Variation and Change in Modern Received Pronunciation: Understanding interactions between private education and regional accent variation

Supervisors: Dr Ghada Khattab[*], Dr Rory Turnbull [†], *(previously Dr Danielle Turton[‡])*

Caitlin Halfacre

October, 2023

Submitted for the fulfilment of Doctor of Philosophy
School of English Literature, Language, and Linguistics

[*] School of Education, Communication, and Language Science. Newcastle University
[†] School of English Literature, Language, and Linguistics. Newcastle University
[‡] Linguistics and English Language Department. Lancaster University

Abstract

In this project I approach Received Pronunciation (henceforth RP) from a sociolinguistic standpoint, using three groups of recorded speakers to understand the potential for regional variation. CoRP-NE are speakers privately educated in the North East and they are compared to CoRP-SE (speakers privately educated in the South East, canonically the home of RP) and DECTE (state-educated Tyneside speakers from Corrigan et al. 2012), as baselines for regional vs. non-regional behaviour.

Trudgill (2008) suggests that innovations arise in RP as a process of change from below where new variants enter the variety from working class south-eastern accents before diffusing across the country. Alongside comparing synchronic variables, the life cycle of phonological processes (Bermúdez-Otero, 2015) is used to understanding the nature of a diffuse speech community and the regionality of RP within this framework.

I find that regional vs. non-regional behaviour depends on the variable. In the FOOT-STRUT split, male CoRP-NE speakers behave regionally, not producing a split, whereas female speakers have a different pattern to either the regional or 'RP' version, creating a split with a STRUT vowel different to the CoRP-SE version. In the TRAP-BATH split CoRP-NE speakers behave broadly regionally with no split in vowel frontness.

GOAT allophony is more complex. The CoRP-NE speakers show a similar GOAT vowel and GOAT-GOAL split in the monosyllabic context to the CoRP-SE speakers (DECTE speakers show a monophthong with no split), demonstrating non-regional behaviour. However, in analysis the morphological conditioning of the pre-/l/ position of the GOAT vowel, I found that the CoRP-NE speakers show a different pattern to the CoRP-SE speakers, reaching stage 3 of the life cycle of phonological processes. The pattern appears to be either a simplification of the rule from the diffusion process, or a further progression of the change moving through the grammar. The data cannot show which of these is the case but either case demonstrates difference to the non-regional pattern.

Overall, results show that speakers in the CoRP-NE category are a unique speech community. There are two possible conclusions from these results. The first is that there is a non-regional accent in the North East but the speaker group recorded here is not of a high enough social class to have it. This implies that the non-regional variety can only be found in a higher social class group in the North East than in the South East. The second possible conclusion is that if a non-regional accent ever did exist, it does not any more.

Acknowledgements

This work and my MA were funded by a NINE DTP ESRC studentship, and I am grateful for their financial and training support.

I often say that I'm borderline too much of an extravert for academia (and then a lockdown happened), and as such there are many people without whom I'm not sure I'd still be here, let alone this thesis would exist, so many thanks are due.

I have been so fortunate to find a department in the School of English Literature, Language, and Linguistics that has supported me and looked out for me since day one of my undergraduate degree. There are so many staff, past and present who have contributed to the last nine years. Particular thanks need to go to Heike Pichler for encouraging me into opportunities I never would have thought of, to Joel Wallenberg for pulling my initial BA application out and encouraging me to start, and Scott Burdon for dealing with endless admin emails. Danielle Turton, who helped create the idea for this PhD and without whom it would never have started, our loss was Lancaster's gain.

In stage one of this PhD I started a writing group to help with motivation and never imagined that it would end up being such a source of support and encouragement, so thank you to writing group members past and present: Jasmine, Sheradan, Joaquin, Bella, Andreas, Damar, Olivia, Yanyu, Carol-Ann, there are literally words on these pages that wouldn't have been without you.

To many other staff and fellow students. From Georgetown; Jen Nycz for supervising me at Georgetown, Lisa Zsiga for welcoming me into her class, all the students there who pushed me in discussion, and particularly to Kerri-Ann Butcher, for a friendship that started with working out accommodation and visas but continued into debugging GAMMs across WhatsApp. At Newcastle, Jalal Al-Tamimi for teaching me how to use Rocket, the staff and students of the Phonetics & Phonology, and Variation Lab, Gary Taylor-Raebel for being interim supervisor, UCU colleagues for reminding me that academia is worth fighting for. And of course, my wonderful supervisors, Ghada Khattab and Rory Turnbull. Liam Keeble, for walks around Exhibition Park in lockdown, discussing the principles of science and reproducibility, and for excellent proof-reading skills.

I owe such a debt of gratitude to my parents, David and Karen Halfacre who always believed in my abilities and encouraged me through education, even when the route was a little untraditional. My housemate Catherine, who is ever patient with me, and put up with the complete life chaos of the last few weeks of write-up, all the friends who offered coffee, meals, chocolate, ranting support and were generally there all the way through, and lastly, my church family at Jesmond Parish Church, particularly my home group and focus groups, you really are my family.

~ *Soli Deo Gloria* ~

Contents

List of Figures

List of Tables

Chapter 1.

Introduction

Despite the high profile of Received Pronunciation (henceforth RP) in both the academic and popular spheres, very little recent research has focused on changes in the accent itself or the specifics of variation in pronunciation. The unique position that RP holds as a non-regional sociolect, rather than an region-based accent or dialect is accepted and taught but has not been considered in depth, with much of the existing research accepting southern English standards as exemplary of RP. The existence of this accent raises the question of how a language variety can exist when its speakers are hundreds of miles apart. This thesis approaches RP from a sociolinguistic perspective (Fabricius, 2000*a*), as a variety spoken by real speakers, with all the potential for variation and change that brings. Further information on defining the speaker group and how that is used to understand regional variation is given in the Methodology in chapter 3. The aims of the project include understanding the potential for regional variation as well as understanding potential mechanisms that may underpin variation and change, both from a sociolinguistic and phonological perspective. Section 1.1 below will outline the research questions and results found.

1.1. Research Questions

The research questions address the descriptive, social and grammatical aims of the project respectively.

1. In terms of pronunciation, do British north-eastern 'RP' speakers (CoRP-NE) behave regionally or non-regionally? The variables to be considered are:
   - FOOT-STRUT split
   - TRAP-BATH split
   - GOAT allophony
2. What insights do these differences give us into the nature of a non-regional sociolect comprising diffuse speech communities?
3. If the results of 1 vary depending on the variable in question, what phonological or social factors control this?

Based on pilot studies conducted during my BA and MA I can make hypotheses about the results of question 1, and also suggest a theoretical framework for understanding the variation. Halfacre and Khattab (2019) discuss the FOOT-STRUT and TRAP-BATH splits, observing non-regional behaviour for FOOT-STRUT and regional behaviour (with possible rule simplification) for TRAP-BATH. Halfacre (2017) does not look at regional variation but demonstrates a possible morphological effect for GOAT allophony. All of these are discussed further in the literature review and alongside other related literature that provide grounding for the hypotheses and theoretical frameworks within which to understand the data that will be collected for this dissertation.

The results found show that with respect to the FOOT-STRUT split male CoRP-NE speakers behave regionally and female CoRP speakers behave neither regionally nor non-regionally; with respect to the TRAP-BATH split, the CoRP-NE speakers behave broadly regionally with some indications of different behaviour or a change in progress. For research question 2 results lead to the conclusion that either evidence of a 'non-regional' accent can only be found in the North East by going further up the socio-economic spectrum than is needed in the South East, or that a non-regional accent does not exist. Evidence is also found against change diffusing across the country, as suggested by Trudgill (2008).For research question 3, a clear pattern is not found that governs the overtly regional or non-regional behaviour.

Chapter 2.

Literature Review

2.1. Introduction

This chapter will provide an overview of the literature relevant to the research questions, including the origins and definitions of RP, as well as the history of the variables under study. Section 2.2 will overview the origins of RP that can be found in the literature, including the historical context and different writers' views on the the scope for change in the accent. Definitions of RP are also given from past studies on the accent. Section 2.3 brings in various ways in which social class has been used and studied as a predictor for variation in previous linguistic studies. Sections 2.4, 2.5, and 2.6 overview the splits that will be analysed, including the potential variants and the history of the split in the South East and North East. Section 2.6 also includes the phonological background to the morpho-phonological conditioning of the split.

2.2. Received Pronunciation

Received Pronunciation (henceforth RP), an accent of English, is generally considered to be a non-regional sociolect (Honey, 1985; Kerswill, 2007) and is often used as a reference variety of English, particularly to define the phonological inventory of English, and as a point of comparison for regional varieties (Wells, 1982*a*,*b*; Davenport and Hannahs, 2010)).

This has led to entangled and conflicting usage of the term in research. The first usage is those who are broadly describing English phonetics or creating pronunciation dictionaries (for example Jones 1917, 1972; Gimson and Cruttenden 1994). This usage also influences and is influenced by standard language ideology. The second is dialectologists and sociolinguists who describe and analyse the speech of existing speakers (for example Wells 1982*b*; Fabricius 2002*c*,*a*; Trudgill 2008).

This thesis is directly related to past sociolinguistic study of RP. However, in the endeavour of coming to an understanding of the history and current state of the accent, the contributions of phonetics, pronunciation guides and other EFL resources cannot be disregarded. The rest of this section attempts to bring together literature surrounding the origins, history, and predictions regarding the future of RP.

2.2.1. Origins and History

The first use of the term 'Received Pronunciation' was traced by Fabricius (2000*a*) and Hannisdal (2006) to Ellis (1889), who used it to describe a particular model of speech that was socially acceptable at court and in the London professional circles. Macaulay (1988) describes 'Received' as both alluding to 'received wisdom' and implying that a person would be received at court. A few decades later Jones published a model of pronunciation sourced from the speech of families from the south of England "whose men-folk have been educated at the great public boarding schools"(1917, viii). Generally these descriptions and definitions (which will be further explored in section 2.2.3) begin at the nineteenth century.

However, there were prestige accents and perceived standards for a long time before that, some of which may have been ancestors of RP. According to Cruttenden (2001) there was possibly a spoken English standard as far back as Chaucer, and according to Strang (1970), at least as far back as the last five centuries. This history is cited by Trudgill (2008), who describes a ruling class variety of English from London (particularly the court), which then disseminated out across the country, coinciding with the establishment of the English Public Schools (*Public Schools Act* 1868). These schools were residential, and it is this point which Trudgill (2008) identifies as the beginning of the 'non-regionality' of RP, and also specifies that it distinctly contained “no regional features whatsoever" (2008, p.5) (more details on this in section 2.2.2). We can see from these accounts that a standardised, or prestige accent has always been tied to both social class and schooling. Fabricius (2000*a*) also puts the origins at this time, within a particular set of sociological and ideological circumstances. The beginnings of accent ideology can be traced back to at least the 1700s, but in the nineteenth century it was growing and the ideas of ‘correct’ and ‘proper’ speech were on the rise (Mugglestone, 2007).

In 1870 the Elementary Education Act (HM Government, 1870) promised elementary education for all, and with this rose the idea of accent reflecting education and status, with schools pushing to discourage regional accents. Not long before this the public schools had shifted from their earlier pattern of educating children from families with little financial resources, with some children of more affluent classes, to being almost completely dominated by the latter group. Mugglestone (2007) gives, as an example, the proportions of Eton College’s intake. They originated in 1440 with seventy of the ‘poor and needy’ as free students and twenty fee-paying noblemen’s son’s, whereas by the late eighteenth century, this balance had shifted dramatically, between 1755 and 1790 only 1.3% of their intake were sons of tradesmen (thirty-eight out of 3,500), and between 1821 and 1830 only two pupils represented lower social classes, before this dropped to zero in the next decade. As the increase of fee-paying students grew, these were coming from more varied areas of the country, at the end of the eighteenth century, 80% of the students at Rugby School were not local. This trend created the context in which the accent referred to by Trudgill above was formed, and the minority of the ‘local scholars’ allowed for an approach where schools sought to impose “polish" (Mugglestone, 2007) on all their pupils. Other schools were also set up, aiming at this upper and middle class majority, rather than having the initial aims of the public schools. The new public school image included a social process and all produced similar outputs, of young men that spoke in a similar way, creating what Mugglestone describes as “one of the most enduring images of ‘talking proper’ that England has ever known" (2007, p. 275). As the social and linguistic ‘purism’ endured, these schools became an isolated linguistic environment, avoiding any marker of regionality that could be perceived as ‘rustic’ or ‘provincial’. This was achieved as much by peer pressure as by instruction.

The nineteenth century created a perfect storm where, the rise of accent ideology in general (Mugglestone, 2007), and the two government acts (HM Government, 1868, 1870), dilineated a non-regional accent as something separate and to be aspired to. An RP accent was produced by a elite education system and thus a non-regional accent was to be aspired to. Regional variation began to be seen as a mark of less education, social value, or aspiration. Aspects of these attitudes continue to the present day, as seen in reports from Parveen (2020) that students at Durham with northern accents faced ridicule from their peers.

According to Lindsey (2019), the spread of an accent that arose in London and the South East, is also rooted in the hierarchy of power and authority needed to control the British Empire, a hierarchy which had the monarchy and the London court at the top. The process of controlling the British Empire relied on stratification and rank, and also required a large number of the ruling classes to live abroad, leaving their sons in the public schools, where they were taught manners, behaviour, and certain patterns of speech.

Trudgill (2002, 2008) considers RP as typologically originating in the South East, despite the lack of regional features. He states that even a trained linguist cannot locate a speaker of RP to a particular region of England (2002), a view also supported by Kerswill (2007). A visualisation of the non-regional model can be seen in figure 2.1. The typological origins are seen in features such as lack of rhoticity and the existence of FOOT-STRUT and TRAP-BATH distinctions. However, these distinctions also demonstrate that the non-regional variety is being influenced by changes in working class varieties of London and Southern English. All three of the features mentioned above arose in English after at least the pre-cursor to RP existed. Therefore, Trudgill (2008) suggests a process of ongoing change from below (Labov, 2001*b*), that RP is susceptible to because the speakers are a very small minority of the population.

