (the instances of [h] were shorter than those of [ç]) without great constriction, which is useful in casual speech where speakers are less focused on the modulation of words. It could even be proposed that [h] seems to be the result of a corresponding lack of articulatory effort when compared to [ç].

It cannot be proclaimed with certainty that the palatal fricative is chosen on purpose by speakers, given that the frequency values before /i/ only varied slightly in both styles. As palatalisation seems to be linked with controlled speech, the use of less palatalised variants seems to be linked with speed and

Received: 31/07/2020
Revisions: 29/10/2020
Accepted: 08/03/2021

lack of caution in speech. It could be suggested that these variants are not articulated on purpose, but that they are merely an effect of casual, fast, and colloquial speech. These results together would imply that the degree of palatalisation is not chosen deliberately, but by default depending on the speech style. This change according to style could demonstrate that the process is also tied to the sociocultural factors that surround the speech of both genders, discussed in greater depth in the following chapter.

5.2 Discussion of Results by Gender

Bearing in mind the third hypothesis that gender will affect the degree of palatalisation, it was noted, above all, that both men and women tended to front their realisations of /x/ in controlled speech. However, an intersection between gender and style is also apparent. Female speakers maintained a reasonably constant degree of palatalisation in both speech styles, whereas the male speakers exhibited a significantly lesser degree of palatalisation in casual speech. Therefore, if palatalisation is linked to articulatory effort, it can be inferred that the women maintain a constant level of effort, while the men take advantage of the informal situation to relax their effort levels. To an extent, these results coincide with sociolinguistic patterns discovered by variationists.

Silva Corvalán (1987 [2001]) suggests that women try to stand out in their environment by employing more 'correct' forms of speech, lacking opportunities to stand out in other environments such as business, politics, etc. (although women now enjoy more opportunities in employment, it is possible that this careful linguistic behaviour has been maintained). Extrapolating this information to this context, as production of highly palatal fricatives is most common when the speaker is most conscious of their speech, it is therefore logical that palatalisation increases when a speaker is culturally and socially inclined to pay attention to their speech, and even more so during a situation which requires a high level of consciousness of speech and articulation.

Women tend to dictate the trajectory and results of linguistic variation. There exist points in favour and against the notion that the presence of [ç] is the result of change 'from below.' Supposedly, the allophone is derived in part from assimilation (an idea that the differences in frequencies before different vowels would support), an internal linguistic factor. Women tend to be more innovating in terms of this type of variation, being the first to adopt new variants which result from systematic linguistic change (Labov, 2006). However, 'from below' change is carried out with a low level of consciousness of the process and the variant tends not to demonstrate stylistic change (Labov, 1972; Hawkey, 2016). This appears to contradict the greater degree of palatalisation in the reading test, where speakers paid more attention to their speech. It could be the case that a lesser degree of palatalisation appeared first, paving the way for the generation of [ç], which subsequently became a signal of articulatory effort. In that case, the more conscious 'from above' change also tends to be spearheaded by women, who calculatedly choose the most prestigious new variants (Labov, 2001). Although a high degree of palatalisation may not be chosen on purpose, it seems reasonable to say that a high level of articulatory effort (which could have some prestige) could indeed be chosen consciously. Similarly, women tend to be the first to reject new variants which they do not consider acceptable for social situations (Labov, 1990). Independently of their origin, if it were the case that palatal variants of /x/ were not considered acceptable, it is probable that they would have been somewhat less frequent or even eradicated in female speech.

Contrary to the tendencies of the women, male speakers registered a lower degree of palatalisation before /e/, especially during the casual interview. This trend indicates that when a careful speech style was not necessary, the men were less predisposed to palatalise their realisations of /x/ when the following vowel did not (inevitably) physically trigger a greater degree of palatalisation. As a result, there exists a need to search for reasons as to why [ç] lacks popularity among men.

It is a popular view that language (consequently, use of certain allophones) plays a key role in the expression of gender identity (Cheshire, 2004; Lawson, 2014). At the same time, men tend to employ more non-normative colloquial variants in speech (see Labov, 2001). Kiesling (1998) suggests that men, focusing on the acquisition of physical strength as a source of power, do not have to concern themselves with the search for status through language use in the same way as women and do not have to modify their speech, or expend as much effort in articulation, in order to increase their social prestige. It is possible that masculine idiolects are dictated by the desire to project a relaxed image in terms of speech, implying that the desired power already comes from other sources. As a result, this masculine identity and self-image could be demonstrated by apathy towards articulatory effort. This preference would be exacerbated in informal situations, where the result of the conversation does not have significant implications for an important sphere such as the working world, or in politics, for example, and a high level of attention to speech is simply not necessary in their eyes.

Coates (2016) affirms that linguistic variation is most probably started and dictated by differences in linguistic behaviour between genders, while Tagliamonte (2011) concludes, as a result of several studies, that men can consciously differentiate their speech from the speech of women. Extending these notions to this context, it is possible that a highly palatalised allophone is a marker of formality and articulatory effort, causing men to avoid it consciously in order to differentiate themselves from women in their speech. However, according to the sets of results specifically before /i/, it seems probable that the apathy is not towards the palatalised allophone specifically, but towards the articulatory effort that generates it. As such, patterns of effort and caution in speech between genders reflect sociolinguistic patterns observed by variationists in the past, but it is doubtful that there is a conscious choice of exact allophones in this specific case.

5.3 Comparison with Other Studies

Huskey (2010) found no significant difference between centre of gravity values before /e/ and /i/, while the results of this experiment show a clear disparity between differing prevocalic centre of gravity values. Considering that frequencies before /i/ were similar in both genders, the convergence in Huskey (2010) could indicate a greater prevalence of high centre of gravity frequencies before /e/. This result may be explained by the elicitation method: Huskey used a series of short phrases with the template 'I say *x*… because I do' which would have generated an extremely controlled speech style due to the minimal quantity of words in each phrase. The positive correlation between speech control and more extreme palatalisation appears to support the results of this study.

However, the results of this investigation do not completely match results from previous literature. Tapia Ladino and Valdivieso (1997) found that the palatal fricative was not significantly more prevalent in one style or speaker gender. Although this investigation and that of Flores (2016) coincide in that palatal fricatives were more favoured in more formal conversational styles (in Flores, the formal discussion programmes broadcast on radio channels), it would be a leap of the imagination to compare a radio programme to a linguistic reading task. Flores did not find that the speaker gender influenced palatalisation. However, it is worth mentioning that in the Flores study, female speakers tended to palatalise more when speaking with an interlocutor, which may indicate a greater desire to speak in a more prestigious fashion, or more carefully, in the presence of others.

The difference between the previous two sets of results and this paper may be explained by the change in elicitation methods. The results from previous literature were taken from speakers likely to have undergone some form of training concerning speech styles in mass media communication. Contrarily, in this experiment, data was taken from linguistic interviews. Even when accounting for some possible lingering effects of the observer's paradox, it can be imagined that these participants

Received: 31/07/2020
Revisions: 29/10/2020
Accepted: 08/03/2021

were more likely to employ a more naturalistic speech style outside of a publicly-broadcast environment.

6 Conclusions and Projections

Firstly, the null hypothesis that the following vowel does not influence the degree of /x/ palatalisation can be rejected. Centre of gravity values were notably higher before /i/ than before /e/. This result is not surprising: as /i/ is a fronter vowel than /e/, it stands to reason that /i/ impulses a greater degree of palatalisation to better facilitate the articulation of juxtaposed front and back phonemes.

Secondly, it can be provisionally concluded that speech style influences the degree of palatalisation, especially before /e/. However, this experiment generated the unexpected result of an increase in palatalisation in careful speech. This result appears to contradict the conclusions of Browman and Goldstein (1987) who proposed that coarticulation and resulting assimilation were more frequent in more casual speech styles. An explanation of this apparently discordant set of results could stem from the notion that velar palatalisation is not entirely due to coarticulation. Instead, velar palatalisation can be considered a separate and not entirely inevitable process, which can be influenced to a certain degree by anticipatory coarticulation stemming from the place of articulation of the following vowel.

Thirdly, the results here suggest a possible interaction of style and gender. In this case, the centre of gravity figures indicate that the degree of palatalisation by male speakers dropped significantly when the formality of the speech style dropped, while this change was less pronounced in female speakers. However, this phenomenon only took place before /e/, suggesting that assimilation has a larger role in governing palatalisation before /i/. It can be proposed that male speakers, in casual speech, have a lower propensity to engage in articulatory effort. Therefore, they tend to follow patterns of coarticulation, submitting to the physical and coarticulatory effects of the less front vowel and producing less palatalised fricatives. On the other hand, productions of female speakers showed less significant differences in frequencies between careful and casual speech, implying that they exhibited a more constant articulatory effort.

With regards to the possibility of predicting the propensity of certain speakers to palatalise the velar fricative, we can see a potential interaction between linguistic variation dictated by social factors and linguistic variation dictated by articulatory factors. It can be concluded that, to a certain point, social factors have an influence over the degree of /x/ palatalisation, given its positive correlation with a careful style of speech and its more frequent employ by female speakers. However, it would be a bold statement to propose that highly-palatalised variants enjoy overt prestige in Chilean Spanish, given that the degree of palatalisation was effectively equal in both genders and both styles before /i/. In this case, it appears that the effects of coarticulation are inevitable and that palatalisation is not totally controlled by social and stylistic motivations: assimilation plays a role in the centre of gravity of /x/. The presence of what appears to be variation dictated by multiple sources (as proposed by López Morales (2004)) generates questions surrounding the level of speaker consciousness of the variation. The most logical proposition, in this case, would be that the degree of /x/ palatalisation is not controlled by conscious choice, but rather, that its presence is controlled by articulatory effort and caution in speech: as a result, it is inevitably tied to sociocultural factors surrounding linguistic variation and imitates patterns previously seen in the variationist paradigm.

It is worth mentioning, briefly, that variation in /e/ and /i/ in Chilean Spanish also tends to carry connotations of prestige or stigmatisation (Soto Barba, 2007). Therefore, if palatalisation is physically linked to vowel place of articulation, it stands to reason that a speaker who exhibits prestigious variants

of these vowels would also exhibit more prestigious centre of gravity values, by default, and without consciousness of the process.