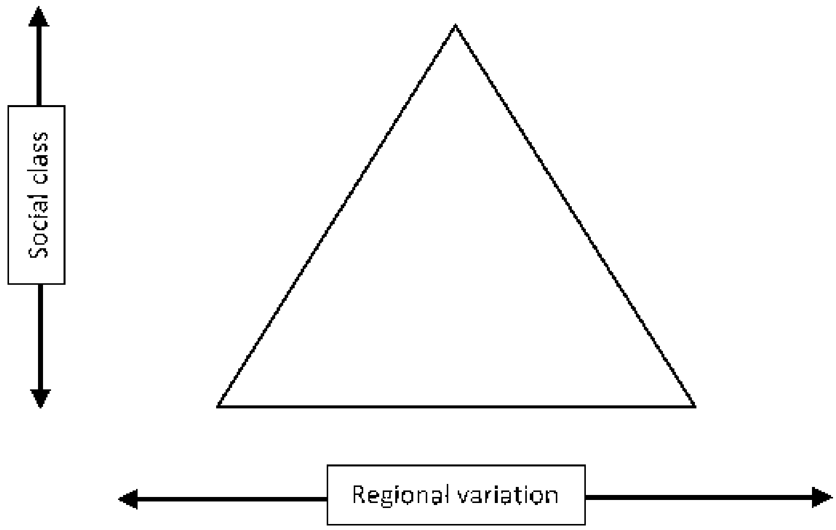

Figure 2.1.: Relationship between social and regional accents in England (adapted from Wells (1982*a*), also reported by Ward (1929) from Daniel Jones)

2.2.2. Current views on the Future of RP

There is disagreement in the literature as to whether RP is changing or disappearing. The idea of disappearance is often drawn from non-RP accents appearing in settings that would previously have been exclusively RP, for example, BBC radio, and due to the fact that adoptive RP (Wells 1982*b*, defined in section 2.2.3) is far less common. However, Trudgill's claim is that there are "no fewer *native* speakers" (2008, p. 4, emphasis in original).

Within Trudgill's model of change from below, described in section 2.2.1 above, where changes enter RP from south-eastern accents, the question of how RP remains unique becomes an issue. This is also raised by Kerswill (2007); if remaining non-regional is a criterion for RP to exist, the issue arises because any entering feature is usually present in a region already. This is answered with three factors:

1. RP lags behind the accents from which changes originate, and by the time it catches up, the others have changed again.

2. Not all innovations from the surrounding south-eastern accents are adopted into RP,

for example, diphthong shift (Wells, 1982*b*).

3. RP has innovations of its own that do not come from other accents. For example, the onset of the GOAT vowel has fronted and unrounded (Trudgill, 2008).

This model of change is key to Trudgill's argument that RP is not disappearing. The criterion for inclusion of a feature in RP, is that it is non-regional; he works through multiple features that have transitioned from regional to non-regional, in adoption by younger speakers, and taking into account the factors above, have become RP features. Therefore, those who were RP speakers are not speaking Cockney, Estuary English, or any other south-eastern variety, they're still speaking RP, but RP has changed. This model is supported by Kerswill (2007), who also works through changes in RP based on Trudgill's and Wells' descriptions, including almost completed changes such as fronting of GOAT and lowering of TRAP, and well established changes such as fronting of GOOSE and allophonic variation of the GOAT vowel before syllable final /l/.

More recently some writers have declared that RP no longer exists at all. These include Lindsey (2019), who discusses the change in accent ideology through the late 20th century, with shift in social hierarchies and acceptance of those from different backgrounds in visible positions such as newscasters and politicians. Lindsey also ties RP more tightly to the British empire than many others, and takes its end point as the end of the 20th century, when the British empire was finally closed by the handover of Hong Kong in 1997.

2.2.3. Terms and Definitions

There has been much debate over the term 'Received Pronunciation', with even Jones (1972) expressing dissatisfaction with the term and Macaulay (1988) suggesting its use be discontinued. However, it has persevered in both English Language teaching (TEFL, TESOL, etc.) and sociolinguistics, with some variation. Over the decades of study of British accents an unusual phenomenon has taken place, whereby if variation or change is uncovered, a different variety is defined instead of describing RP itself as having changed. A few of these categories and descriptions are covered below, organised by author.

Wells (1982*b*) defines mainstream RP, U-RP (upper-crust RP), adoptive RP, and Near-RP. Mainstream RP is only described as a central tendency, which can be isolated by defining U-RP and adoptive RP. U-RP is described as the RP, not of the middle and upper-middle classes, but of the dowager duchess and "jolly-hockey-sticks schoolmistress at an expensive private girls school" (1982*b*, p. 280). Typical features of this variety's vowels include an opening diphthongal realisation of the TRAP /æ/ vowel, as [ɛæ] or [eæ], relatively front starting points of the MOUTH and PRICE diphthongs, a relatively front and unrounded endpoint for MOUTH [aɨ], relatively back STRUT and BATH-PALM-START vowels and /ɔː/ in CLOTH words. Typical consonantal features include no glottalling of plosives and a tapped /*r/, [ɾ].

Adoptive RP is the accent of adults who did not speak RP as children, but have acquired it later in life, usually due to social circumstances. If acquired successfully, it merges with mainstream RP but is often characterised by under use of informal characteristics of RP which could be perceived as careless or incorrect, such as /r/ sandhi (linking or intrusive /r/).

Fabricius (2000*b*,*a*, 2002*c*,*a*) divides the varieties slightly differently, with c-RP (constructed RP) being the norm described in pronunciation dictionaries and n-RP (native RP) as the accent of people who have acquired it as native speakers; in contrast to Wells (1982*b*), Fabricius does not distinguish a form of adoptive-RP but discusses an unnamed variety equivalent to Wells' U-RP and a set of accents that are in-between RP and non-RP accents in Britain, equivalent to Wells' adoptive-RP and near-RP. In later papers (Fabricius, 2006) varieties of near -RP such as 'London-near-RP' and 'Regional-near-RP' are suggested.

Lindsey (2019) (as discussed in section 2.2.1 above) claims that RP can now be referred to in the past tense because the sociological environment which created a top-down standard no longer exists. However, Lindsey is only discussing RP as a standard, not as a sociolinguistic object (Fabricius, 2002*c*); at no point does Lindsey handle key details to understanding a language variety from a variationist, or sociolinguistic perspective. As discussed in Halfacre (2020), there is no consideration of who the speakers of RP were, what happened to them, and what they are speaking now. Instead, the book moves to a pronunciation guide, based on what the author considers to be the accent that is currently acceptable and/or the most useful for learner. Therefore, a new standard is defined, called 'Standard Southern British (SSB)'. An important point to note is that, by using 'southern', this term actively ties the accent to a particular region, reducing possibilities for an English non-regional accent.

A repeated theme from many of these authors is the tension between potential change in RP, and the disappearance of RP. This tension reflects not just a different use in the meaning of change versus disappearance but also different standpoints from which the variety is viewed. Those for whom RP is a standardised variety at a specific time (for example, Lindsey 2019) tend to describe RP as disappearing or being lost, whereas those who take a more sociolinguistic approach by considering the equivalent set of speakers in the modern speech community (for example, Trudgill 2008) tend to discuss changes occurring within the accent. This mixed situation is commented on by Britain, who notes that there is a large body of literature on an accent cited as RP, that is *assumed* (2017, p. 288) to be an accent of the elite (i.e. descriptions of standard and educated speech, used for things like pronunciation dictionaries) but there is very little empirical study of the unscripted speech of these people. This call for study echoes Fabricius (2002*a*) who cites Schneider's (1999) statement that "for sociolinguistic modelling, a continuum of which one pole just does not exist, would not be very convincing" and describes RP as a 'sociolinguistic object'. If we want to understand the full sociolinguistic landscape of England, we cannot, as Schneider points out, take a model which doesn't include both ends of the social spectrum. RP speakers are part of the British speech community, and therefore, display social characteristics and linguistic features which can form the basis for empirical research. While in the older field of dialectology RP has always been present as a background comparator, in modern sociolinguistics it should be one of the varieties of English that can be studied, and in fact must be in order to fully model accent variation in England.

Badia Barrera (2015) is one of few recent studies that approaches RP as a current variety and looks for variation within it. As such it warrants some time spent considering its approaches and results. Badia Barrera studies the incorporation of a non-standard feature, /t/-glottalling, into the speech of young RP speakers, focussing on possible social variation within RP by comparing teenage speakers from three different types of schools in the South of England (a major public boarding school (see section 2.2.1), a private non-boarding school and an outstanding rated comprehensive school in a prosperous rural area) and comparing them to older speakers who are alumni of the same schools. The three school types are operationalised as proxies for social class.

The study finds different social patterns for word-final and word-medial positions. In word-medial position (e.g. *water, butter*) the speakers from the London private non-boarding school favour glottalling the most, closely followed by the comprehensive school, and the private boarding school are a lot further behind. In word-final position speakers from the comprehensive school favour glottalling the most, followed by the private non-boarding school and then the private boarding school. If the change is from below it would be expected that the group considered the lowest social class (the comprehensive school) would be leading in both contexts but the difference is suggested to be because word-medial /t/-glottalling is particularly associated with London and so the location of that school is having a greater effect than the social class. When asked about friends and activities outside of school the students from the London school said that they had friends from a nearby sixth form college as well so it is likely that due to their social networks interacting with other schools, these students are taking on non-standard London features faster.

2.3. Social Class in Linguistics

The study of social class within linguistics is well developed, reaching as far back as Labov's (1966) study of rhoticity in New York City English, and the relationship between social class and language has long been accepted, both academically and popularly. In 1972, National Opinion Polls carried out a survey of a random sample of the British Public asking what factor was the most important in being able to tell what class someone was; the highest scored answer was 'the way they speak' (Wells, 1982*a*; Reid, 1981).

Social (or socio-economic) class is a form of socially constructed stratification, an arrangement of society into layers. Social stratifications of one form or another exist in almost every society in the world. They are not pre-determined facts about people but have very real impact in society. Social classs is one of the most complex social stratifications to describe and define (Reid, 1981).

Studies of social class in linguistics have measured social class in various ways, including education (Labov et al., 2016; Lawson et al., 2011), deprivation index by postcode (Alderton, 2019) and occupation (Baranowski, 2017). Trudgill (1974) uses a composite measure calculated from occupation, parents' occupation, education, income, locality, and housing. According to Meyerhoff (2006), Max Weber's work provides a conceptualisation of class that is most appropriate to linguistics (for example 1978; 2012). This is because Weber theorises class based not only on an individual's current circumstance but on their social actions, including factors such as economic situation, lifestyle, and life chances. This approach can explain the disconnect that sometimes occurs between a person's current situation and how they identify their own socio-economic class. Macy (2001) explains, with an example of blue-collar workers in the United States, that variation within an occupational group, who would in other studies be considered the same socio-economic class, can be explained by their background. College educated people in a blue-collar job are more likely to identify as middle or upper class than others in similar jobs. This type of variation, which includes current situation, past experiences, and expectations, is particularly important to linguists because using these factors makes it possible to include an individual's participation in various social behaviours, their aspirations and their attitudes, all of which affect their language use (Meyerhoff, 2006). This tradition of sociology, which takes into account both similarities and differences between groups along with individual attitudes, perceptions and aspirations, allows for mobility in a class system, which affects both language and attitudes towards it (Mugglestone, 2007).

Despite the above discussions by Meyerhoff and others, linguistic studies often categorise social class very simply, usually by one factor, which in sociolinguistic studies is most often occupation (for example Labov 1990; Baranowski 2017).

In the early stages of linguistic study upper middle class speech was well studied under the guise of phonetics, particularly amongst those who focussed on describing educated or standardised speech. Amongst more recent study of social class and language, the upper and middle classes are generally left aside in favour of working class speech, and variationist studies that do consider social class generally stop at the middle class (for example Baranowski 2017). Within American English Kroch (1995) still stands alone as one of very few studies of upper class speech, and in the UK, Britain (2017) has called for further study of the speech of 'the elite'.

2.4. The FOOT-STRUT split

2.4.1. Overview and History

The FOOT-STRUT split is defined by Wells (1982*a*) as the phonemic distinction between the short vowels found in, RP and southern English accents in England that distinguishes pairs such as *put*∼*putt* and *look*∼*luck*. This distinction did not exist in Middle English (Wells, 1982*a*). The short vowel /u/ found in Middle English split into two separate phonemes /ʊ/ and /ʌ/, except in northern English accents where they remain one phoneme, /ʊ/. The process was that the vowels in some words unrounded and became more open, creating the [ʌ], whereas others kept a short vowel which centralised creating the modern day [ʊ]. Wells (1982*a*) times the split as being established by the middle of the seventeenth century, whereas Beal (2012) places it later, in the eighteenth century. Figure 2.2 presents the development of the split in RP and GenAm, as proposed by Wells, who suggests that it could have begun as an allophonic split with [ɤ] (forerunner to [ʌ])in the majority of environments but a rounded quality (modern [ʊ]) after labials. However, there are multiple exceptions to this rule, *vulture* and *fun* both have /ʌ/ despite beginning with labials and *sugar* has /ʊ/, despite no preceding labial. Turton and Baranowski (2020) suggest a split due to more complex phonetic factors, which is discussed further below.