Although measures were taken to ensure the quality of the investigation and naturalistic qualities of the speech samples, there were some obstacles to research which should be mentioned here. Firstly, tokens of /x/ before /i/ appeared only infrequently during the conversational interview. This problem could have been solved by extending the duration of the interview, but this measure was not taken due to the uncomfortable nature of the microphone after too long a time. Secondly, it would have been beneficial to interview more speakers, but it was necessary to balance the size of the participant pool with the time available. Ideally, a further study with more participants and more powerful statistical testing would be compared to the conclusions here.

The cause of extreme palatalisation in Chile is still not known with certainty. One explanation could be the geographical isolation of Chile: surrounded by ocean and the Andes mountains, it is more likely that linguistic variation will take hold more intensely if there is less possibility of dialect contact in border areas. However, dialect contact is surely established by contemporary immigration of speakers of different varieties of Spanish, who settle in Chile. Alternatively, it is possible that palatalisation is caused by contact with a substrate language (in this case, Mapudungun) where speakers palatalise velar consonants strongly before front vowels (Pérez & Salamanca, 2016). Further investigation of the contact between these two languages would surely illuminate more aspects of this phenomenon. Equally, an expansion of the experiment to involve different geographical regions and socioeconomic strata would give rise to further useful insights surrounding palatalisation.

7 References

Aleza Izquierdo, M. (2010). Fonética y fonología. In M. Aleza Izquiero & J.M. Enguita Utrila, (Eds.), *La lengua española en américa: normas y usos actuales* (pp. 51–94). Valencia: Universitat de Valencia.

Ariza, M. (2004). Revisión del cambio fonético y fonológico. *Lexis,* XXVIII(1–2), 7–27.

Ávila, R. (2003). La pronunciación del español: medios de difusión masiva y norma culta. *Nueva Revista de Filología Hispánica*, 51(1), 57–79.

Barrett, R. (1999). Indexing polyphonous identity in the speech of African American drag queens. In M. Bucholtz, A.C. Liang, & L. A. Sutton (Eds.). *Reinventing identities: The gendered self in discourse* (pp. 313–331). New York: Oxford University Press.

Boersma, P. & Weenink, D. (2018). Praat: doing phonetics by computer [Computer program]. Version 6.0.37. Retrieved from: <http://www.praat.org/>.

Boersma, P. & Weenink, D. (2020). Praat: doing phonetics by computer [Computer program]. Version 6.1.16. Retrieved 30/07/2020, from: <http://www.praat.org/>.

Browman, C. P. & Goldstein, L. (1987). Tiers in articulatory phonology, with some implications for casual speech. In J. Kingston, & M.E. Beckman. (Eds.). *Papers in Laboratory Phonology I: Between the Grammar and the Physics of Speech* (pp. 341–376). Cambridge: Cambridge University Press.

Butragueño, P. M. (2014). Distribución dialectal de /x/ en datos del Atlas Lingüístico de México. In P.M. Butragueño. & L. Orozco (Eds.). *Argumentos cuantitativos y cualitativos en sociolingüística: Segundo coloquio de cambio y variación lingüística* (pp. 107–142). D.F., México: El Colegio de México.

Cartagena, N. (2002). La evolución fonética y gramatical del español en Chile. *Boletín de Filología,* 39(1), 339–361.

Cepeda, G. (1991). *Las Consonantes de Valdivia*. Valdivia: Universidad Austral de Chile: CONICYT.

Received: 31/07/2020
Revisions: 29/10/2020
Accepted: 08/03/2021

Cheshire, J. (2004). Sex and gender in variationist research. In J.K. Chambers, P. Trudgill & N. Schilling-Estes (Eds.) *The Handbook of Language Variation & Change* (pp. 423–443). Oxford: Blackwell.

Coates, J. (2016). *Women, Men and Language: A Sociolinguistic Account of Gender Differences in Language* (3rd ed). Oxford: Routledge.

Coloma, G. (2011). Caracterización fonética de las variedades regionales del español y propuesta de transcripción simplificada. *Revista de Filología Románica*, 28, 11–27.

Eckert, P. (1989). The whole woman: sex and gender differences in variation. *Language Variation and Change*, 1, 245–267.

Flores, T. L. (2016). Velar palatalization in Chilean public speech. *Glossa: a journal of general linguistics*, 1(1), 1–6.

González, C. (2014). Prevocalic velar advancement in Chilean Spanish and Proto-Romance. In M. Coté & E. Mathieu (Eds.). *Variation within and across Romance languages* (pp. 277–296). Amsterdam: John Benjamins Publishing Company.

Haska, C. (2016). La percepción fonético-fonológica del fonema /tʃ/ del español de Chile: un estudio sociofonético experimental. *Estudios filológicos*, 57, 65–78.

Hawkey, J. (2016). Developing discussion of language change into a three-dimensional model of linguistic phenomena. *Language and Linguistics Compass*, 10(4), 176–190.

Hualde, J. I. (2014). *Los sonidos del español*. Cambridge: Cambridge University Press.

Hunter, J. D. (2007). Matplotlib: A 2D graphics environment. *Computing in Science & Engineering*, 9(3), 90–95.

Huskey, L. C. (2010). *Velar Palatalization: A cross-dialectic study of Chilean, Mexican, and Castilian Spanish*. Honors thesis, University of Arizona, Tucson, AZ, USA. Retrieved 28/02/2018, from: <http://arizona.openrepository.com/arizona/handle/10150/146905>.

Ipsos MediaCT (2009). *Social Grade: a classification tool* [online]. Retrieved 28/10/2020, from: < https://web.archive.org/web/20160315075958/https://www.ipsos-mori.com/DownloadPublication/1285_MediaCT_thoughtpiece_Social_Grade_July09_V3_WEB.pdf>.

Jongman, A., Wayland, R., & Wong, S. (2000). Acoustic characteristics of English fricatives. *Journal of the Acoustical Society of America*, 108(3), 1252–1263.

Kiesling, S.F. (1998). Men's identities and sociolinguistic variation: the case of fraternity men. *Journal of Sociolinguistics*, 2(1), 69–99.

Kiesling, S. F. (2011). *Linguistic Variation and Change*. Edinburgh: Edinburgh University Press.

Krivoshein de Canese, N., & Corvalán, G. (1987). *El español del Paraguay*. Asunción: Centro Paraguayo de Estudios Sociológicos.

Labov, W. (1966). The social stratification of (r) in New York City department stores. In N. Coupland & A. Jaworski (Eds.). *Sociolinguistics. Modern Linguistics Series* (pp. 168–177). London: Palgrave.

Labov, W. (1972). *Sociolinguistic Patterns*. Philadelphia: University of Pennsylvania Press.

Labov, W. (1990). The intersection of sex and social class in the course of linguistic change. *Language Variation & Change*, 2, 205–254.

Labov, W. (2001). *Principles of Linguistic Change, Vol. 2: Social Factors*. Hoboken, NJ: Wiley-Blackwell.

Labov, W. (2006). *The Social Stratification of English in New York City* (2nd ed.). Cambridge: Cambridge University Press.

Lawson, R. (2014). Fight narratives, covert prestige, and performances of 'tough' masculinity: some insights from an urban center. In T. M. Milani, (Ed.). *Language and Masculinities: Performances, Intersections, Dislocations* (pp. 53–76). Oxford: Routledge.

Lenz, R., Bello, A. & Oroz, R. (1940). *El español en Chile*. Buenos Aires: Instituto de Filología.
Lipski, J.M. & Iglesias Recuero, S. (1996). *El español de América*. Madrid: Catédra.
López Morales, H. (2004). *Sociolinguística* (3rd Ed.). Madrid: Gredos.
Maniwa, K., Jongman, A., & Wade, T. (2009). Acoustic characteristics of clearly spoken English fricatives. *The Journal of the Acoustical Society of America*, 125(6), 3962-3973.
Milroy, L. & Gordon, M. (2003). *Sociolinguistics: Method and Interpretation*. Oxford: Blackwell.
Moreno Fernández, F. (2009). *Principios de sociolingüística y sociología del lenguaje* (4th ed.). Barcelona: Editorial Ariel.
O Brien, J. (2012). *An Experimental Approach to Debuccalization and Supplementary Gestures*, Ph.D. thesis, University of California, Santa Cruz. Retrieved 02/03/2018, from: <http://jeremypobrien.nfshost.com/papers/obrien_qe.pdf>.
Pérez, C. & Salamanca, G. (2016). El mapuche hablado en Curarrehue: fonemas segmentales, fonotaxis y comparación con otras variedades. *Literatura y Lingüística*, 35, 315–336.
Sadowsky, S. (2015). Variación sociofonética de las consonantes del castellano chileno. *Sociolinguistic Studies*, 9(1), 71–92.
Sadowsky, S. & Salamanca, G. (2011). El inventario fonético del español de Chile: principios orientadores, inventario provisorio de consonantes y sistema de representación (AFI-CL). *Onomázein*, 24(2), 61–84.
Silva Corvalán, C. (1987 [2001]). *Sociolingüística y pragmática del español.* Washington: Georgetown University Press.
Soto Barba, J. (2007). Variación del F1 y del F2 en las vocales del español urbano y rural de la provincia de Ñuble. *RLA*, 45(2), 143–165.
Spolsky, B. (2003). *Sociolinguistics*, (4th ed). Oxford: Oxford University Press.
Strevens, P. (1960). Spectra of fricative noise in human speech. In D.B. Fry (Ed.). *Acoustic Phonetics: a course of basic readings* (pp. 132–150). Cambridge: Cambridge University Press.
Tagliamonte, S. (2011). *Variationist Sociolinguistics: Change, Observation, Interpretation*. Oxford: Wiley-Blackwell.
Tapia Ladino, M. & Valdivieso, H. (1997). La palatalización de las velares: análisis acústico. *Onomázein*, 2, 135–149.
Trudgill, P. (1972). Sex, covert prestige and linguistic change in the urban British English of Norwich. *Language in Society*, 1, 179–195.
Trudgill, P. (2000). *Sociolinguistics: an introduction to language and society* (4th Ed). London: Penguin Books Ltd.
Van Rossum, G., & Drake, F. L. (2009). *Python 3 Reference Manual*. Scotts Valley, CA: CreateSpace.
Waskom, M., Botvinnik, O., O'Kane, D., Hobson, P., Lukauskas, S., Gemperline, D. C., ... & de Ruiter, J. B., Chris Fonnesbeck, C., Lee, A. & Qalieh, A. (2017). *mwaskom/seaborn: v0.8.1*. Retrieved 06/08/2020, from: <https://zenodo.org/record/883859#.Xt4PiDpKhPZ>.
Weinreich, U., Labov, W., & Herzog, M. (1968). Empirical foundations for a theory of language change. In E.W. Lehman & Y. Malkiel (Eds.). *Directions for Historical Linguistics* (pp. 95–195). Austin: University of Texas.