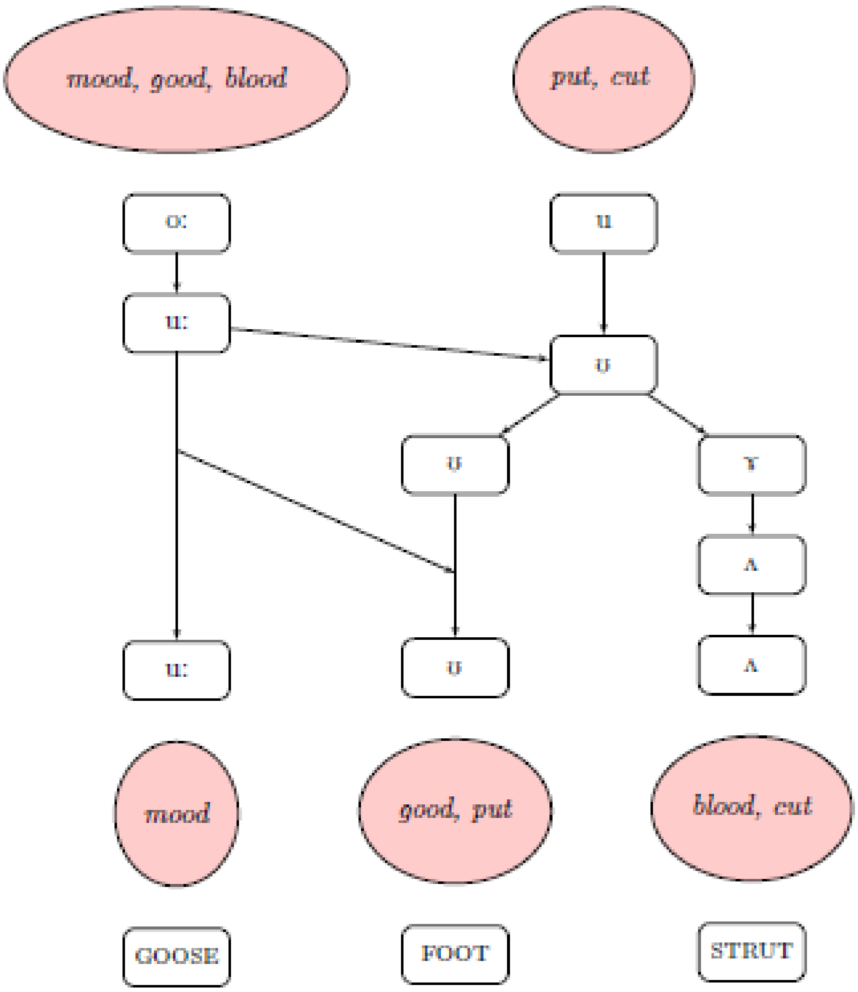

Figure 2.2.: FOOT∼STRUT split development (Turton and Baranowski, 2020, p. 5) adapted from (Wells, 1982*a*, p. 198)

2.4.2. Present day situation of the FOOT∼STRUT Distinction

In present day England, the FOOT-STRUT distinction is one of two markers popularly used to split the country into the (linguistic) North and South (Wells, 1982*b*), dividing the population approximately in two (the other being the TRAP-BATH distinction discussed in section 2.5 above). Therefore, aside from RP speakers, or in line with the definitions of this project, aside from speakers without regional features, approximately half the population of England do not have FOOT-STRUT distinction, and hence only have five short vowels in their system (Wells, 1982*b*) rather than six. However, Britain et al. (2016) provide possible evidence for the split spreading, though this could be due to social class-based variation that was not accounted for in their data collection method. Wells (1982*b*) proposes that the further north one goes the higher up the social spectrum one has to go to find the distinction. Speakers without the [ʌ] in their phonology, who try to acquire the vowel when attempting to shift their accent towards the standard, or what they perceive as more prestigious, struggle to either reach an [ʌ] realisation and produce a more schwa-like vowel, or to predict the split and can end up hyper-correcting and putting [ʌ] vowels in FOOT words. Evans and Iverson (2007) show the difficulty that speakers have in gaining a vowel that they do not have natively; they find that speakers with northern accents living in London for university not only changed their vowel in the words *bud* and *cud* but also in the word *could*.

Turton and Baranowski (2020) found that speakers from Manchester who identified a split in a minimal pair (even if it was not audibly perceptible) were majority lower middle class (the highest class of their 5 point occupation based scale). In the south the split could be found in all speakers, whereas in Manchester it can only be found in lower middle class speakers. This lends support to Wells' theory that the further north speakers are from the higher up the socio-economic scale one has to go in order to find the distinction. Therefore, it is reasonable to assume that in the North East the split would only be found in speakers of an even higher socio-economic class. Even in speakers without a phonemic distinction, STRUT words were found to have statistically higher F1 values than FOOT words. They suggest that this is due to a combination of phonetic conditioning and articulatory factors and could be an indication of the diachronic precursor to the original split. While considering

	STRUT	FOOT	
Broad	[ʊ]	[ʊ]	One phoneme, /ʊ/
Intermediate	[ə∼ʌ]	[ə∼ʌ]	One phoneme; realisation modified
	[ə∼ʌ]	[ʊ]	Two phonemes; incidence may be erratic
RP	[ʌ]	[ʊ]	Two phonemes; /ʌ/ and /ʊ/

Table 2.1.: Possibilities for FOOT and STRUT in the north of England, adapted from (Wells, 1982*b*, p. 353).

social class/socio-economic class at this point, it is important to note that class in the UK has changed a lot over the last 40-100 years. However, Wells' model as shown in figure 2.1, is still a helpful starting point, particularly since public schools and private education still play a part in the political, social, and linguistic landscape of the country. Evidence such as Turton and Baranowski (2020) referenced above demonstrates that language variation based on at least some definition of social class still interacts with regional variation. How class will be defined and speakers selected for this study will be discussed in chapter 3.

When a distinction is present it may be smaller in younger speakers due to the change in height of the STRUT vowel as described by Fabricius (2007).

2.5. The TRAP-BATH distinction

The TRAP-BATH distinction was a split that occurred in the mid-eighteenth century, and is often generalised as lengthening of the TRAP vowel in pre-fricative (specifically voiceless fricative) and pre-nasal position. The resulting (BATH) vowel is usually considered to be the same as the PALM (and START in non-rhotic accents) vowel, which is actually a low back vowel, as well as longer than the TRAP vowel. The details of the distinction will be discussed further below.

2.5.1. TRAP

The TRAP lexical set is defined by Wells (1982*a*) as the words whose citation form (in RP and General American) have the stressed vowel /æ/. The phonetic realisation is between cardinal vowels 3 and 4, front and unrounded. However, there is evidence that the vowel in RP is lowering, tending towards cardinal vowel 4. Fabricius (2007) shows that in speakers born between 1945 and the late 1970s the positions of the TRAP and STRUT vowels have rotated, with the TRAP vowel becoming the lowest open vowel in the younger speakers' vowel systems. This movement is further confirmed by Fabricius (2019) and Lindsey (2019) who show lowering of the TRAP vowel within a system of anti-clockwise change in the entire vowel system.

2.5.1.1. BAD-LAD split

In some British speakers there is variation within the TRAP lexical set. Some words with a voiced consonant following the vowel show lengthening, causing a distinction between pairs such as *bad* (lengthened) and *lad* (not lengthened). According to Wells (1982*b*) the split is lexically specific but Kettig (2016) shows that evidence for this is mixed, with different speakers varying which words are lengthened. Phonetic predictors for lengthening include voiced following segment and fricative following segment (Kettig, 2016).

2.5.2. BATH

The BATH vowel has been qualitatively described at length by writers such as Lass (1976) and Beal (1999). It is defined by Wells (1982*a*) as the words which are pronounced with /æ/ in GenAm, but /ɑː/ in RP. The /ɑː/ realisation is phonetically the same as PALM and START (Kettig, 2016): fully unrounded, between back and central (see section 2.5.3 for more details). Outside of the PALM and START lexical sets, it only occurs in words that have been subject to the TRAP-BATH split (2.5.4).

2.5.3. PALM

Wells (1982*a*) defines the PALM lexical set as the words containing /ɑː/ in RP and /ɑ/ in General American (excluding those with a following /ɹ/ in GenAm, which are under START). It is a fully open unrounded vowel, which occurs in both open and closed syllables. The PALM vowel derived from Middle English /au/ or /a/, with lengthening. The lengthening is the same process as occurred in the BATH words (see above 2.5.2 and below 2.5.4), but according to Wells (1982*a*) an explanation has not been found as to how GenAm has /ɑ/ in *father, palm* but not in *half, halve*.

2.5.4. The TRAP-BATH split

2.5.4.1. History

The TRAP-BATH split is described by Wells (1982*a*) as a 'half completed sound change'. In the mid-eighteenth century, the precursor to RP had [aː] in PALM, [aːr] in START, and [æː] (an allophone of /æ/; Barber et al. (2010)) in BATH words. Therefore, to create the modern day situation, two changes occurred, the phonemic split of TRAP and BATH and the backing of the PALM/START/BATH vowel. The phonemic split was established via lexical diffusion when some lexical items lengthened to /æː/ (which later became /ɑː/). The original lengthening occurred in positions before a voiceless fricative, in a rule that is approximately:

(1) [Open V] → [long V] / _Voiceless Fricative (adapted from Wells 1982*a*, p.204)

Or more formally:

$$(2)\quad \begin{bmatrix} \text{- high} \\ \text{+ low} \\ \text{+ front} \\ \text{- back} \\ \text{- long} \end{bmatrix} \rightarrow \begin{bmatrix} \text{- high} \\ \text{+ low} \\ \text{- front} \\ \text{+ back} \\ \text{+ long} \end{bmatrix} / \_ \begin{bmatrix} \text{+ cons} \\ \text{- son} \\ \text{+ cont} \\ \text{- voice} \end{bmatrix}$$

The pre-fricative lengthening described above applied to both /æ/ and /ɒ/, but only phonologised in the former in RP and in the latter in GenAm. Hence the lexical sets LOT and

CLOTH are realised differently in GenAm but not in RP, whereas the lexical sets TRAP and BATH are realised differently in RP but not in GenAm. For some time there was distinction between LOT and CLOTH in RP, with the presence of [ɔː] in Wells' upper-crust RP (now considered old-fashioned or upper class; Barber et al. 2010), which must be recognised but no longer exists in the majority of speakers (Wells, 1997).

Accounts of when the TRAP-BATH split occurred vary, with Wells (1982*a*) citing it as an eighteenth century split, Lass (1976) places the early stages in the late seventeenth century and MacMahon (1998) describes one eighteenth century writer as 'unusual' in not making the distinction.

As shown above, historically, the most favourable environment for TRAP-lengthening is the pre-fricative environment (Fudge, 1976; Harris, 1989) but it had not until recently been described as an environment that allows for secondary /æ/-lengthening), which is the more modern TRAP-lengthening described as the BAD∼LAD split above. However, Kettig's (2016) data shows the fricative environment as a strong predictor of lengthening in the TRAP vowel of SSBE (southern standard British English) speakers and suggests that these results could point to an explanation of the historical causes of the TRAP-BATH split.

However, there is still variation in words with a following fricative or nasal, with *grass, glass, staff*, and *plant* having lengthened vowels but words such as *gas, asp*, and *rant* still having short vowels. It is suggested that despite the phonologisation of the pre-fricative lengthening, lexical diffusion did not complete (Wells, 1982*a*; Beal, 1999). This scenario could lead to a lexically specific aspect to the rule system controlling the BATH vowel. The current regional variation in England is an isogloss running approximately horizontally east to west, north of Northampton and south of Birmingham and Leicester (Wells, 1982*b*).

2.5.4.2. TRAP tensing in American English

Research into American ENglish (particularly Philadelphia and New York Labov 2001*a*; Payne 1980; Labov et al. 1972) has shown a TRAP tensing system not disimilar to the TRAP-BATH found in English spoken in England. Descriptions. Four systems for the tensing of TRAP (or /æ/ as given in Labovian notation) have been described in accents in the

Eastern United States. The first system displays only lax variants in all contexts. The second system shows a nasal pattern, where the /æ/ is raised and tensed before nasals, but remains lax elsewhere. The third system, which is found in Philadelphia and New York, is a complex system, where tense and lax variants are controlled by a rule system. The Philadelphia systems incudes lexical specificity, following nasals, and following fricatives (a summary can be found in Payne 1980, p. 158). This system leads to phonetic contexts where TRAP is always lax, contexts where it is tense, and contexts where it varies. The New York system is similar, but with more tensing environments. The fourth system is found in the Northern Cities, where all variants of TRAP are raised and tense. (Labov et al., 2013) and Labov et al. (2016) and others show that more recently competing systems have been found in speech communities, with evidence of simplification to the straightforward nasal system. Without lengthy analysis these systems could look like simple lexical diffusion, as the TRAP-BATH split does. Therefore, it is possible that the TRAP-BATH split is controlled by a more complex rule system than simple lexical specificity.

2.6. GOAT Allophony

2.6.1. The GOAT vowel

The GOAT lexical set is the set of words realised with a stressed /əʊ/in RP and /o/ in General American (specified to [o ~ oʊ]) (Wells, 1982*a*). Lexical occurrence of the vowel is generally similar across the two varieties. In RP the vowel is a diphthong, mid central unrounded starting point (similar to /ɜː/), moving to a closer and backer 'lightly rounded' glide (more detail on the RP vowel in section 2.6.1.2). Traditionally named 'long O', this vowel derives from /ɔː/ via Great Vowel Shift, or from /ɔu/ via the GOAT Merger. The GOAT Merger, together with the FACE Merger was part of the Long Mid Mergers (Wells, 1982*a*) that took place around the seventeenth century merging the monophthongal sets including *mane* and *toe* with the diphthongal sets including *main* and *tow*. The words which were monophthongal /ɔː/ in Middle English (e.g. *soap, boat, both*) raised to /oː/ by 1600 (Great Vowel Shift), and then the GOAT Merger brought those that were previously /ɔu/

into this set also. The whole group then diphthongised in 'polite usage' (Wells, 1982*a*, p. 193) at about the beginning of the nineteenth century. This 'Long Mid Diphthonging' (Wells, 1982*a*) added a closing offglide to long mid vowels. Wells (1982*a*) suggests that it began in open syllable words such as *day*, which would match Lindsey's (2019) system of long monophthongs being realisationally closer to diphthongs. The change seems to have occurred in the precursor to RP by the turn of the nineteenth century but likely was present in other accents before 1700 (Dobson, 1968, p. 1020). However, Long Mid Diphthonging did occur across all regions, and Wells (1982*a*) suggests that the Long Mid Mergers are a precondition for that change to take place. Therefore, in varieties such as Tyneside English, which distinguish between diphthongal *pain* and *pane*, the latter will be a monophthong or centring diphthong. This is also evident in the high level of variation found in the Tyneside GOAT vowel by Warburton (2020).