8 Appendix: Reading Test Slides

En la playa se ven todo tipo de palmeras; en el mar, cardúmenes de peces multicolores y algunos delfines muy graciosos.

Received: 31/07/2020
Revisions: 29/10/2020
Accepted: 08/03/2021

Me vendieron un embudo sin boleta en Chimbarongo.
Después de registrar una visa, hay que ir al Registro Civil y esperar.
Ella era gorda, pecosa y de cabello de color amarillento muy crespo.
Mi hermana estudió ingeniería en la universidad, pero ahora es profesora de biología, lo que no afecta su fascinación con todo lo tecnológico.
Podrán vivir sin guerra en Angola.
Te mandé un mensaje diciendo que trajeras dinero en efectivo para pagar el viaje, pero trajiste tu tarjeta de crédito.
Los furgones de carga dan mayor rentabilidad, porque presentan mayor capacidad.
Hace unas horas, yo, el líder de la comunidad británica en Chile, elegí personalmente al encargado de traer ginebra a los carretes de la organización.
En Lota, las actividades comunitarias de invierno contribuyeron a paliar la pobreza de la gente.
"Ya te dije que vayas a la página 394, Sr. Potter, y quiero que vayas a esa página antes de que me enoje," dijo Snape.
Vi con angustia como un gato angora se comía una anguila.
Los abogados tuvieron que corregir los errores de tipeo en los documentos legislativos, sin hacer grandes cambios al contenido original.
Se dice que el tordo, por su brillante color negro, atrae la atención de los turcos.
La gente se mantiene en forma mediante ejercicio físico: por eso, es útil tener un gimnasio cerca de la casa.
Existen muchas mentiras que parecen verdaderas y agradan más cuando tienen un origen dudoso, pero posible.
La gira del grupo de metal fue un gran éxito; los artistas recibieron muchos halagos acerca de la calidad y la complejidad de su música, siendo las habilidades del bajista las más comentadas.
Ángel Fernández señaló que será la empresa privada la que hará los proyectos económicos.
Escogiste la mejor temporada para pasar tiempo en las playas francesas, porque en verano se cosechan los mejillones.
La esbelta joven era tan terca que nunca estudiaba para los certámenes.
La inteligencia de esta nueva generación de estudiantes se nota en sus capacidades de buscar soluciones logísticas que nos mejoran la vida.
Me invitaron a pasear en un bote sin combustible.
Me imagino que a la gente le enganchan las series porque acaban sintiendo cariño por los personajes.
Antofagasta lanza su apuesta para enfrentar a Calama.
Aguirre sonó su nariz y la balsa en la que viajaba empezó a hundirse rápidamente.
El ají chileno no pica tan fuerte como el mexicano, que se usa en las fajitas y toda esa comida.
Seguro que va a haber intercambio de bombas: molotovs para allá, lacrimógenas para acá.
Ya es hora de exigir que nos dejen acceder a la plataforma digital, porque tenemos una cantidad de lecturas pendientes.
Claro que me gusta comer caldo de congrio, con harto ajo y un pichintún de ají.
Un exceso de lluvia puede arruinar el día, aquí y en la quebrada del ají.
Más vale un largo etcétera que una lista latera.
Allí resonó por última vez el fusil del guerrillero abatido por la mujer del general Javier Grandón Jiménez.
La respiración es un proceso involuntario que permite la obtención de oxígeno y la eliminación de desechos en estado gaseoso.
Los glóbulos rojos, apoyados por la hemoglobina, expulsan las moléculas de dióxido de carbono y reciben las de oxígeno, que luego llevarán hasta cada una de las células del cuerpo.

En el diccionario de la Real Academia Española, se define tribu como un grupo social primitivo de un mismo origen, real o supuesto, cuyos miembros suelen tener en común usos y costumbres, y también como grupo grande de personas con alguna característica común.

Ya está dicho: nos invaden los extranjerismos. Ir de camping, divertirse con stickers, pertenecer a un club, ir a un casting, dedicarse al marketing, probarse las panties, tener un partner, comprar los tickets...

Se comenta que los habitantes de esta ciudad, así como los de todo el orbe, apetecen más el vino tinto que el blanco, debido a la presencia de sustancias que retardan el envejecimiento, favorecen la digestión e inhiben la formación de trombos, disminuyendo de este modo el riesgo de accidentes cardiovasculares.

Mientras cada uno se esforzaba por presentar argumentos más ingeniosos que el otro, escucharon pasar a un extraño viajero envuelto en unas ropas muy abrigadas, hechas de un fino tejido ovejuno.

Urgido como nunca antes en su vida, Carlos trató de apagar el notebook, pero el botón de encendido no hizo nada.

About the Author

Madeleine Rees is a PhD candidate at the Department of Theoretical & Applied Linguistics of the University of Cambridge. She graduated from the University of Southampton with a BA in Spanish & Linguistics in 2019 and from the University of Cambridge in 2020 with an MPhil in Linguistics. Her research interests include speech production, both from a cognitive and sociolinguistic standpoint, speech perception, and language variation and change.

Acknowledgements

I would like to thank all participants for their time and willingness to be part of the experiment, especially during a 30-degree summer in Santiago.

Received: 31/07/2020
Revisions: 29/10/2020
Accepted: 08/03/2021

The Graded Co-Salience Hypothesis for Polysemous Ambiguity

T. R. Williamson
University of Cambridge

Abstract. It is well established in the theoretical (see Weinreich 1964) and empirical study (see Jastrzembski, 1981, Williams, 1992) of polysemy that its comprehension in context can lead to ambiguities arising. The present paper aims, by reviewing current literature and employing corpus methods, to determine whether the pragmatic theory of the Graded Salience Hypothesis (Giora, 1997; 2003) may be used outside of the original scope of its application to determine precisely why polysemes in context can be ambiguous. Using data from the ARCHER 3.2 (2013) corpus to analyse frequency as an input factor to mental lexica structure (see Bybee, 2006; 2010), 2,761 token instances of nine polysemes — 'hand', 'head', 'door', 'once', 'book', 'run', 'cut', 'stop', and 'court' — are manually tagged in a semantic decision task (from Glynn, 2016) as being either of a particular sense, or ambiguous (based on the biasing/priming effects of their context). It is found that significant incidences of polysemous ambiguity can be explained as a function of the plurality of salient senses per lexical item, among other observable characteristics. The hypothesis this paper brings is that polysemous ambiguity might be explained, following further research, with reference to the Graded Salience Hypothesis; that it may be the result of the co-activation of co-salient senses: the Graded Co-Salience Hypothesis.

Plain English Abstract. Often, one word can have multiple meanings. The word 'door', for instance, might be said to have two prominent meanings: the object one opens and closes, and the empty space through which one walks. Because these two meanings are similar, 'door' can be known as an example of *polysemy*. In both theoretical and empirical approaches to polysemy to date, it has been established that, sometimes, comprehending a polysemous word, or polyseme, can lead to ambiguities arising. This paper aims to suggest that the Graded Salience Hypothesis (Giora, 1999; 2003), a proposal about how accessing words in our minds works, can also be used to explain why polysemes are ambiguous. Using methods from corpus linguistics, 2,761 instances of nine polysemes ('hand', 'head', 'door', 'once', 'book', 'run', 'cut', 'stop', and 'court') are analysed to determine how frequently their placement in a particular context either has a specific meaning or is ambiguous. This study's findings suggest that significant incidences of ambiguous polysemes can be explained by observing various characteristics about the frequency of their senses, such as the plurality of salient senses. It is posited that this data indicates, in ambiguous contexts, that a polyseme's most salient senses are activated simultaneously in a manner that might prevent disambiguation, with further psycholinguistic study required to verify this claim. In this way, it is suggested that an alternate method of studying polysemous ambiguity, through the lens of the Graded Salience Hypothesis, might do well to be considered. Thus, this paper suggests the Graded Co-Salience Hypothesis.

Keywords: corpus pragmatics; polysemy; ambiguity; salience; cognition

---

1 Introduction

1.1 Polysemy

In ordinary language use, the words with which we communicate can quite plainly be seen to possess, relate to, or represent multiple meanings. When one person tells another, 'I went to the bank on Wednesday', without sufficient context or supplementary information, it might be potentially ambiguous as to whether this person visited a build-up of earth by the side of a flowing body of water or a building that contains a financial institution whose role it is to protect, save, and invest its clients' assets. Should it have been the case that this person was recently experiencing financial difficulties, the

hearer might have been safe in the assumption that the latter sense of the word 'bank' had been intended; had the speaker been a potamologist, however, one might be excused for assuming the former.

This problem, represented in the ambiguity of the lexical item 'bank', has been of interest to linguists for well over a century. Although the notion of the multiplicity of word meanings was first broached in Bréal (1897), an introduction into the underlying theory is best given with reference to Weinreich (1964). Lexical ambiguity can be thought of as a blanket or family term for two interrelated linguistic phenomena between which Weinreich (1964) is able to distinguish: *polysemy* and *homonymy*. This division is chiefly rooted in the extent to which the alternate meanings, or senses, of lexical items are interrelated: if a word's senses seem unrelated, it is homonymous; if its senses appear semantically similar, it is polysemous. So, homonymy is exemplified with reference to the example involving 'bank' above, as financial institutions and riversides do not seem semantically related. For polysemy, consider the example 'I made a door', paying particular attention to the possible readings of 'door'. One can observe that this verb phrase might refer either to the making or opening of a passageway or to the construction of a physical object that functions as a moveable barrier between areas

[1]. These two readings are spelled out with examples (1) and (2).

(1) I made an entryway by demolishing the wall.
(2) I made a door by sawing a piece of wood and fixing it to an entrance with screws and hinges.

The analysis of Weinreich (1964) goes deeper into this issue of sense interrelatedness in the case of polysemy. It is proposed that this subtype of lexical ambiguity can be further classified by the extent of the interrelatedness of polysemes' senses: polysemes with related senses are *complimentary*, whereas those with disparate senses are *contrastive*. Nerlich and Clarke (2003) expand on this further, understanding this distinction as a measure for the gradedness of the extent of polyseme sense interrelatedness by placing significance on both synchronic and diachronic studies of polysemy. Further to this, Nerlich and Clarke (2003) interpret Weinreich (1964) with reference to the semantic structure of the mental lexicon, suggesting that what underlies polysemy in the mind are complex networks of interconnected meanings that form around prototypes (see Rosch, 1975) and family resemblances (see Wittgenstein, 1974). It is this psychosemantic interpretation of polysemy theory and the place of polyseme sense interrelatedness that will become crucial for the discussion of this paper's empirical findings.

1.2 Psycholinguistic Approaches to Polysemy

Indeed, moving away from theoretical approaches to polysemy, there have been many empirical studies in the last few decades that aim to examine the structure of polysemy in the mental lexicon. In some sense, the precedent for closer examination of polysemous structure was set in Jastrzembski (1981), where it was argued that polysemous words have a kind of special structuring based on findings that words with many meanings are accessed more quickly than words with fewer. This said, discussions are ongoing in psycholinguistics regarding the exact nature of this structure. When contrasted with homonymous words' mental representations, some say that polysemous senses share one core

[1] A keen reader may observe that the polysemy at issue can be analysed on the morphosyntactic level as well as on the semantic/conceptual level. With 'door' being 'object' or 'entrance', there is a distinction between two different readings that trade on alternate semantics whilst maintaining the same syntactic category. However, with the example 'milk' being either 'liquid *x*' as a noun or 'to extract liquid *x*' as a verb, there is both alternate semantics and alternate morphosyntax (with two different morphemes of two different syntactic categories).

Received:17/08/2020
Revisions: 27/03/2020
Accepted: 10/04/2021

representation whereas homonyms have multiple (see Frazier and Rayner, 1990; Rodd et al., 2002), yet others argue that the multiple senses of both polysemes and homonyms have similarly separate representations in the mind (Klein and Murphy, 2001; 2002).

One psycholinguistic study that is of particular interest to the present paper is Williams (1992). In supporting the suggestion that polysemous words are processed and represented differently to homonyms, one key finding is brought: sense activation for polysemous lexical items can be observed as a function of those senses' *salience* in the minds of speakers. In this study, Williams (1992) employed a lexical decision task with the polyseme 'firm' – analysed as having senses 'strict' (disciplinarily) and 'hard' (physically); of which the former was proposed to be more salient – and found that the more salient sense was activated regardless of whether it was embedded within a biasing context. In other words, Williams (1992) provides evidence for the Graded Salience Hypothesis (Giora, 1997; 2003), a model that proposes that salient senses in the semantic networks of lexical items are always processed first, and in order of their salience. This finding lays the stage for a pragmatic theory of polysemy that approaches the distinctions between polysemes' senses, and resultantly the ambiguity that can arise as a result of their co-existence, using the framework that Giora provides.

1.3 Corpus Approaches to Polysemy

While giving an overview for psycholinguistic studies into polysemous structure is useful for theoretical background, an outline of empirical work in corpus approaches to polysemy is useful to frame the methodology of the present paper. To this end, a number of important corpus investigations into polysemy are relevant (Kishner & Gibbs, 1996; Raukko, 2003, see also Berez & Gries, 2008; Navarro, 2000; Fillmore & Atkins, 2000; Gries, 2006, see also Glynn, 2014). An important point to mention in the consideration of these works is that, whilst their endeavours are of course useful in and of themselves, the extent of their respective generalisabilities can be scrutinised because all of their empirical efforts feature analyses of only one polyseme (although in great depth). Given the aims of the present paper to yield results generalisable to language more broadly, multiple polysemes are included in this study. Regardless, the precedent set by the aforementioned papers for the corpus study of polysemy is highly important.

One of the most prominent of these was Gries (2006), who, using the International Corpus of English and the Brown Corpus of American English, collected 851 instances of the verb 'to run', and analysed them based on its different senses. A total of 252 different senses were found, where 40% were manually tagged and 60% arose from collocates, and it was determined that the prototypical sense of 'to run' was 'fast pedestrian motion' given that it occurred in 203 of the instances (approximately 25%). This determination, along with the identification of other metaphorical and extended uses of 'to run' (such as in 'to flow', 'to manage', or 'to escape'), enables Gries (2006) to develop an image-schema representation of the senses of 'to run' and their interconnection.

Very useful insights also come in Glynn (2016), who collected 500 instances of the polyseme 'annoy' from the LiveJournal Corpus (Speelman & Glynn, 2005) and categorised them in terms of three usage-features, or ID-tags, that relate to Rudzka-Ostyn's (1988; 1989; 1995) work in Usage-Feature Analysis in the description of communicative verbs: cause, patient, and agent. Using a number of statistical measures (specifically a chi-square distance measure, a hierarchical cluster analysis, and a partition cluster analysis), a wide array of usage-based features of 'annoy', both semantic and morpho-syntactic, are found that attempt to paint a very full picture of the profile of usage for the word. The reason why Glynn (2016) is particularly useful for the present paper is in its methodology. In an effort to rectify a problem of unfalsifiability he identifies in cognitive linguistics, Glynn (2016) sets out an alternate method to operationalise polyseme senses: the manual annotation of multidimensional

clusterings of usage-based features. In other words, Glynn (2016) proposes that the best way to understand the complex semantic networks that underpin polysemy is by doing the hard work of taking a large body of corpus data and manually going through it to annotate each token instance based on their correspondence with particular meanings.

There are two particularly intriguing features of this methodology. The first is that it is argued one can establish the senses of a polyseme a posteriori rather than speculating and prescribing a priori to ensure that findings are reliable, and one can demonstrate adequate quantification in one's method to allow for inductive and repeatable analyses. The second is the implication that one can produce a useful account of the mental representations of polyseme networks on the basis of large-scale corpus analysis. Understanding where the mental lexicons of speakers of a particular language intersect in their shared knowledge of polysemes' senses in this way, as represented by usage, corpora can arguably be taken as productive tools for theorising about the psycholinguistic semantic structures we hold for our languages. For these reasons, Glynn (2016) acts as methodological precedent for the research undertaken in the present paper, and inspiration in large part is drawn from it.

1.4 The Graded Salience Hypothesis and Frequency

As mentioned above, there is some link to be found between empirical studies of polysemy and the Graded Salience Hypothesis of Giora (1997; 2003). To expand further, the Graded Salience Hypothesis asserts that, in the employment of the mental mechanisms responsible for lexical access, more salient senses of lexical items are accessed and activated faster than less salient senses (Giora, 2003). This activation is proposed to occur in a sequential order that represents the process of our faculties for lexical access activating, evaluating, and consequentially accepting or rejecting a particular sense on the basis of its relevance to the discourse environment in order of the salience of senses in question (Giora, 1997). In turn, the concept of salience is constituted by Giora fourfold: the conventionality, frequency, familiarity, and prototypicality (Giora, 2003) of senses corresponding to lexical items all contribute to each individual sense's salience.

It will be useful to dissect salience's fourfold definition to understand how a sense can become salient, from Giora (2003). First, conventionality is explained as an implicitly agreed relation within a population between a situation and a linguistic form that has some kind of regularity of use. This relation, it is suggested, arises just because of a general preference for uniformity rather than because of some explicit reason. No preference is made for the motivations of conventionalisation like, say, Lewis' (1969) theory that conventions arise out of interlocutors' desire to coordinate intentions between themselves so as to economise their communicative and interpretative efforts. Though, clearly, for the mental lexicon, it does not matter *how* a sense is conventionalised, just *whether*.

Second, frequency is cashed out in two ways. Giora (2003) suggests that it might bear relations to a kind of probability of occurrence in language use (that is, if a word is more likely to appear then it possesses greater frequency), but that it also might relate to the frequency of cooccurrence of particular items (that is, 'piece' might have a more strengthened node in a lexical network because it frequently cooccurs with 'paper', even though 'piece' may not have a high frequency without). The difference between frequency and familiarity, as the third component of salience, is slightly tricky. Giora (2003) explains familiarity as a product of experience; more precisely, of encounters with a particular intended use as a result of its existence within a certain domain. So, because linguists are more familiar with 'tree' in the formal syntactic, complex node-branch-network structure sense than non-linguists might be familiar with 'tree' perhaps in the similar 'family tree' sense, 'tree' qua complex structure will have higher salience within the mental lexica of linguists. In this way, it is distinct from frequency: 'tree' qua

Received:17/08/2020
Revisions: 27/03/2020
Accepted: 10/04/2021

syntax is globally low-frequency (insofar as not everyone in the world does syntax) but linguistically high-familiarity, for instance.

Fourth, an explanation of prototypicality relies on an understanding of prototype theory from psychology (e.g., Rosch, 1975). Simply, Giora (2003) suggests that the more prototypical an item within a particular category, given its proven lexical access priority, the more salient it is. Of course, the salience of a prototypically-high entry into the mental lexicon is contingent upon the context under which its prototypicality would be salient; that is, 'robin' is salient in the context of birds, but might not be in the context of comic superheroes ('Superman', 'Batman', etc. might be prototypical, whereas 'Robin', as Batman's sidekick, might be peripheral).

To analyse the role of frequency in the creation of inner linguistic constructions in more depth (as it is the main component of salience analysable with corpora), reference can be made to the Usage-Based Grammar advocated by Bybee (2006; 2010) that descends from Construction Grammar (Fillmore et al., 1988). In Bybee's (2006) account, linguistic (particularly grammatical) representations are formed through the process defined by exemplar theory. In this, token experiences of linguistic forms (produced and perceived) are compared and contrasted with nodes on massive network of representations. If a certain experienced form is identical to a pre-existing exemplar, then that form is mapped onto that exemplar and the overall representation is strengthened in the mental lexicon. If, however, there are only strong similarities, that experienced form is stored as a separate but similar exemplar that then ends up making larger clusters or categories (Bybee, 2006). In this way, the more one experiences a particular linguistic form, the stronger our mental representations for those forms become, the more salient they might be during the processes of lexical access.