2.6.1.1. Tyneside

Wells (1982*b*) describes the Tyneside GOAT vowel as either a monophthong or a centring diphthong (in symmetry with FACE). In more recent literature the realisation of the GOAT vowel in Tyneside English has been found to be very variable but stratified along classic sociolinguistic boundaries including speaker sex and socio-economic class. The most common variant in the early 2000s was [oː] (Watt and Allen, 2003), which was found across the North of England, from Liverpool to Grimsby, including Sheffield and the rest of Yorkshire (Watt and Milroy, 1999). Those less likely to use this variant included older working class men, who favoured the traditional centring diphthong [ʊə]. Watt (1998) describes four variants: a peripheral monophthong [oː], a centring diphthong [ʊə], a closing diphthong [oʊ] and a central monophthong [ɵː] (The transcription [oʊ] is used exclusively for the closing diphthong but is likely the same as what is commonly transcribed as [əʊ]). The general preferences was for the peripheral monophthong, apart from old working class and young middle class males who only used 31.6% and 44.7% of [oː] respectively. Old working class male speakers showed an approximately equal spread of all variants except the closing diphthong, which is the form found in much of the rest of the UK, including RP. The other trend that Watt notes is that the localised forms [ʊə] and [ɵː] were in approximately complementary distribution with the prestige form [oʊ], and this distribution correlates with male vs. female speakers. While working class speakers avoided the [oʊ] variant almost completely, it was increasing in use in the middle class speakers, with young speakers using it more than old speakers.

The fronting of the monophthongal variant to [ɵː], may be a similar process to the fronting of the [əʊ] variant discussed above. The nationally more common raising diphthong ([əʊ]) was also found in Tyneside (transcribed by Watt and Milroy (1999) as [oʊ]), in exclusively younger middle class speakers, and especially in females. The age patterning found by Watt and Milroy points to a change towards this national prestige variant in Tyneside speakers, led by middle class female speakers. More recently, as found by Warburton (2020), there are still high levels of variation in the GOAT vowel in Tyneside. Warburton found a similar pattern to Watt, perhaps showing more of a move away from the localised [ʊə] and [ɵː] variants, which made up less than 4% of the data, and finding that middle class speakers were the predominant users of the [əʊ] variant.

In older accounts of Tyneside English there is little to no mention of an effect of following /l/ or a a GOAT-GOAL split, though there is some discussion of archaic variants that may point to an effect; this is discussed further in section 2.6.2.1 below. Warburton (2020) finds that the presence of an ɫ coda leads to more productions that sound like the [ʊə] variants, and concludes (in line with Watt 1998) that this is due to the tendency of English speakers to diphthongise monophthongs in pre-/l/ position rather than true realisations of the centring monophthong variant. In the analysis of the GOAT-GOAL split this may create a trajectory difference between GOAT and GOAL-type words in the Tyneside speakers. However, the difference is not expected to be the same as in the CoRP-SE speakers.

2.6.1.2. RP

According to Gimson and Cruttenden's 1994 description of RP, the GOAT vowel is a diphthong that moves from a central vowel to /ʊ/. However, older speakers tend towards a rounder nucleus. The variety referred to as "[r]efined RP" (1994, p. 125) has an unround nucleus, produced further forward in the mouth. Jones's (1972) description has a similar movement, though describes the nucleus as closer to the cardinal vowel [o], with medium rounding, which may be the form of the diphthong that preceded GOAT fronting. Long Mid Diphthonging (above) brought the realisation of the GOAT vowel to [oʊ], (where it remains in most accents of American English). GOAT fronting in RP, an advancement of the nucleus from a [o̩+ʊ] to the more recent [ɜʊ∼ɵʊ] (phonemically transcribed by Gimson and Cruttenden (1994) as /əʊ/), is located by Wells (1982*b*) as occurring in the early 20th century, taking the First World War as a dividing point, adults before 1914 would have the backed variant whereas speakers who grew up after that would have the more central variant. The modern RP GOAT vowel is described by Wells as starting with the same quality as the NURSE vowel, /ɜː/ and therefore can be transcribed as /ɜʊ/, but despite the RP vowel being fairly open, the more common transcription is /əʊ/.

2.6.2. The GOAT split

2.6.2.1. In General

The GOAT split is the (generally allophonic Wells 1982*a*) difference in realisation of the goat vowel in pre-/l/ (usually dark [ɫ]) position. Both realisations are diphthongal. Therefore, since diphthongal realisations of the GOAT vowel only appeared in the late 18th century (Wells, 1982*b*; Sampson, 1985), the earliest the split could have occurred is the early 19th century (Sampson, 1985; Luick et al., 1921).

As mentioned above, in the 19th century the GOAT diphthong fronted and lowered, and allophonic differences appeared in various south-eastern dialects (Altendorf and Watt, 2008), though were particularly prevalent in London (Sampson, 1985). The difference of interest is that as this fronting occurred, the GOAT vowel that appeared before lateral consonants, (particularly coda /l/ - but later literature suggests this is variable, see section 2.6.2.3), was blocked from fronting. The blocking led to different realisations of the vowel in pairs of words such as *goat* and *goal* or *hope* and *hole*, where *goat* and *hope* had a fronted diphthong and *goal* and *hole* retained the back diphthong. Sampson (1985) explains this effect with assimilatory difference: a velarised (dark) [ɫ] preserves a back diphthong.

Gimson and Cruttenden (1994) mention that [ɫ] can affect vowel quality but do not mention the blocked fronting of the GOAT vowel, or subsequent backing in RP. The GOAT split is also not mentioned in their summary of current changes in RP. Jones references the split found in "[s]ome Southern English people" (1972, p. 104), who produce a ɔu preceding a [ɫ] and also refers to the effect of [ɫ] when teaching the diphthong to English learners, referring to the difficulty of producing a front diphthong when followed by [ɫ].

As mentioned above there is little reference to the GOAT split in the literature on Tyneside English, however, some archaic realisations may point to an effect of coda /l/. Some GOAT words are historically derived from words realised as [aː], and this is retained in some lexical items such as *know* (particularly in the phrase *y'knaa*) and *snow*. However, older speakers can also have a diphthongal realisation, [aʊ], in words such as *old, cold, soldier* and *shoulder* (Watt and Milroy, 1999). The development of a diphthongal realisation, beginning with the low back vowel could point to the coda /l/ causing an offglide. This would be defended by the difference found in the Sheffield GOAT vowel by Stoddart et al. (1999). In Sheffield, similar patterning to Tyneside is found with the majority variant being a monophthongal [ʉː], but there is a split occurring with GOAL words more likely to take a diphthongal realisation such as [oʊ ∼ ʉʊ], [əʊ] or [aʊ]. However, /l/ is noticeably clearer in north-eastern varieties, including in word-final position (Wells, 1982*b*), and consequently the dark /l/ environment that triggered the GOAT split elsewhere is less likely to be present. Instead, there is more of a tendency to find the [ʊə] variant in pre-/l/ position, which is perhaps due to the tendency of Tyneside speakers to append an offglide to the monophthongal [o] variant Watt (2000).

Due to the variation in disyllabic contexts (see section 2.6.2.3), some (Wells, 1970, including) consider the GOAT split a phonemic split rather than merely allophonic variation. This will be discussed further below.

2.6.2.2. The GOAT Split in RP

As was discussed above, the GOAT vowel underwent fronting in the 19th century, and in most varieties of British English, this occurred in all contexts except before /l/.

However, RP generalised the fronting, and hence did not have allophonic variation between words such as *hope* and *hole*, or *goat* and *goal* Wells (1982*b*); the GOAT vowel was a fronted diphthong in all contexts. According to Sampson (1985), this generalisation was unique to RP, and hence was a truly non-regional feature.

Maintaining a front vowel before an /l/ (particularly velarised [ɫ])is both cross-dialectally and cross-linguistically unusual (Jansen 2017; Baranowski 2017 and many others), because the velar gesture of the [ɫ] impacts tongue position in the preceding vowel, causing retraction which is then phonologised (Bermúdez-Otero, 2015). Therefore is it not surprising that recent studies, including (Hannisdal, 2006) and (Hughes et al., 2012) observe the presence of a back allophone in pre-/l/ position in RP. Hannisdal (2006) takes 30 news presenters from three different news stations, and treats these as exemplary of modern RP. This study was the first investigation of GOAT allophony in RP, and is quantitative (though it codes the allophony auditorily rather than acoustically). The study defines the GOAL subset as those words where "stressed /əʊ/ is followed by final or preconsonantal /l/" (2006, p. 154). The fronted diphthong is described with a central unrounded onset [ɜ∼ə] and the backed one with a back round onset ɒ∼ɔ. It also suggests that while the nucleus of the diphthong is the primary place of the difference, the glide can also have a change in quality, appearing as [ʊ̹̈∼ʉ̹] rather than [u∼ʊ∼o]. The phonetic conditioning of the appearance of the allophone is categorised as the backed allophone only occurring before “non-prevocalic /l/". She finds that 24 speakers had GOAT allophony, 4 were variable, and 2 had no allophony, suggesting that this feature is “becoming firmly established in modern RP" (2006, p. 155). Hannisdal also observes GOAT allophony in regional varieties of English and hence justifies its inclusion in a non-localisable RP. The two speakers with variable allophony did not show any pattern, and are suggested to either be in the middle of a change, or are not natively RP and are shifting towards traditional RP by suppressing the back allophone. Hannisdal's study did not reveal gender differences, implying social neutrality. However, there may be some formality effect, with all the speakers from Sky News (which Hannisdal considers to be more informal than BBC News) having GOAT allophony and 4 of the 6 speakers who have no, or variable allophony are from the BBC World channel, which is considered more formal.

2.6.2.3. In Disyllabic Contexts

The effect of the disyllabic context on the realisation of the GOAT vowel is described differently in different sources. Some of these merely note the difference (for exampleJones 1972), who observes a distinction between *bowl* and *bowling*, the former being backed), whereas others make steps towards explaining the variation (for example Sampson 1985)

Jones (1972) notes that "[s]ome Southern English people" have an allophone ('subsidiary member') in pre-[ɫ] position. This allophone is described as retracted and is only in contexts where /l/ is in the coda ([ɫ] is conditioned by word position Turton 2014*a*); examples of the retracted realisation include *bowl* and *bolt*. Examples of contexts where dark /l/ and GOAT retraction are not found include *bowling*, demonstrating the effect of syllabification and *roll it*, which shows that cross-word syllabification also affects the variation.

Sampson's 1985 study investigates the GOAT split, as discussed above, and also notes that there is variation in contexts where /l/ is followed by a vowel. It is also suggested that the /l/ can be clear and still produce the backed allophone. An explanation given for the variation begins by considering the morpheme boundary (e.g. *holy* is front and *wholly* (*whole* + *-y*) is back) but this does not hold for words such as *Polish*, which invariably show the front allophone. Possible explanations for this include variation based on suffix identity rather than merely the presence of a morpheme boundary, or a suggestion that some words that show the fronted allophone are etymologically bimorphemic but speakers analyse them as monomorphemic. However, it is difficult to come to a phonological analysis based on these explanations. (Sampson, 1985, p. 293-4). Sampson's phonological analysis begins with treating the two variants as allophones, which would normally be handled by phonetic rules. In most theoretical frameworks these are at the end of the phonological component (Zsiga, 2020) but are relatively understudied. It is also assumed that items such as *Polish, solo*, and *polar* are interpreted as described above, as if they do not contain a morpheme boundary. Sampson's first two suggested rules are shown in figure 2.3 (transcriptions as in the original paper, not updated to modern conventions).

Importantly, rule 1 of this analysis does not include that the retracted GOAT vowel is caused by the effect of a following dark /l/, this is not covered till rule 2, which, as put by

$$\textbf{R1} \quad \bar{\text{o}} \quad \rightarrow \quad \left\{ \begin{array}{l} \text{ɔ̈ʊ} \ / - \text{l} \left\{ \begin{array}{c} \text{C} \\ + \end{array} \right\} \\ \text{əʊ} \ / \ \textbf{elsewhere} \end{array} \right\}$$

$$\textbf{R2} \quad \text{l} \quad \rightarrow \quad \text{ɫ} \ / - \left\{ \begin{array}{c} \text{C} \\ \# \# \end{array} \right\}$$

Figure 2.3.: Rules controlling the GOAT-GOAL split according to Sampson (1985, p. 294) (1)

$$\textbf{R1} \quad \text{l} \quad \rightarrow \quad \text{ɫ} \ / - \left\{ \begin{array}{c} \text{C} \\ + \end{array} \right\}$$

$$\textbf{R2} \quad \bar{\text{o}} \quad \rightarrow \quad \left\{ \begin{array}{l} \text{ɔ̈ʊ} \ / - \text{ɫ} \\ \text{əʊ} \ / \ \textbf{elsewhere} \end{array} \right\}$$

$$\textbf{R3} \quad \text{ɫ} \quad \rightarrow \quad \text{l} \ / - + \text{V}$$

Figure 2.4.: Rules controlling the GOAT-GOAL split according to Sampson (1985, p. 295) (2)

Sampson could be considered “putting the phonetic cart before the horse" 1985. The second analysis described by Sampson (figure 2.4) is more complex, and, as admitted by the author, is a Duke of York Gambit (generally considered a problematic form of explanation (Pullum, 1976)) in which a surface form is the same as an underlying form but created via another form, in the method of A→B→A. Sampson does not reach a descriptive conclusion without needing to describe exceptions, and suggests a governing principle of ‘vowel projection’ where any lexical item with the GOAT vowel before [ɫ]“which can appear as an independent word will normally have the vowel quality found in that independent word projected onto the pronunciation of its derivational offshoot". Sampson’s assessment does not include perceptual data or the potential for change in speakers’ processing of bimorphomic words. It is possible that words such as *Polish* and *solar* are no longer processed as bimorphemic and hence do not require the above exception.

Of 13 speakers in Hannisdal’s 2006 study, 7 showed what is termed ‘morphological regularisation’. That is, producing the back allophone in disyllabic (e.g. *bowler*) as well as monosyllabic contexts.

2.6.2.4. Summary of the GOAT vowel

As seen above there is a high level of variation in the GOAT vowel, both regionally and socially. For the sake of consistency, further analysis and discussion will assume that the underlying representation that speakers maintain is a front diphthong or vowel, with the backing environment causing the change. For the CoRP-SE speakers this would be /əʊ/, for the CoRP-NE speaker it could be /əʊ/, /oʊ/, or /o/, and for DECTE speakers it is likely to be /o/ or /ʊə/, but could be /oʊ/ or /əʊ/.

2.6.3. The Life Cycle of Phonological Processes (The GOAT split)

2.6.3.1. Variation in the GOAT split

As discussed above, if the GOAT split as seen in the words *goat* and *goal* or *hope* and *hole* exists in RP speakers, it has arisen within the last 30-40 years. There is a large amount of variation in the change, and various theories have been put forward to explain this. These accounts include, phonemic and allophonic categorisations as well as more complex processes accounting for the disyllabic contexts. This section will give an overview of these theories before focussing on Bermúdez-Otero's (2007; 2015) *Life Cycle of Phonological Processes* as a proposed explanation for the variation found.

Allophony within the GOAT lexical set, conditioned by coda /l/, is attested by various writers, but Wells (1982*b*) suggests a phonemic split, rather than just allophony, particularly if the analysis is focussed on London speech (from which RP possibly gained the split). This analysis is based on the interaction between the retracted diphthong and /l/-vocalisation, which creates minimal pairs such as *so*∼*soul* and *row*∼*roll/role*. However, while /l/-vocalisation is widespread in the South East, it has not been reported in RP. Sampson (1985) disagrees with the phonemic categorisation, utilising disyllabic contexts. The only disyllabic minimal pair is *holy*∼*holey*, and Sampson argues that this is not enough to demonstrate phonemic contrast. However, this account does not include /l/-vocalisation. /l/-vocalisation is very common in London, but not advanced in many RP speakers, so likely will not affect the speakers under study here. These two accounts demonstrate the interactions between mono and disyllabic contexts in the the spread of the GOAT split. The rest of this section will consider the source and progress of sound change according to the life cycle of phonological processes (Bermúdez-Otero, 2007, 2015). Instead of contrasting the variation found in monosyllabic and disyllabic contexts, the life cycle framework takes a sound change through the grammar, hence allowing for and explaining variation between these contexts.

2.6.3.2. Diachronic Sound Change

Ohala (1981) presents the listener as the source of sound change, by either failing to apply a reconstructive rule, or by reconstructing that which does not need to be. This is described by (Bermúdez-Otero, 2007, p. 1) as a speaker and listener (e.g. a child and a caregiver), not solving the coordination problem posed by speech, where the speaker must produce a phonetic stimulus, which the listener can interpret as the intended phonological representation, by deciding which parts of the incoming stimulus are signal and which are accidental noise.

When sound changes occur they are always gradual in one or more of the sociolinguistic, phonetic or lexical domains. Gradual phonetic change is where a change is not abruptly from one phonemic category to another but for example gradual stop lenition or a gradual movement of a vowel. Gradual lexical change is where not all words are affected at once but the change spreads gradually through the lexicon. These changes are often divided into two categories First, Neogrammarian sound changes: these are are regular but gradient, with gradient phonetic effects but abrupt lexical changes. Secondly, classical lexical diffusion: this is phonetically (and hence phonologically) abrupt but lexically gradual (Bermúdez-Otero, 2007). This division is supported by Kiparsky's theory of lexical phonology but has been challenged by Bybee (2001), who claims that all sound changes are both lexically and phonetically gradual. This view affects discussion down to the structure of the internal grammar, and whether or not lexical representations contain gradient phonetic details. However, as further discussed by Bermúdez-Otero (2007), the number of potential conditioning factors (phonological context, morphological context, sociolinguistic factors) which may affect the appearance of an innovative variant is so large that determining whether a particular change is regular or diffusing can require large data sets. Despite the GOAT-GOAL split being a recent (and potentially ongoing Hannisdal (2006)) change, it is unlikely that the data set in this project will be large enough to make a conclusive decision on the status of the sound change. However, the morphological conditioning suggested by Wells (1982*b*), Sampson (1985), and Hannisdal (2006) suggests classical lexical diffusion, and hence the framework of generative phonology and Bermúdez-Otero's (2015) life cycle of phonological processes will be used to understand the variation within the split.

2.6.3.3. Generative Phonology

Within generative phonology there is a distinction between phonetic and phonological rules. This is situated within a modular feed-forward module. That is a structured architecture of the grammar that depends on (a) lexical and phonological discreteness and (b) modularity. (a) means that lexical and phonological representations cannot encode fine phonetic detail and their attributes (any phonological property) have discrete values. (b) means that phonetic rules cannot refer to lexical representations, i.e. each level of the grammar cannot be 'seen' by non-adjacent modules, the interfaces are between phonetics and phonology, phonology and morphology, hence morphological structure cannot affect phonetics. According to lexical and phonological discreteness, phonetically gradual changes can only take place through change in the phonetic rules that assign realisations to phonological categories; but according to modularity there cannot be any lexical influence on these changes.

The principles above allow for both the Neogrammarian and classical lexical diffusion-type changes discussed above (section 2.6.3.2). From these principles Bermúdez-Otero (2007) proposes three possible modes of implementation, shown in table 2.2

Mode of implementation		**Possible?**	**Innovation in what component of the grammar?**
phonetic dimension	lexical dimension		
abrupt	gradual	Yes	lexical representations
abrupt	abrupt	Yes	phonological representation
gradual	abrupt	Yes	phonetic rules
gradual	gradual	No	N/A

Table 2.2.: Modes of implementation predicted by the classical architecture (Bermúdez-Otero, 2007)

2.6.3.4. The Life Cycle

Bermúdez-Otero's (2015) life cycle of phonological processes approaches sound change is gradual or abrupt over time by considering the internal structure of the grammar. As stated above (section 2.6.3.2), Bermúdez-Otero (2007) argues that sound changes are always gradual in one or more of the sociolinguistic, phonetic or lexical domains; in the sociolinguistic domain this can include idiolectal variation and generational change. For a change to be phonetically abrupt features must be gained or lost (e.g. [+front] $\rightarrow$ [-front]) in all words with the identical context.

All of this change takes place within the modular feedforward model described above, and also, based on Lexical Phonology (Kiparsky, 1982), assumes stratification within the phonological module (relevant within stage III, reanalysis), as described below.

1. Stem Level: phonological rules apply to the stem

2. Word Level: phonological rules apply to words (after addition of any inflectional morphology) but not accross word boundaries

3. Phrase Level: rules can apply across word boundaries.

The modular feedfoward model lends itself well to a clear understanding of the characteristic life cycle of historical sound changes (Baudouin de Courtenay, 1972, translated from 1895) that has long been accepted. A theory neutral (with respect to how the phonology functions, this has been implemented by others using both Optimality Theory and Lexical Phonology) summary of the life cycle is as follows (Bermúdez-Otero, 2007):

- *Phase I:*
  Phonologisation - the addition of a new phonetic rule to the grammar, due to some phenomenon causing a new pattern of phonetic implementation.

- *Phase II:*
  Restructuring - (also known as stabilisation Turton 2014*b*; Ramsammy 2015) the new sound pattern becomes categorical, restructuring the phonological representations

- *Phase III:*
  Reanalysis (or domain narrowing Ramsammy 2015) - categorical patterns can change, often becoming sensitive to morphosyntactic structure, reducing their domain of application.

- *Phase IV:*
  Morphologisation/Lexicalisation - sound patterns no longer being phonologically controlled.

Within a modular understanding of the grammar of language, this structure shows a process moving from the phonetics to the phonology, (Zsiga, 2020), and deeper into the grammar; it displays as Neogrammarian sound change (regular and gradient).

The life cycle allows for the variation in the GOAT vowel found in disyllabic contexts by taking into account the effect of morphological structure and suffixes. The sections below give more detailed explanation of these processes. It is worth noting that Phase II (Restructuring) is also often known as Stabilisation, and Phase III (Reanalysis) is often split into two stages of domain narrowing, henceforth these will be referred to as IIIa and IIIb (etc.). Below the phases of the life cycle are explained in more detail, referring to the case of post-nasal stop deletion as laid out by Bermúdez-Otero (2006, 2011) and Turton (2014*b*), and with examples of how this might apply to GOAT backing.

	finger	sing-er	sing it	sing
stage 0	[ŋg]	[ŋg]	[ŋg]	[ŋg]
stage 1	[ŋg]	[ŋg]	[ŋg]	[ŋ]
stage 2	[ŋg]	[ŋg]	[ŋ]	[ŋ]
stage 3	[ŋg]	[ŋ]	[ŋ]	[ŋ]

Table 2.3.: Life Cycle of Phonological Processes, adapted from Bermúdez-Otero (2011)

2.6.3.4.1. Phase I: Phonologisation

The changes that can be understood under the life cycle begin at stage 1 of table 2.3 with a phonetic effect that is beyond conscious control of the speaker (Turton, 2014*b*; Ramsammy, 2015), for post-nasal stop deletion this is likely to have been some form of gradient lenition of the velar stop. In order to become a phonological rule this must then undergo a change in status (Ramsammy, 2015; Anderson, 1981) where it is reinterpreted, from being controlled by articulatory, or phonetic processes, to being systematic and under control of the grammatical system.

For GOAT allophony, this would be a coarticulatory effect, where the retraction of the tongue for following [ɫ] affects the frontness of the vowel

2.6.3.4.2. Phase II: Restructuring(/Stabilisation)

At this point a gradient phonologised process becomes a categorical phonological rule, another change in status. This would occur by a learner hearing a pronunciation that has undergone gradient change, and posits the resulting realisation as their underlying form. The crucial factor in this stage is that the rule applies at phrase level; in *sing it*, the following word is 'visible' to the segment to which the rule is applying (Turton, 2014*b*) so the [g] can resyllabify to the beginning of the next syllable . This has an impact on syllabification processes, and hence which segments are in an appropriate position for the rule to apply, giving rise to stage 1 of table 2.3 For a phrase such as *roll it*, the /l/ would be re-syllabifying to the beginning of *it*.

2.6.3.4.3. Phase III: Reanalysis

The process is now under the control of categorical phonology but only applying at the phrase level, so is not sensitive to the morphology. At this point it is possible for the rule to move up within stratified phonology.

2.6.3.4.3.1. IIIa: Domain Narrowing 1

The first phase is that the domain of application of the rule gets smaller as the pattern is reanalysed (Ramsammy, 2015) becoming sensitive to the edges of grammatical words. This occurs when a learner is exposed to tokens where the rule has applied at phrase level but reinterprets the final realisation as the product of a word level rule. The rule now applies to sing, preceding cross-word syllabification, and produces stage 2 in table 2.3.

2.6.3.4.3.2. IIIb: Domain Narrowing 2

A rule applying at word level means that it is now sensitive to grammatical word boundaries and morphological operations that occur at the word level. The next stage of narrowing again shrinks the domain of application, to the stem level, meaning that the rule applies before any inflectional morphology is added. This occurs by restructuring of the word-level input modifying the phonological processes (Ramsammy, 2015). This reaches stage 3 in table 2.3, where the rule applies to *singer* before the -er suffix.

The domain narrowing processes described above take place diachronically, from speaker to learner. Therefore, it can be inferred that a rule which has reached stem-level of application is diachronically advanced (Turton, 2014*b*), and also that speakers with higher or lower levels of application are more advanced or further behind in a change respectively.

2.6.3.4.4. Phase IV: Morphologisation/Lexicalisation

According to Ramsammy (2015) one final stage in this process is possible. When a rule is applying at the stem level the output realisation can be interpreted by a listener as underlying, causing lexicalisation. At this point the rule is no longer producing alternations, it is not phonologically active and the life cycle has reached its end.

The key part of the above process for diachronic understanding is Reanalysis, where a process reduces its domain of application. This can be seen in GOAT backing in disyllabic contexts (Sampson, 1985), where for some speakers both *Roland* (monomorphemic) and *rolling* (bimorphemic) have a front GOAT vowel, showing that the backing rule applies at word level, but other speakers have a front vowel in *Roland* but a back vowel in *rolling* showing that the domain of application has reduced to stem level.

2.6.3.5. Summary of the use of the life cycle

While the life cycle is not unopposed in literature on sound change, for example Strycharczuk and Scobbie (2017) and Strycharczuk and Scobbie (2017) suggest that phonetic rules can apply at the morphological level, it has strong explanatory power, and will be a helpful too to understand the progress of the GOAT split in the speaker groups under study here.

2.6.3.6. A note on syllabification

Discussion so far has assumed onset-maximal syllabification at all levels, as supported by Kiparsky (1979), and shown experimentally by Fallows (1981), and will continue to do so in line with Kiparsky's (1979) sonority hierarchy in the onset, and other work describing syllable based sound variation (Turton, 2014*b*; Bermúdez-Otero, 2011). Ambisyllabicity (proposed by Kahn 1976, opposed by Kiparsky 1979; Fallows 1981; Nespor and Vogel 1984) will not be used.

Onset maximal syllabification at the phrase level (across word boundaries) will also be used. This leads to consonants such as the /g/ in *sing it* or the /l/ in *roll it*, which are in the coda at word level, are resyllabified to the onset when followed by a vowel-initial word. Minkova (2003) provides evidence for Middle English developing onset maximal syllabification at the phrase level, and this is supported in further work such as Kiparsky (1979); Bermúdez-Otero (2011).