Further to this, Usage-Based Grammar argues that the effects of frequency on the structure of linguistic representations can be seen in a number of observable phenomena found in language use. One example of this is phonological reduction, wherein the automisation of neuromotor processes that arises from repeated use of the same, most frequent, phonological representations ends up reducing the magnitude of articulatory gestures and increases how much those gestures overlap (Bybee & Hopper, 2001). With this comparison, one can complement the claims of the Graded Salience Hypothesis (Giora, 1997; 2003) regarding the influence of frequency on lexical access with the theoretical and experimental efforts of functionalist-based approaches to understanding processes underpinning the grammars of speakers. The corpus methods employed in the present study aim to exploit this clear importance of frequency for an understanding of linguistic structure by utilising data pertaining to frequency of usage available from corpora.

Prima facie, the Graded Salience Hypothesis seems related just to the psycholinguistic processes underpinning lexical access, or, more generally, to language comprehension. Indeed, in the work of Giora and Fein (1999) on irony comprehension, one clear application of the theory arises: they use a fragment completion test to prime a particular literal or ironic interpretation of a given phrase and find that the comprehension of ironic utterances involves the activation of concepts related both to the ironic sense and the literal sense, whilst comprehension of equivalent literal utterances only involves activating concepts relating to the literal sense. The theory is applied further in Giora et al. (2012), where they test the common presupposition that individuals with Asperger's Syndrome are less sensitive to contextual cues and more to literal senses in social communication (in which no bias towards literal senses is found).

However, this paper argues that the Graded Salience Hypothesis has uses beyond the original designation of its scope. Whilst at the outset, and in applications since its inception, the theory has been used to examine the mechanisms underpinning lexical access with respect to words whose meanings shift depending on the intention behind their usage and their corresponding contexts (see figurative language, Giora, 1997; irony, Giora & Fein, 1999; idioms, jokes, Giora, 2003; sarcasm, Giora et al.,

2014), there also seems to be scope for application of the theory with respect to words with multiple meanings tout court; or, more exactly, with respect to polysemy. The Graded Salience Hypothesis so far has been used to understand how we come to understand the meaning of a particular utterance in light of the potential ambiguity caused by external factors, but not yet has the theory been applied in the context of words whose ambiguities arise just by the interrelatedness of their senses.

1.5 Aims

In this study, whether the Graded Salience Hypothesis has any scope to be applied to the investigation of the structure and processing of polysemy networks in the mental lexicon will be examined. Understanding frequency effects as central in the construction of linguistic representation as per the definition of salience in Giora (1997; 2003) and the role of frequency in Usage-Based Grammar (Bybee, 2006; 2010), 2,761 tokens of nine frequently-occurring polysemes in the ARCHER 3.2 corpus will be manually tagged as either ambiguous or relating to a particular sense from their respective contexts. The statistical significance of the incidence of ambiguous sense token polysemes will be ascertained to determine the extent of each polyseme's individual ambiguity, and data from frequency will be analysed to work out what caused this significance. To this end, the following questions will be asked. Can the notion that lexical access is mediated by the salience of senses be applied to the study of polysemy? Could senses' salience in their respective polysemy networks give any indication as to why ambiguity arises? And, if so, what might this tell us about how polysemous words are stored psycholinguistically?

2 Methodology

2.1 Data Selection

All data used in this paper was collected from the ARCHER (A Representative Corpus of Historical English Registers) 3.2 corpus (2013), which is a multi-genre, diachronic corpus of British and American English covering the period 1600-1999. It consists of 3,298,080 words (1,957,499 British English and 1,340,581 American English). ARCHER 3.2 was selected for its size and variety in dialect and kinds of contexts (see Section 2.2) featured — to paint the broadest picture possible of polysemy across the English language, such types of variety are essential. To ensure that historical semantic change did not affect the findings here, only a short period of time could have been selected. This was balanced with the desire to ensure the possibility for the collection of a sufficient number of materials for analysis, so it was decided that a century's worth of corpus data would be sufficient. It was also noted that the senses of lexical items employed from previous centuries would have been more inaccessible for unambiguous comprehension, so the century 1900–1999 was chosen. This period contains a total of 1,294,244 lexical items in the corpus.

Nine polysemous words were selected for analysis in this study. In past studies on polysemy and corpora (see above) the only rationale for the selection of the polysemes analysed has been that there were distinct senses to scrutinise. As a criticism of those, and to make this methodology more rigorous, the polysemes in this study have been selected against a range of others, and with a specific criterion in mind. First, a list of 26 polysemes was made, chosen from studies with similar methodologies to this one (see Table 1) and by the researcher. For data collected to be as representative of polysemy more broadly, the selected polysemes were required to occur frequently in the corpus. Thus, the frequency per million scores of all 26 polysemes was analysed for their selection in this study, as shown in Table 1.

Received:17/08/2020
Revisions: 27/03/2020
Accepted: 10/04/2021

Table 1: *Frequency per million scores in ARCHER 3.2 of 26 polysemes*

Polysemes	Freq./million
In (Navarro, 2000)	15,913.54
Over (Tyler & Evans, 2001)	1,199.57
Just (Kishner & Gibbs, 1996)	968.13
Man	961.18
Get (Raukko, 2003)	941.09
Make (Kishner & Gibbs, 1996)	679.16
Hand	387.10
Head	345.38
Door	342.28
Once	316.01
Book	220.98
Run	175.39
Cut	139.85
Stop	132.90
Court	108.96
Foot	87.13
Character	84.99
Drink	79.58
Shop	72.63
Milk	60.27
Game (Wittgenstein, 1953)	58.72
Funny	57.95
Path	39.41
Paint	17.77
Crawl (Fillmore & Atkins, 2000)	4.64
Annoy (Glynn, 2016)	3.86

Polysemes already analysed or discussed by other researchers were not included in the present study to prioritise novelty. That said, the study by Gries (2006) focused on *to run* specifically, and as such the polyseme 'run' has been included here as a separate lexical item. Additionally, the polyseme 'man' was not included because it occurred much more frequently in the corpus than any other unresearched polyseme such that it might have skewed the dataset: homogeneity was favoured to preserve reliability. Presented with 16 polysemes remaining, it was decided that the criterion for polyseme selection was a frequency of 100/million or higher in the ARCHER 3.2 corpus, keeping in mind the aforementioned

desire to study a range of frequently used polysemes – thus, nine were chosen: 'hand', 'head', 'door', 'once', 'book', 'run', 'cut', 'stop', and 'court'. Table 2 shows these selections with their frequency per million scores.

Table 2: *Frequency per million scores in ARCHER 3.2 of the nine polysemes selected for this study*

Polysemes	Freq./million
Hand	387.10
Head	345.38
Door	342.28
Once	316.01
Book	220.98
Run	175.39
Cut	139.85
Stop	132.90
Court	108.96

2.2 Data Collection

The ARCHER 3.2 corpus was accessed through the website CQPweb (Hardie, 2012). Data for each polyseme was downloaded as a .txt (Text Document) file, and then copied into an .xlsl (Microsoft Excel Document) file for analysis. By copying over downloaded data to an .xlsl file, it was much easier to manually annotate the corpus data. Search parameters remained the same throughout data collection: all instances of each polyseme from both 1900–1949 and 1950–1999 were included. However, only root forms of polysemes were included in searches: words such as 'runs', 'doorway', or 'stopper' were considered to be separate lexical items, and even though polysemies may be shared between inflected or morphologically adapted forms, they were not included to narrow the scope of this study.

To ensure the reliability of sense disambiguation, the corpus data was analysed with the broadest possible context: polysemes were downloaded with the maximum context of 20 words before and after. To ensure that data here might be generalisable to wider contexts of language use, it was important that the sample represented ordinary discourse as much as possible. To this end, data from all genres available in the corpus was collected for selection and then analysis; these were: *Advertising*, *Drama*, *Fiction*, *Sermons*, *Journal*, *Legal*, *Medicine*, *News and periodicals*, *Early prose*, *Science*, *Letters*, and *Diary*[2]. Potential biases were avoided by opting not to download metadata pertaining to the following categories available in the ARCHER 3.2 corpus: author sex, author, bibliographic info, publication date, genre, and regional variety.

Where polysemes occurred as a person's name (e.g., 'Mrs Hand'), they were excluded as unindicative of polyseme sense. Place names (for example, 'Jefferson Market Court'; 'court', 1906nyt1_n7a) were included because they were considered to have inherited their name through a relation to the particular sense – Jefferson Market Court would not have been named thus without having been an 'enclosed space'. Additionally, some lines in the corpus were reduplicated – these were

[2] Except *Legal* for the polyseme 'court' given that the term, in reports of legal proceedings (which do not really represent organic discourse anyway) almost always refers to a court of law, and as such would have diminished the significance of extra-contextual ambiguous sense tokens 'court' appears 412 times in *Legal* alone. The genre *Medicine* was included in spite of potential overrepresentation of 'head', 'hand', and 'cut' as the as they only appear 7, 11, and 4 times in the corpus within *Medicine* respectively.

Received:17/08/2020
Revisions: 27/03/2020
Accepted: 10/04/2021

eliminated from analysis, and inconsistencies in the numberings of lines in the data can be explained accordingly.

2.3 Data Analysis

In total, 2,761 token senses of the nine polysemes were manually tagged as either relating to a sense category, determined post-hoc, or as being ambiguous. All data analysis was done on Microsoft Excel, including the manual tagging of polyseme sense category and statistical calculations of frequency and mean. Calculations of Pearson's chi-squared test were coded in the software RStudio (2020). The sense category to which each token polyseme corresponded was determined with the help of the Oxford English Dictionary (OED), as well as by the researcher language-internally; that is, by establishing a synonymic, categorical, or metonymic connection between that sense and a related term. For instance, the sense category of '… the inguinal glands when *cut* into were matted together' ('cut', 1905robe_m7b) was analysed as synonymous with 'incise' as a verb; and the sense of '… of Charles the First's *court* …' ('court', 1952whit_f8b) was analysed metonymically as being related to the category 'royal', and thus tagged with respect to this adjective. Some entries in the corpus were present as part of an idiomatic expression (for example, 'hand in hand' in 'hand', 1933hodg_h7b) – these were analysed as having distinct senses, rather than marking them as simply 'idiomatic', to avoid confusions within the sense category classifications.

Each sense category was manually tagged in the .xlsl document in a table adjacent to the corpus data. As more lines of each polyseme were analysed, more unambiguous senses were found and tagged: the mean number of unambiguous, disparate sense categories was 6.78 per polyseme. Although it might have made sense only to tag each polyseme token as 'ambiguous' or 'unambiguous' given the concern of the present study, the decision to distinguish between each sense of any token polyseme, and then whether it was ambiguous, was taken to clarify for the reader the rationale behind the tagging of each polyseme token specifically.