Chapter 3.

Methodology

3.1. Introduction

The project approaches RP from a sociolinguistic standpoint (Fabricius, 2002*b*), analysing the sociophonetics of three sets of speakers via social and phonological variables, including private education and region to gain and understanding of potential patterns of variation and change. The speakers in focus (henceforth CoRP-NE) are privately educated (the background commonly cited as the source of RP (Fabricius, 2000*b*)) but from the North East, an area less commonly discussed in the literature. They can be seen in the yellow box below. These are understood in the context of two comparisons: first, in comparison to speakers privately educated in the South East (blue box below, henceforth CoRP-SE), who are the most common focus of RP study, to investigate potential non-regional patterns; secondly, in comparison to state educated speakers from the same region (speakers taken from DECTE (Corrigan et al., 2012) - green box below), to understand how they vary from the local accent, and whether differences found in the first comparison are consequences of regional influence.

Dividing these

The rest of this chapter will outline the methodological approach of the project, and also the specific methods used to approach speaker selection, data collection, and analysis.

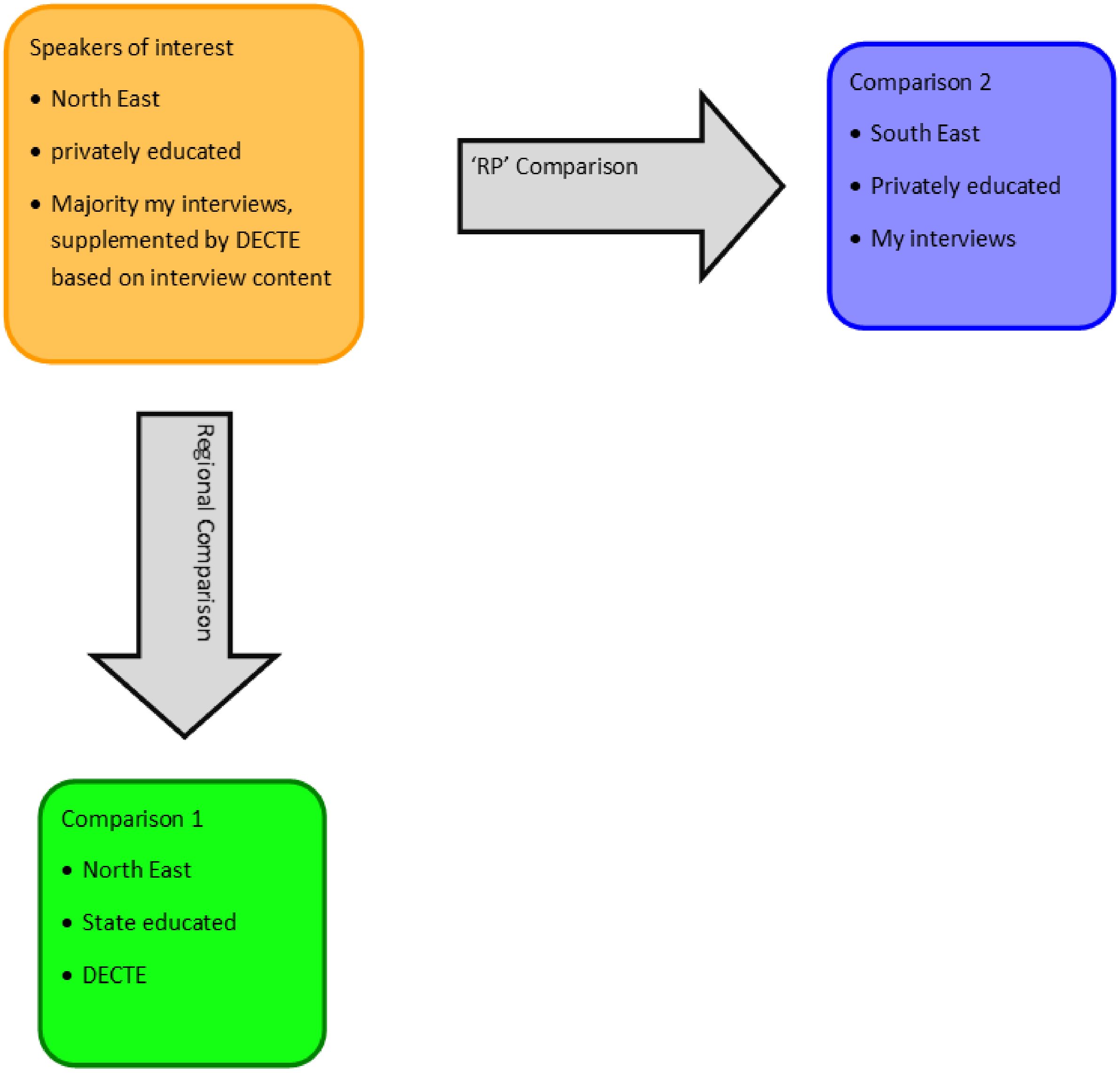

Figure 3.1.: Speaker groups and directions of comparison to understand regional vs. non-regional patterns

3.2. Methodological Background

3.2.1. Sociolinguistics

Modern variationist sociolinguistics is generally considered to have begun with Weinreich et al. (1968), and continued through many studies by Labov (to the extent where it is sometimes known as Labovian sociolinguistics). For many years, the practices and methods, particularly surrounding data collection, were passed on via word of mouth between researchers, including via the supervision of doctoral students (Tagliamonte and Mesthrie, 2006). This tradition has since been recorded by Tagliamonte and Mesthrie, setting the stage for following generations of variationist researchers. Variationist analysis combines techniques from a variety of other disciplines, including statistics, anthropology, and linguistics, aiming to investigate the use and structure of language. This tradition aimed to contribute to the wider field of linguistic study by formulating a model of language that both allows for, and actively includes, variation and change.

3.2.2. Sociophonetics

The term ‘sociophonetics’ is first used to describe a subsection of the field of sociolinguistics in Deshaies-Lafontaine’s (1974) PhD thesis. This thesis is described by Foulkes and Docherty as “squarely within the emergent field of Labovian or variationist sociolinguistics" (1999, p. 703), and demonstrates emphasis on the phonetic variation found in language, rather than syntactic, lexical, or other variation that had previously been popular foci of study. According to Foulkes and Docherty, the aim of sociophonetics is to identify and explain where and how socially structured variation in speech occurs, including how it is learnt, cognitively stored, evaluated, and processed in both speaking and listening; the term ‘sociophonological’ has also been used to describe work with the same intent. The distinction between the phonetic vs. the phonological is clarified by Di Paolo and Yaeger-Dror (2011), by their description of the main aim of sociophonetic research, as considering phonological variation in order to understand how it relates to sound change and understanding its interactions with salient social categories within a speech community.

This project is placed within the methodology described above ; it uses structured pools of subjects to understand phonological variation in fine phonetic detail, including how these differences function socially.

3.3. Data Collection

In the social sciences, research usually aims to use truly random data samples, where every unit of the population under study has an equal chance of being chosen. In order to produce random sampling, it needs to be possible to delineate the boundaries of the population under study. However while this is theoretically ideal, it is impossible to completely attain due to the need for participant consent, and even as an aim it is often practically impossible (Tagliamonte and Mesthrie, 2006).

According to Tagliamonte and Mesthrie, it has been demonstrated that it is possible to account for linguistic variation without strictly random sampling; samples used in linguistic studies that were technically too small to be representative have been shown to account for language variation. They suggest using stratified random sampling, which is the same idea as Foulkes et al.'s (2010) structured pools, to make inferences about a population. Since this thesis aims to understand the nature and location of the speech community of RP speakers, including who this speech community is comprised of, it is impossible to list and randomly sample the community. While the study has clear eligibility criteria (outlined below in section 3.3.1), it is impossible to create a list of every person who is eligible and hence sample them randomly. This study is also a part of wider work that is endeavouring to understand who and where the speech community of RP speakers actually is. Therefore, stratified random sampling, as suggested above, was used.

3.3.1. Speaker Recruitment

A long-standing problem in studies of RP is the circular nature of defining the speaker population and the accent. Often studies state that their speakers are RP speakers by describing features in their speech, but have no social or demographic reasoning. The other type of study (e.g. Hannisdal 2006, Bjelaković 2016) uses a narrow speaker pool, often newsreaders, due to the historical requirements placed on news persons by the BBC and associated organisations. Neither of these approaches is sufficient for this study. The narrow, newsreader-type speaker pool would not be able to answer the social aims of the study, due to stylistic variation caused by the formal environment and the lack of regional information. The approach of defining speakers by accent features would lead to a circularity problem, wherein looking for variation would be impossible if the features of the accent were pre-defined.

One of the most in-depth studies of RP speakers that exists is Fabricius (2000*b*) (and later associated papers). Fabricius focusses on speakers from the upper middle class (based on Wells 1997) and defines this partly by educational background. This is based on the historical association between RP and the public school system, elaborated on in chapter 2. Fabricius' methodology developed to a very narrow speaker sample after advertising for speakers who had been to an independent school but then citing that she was "was dissatisfied with especially the male speakers from the 'Rest of England' area, some of whom had been to selective grammar schools or 'non-public' independent schools and turned out to have localisable non-southern features in their speech." (2000*b*, p. 74). This approach is helpful for decisions on defining the speaker pool because such features appearing would help identify the borders of non-regional speech within the UK's social structure. Therefore, recruitment was based the broad definition of speakers who have been to some form of fee-paying (private) school, but also included collection of detailed educational information.

As discussed in chapter 2, class is socially constructed and difficult to define. I have chosen to use this specific educational category for two reasons. First, since class is socially constructed, it is based on a number of different sociological factors, and rather than categorising individuals broadly into pre-determined categories that don't have a clear reality, understanding how the different factors are combined socially, and ultimately how they affect language variation is a more helpful approach for sociolinguistic study. Secondly, since the private educations system is the original source of RP (section 2.2), investigating its impact into modern accent variation is an important step in understanding the landscape of social variation in England. This approach is not using education as a proxy for social class but investigating one factor that feeds into the construction of the sociological phenomenon that is social class, and specifically one that has a direct impact on what has been considered a non-regional sociolect.

These speakers who have been to private school made up CoRP-NE, the 'speakers of Interest', shown in orange in figure 3.1, and the speakers for the 'RP comparison', shown in blue. To make up the 'regional comparison', shown in green, interviews from the Diachronic Corpus of Tyneside English (DECTE, Corrigan et al. (2012)) were taken as a baseline comparison of non-privately educated Tyneside English, to aid in identifying what constitutes 'local features'. These speakers were selected from the corpus based on those that had the most complete information surrounding social class variables, such as occupation.

In general, the approach of social sciences is to aim for a truly random sample of the population. However, for linguistics studies this is difficult and often impossible. In order to gain a random sample it is necessary to delineate the boundaries of the population, which would be impossible in this study. The study has clear eligibility criteria but it is impossible to create a list of the people that fit this and randomly sample them (Halfacre, 2018).

According to Tagliamonte and Mesthrie (2006), it is possible to account for language variation using samples that are technically too small to be representative. Tagliamonte and Mesthrie suggest using stratified random sampling, (similar to Foulkes et al.'s (2010) structured pools), to make inferences about a population. Sentence referring to table]. The participants in this study were recruited by a friend of a friend and snowball sampling method. I used personal contacts, friends of friends, and asked participants to recommend people they know who would be willing to take part. I also contacted local independent schools and asked them to reach out to alumni communities. Unfortunately the final method only yielded two contacts, neither of whom were currently based in Newcastle so could not be interviewed. My familiarity with the participants varied from family members to those I'd never met before. A full table of the participants can be found in the data folders in the github repositories listed in Appendix A.

3.3.1.1. North East Private Schools

At this point it is worth pausing to comment on the types of private school found in Newcastle and the surrounding area of North East England. Historically, schools in the UK could be approximately divided into the original six public schools HM Government (1868) (majority boarding), other fee-paying 'private' (or independent) schools (which varied in price, culture, academic goals, and proportions of boarding pupils), and state (government-funded) schools. Some of the private schools were far more like the public schools in culture and boarding numbers (particularly the cathedral schools and the countryside boarding schools). In the present day, these schools, the original public schools, and some of the other private schools make up The Heads' Conference (HMC), "a thriving, proactive association of heads at some of the world's leading independent schools" (Conference, 2023), which is now the only functional definition of a 'public school'. This leaves the modern divisions as approximately HMC schools, private schools (both of which may have boarding), and state schools. HMC schools exist across England but are more concentrated in the South East, and the boarding schools are even more so. In Newcastle there were many small private schools, a number of which have closed or merged over the last twenty-thirty years (the most recent being the merging of Central Newcastle High School and Newcastle Church High School to form Newcastle High School for Girls). The schools in Newcastle are now a mixture of HMC and other private schools; there are no boarding schools in Tyne and Wear, and very few in the surrounding North East counties (Northumberland and County Durham). Due to the influence of specifically the boarding environment on the development of the RP accent (see discussion in chapter 2, the speakers in this study were categorised into boarding, private, and state for primary, secondary, and sixth form (data folders in the github repositories listed in Appendix A). However, due to the size of the sample, it was not possible to do fine-grained analysis within the CoRP groups based on these categories.

3.3.2. Sociolinguistic Interviews

Sociolinguistic interviews are a standard method of collecting speech data in sociolinguistic studies (Tagliamonte and Mesthrie, 2006), and have long been considered the best compromise between acoustic quality and casual speech. The basic format is an informal conversation between the researcher and participant, that aims to access the vernacular (Krug and Sell, 2013). The participant knows that they are being recorded but the researcher aims not to highlight this and use various conversational techniques to put the participant at ease and reduce their focus on their speech. Part of this process is asking questions that elicit narratives, putting the speaker into what Tagliamonte and Mesthrie (2006) term 'storytelling mode'.