3 Results

In total, 2,761 polyseme tokens were manually tagged as being ambiguous, or unambiguous and pertaining to a certain sense category. The mean frequency of polyseme tokens per polyseme analysed in this study was 306.8. For all nine polysemes, the frequency of ambiguous sense tokens was 435, with a mean of 48.33 per polyseme. This means that 15.76% of all polyseme tokens were ambiguous in sense, calculated as the percentage of frequency of ambiguous sense tokens over the total frequency of polyseme tokens. The mean instances per million of each polyseme selected was 239.99 and the mean number of sense categories per polyseme was 6.78. The frequencies of all polyseme tokens and ambiguous sense tokens can be found in Table 3.

Table 3: *Frequency of polyseme tokens and ambiguous sense tokens (1900-1999)*

Polysemes	Freq. (total)	Freq. (ambiguous)	Freq. (ambiguous%)
Court	120	9	7.50%
Stop	172	23	13.37%
Hand	496	27	5.44%
Door	442	173	39.14%
Head	442	17	3.85%
Book	286	64	22.38%
Once	400	63	15.75%
Cut	181	27	14.92%
Run	222	32	14.41%
Mean:	306.78	48.33	15.76%

As can be seen, although 'hand' occurred most frequently in the ARCHER 3.2 corpus, the polyseme 'door' was found to have the highest frequency of ambiguous sense tokens at 173. This number was higher than any other polyseme, with the next highest being 'book' with 64 and then 'once' with 63. The frequency of ambiguous sense tokens for 'door' is noteworthy given that they account for 39.14% of all instances of that polyseme in the corpus data. Also interesting is that, despite 'court' having the lowest frequency of ambiguous sense tokens at 9, 'head' has the lowest frequency of ambiguous sense tokens as a percentage of total polyseme tokens at 3.85%.

To assess the significance of these findings, Pearson's chi-squared test was used. This was done because Pearson's chi-squared is a statistical test that assesses the probability for the observed data to have deviated from an expected distribution of data, so it can reliably indicate whether the frequency of incidence of a particular polyseme alongside the others is significant. The frequency of ambiguous sense tokens and unambiguous sense tokens were analysed for each of the nine polysemes, and the results of this test can be found in the association plot depicting Pearson residuals. In an association plot, each column, or box, represents a statistic that might either be greater or smaller than expected where the area of the box is a function of expectedness (Gries, 2009). In Figure 1, 'amb' denotes ambiguous sense tokens, and 'non_amb' denotes unambiguous sense tokens.

Received:17/08/2020
Revisions: 27/03/2020
Accepted: 10/04/2021

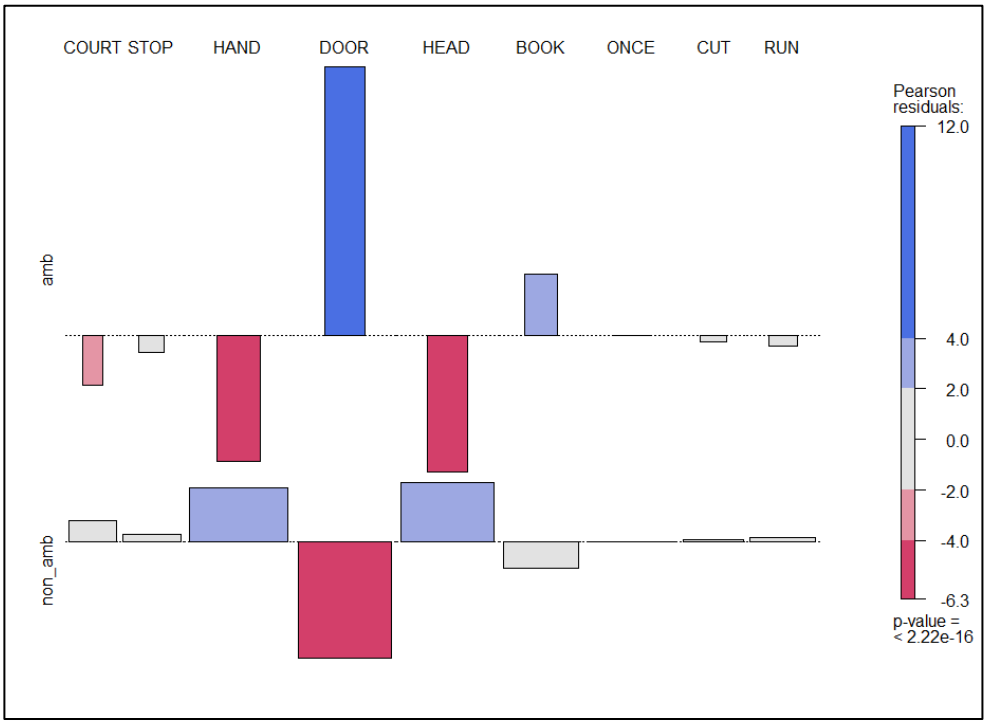


Figure 1: *Association plot of Pearson residuals for the distribution of ambiguous and unambiguous sense tokens.*

From this plot of the chi-squared test, several important deductions can be made. Firstly, the frequency of the ambiguous and unambiguous sense tokens observed had high statistical significance ($p < 2.22 \times 10^{-16}$) in comparison with the test's predictions. Though somewhat self-evident, this finding gives good reason to believe that polysemy can indeed give rise to ambiguity in ordinary discourse.

Moreover, it was also found that the individual frequencies of certain polysemes, either for ambiguous or unambiguous sense tokens, were significant — the colouration of particular bars of Pearson residuals in Figure 1 reflects this prominence. Specifically, it was found that the polysemes 'door' and 'book' had significantly higher Pearson residuals for the presence ambiguous sense tokens, whereas 'court', 'hand', and 'head' had significantly lower residuals for their absence. Similarly, polysemes 'hand' and 'head' were most significant for the presence of unambiguous sense tokens, as was 'door' for their absence. These polyseme-specific findings are illustrative of an important point regarding the nature of polysemous ambiguity. While it seems true that polysemy does give rise to ambiguity, there can be no uniform statement that might that assert all polysemes individually will necessarily turn out ambiguous in some cases of their use. Lexical ambiguity should be understood on a word-to-word basis, taking the semantic networks of polysemes into account: some words are naturally more ambiguous than others. Why this is the case is discussed below.

4 Discussion

To frame these findings within the scope of the Graded Salience Hypothesis (Giora, 1997; 2003), there is a certain burden of proof to be met by the present paper: evidence ought to be found that the significance of the frequency of ambiguous sense tokens could be correlated with the salience of senses within the mental lexicon. In order to do so, a key assumption has to be made: that each instance of a

particular polyseme, analysed as the use of a particular sense (or as being ambiguous), should be taken to represent a token instance of the activation of that sense (and any others depending on the level of salience of senses within that polysemy network) according to Giora's account. As such, the presence of a particular sense in the corpus data implicates that the semantic network of the corresponding polyseme was activated and that sense was selected as appropriate. However, what exactly the presence of ambiguous sense tokens represent for the study of polysemy is less clear.

With this assumption about what frequency represents in mind, the first point for this discussion is an analysis of the polysemes that showed as statistically significant for the presence of ambiguous sense tokens. For data collected here, two polysemes were found to be significantly ambiguous: 'door' and 'book'. To understand why this was the case, closer inspection of the findings from the manually annotated corpus data is required.

Table 4: *Frequency of senses for 'door'*[3]

Sense	Frequency
Object (n)	186
Entrance (n)	62
Place (n)	21
Ambiguous	173

Table 5: *Frequency of senses for 'book'*

Sense	Frequency
Object (n)	105
Content (n)	112
To reserve (v)	5
Ambiguous	64

Four major observations can be made with respect to this data in comparison with the rest of the findings from this study. The first is that the number of senses recorded for both polysemes was quite well below the mean for all nine analysed: each of 'door' and 'book' both were found to pertain to three discrete senses in the corpus data, along with their being ambiguous, whilst the mean number of senses per polyseme was 6.78. The second observation is that there seems to be one discrete sense that is much less relevant in how frequently it was found: the third sense of 'door' represents 0.05% of instances and the third of 'book' represents 0.02%. The irrelevance of these senses seems almost in direct contrast with the third observation to these findings, which is that, for both polysemes, there appear to be two unambiguous senses competing for activation (or, occurring most frequently); albeit this is more so for 'book', whose first and second senses represent 75.87% of all occurrences, than 'door', whose first and second senses represent 56.11%.

Following on from this third observation, the fourth that can be made is that these noticeably high-frequency senses also bear quite strong semantic relations with one another insofar as one might say the similar locations in physical space to which the senses in question refer are quite coordinated. In other words, one can see that the two co-frequent senses for 'door' are quite conceptually similar

[3] The letters in brackets for Tables 4 to 7 represent a designation of syntactic category; (n) designates a noun, (v) designates a verb, (adv) designates an adverb, and (adj) designates an adjective.

Received:17/08/2020
Revisions: 27/03/2020
Accepted: 10/04/2021

given that one might often expect to find the 'object' door in quite the same place as an 'entrance' or doorway. Likewise, it would not be amiss to propose that the 'content' of a book might be found somewhere locatively quite similar to the 'object' of a book. This notion of conceptual similarity is made clearer with reference to the corpus data itself. Below are four instances from the data that represent each of the four most relevant senses being analysed here for 'door' and 'book':

(3) '…Lee?" The old lady chuckled. Her eyes lit up as Marie bent down and opened the oven ***door***. A delicious hot fragrance blew out into the tidy kitchen. "My, somet'ing smell good…' ('door', 'Object' sense, 1913cath_f7a, formatting added)

(4) '…heads. Alec stopped to watch. What fascinated him was how they negotiated the wide trays through the narrow ***door***, just by turning their shoulders. Before he had gone to Corker's he had been offered a…' ('door', 'Entrance' sense, 1964berg_f8b, formatting added)

(5) '…Can barely see to write this. But did enjoy photos. Edith Sorel of Le Monde sent me a ***book*** -- letter from Castellorizo -- very tiny island. You may know it. It's not even on…' ('book', 'Object' sense, 1978mill_x8a, formatting added)

(6) '…When he came to my chapter he laughed. I told him that it was a very wicked ***book***. Growing rice looks like short oats at this time of season. It is hung up in tiny stacks…' ('book', 'Content' sense, 1921nort_j7b, formatting added)

In this way, this paper suggests that four corpus-identifiable characteristics of significantly ambiguous polysemes can be observed:

(1) A low number of discrete senses;
(2) At least one very low-frequency sense;
(3) At least two highly-frequent senses;
(4) And, a high degree of conceptual/semantic similarity between the senses in (3).