Chapter 3 of Tagliamonte and Mesthrie (2006) provides an example of 'interview modules', with questions separated into sections that are based on topics, moving from least to most personal. These are designed to put the participant at ease and encourage them to talk; the questions put the participant in the position of providing information to the interviewer so that they are comfortable and do not feel in a lower position of power. For this project, the modules were re-written to suit the researcher's style of speech and to make the questions appropriate for the demographic of the participants.

The participants were told that the research was related to the interactions between language, social class, and education; the information letter was clear and correct but did not overtly highlight the linguistic and phonetic nature of the project so as to prevent the participants becoming hyper-aware of their speech. Since the initial recruitment criteria required that the participant had been privately educated, it was found that this was a natural place to start the conversation and in practice the majority of the interviews did not require more than the first few pages of the interview schedule. The original interview modules have many questions relating to area and neighbourhood, but many people who have been privately educated have been to school outside of their neighbourhood, or have boarded further away. This meant that they tended to have more social and emotional ties to their school than their neighbourhood. Through the process of interviewing it was found that for many of the participants asking about school life, and for stories and experiences related to that often yielded more natural conversation. All of the interviews were recorded with a Zoom H4n Pro Handy Recorder, using the inbuilt microphone.

The first set of interviews (speakers 001-010) were recorded in the Summer of 2018, as part of the researchers' MA dissertation (Halfacre, 2018) and included six speakers educated in the South East and four in the North East. Sufficient demographic information was obtained to make these comparable with the rest of the data. The second set of interviews were recorded between June and December 2019. This set included eight speakers educated in the North East and two in the South East, bringing the total to twelve North East speakers and eight South East speakers. These recordings make up a small corpus named 'Corpus for investigation of Received Pronunciation' (henceforth CoRP). Supplemental interviews as needed were planned for Spring 2020 but postponed due to the COVID-19 outbreak and the following social distancing guidance and travel ban.

3.3.3. Demographical Information

Demographic information was collected from the speaker using a form after the interview. This included educational information from both the participant and their parents, and occupation information for both (Baranowski, 2013). A high level of detail on education, including how many years spent in state vs. private schooling was discussed here, if not covered in the interview already. The data from this form was joined to the whole data set before analysis (Wickham et al., 2022).

3.3.4. Data Extraction

The interviews were transcribed by either trained students, the author, or using the automated transcription software CLOx (Wassink et al., 2018), which creates a rough orthographic transcription by running the sound file through Microsoft's Speech Service API Client Libraries. These were then edited by the author. All transcriptions were checked by the author in order to maintained comparability. The transcriptions were created in ELAN (Max Planck Institute for Psycholinguistics, 2008) producing time-aligned text in breath groups, which were then exported to a Praat TextGrid. The 10 interviews collected in 2018 had already been aligned with FAVE-align (Rosenfelder et al., 2014). FAVE creates a phonemic transcription using the CMU Pronouncing Dictionary (Weide, 1998), and asks for manual input for any words that are not recognised. The programme then uses an acoustic model to match the transcription to the sound file. The more recent interviews and DECTE interviews were aligned using the Montreal Forced Aligner (henceforth MFA) McAuliffe et al. (2017) version 1.1. MFA works in a similar way to FAVE, using a pronouncing dictionary and acoustic model to match the transcription to the sound file, producing an aligned Praat TextGrid. The MFA pronouncing dictionary used was an adapted version of the MFA librispeech dictionary (based on the CMU pronouncing dictionary, with American English phonemes) with alternatives to the rhotic pronunciations included by using a find and replace function with the following regular expression commands and adding all resulting pronunciation variants (including coda /r/) to the original dictionary. This was important because including /r/ phoneme in words such as *farm*, where it is not produced English English will mean that the whole vowel that is present in the recording will not be measured. It was less important to adjust the vowels (e.g. American English does not have low back rounded vowel, such as in LOT) because these were later recoded to lexical sets in the data processing step so no information was lost.

- “ER(\d) ([AEIOU])" → “AH$1 R $2"
- “ER(\d)" → “AH$1"
- “(\d) R ([BCDFGHJKLMNPSTVZ])" → “$1 $2"
- “R\n" → “\n"

Once alignment was complete, measurements were extracted using a version of FAVE-extract (Rosenfelder et al., 2014), adapted by Warburton (2020) to give normalised F1 and F2 measurements at 10% intervals along the vowel length as well as the usual single measurements. The single measurements are at one third duration for /a,æ,ɑ,ʊ,ʌ/ and halfway between onset and F1 maximum for /əu/ (Rosenfelder et al., 2014). The measurements were normalised using the Lobanov method to give normalised values for the 10% interval measurements. The vowels are also coded for their phonological environment using the Plotnik (Labov, n.d.) method.

3.3.5. Data Processing

All extracted data was read into R and joined with the demographic information and a lexical set reference list (every word in the data set coded to a lexical set by the author based on Wells 1982*a*,*b*). All data cleaning was done using R in RStudio (RStudio Team, 2016), and the code can be viewed via the link in Appendix A. The initial process, performed on all data, before separating out lexical sets for individual analysis for the separate research questions, included the below steps:

- filtered so that only primary stressed vowels are included
- any tokens with duration less than 50ms removed
- a list of stop words removed due to the probability of reduction (stop words defined as words removed from a search list, in this context an example would be ‘an’, which is canonically a TRAP word but is often realised as a schwa)
- Any words that the speaker did not produce in their own accent were marked in the transcript as ‘xxwordxx’, for example. “I say /ant/ but may family say /ɑːnt/" would be transcribed as ‘I say aunt but my family say xxauntxx’. The final data set then had any words with “xx" in them filtered out.
- Outliers for each formant value and lexical set were filtered out based on being outside a boxplot maximum and minimum (greater than Q3 + 1.5*IQR, or less than Q1 - 1.5*IQR). This process was performed on the dataset as a whole.

3.4. Analysis

The variables under discussion require different analysis techniques to full understand the variation present. The FOOT-STRUT and TRAP-BATH splits, as monophthongs that mostly vary along one dimension, can generally be considered as a merged or split, and analysed using single point formant measurements (from FAVE these are at one third of the vowel duration, which was found to mostly closely approximate human annotators' behaviour), the methods for which are discussed below (section 3.4.1). However, GOAT allophony is more complex, most realisations are diphthongal, and there is variation in reports of where in the diphthong the change has occurred and changes are found in both the F1 and F2 directions. Methods for analysing this will be discussed in section 3.4.1.2. An explanation of all the variables in the models can be found in appendix D.1.

3.4.1. Measuring Vowel Mergedness

Nycz and Hall-Lew (2014) discuss best practices in measuring vowel mergers and splits. Much of the variation covered in this project is related to splits rather than mergers but many of the methods used to describe the relationship between word classes can be applied in reverse, to understand whether a person still has merged vowels or if they have split. e.g. The FOOT-STRUT split was, as discussed in chapter 2 historically a splitting process but is now a set of words that either all have the same vowel (i.e. are not split or 'merged') for the majority of northern speakers, or have two different vowel (i.e. are split or 'not merged') for southerners (Baranowski and Turton, 2018).e Using merger methods to analyse a split is also shown by Nycz and Hall-Lew (2014) in consideration of Canadians in New York acquiring a COT-CAUGHT split.

According to Nycz and Hall-Lew (2014) a method for measuring mergedness needs to take into account the following (p. 2-3):

1. **Capture the distance between word classes in acoustic space.** That is, to quantify the difference between the central tendencies of the two word classes, according to F1 & F2, but also to identify which dimension accounts for the majority of the difference. Also determine whether this difference is significant.

2. **Capture the degree of overlap between word classes in acoustic space.** While degree of overlap is dependent on how variable each word class is, and the distance between the classes, in principle it relates to how easy it is for a listener to perceive a difference between the categories. It can also reveal a change in progress.

3. **Take into account the unbalanced nature of naturalistic data.** Spontaneous speech data means that words from one class may appear more within a certain phonological context than others, and individual words may appear with a higher frequency than others.

4. **Enable a comparison between speakers in a corpus.** The measure of difference needs to be able to function as a dependent variable in statistical analysis, in order to compare social factors.

The following methods are compared:

1. **Euclidean Distance** - The diagonal distance between the mean of each vowel distribution

2. **Mixed Effects Regression and Adjusted Euclidean Distance** - This first models F1 and F2 each separately via difference between models including and not including Word Class (/Lexical Set). The adjusted Euclidean difference can then be calculated using the two effect sizes.

3. **Pillai-Bartlett Trace** (also known as Pillai score) - This is the result of MANOVA models of F1 and F2 as dependent variables simultaneously, the higher the value the greater the difference between the sets.

4. **Spectral Overlap Assessment Method (SOAM)**. The scatter for each vowel distribution is modelled as ellipses (best-fit and weighted), angled with respect to F1 and F2. The output is an overlap fraction representing the degree of overlap between the two ellipses.

Below are the factors that need to be considered for any particular method, and table 3.1 summarises the methods discussed by Nycz and Hall-Lew (2014), highlighting the advantages and disadvantages in capturing these.

1. Capture **distance** - the distance in the F1/F2 pane between the central tendencies of the two sets, and which direction this is in.

2. Capture **overlap** - the size of overlap between the areas the vowel is found in, related to both distance (above), and the range a vowel class covers.

3. Take into account **naturalistic data** - spontaneous speech data is often both unbalanced and multivariate, to analyse it a method needs to be able to include multi-factor analysis to capture this.

4. Enable **comparison of speakers** within a corpus - needs to be a measure that is comparable across speakers so that it can be used as a a dependent variable in statistical analysis in order to include social factors and other factors of interest.

Method	1. Distance	2. Overlap	3. Naturalistic Data	4. Comparison of Speakers	Notes
Euclidean Distance	Yes, but not whether it is significant or not	No	No	Yes if based on normalised data	Mean may not be best indication of central tendency, particularly in the case of changes in progress which may cause non-normal distributions. Median can help.
Mixed Effects Regression	Yes	No	Yes, by using fixed effects	Yes	
Pillai-Bartlett Trace	An abstracted 'difference' as a score from 0 to 1.	*See distance*	Can account for phonological environment but not random effects	Yes	Not an easily interpretable value, such as Hz.
Spectral Overlap	No	Yes, in a more direct way than any of the above methods	No	No	

Table 3.1.: Summary of methods for measuring mergers, adapted from Nycz and Hall-Lew (2014).

For the TRAP-BATH and FOOT-STRUT splits in this project, linear mixed effects models Bates et al. (2018) were utilised, using F1 and F2 measurements at one third duration (Rosenfelder et al., 2014). A major factor in this decision was the capability of this method to deal with naturalistic data. The data in this project has many independent variables that need to be captured, including phonological environment and various social factors. Another important factor was the ability to include random effects; the TRAP-BATH split did not complete in English, therefore, word specific variation is important to capture; as shown by Nycz and Hall-Lew (2014), not including such effects can lead to overestimation of a vowel difference. More information on modelling random effects is below.

3.4.1.1. Linear Mixed Effects Models

Linear Mixed Effects Models are an advancement on Linear Regression Models. Linear Regression Models can include one (simple models) or more (multiple models) predictors but cannot model random effects or interactions. Random effects are those that are not repeatable, for example word, or speaker, rather than, for example, age group, which has a fixed set of possibilities. Within this data set this included speaker and word. Another advantage to mixed effects models are random slopes which vary more than just the intercept, like an interaction would. For example if different speakers vary by different amounts between speaking and reading styles, using a random slope adds this possible interaction in to consideration. These were used to test relationships between speaker and time through the interview (to check for accommodation) and between speaker and style. If a speaker accommodates to the interviewer, their accent will change through the interview, but this will happen at a different rate, if at all, for each speaker. Therefore, a random slope can model the variation. Once a models with random slopes are fitted they can be compared to a model without one by using AIC (Akaike Information Criterion) values. AIC is a measure of fit, where a smaller value indicates a better fit, including taking into account unnecessary predictors that do not explain a good proportion of the variance present. Linear mixed effects models were fitted using the R package lme4 (Bates et al., 2018), specifically the *lmer* command for linear mixed effects regression models. As discussed above, these models could be fitted using both fixed and random effect terms. Specific details on the details of model fitting for each variable is included at the beginning of each analysis chapter (chapters 4, 5, and 6).

3.4.1.2. Generalised Additive Mixed Models

As mentioned above, analysis of the GOAT vowel is more complex than the monophthongs, the entire formant trajectories are under question and dynamic analysis is required. Therefore, generalised additive mixed models were used to compare formant trajectories. For more information on GAMMs the reader is directed to Sóskuthy (2017) and Wood (2017), which explains and expands on much of the below.

Generalised Additive Mixed Models (henceforth GAMMs) are used for dynamic speech analysis, that is, aspects of speech that vary in space or time, particularly in this project, the short term temporal variation of a formant contour. They have recently come to the fore in use for modelling sound variation and change, and specifically have been used by Warburton (2020) to model the GOAT and THOUGHT vowels in Tyneside English, testing for a possible merger. Other recent research using this technique to compare and understand diphthong variation include Sóskuthy et al. (2019); Cole and Strycharczuk (2019); Stanley et al. (2021) and have found that using GAMM modelling gives more information about variation in a diphthong and capture patterns that would otherwise be missed. These models allow consideration the shape of the formant trajectory across the length of the vowel, and to compare effects of social and phonological factors (for example, education and surrounding environment) on that shape, rather than a point within it, as would be seen from a linear mixed effects model.

In a similar way to how Linear Mixed Effects Models are related to Linear Regression Models, GAMMs are developed from GAMs (Generalised Additive Model), which fit a regression curve to (in the case of this data) a formant trajectory . A GAM accounts for non-linearity by allowing a *smooth* term alongside the intercepts and slopes; the smooth term is a function of one or more of the variables. When using the mgcv package (Wood, 2017), the smoothing parameter is estimated directly from the data.