It should be noted that this list is not meant to be prescriptive, normative, or exhaustive. At most, these four characteristics attempt at being representative of ambiguous polysemes, but only on the level of contingent (i.e., not necessary) properties. The observations that these four allow might perhaps permit a level of polyseme-spotting for future research.

Yet, why do these characteristics help us interpret the data? For the two significantly ambiguous polysemes found in this study, one can first notice (1); that the number of (relevant) sense options for the interpretation of each polyseme in context is very limited, and thus (3); that each polyseme only has two senses that seem most reflective of the majority of its occurrences. To expand on this further, a useful question can be asked: why should this impact the ambiguity of a given polyseme? In each instance of use, the process for comprehending a polyseme can be understood in terms of a sequence of semantic decisions for what that term might have meant in that context (or, might have been intended to mean; see Giora, 1997; 2003). The hearer or reader receives as input the polyseme, and, taking context into account, checks to see which of the senses they know to which it might correspond. If there are multiple senses of similar relevance in a particular context, interpreting a particular instance of a polyseme becomes more challenging from a processing perspective.

Furthermore, noticing (4), the conceptual similarity between the two polysemes' respective senses, helps to understand the extent to which context is useful for semantic interpretation. Were the competing senses for each polyseme quite conceptually disparate, one would expect that contextual

factors would help to eliminate ambiguity. This absence of the helpfulness of contextual factors (in this case, the rest of the sentence) is made apparent in the following two examples of ambiguous sense tokens for 'door' and 'book':

(7) '…Griselda was sitting on the rug taking off her stockings. They looked up when Ty Ty stopped at the ***door***." What do you want, Pa?" Buck asked irritatedly." Son," he said…' ('door', 1933cald_f7a, formatting added)

(8) '…many complications as might have been expected. But I foresee the need for some very intelligent editing before the ***book*** sees the light. Of course I shall go on working for penal reform and individual prisoners. I much…' ('book', 1981long_y8b, formatting added)

With the assumption above, that every instance of a polyseme in the corpus represents the activation of a sense in the mental lexicon, it would necessarily follow that every instance of a polyseme token found to be ambiguous also represents the occurrence of some kind of activation. From the data presented here, given the co-relevance of two particular senses for the interpretation of each of the significantly ambiguous polysemes, and those senses' conceptual similarity for which context seems unhelpful in distinguishing between, this paper proposes that the ambiguity that arises from polysemy can be explained in terms of the co-activation of two co-salient senses. In this process, a hearer or reader of a word that bears polysemous ambiguity attempts to interpret that word's meaning in line with the proposals of Giora (1997; 2003), wherein the access of sense nodes in the semantic network for that lexical item is ordered depending on the salience of those senses.

Understanding findings pertaining to frequency from this study as input factors into the semantic structure of polysemous networks, this paper therefore proposes that a new way of thinking about polysemous ambiguity might be explored. Simply, it is that the co-relevant senses observed for significantly ambiguous polysemes might represent *co-salient* senses within the mental lexicon. Moreover, under the assumption that each token instance from the corpus represents some form of activation in our minds, it is also proposed that these co-salient senses may be activated simultaneously within the process of interpretation in such a way that derives no result helpful for the disambiguation of that polyseme in context in cases where it becomes ambiguous. In other words, the hypothesis this paper brings for an understanding of polysemous ambiguity is that it might be the result of co-salient senses being co-activated such that the polyseme in question is rendered ambiguous. Further psycholinguistic study will be necessary to substantiate this hypothesis.

Following this assertion, an additional line of inquiry can be pursued by the present discussion to determine its credibility more certainly. Should it be the case that polysemous ambiguity can be explained as a function of the co-salience of senses being co-activated, one would expect to find in the data presented here that polysemes found to be significantly unambiguous ('hand' and 'head') would show converse observations to those noted in the given four characteristics for the polysemes 'door' and 'book' above[4]. In this way, one would expect 'hand' and 'head' to present higher numbers of discrete senses than the mean (the opposite of (1) above, henceforth ¬1), without one noticeably infrequent sense (¬2), with no two co-salient senses (¬3), and without any noticeable conceptual similarity between senses (¬4). To investigate this, the findings from the manually annotated corpus data are required:

[4] Insofar as those observations represent a cause for those lexical items' polysemous ambiguity.

Received:17/08/2020
Revisions: 27/03/2020
Accepted: 10/04/2021

Table 6: *Frequency of senses for 'hand'*

Sense Categories	Frequency
Limb (n)	315
Location (n)	12
Assistant (n)	2
Control (n)	8
Owner (n)	1
To pass (v)	8
To participate (v)	2
To help (v)	3
Contrastingly (adv)	66
Together (adv)	12
Available (adj)	31
Relevant (adj)	9
Ambiguous	27

Table 7: *Frequency of senses for 'head'*

Sense Categories	Frequency
Body part (n)	317
Top-most part (n)	28
Heading (n)	3
Leader (n)	65
Temperament (n)	1
To lead (v)	3
To aim (v)	1
To fight (v)	7
Ambiguous	17

As can readily be seen, neither of these significantly unambiguous polysemes correlate with the observations presented for the significantly ambiguous ones as listed above. First, the polyseme 'hand' was observed to have 12 discrete senses, and 'head' 8, both higher than the mean number of senses per polyseme of 6.78 (¬1). Second, many senses for both were noticeably infrequent; of the 12 senses of 'hand' had under 10 instances and 5 of the 8 senses for 'head' were as well (¬2). Third, no two senses for either appear co-salient (insofar as there might be two that represent a high, relevant percentage of total instances) for either polyseme as well; in fact, they both seem to have one most salient sense, as 'limb' for 'hand' represents 63.51% of all its instances and 'body part' for 'head' represents 71.72% (¬3). Finally, little apparent conceptual similarity between any of the two polysemes' senses is observable, in part due to the diversity of syntactic categories designated to discrete senses found (¬4). The only noticeable semantic parallels exist metaphorically between the noun senses for 'head', as the senses of 'body part', 'top-most part', 'heading', and 'leader' all seem to refer to an entity at the top of a vertically-construed structure or hierarchy.

5 Conclusion

5.1 Reflections, Issues, and Further Study

One clear route for further empirical work is to employ behavioural, rather than corpus, psycholinguistic methods in assessing how we deal with polysemous ambiguity under experimental conditions[5]. The Graded Co-Salience Hypothesis might predict that, for example, disambiguating a polyseme with co-salient senses in an ambiguous context might incur greater processing costs through increased time taken to identify the sense of the polyseme. Conversely, polysemes without co-salient senses might be predicted to result in lower processing costs insofar as no two or more equally salient senses would be competing for activation from a given context.

Another issue for reflection is the extent of the import of these findings for a clearer picture of ordinary language use. Some key questions one might ask are the following: does polysemous ambiguity ever actually matter on the outcomes of ordinary discourse? Could the subtle semantic distinctions proposed here to exist within an instance of ambiguity have an effect on a conversation? Does it matter all the time whether we know what a person is referring to? Answering these questions to some extent will involve future study into this area; given that findings here point to the incidence of polysemous ambiguity within the polysemes studied was statistically significant, there is some scope for its analysis as impactful upon the outcomes of discourse.

However, some recourse can be taken in answering these questions to the concept of a 'context set' in Stalnaker (1979). In this, it is proposed that conversational participants will have a set of beliefs in any context concerning the precise topic of discussion at different stages of that conversation. The further a conversation progresses, the closer the mutual understandings of participants will end up. With this in mind, it might be proposed that further avenues for the study of polysemous ambiguity could involve the extent to which its incidence affects the construction of interlocutors' context sets in real time.

A final issue worth exploring with regard to the methodology of this paper concerns what precisely underpins the ability to disambiguate the senses of a polyseme and interpret it in accordance with its intended meaning. How do we end up knowing what a specific use of a specific word means? In this study, only the linguistic context in which polysemes were found to be situated from the corpus data was relevant for determining meaning, but in ordinary discourse there are a plethora of other factors involved. Examples of these might range from gesture (see Schlenker, 2019) in a physical fashion, but also the kinds of sources of information that compile a merger representation in Jaszczolt's Default Semantics (2005; 2010) such as world knowledge or presumptions about society and culture. As such, it might be the case that findings in the present research would be enriched with the contributions of these factors unlimited.

5.2 Limitations of the Methodology and Data

This paper's methodology could first be criticised for the manual tagging of polysemes' senses. Although the potentially constrictive effects of using the OED to define the senses of each selected polyseme precisely were slightly offset by the researcher's input, nevertheless using a dictionary to determine the meaning of any word will be slightly problematic in any experimental setting given its

[5] This is not to say, however, that corpus studies like this one investigating frequency effects as if cognitively representative should be discarded. For an interesting back-and-forth discussion on this matter, see Arppe et al. (2010).

Received:17/08/2020
Revisions: 27/03/2020
Accepted: 10/04/2021

prescriptive nature. Conversely, criticism could also be aimed at the researcher's own input into determining to which particular sense category each token polyseme referred.

A third criticism might argue that the methodology for data analysis conflates ambiguity and vagueness. This critique suggests that there is an issue in how polyseme tokens were marked as ambiguous when the sense was unknown (or where there was insufficient context within the corpus data to tag as unambiguous) *and* when the term seemed to pertain to more than one of the discrete senses observed. Although the observation in this critique is true, it has been taken as unproblematic for the purposes of this study given how it must surely be the case that each vague token polyseme would have been used by its author with a specific sense in mind, and as such will still serve to represent some node on the sense network of that polyseme. Therefore, this critique can be met with the suggestion that vagueness does not arise enough as the result of speaker intentions for its possible existence in the corpus data to be potentially worrying in the attempt to understand the semantic networks of polysemes.

5.3 Concluding Remarks

In this paper, it has been argued that the occurrence of polysemous ambiguity can be understood as a function of the co-activation of co-salient senses within the polysemy network of a certain polyseme. It is proposed here that, in line with evidence for the Graded Salience Hypothesis (Giora, 1997; 2003), lexical access for polysemy should not be taken as particularly special in comparison with processing for any other kind of word; that the mental mechanisms underpinning such operations function in an ordered fashion during which the most salient senses for each lexical item are prioritised. However, the case of ambiguous polysemes differs in the results of this process. The corpus data collected in this study suggest that the ambiguity arising from the occurrence of polysemous words in unbiasing contexts can be explained as the result of senses with similar levels of salience in those words' polysemy networks competing for activation[6].

Finally, this paper concludes with the suggestion that its findings can be adequately explained with an adaptation of the Graded Salience Hypothesis (Giora 1997; 2003); findings in this corpus-led study indicate that polysemous ambiguity occurs as a result of the co-activation of co-salient senses when polysemes are comprehended in context. Statistically significant data shows some evidence that ambiguous polysemes have a smaller number of discrete senses, and that the two most conceptually similar of these senses occupy co-salient positions in their respective polysemy networks. Conversely, it is found that polysemes that were significantly unambiguous display the reverse of these properties. This paper's proposal for an understanding of the psycholinguistic factors underpinning polysemous ambiguity is thus the *Graded Co-Salience Hypothesis*: while lexical access is ordered by a function of senses' salience, some equally salient senses conflict and compete in the interpretation of some polysemes such that they become rendered ambiguous.

6 References

Arppe, A., Gilquin, G., Glynn, D., Hilpert, M., & Zerschel, A. (2010). Cognitive Corpus Linguistics: five points of debate on current theory and methodology. *Corpora*, 5(1), 1–27.

Berez, A., & Gries, S. T. (2009). In defence of corpus-based methods: A behavioral profile analysis of polysemous 'get' in English. *University of Washington Working Papers in Linguistics*, 27, 57–116.

[6] In other words, contexts that prove unhelpful for the interpretation of the precise meaning of that polyseme, determined insofar as that particular instance was deemed ambiguous.

Biber, D., & Finegan, E. (2013–). *ARCHER 3.2. A Representative Corpus of Historical English Registers*. Retrieved 01/02/2020, from <http://www.manchester.ac.uk/archer/>.

Bréal, M. (1897). *Essai de sémantique (Science des significations).* Paris: Gérard Monfort.

Bybee, J. (2006). From Usage to Grammar: The Mind's Response to Repetition. *Language*, 82(4), 711–733.

Bybee, J. (2010). *Language, use, and cognition.* Cambridge: Cambridge University Press.

Bybee, J., & Hopper, P. (2001). *Frequency and the Emergence of Linguistic Structure.* Amsterdam: John Benjamins.

Fillmore, C. J., & Atkins, B. T. (2000). Describing polysemy: the case of 'crawl'. In Y. Ravin, & C. Leacock (Eds.), *Polysemy: Theoretical and Computational Approaches.* Oxford: Oxford University Press.

Fillmore, C. J., Kay, P., & O'Connor, M. K. (1988). Regularity and idiomaticity in grammatical constructions: the case of let alone. *Language*, 64, 501–538.

Frazier, L., & Rayner, K. (1990). Taking on semantic commitments: Processing multiple meanings vs. multiple senses. *Journal of Memory and Language*, 29, 181–200.

Giora, R. (1997). Understanding figurative and literal language: The graded salience hypothesis. *Cognitive Linguistics*, 8(3), 183–206.

Giora, R. (1999). On the priority of salient meanings: Studies of literal and figurative language. *Journal of Pragmatics*, 31, 919–929.

Giora, R. (2003). *On Our Mind: Salience, Context, and Figurative Language.* Oxford: Oxford University Press.

Giora, R., & Fein, O. (1999). Irony: Context and salience. *Metaphor and Symbol*, 14(4), 241–257.

Giora, R., Drucker, A., & Fein, O. (2014). Resonating with default nonsalient interpretations: A corpus-based study of negative sarcasm. *Belgian Journal of Linguistics*, 28, 3–18.

Giora, R., Gazal, O., Goldstein, I., Fein, O., & Argyris, S. K. (2012). Salience and context: Interpretation of metaphorical and literal language by young adults diagnosed with Asperger's syndrome. *Metaphor and Symbol*, 27, 22–54.

Glynn, D. (2014). The many uses of run: Corpus methods and Socio-Cognitive Semantics. In D. Glynn, & J. Robinson (Eds.), *Corpus methods for semantics: Quantitative studies in polysemy and synonymy* (pp. 117–144). Amsterdam: John Benjamins.

Glynn, D. (2016). Quantifying polysemy: Corpus methodology for prototype theory. *Folia Linguistica*, 50(2), 413–447.

Gries, S. T. (2006). Corpus-based methods and Cognitive Semantics: The many senses of 'to run'. In S. T. Gries, & A. Stefanowitsch (Eds.), *Corpora in Cognitive Linguistics: Corpus-based approaches to syntax and lexis* (pp. 57–99). Berlin: Mouton de Gruyter.

Gries, S. T. (2009). *Quantitative Corpus Linguistics with R: A Practical Introduction.* London: Routledge.

Hardie, A. (2012). CQPweb – combining power, flexibility and usability in a corpus analysis tool. *International Journal of Corpus Linguistics*, 17(3): 380–409.

Jastrzembski, J. E. (1981). Multiple Meanings, Number of Related Meanings, Frequency of Occurrence, and the Lexicon. *Cognitive Psychology*, 13, 278–305.

Jaszczolt, K. (2005). *Default Semantics: Foundations of a Compositional Theory of Acts of Communication.* Oxford: Oxford University Press.

Jaszczolt, K. (2010). Default Semantics. In B. Heine, & H. Narrog (Eds.), *The Oxford Handbook of Linguistic Analysis* (pp. 193–221). Oxford: Oxford University Press.

Kishner, J. M., & Gibbs, R. W. (1996). How 'just' gets its meanings: Polysemy and context in psychological semantics. *Language and Speech*, 39(1), 19–36.

Received:17/08/2020
Revisions: 27/03/2020
Accepted: 10/04/2021

Klein, D. K., & Murphy, G. (2001). The representation of polysemous words. *Journal of Memory and Language,* 45, 259–282.

Klein, D. K., & Murphy, G. (2002). Paper has been my ruin: conceptual relations of polysemous senses. *Journal of Memory and Language,* 47, 548–570.

Lewis, D. (1969) *Convention*. Cambridge: Harvard University Press.

Navarro, I. (2000). A Cognitive–Semantic Analysis of the English Lexical Unit 'in'. *Cuadernos de Investigación Filológica,* 26, 189–220.

Nerlich, B., & Clarke, D. D. (2003). Polysemy and flexibility: introduction and overview. In B. Nerlich, Z. Todd, V. Herman, & D. Clarke (Eds.), *Polysemy: Flexible Patterns of Meanings in Language and Mind* (pp. 3–30). Berlin: Mouton de Gruyter.

Raukko, J. (2003). Polysemy as flexible meaning: experiments with English 'get' and Finnish 'pitää'. In B. Nerlich, Z. Todd, V. Herman, & D. Clarke (Eds.), *Polysemy: Flexible Patterns of Meanings in Language and Mind* (pp. 161–194). Berlin: Mouton de Gruyter.

Rodd, J., Gaskell, G., & Marslen-Wilson, W. (2002). Making sense of semantic ambiguity: Semantic competition in lexical access. *Journal of Memory and Language,* 46, 245–266.

Rosch, E. (1975). Cognitive Representation of Semantic Categories. *Journal of Experimental Psychology,* 104(3), 192–233.

RStudio Team. (2020). *RStudio: Integrated Development for R*. Boston: RStudio PBC.

Rudzka-Ostyn, B. (1988). Semantic extensions into the domain of verbal communication. In B. Rudzka-Ostyn (Ed.), *Topics in Cognitive Linguistics* (pp. 507–553). Amsterdam: John Benjamins.

Rudzka-Ostyn, B. (1989). Prototypes, schemas, and cross-category correspondences: The case of 'ask'. *Linguistics,* 27, 613–661.

Rudzka-Ostyn, B. (1995). Metaphor, schema, invariance: The case of verbs of answering. In L. Goossens, P. Pauwels, B. Rudzka-Ostyn, A.-M. Simon-Vandenbergen, & J. Vanparys (Eds.), *By word of mouth: Metaphor, metonymy, and linguistic action from a cognitive perspective* (pp., 205–244). Amsterdam: John Benjamins.

Schlenker, P. (2019). What is Super Semantics? *Philosophical Perspectives,* 32(1), 365–453.

Speelman, D., & Glynn, D. (2005). LiveJournal corpus of American and British English. Leuven: University of Leuven, Department of Linguistics.

Stalnaker, R. C. (1979 [1999]). Assertion. *Syntax and Semantics,* 9, 78–97.

Tyler, A., & Evans, V. (2001). Reconsidering Prepositional Polysemy Networks: The Case of 'over'. *Language,* 77(4), 724–765.

Weinreich, U. (1964). Webster's Third: A Critique of Its Semantics. *International Journal of American Linguistics,* 30(4), 405–409.

Williams, J. N. (1992). Processing polysemous words in context: Evidence for interrelated meanings. *Journal of Psycholinguistic Research,* 21, 193–218.

Wittgenstein, L. (1953). *Philosophical Investigations.* New York: Macmillan Publishing Company.

Wittgenstein, L. (1974). *Philosophical Grammar.* Oxford: Blackwell.

About the Author

T. R. Williamson is a current MPhil (by Thesis) student at the University of Cambridge, though this paper was completed during his time at Lancaster. His current research aims to bring together work on experimental pragmatics and embodied cognition to investigate how we process idiomatic phrases.

Acknowledgements

I give thanks to two anonymous reviewers for the useful comments they made on previous manuscripts of this paper, as well as to the Journal's Editorial Committee for further comments at later review stages. I also thank my supervisor at Lancaster University, Dr Vittorio Tantucci, for his invaluable support on this project.

Received:17/08/2020
Revisions: 27/03/2020
Accepted: 10/04/2021

Journal of the Undergraduate Linguistics Association of Britain

Vol. 1, Issue 1 | Spring 2021

Articles

@ULAB_Journal
https://www.ulab.org.uk/