Within a GAMM, three types of random factors can be included (Sóskuthy, 2017); a random intercept fits parallel lines with identical slopes, random slopes change the slope (angle) of the line, and random smooths allow a different curve to be fitted to each individual trajectory. A way of understanding the best type of fit for a model is residual autocorrelation. The residuals are the data not captured by the model, which ideally is just noise. However, when the wrong type of model is fitted to data (for example a linear model fitted to a wiggly trajectory), patterns are left in the residuals, seen in the lag values. An example can be seen in figure 3.2, which shows the residual autocorrelation of a GAMM fitted without including the individual trajectories (random smooth above). A formant trajectory moves smoothly over time, that is, the formant value at each point in time is relatively close to the one before. This pattern remains in the residuals, showing that there are patterns in the data that have not been captured by the model. Alongside poor fit, models with residual autocorrelation will be overconfident and lead to under-estimated p-values, creating type one errors. In order to avoid this, models will be fitted either including a value to account for individual trajectories, or using an autocorrelation model.

When fitting GAMMs there are a few ways in which models can be compared for significance and how well they fit the data.

3.4.1.2.1. Nested models are built to exclude the terms of interest and then can be compared using compareML() (a form of ANOVA van Rij et al. 2020), see 3.4.1.2.1. This gives a p-value for the difference between the models.

```
w.gamm <- bam(f2 ~ word + s(measurement.no) + s(measurement.no,
    by=word) + s(measurement.no, traj, bs="fs", m=1), data=dat.
    words, method="ML")

w.gamm2 <- bam(f2 ~ s(measurement.no) + s(measurement.no, traj,
    bs="fs", m=1), data=dat.words, method="ML")

compareML(w.gamm, w.gamm2, print.output=F)$table
```

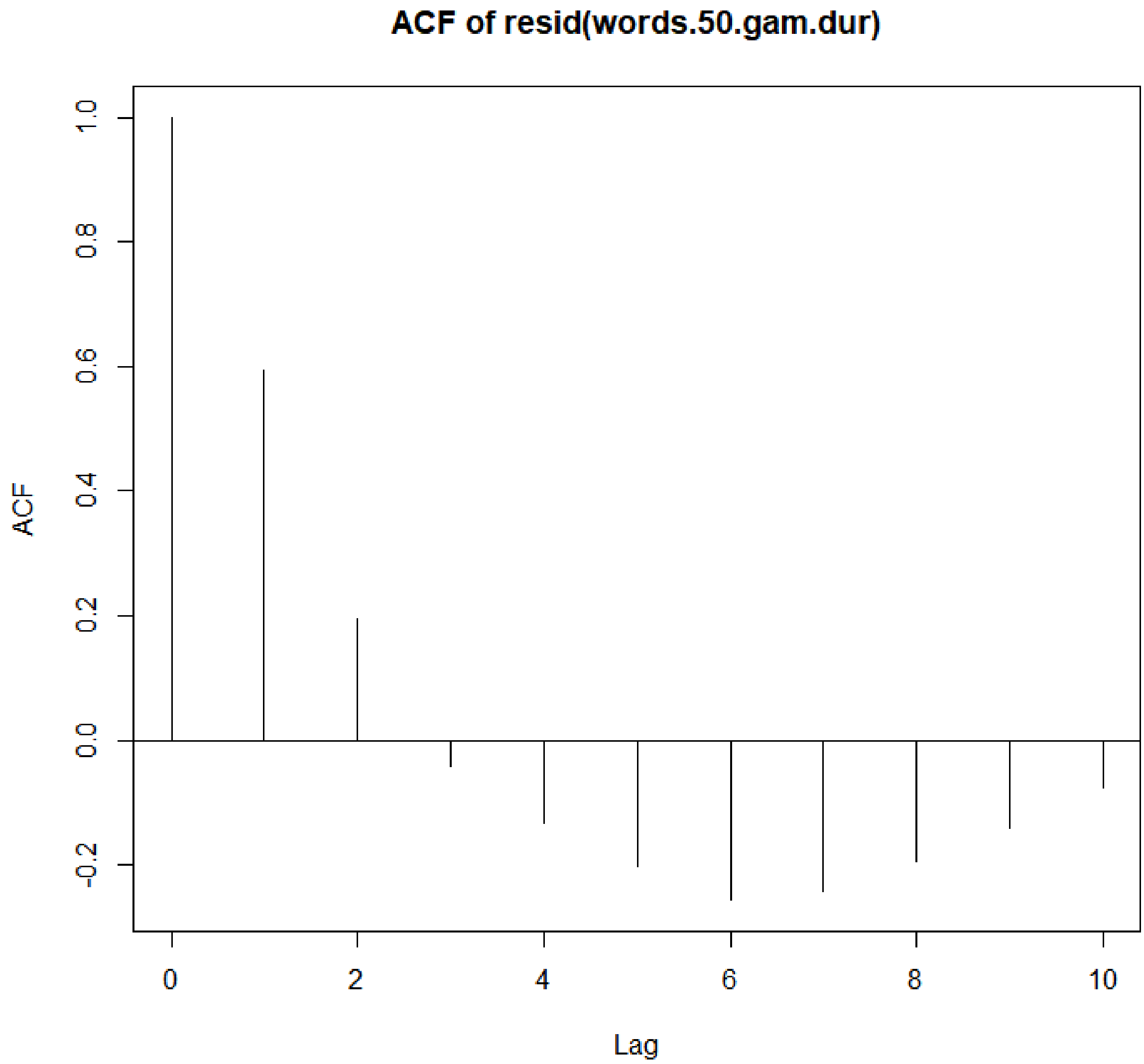

Figure 3.2.: Residual autocorrelation of GAMM produced from example data in Sóskuthy (2017)

```
## Model Score Edf Chisq Df p.value Sig.
## 1 w.gamm2 5923.394 5
## 2 demo.w.gamm 5908.614 8 14.780 3.000 1.707e-06 ***
```

Listing 3.1: example of comparing nested models using compareML()(Sóskuthy, 2017)

3.4.1.2.2. Visual methods for significance testing rely on confidence intervals. Two types of plot can be made, either plotting smooths with confidence intervals for the levels of a factor, and looking at the overlap, or plotting a difference smooth and looking at where the confidence interval includes 0, generally this project will use the confidence interval method.

3.4.1.2.3. AIC (discussed above relating to Linear Mixed Effect Models, section3.4.1.1) was used to compare models which could not be nested, as stated above the lowest AIC is the preferred model.

3.5. Drawing Conclusions

The methods discussed above will give answers as the the distribution of vowel variation between the speakers in the corpus. The comparisons discussed in section 3.1 will be performed by including the corpus information, and demographic information about individual speakers into the models constructed. For example, a significant effect of region in a model of F2 of TRAP-BATH including CoRP-NE and CoRP-SE speakers would demonstrate that the two sets of speakers do the TRAP-BATH split differently, demonstrating evidence for an answer to research question 1, that north-eastern 'RP' speakers behave regionally with respect to this variable.

Chapter 4.

Analysis - FOOT-STRUT split

4.1. Introduction

This chapter will consider the FOOT-STRUT split in all three speakers groups (CoRP-SE, DECTE, and CoRP-NE, henceforth referred to as corpora), first, in section 4.2 by modelling FOOT and STRUT together in each corpus (in both F1 and F2 dimensions), then in section 4.3 by modelling the STRUT vowel alone in all three corpora together. Each section is ordered to first consider the 'typical' RP behaviour in CoRP-SE, followed by the 'regional' behaviour of the DECTE speakers, before analysing the CoRP-NE speakers to compare their behaviour to the two previous groups.

As summarised in chapter 3 the FOOT and STRUT vowels were measured once, at one third of the duration (Rosenfelder et al., 2014). Data cleaning methods are also explained in chapter 3. In order to prevent over fitting of models the linguistic predictors were plotted with F1 and F2, and only included if variation was seen between the levels of the predictor. Models were compared using CAIC, including using the stepCAIC() function from the cAIC4 package (Saefken and Ruegamer, 2018) to determine the model with the best fit. Unless involved in an interaction, categorical predictors were sum coded (using contr.sum() from the stats package R Core Team 2021) in order to understand the intercept as a mean in real terms rather than at a combination of single levels of the predictors Winter (2019), those that were sum-coded are marked in the model tables as 'predictor*Sum*'. Continuous predictors were scaled to a z-score using the scale() function (R Core Team, 2021).

The chapter finds that CoRP-NE speakers do not behave in the same way as either CoRP-SE or DECTE speakers. Unlike the DECTE speakers, they show a split but it is not the same as the CoRP-SE speakers. The height of their STRUT vowel (as measured by F1)is also in between the other two speakers groups.

4.2. The FOOT-STRUT Split

This section will investigate the nature of the FOOT-STRUTsplit. The three groups (CoRP-SE, DECTE, CoRP-NE) are analysed in separate models and the patterning of the FOOT-STRUTsplit compared. The CoRP-NE speakers are discussed last in order to compare the patterning to the other two groups.

4.2.1. The FOOT-STRUT Split in CoRP-SE speakers

Analysis of their vowels shows that the split is found mostly in height (STRUT is lower in the acoustic space than FOOT, F1 = +199Hz) and very slightly in frontness (STRUTis further back in the acoustic space than FOOT, F2 = -86Hz). As discussed below and in section 4.2.1.2, the frontness is a difference in mean but there is full overlap between the ranges. From figure 4.1 it can be seen that there is overlap of individual tokens but the mean position of the STRUT words is lower in the acoustic vowel space than the FOOT words. They are also on average further back but the range falls within a subsection of the total F2 range of the FOOT words, which are more spread out. Full analysis of the F1 and F2 difference is in sections 4.2.1.1 and 4.2.1.2.

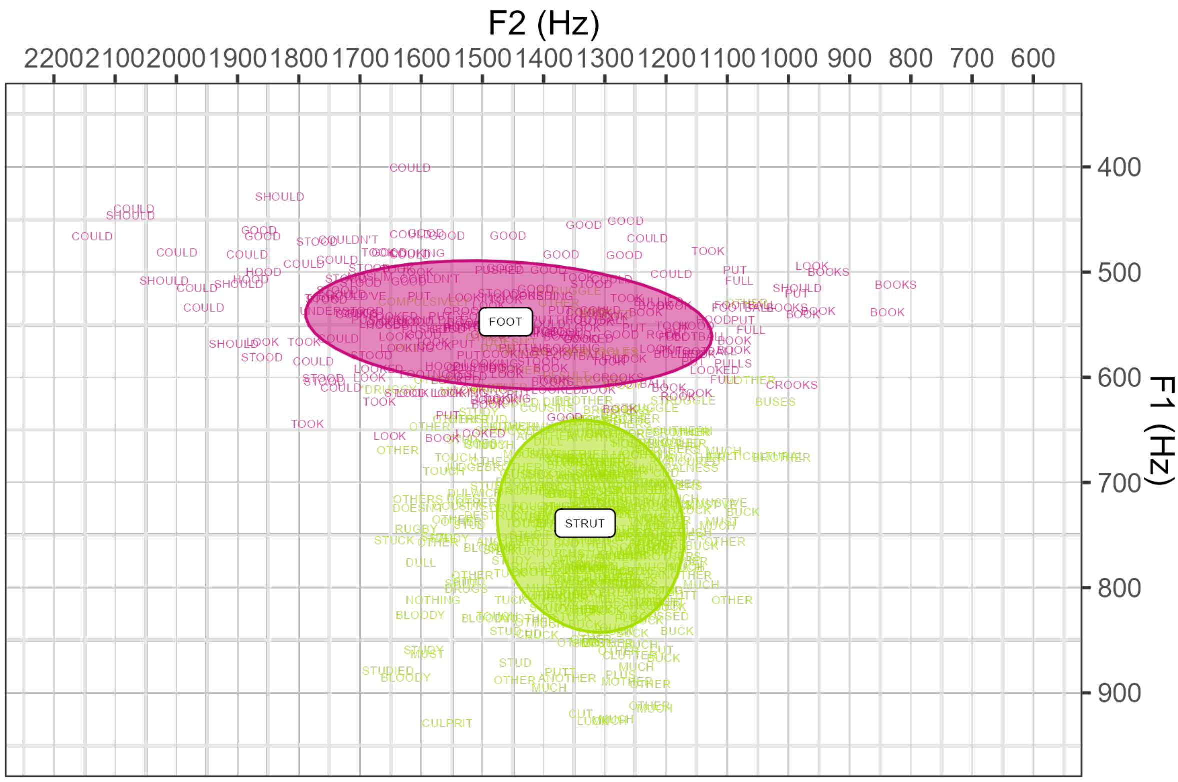

Figure 4.1.: Vowel Space plot of FOOT and STRUT in the CoRP-SE speakers

4.2.1.1. F1 of the CoRP-SE speakers

The best fit model of the normalised F1 of the FOOT and STRUT words is shown in table 4.1; the model also includes random intercepts for speaker and word. The model intercept is 536Hz, which is the mean F1 of the FOOT words. The effect size of lexical Set is +199Hz (t=17.28). Therefore, the mean F1 for the vowels in the STRUT words is 735Hz. There is an effect of speaker sex (t=3.77) but the effect size is only 21Hz, which is not large in the context of the lexical set variation. This model demonstrates that the vowels of the two lexical sets are distinct in height in CoRP-SE speakers, with the mean of the STRUT lexical set lower in acoustic vowel space than the mean of the FOOT lexical set, by 199Hz. The difference is visualised in figure 4.2 (based on raw data, not the model predictions), where the distinction between the vowel measurements in the two lexical sets can be seen clearly, with no overlap between the interquartile ranges.

fixed effect	estimate	t-value
(Intercept)	536.37	35.36
lexical set (baseline FOOT*)*		
STRUT	199.08	17.28
speaker sex (sum-coded)		
Sum1	20.57	3.77
age group (sum-coded)		
Sum1	3.82	0.71
following manner (sum-coded)		
Sum1	12.17	0.76
Sum2	0.15	0.02
Sum3	-20.85	-1.56
preceding segment (sum-coded)		
Sum1	19.32	1.50
Sum2	11.43	1.32
Sum3	-12.04	-0.96
Sum4	-19.77	-0.95
z-scored Zipf frequency (continuous)		
	0.61	0.11
style (sum-coded)		
Sum1	-16.14	-2.18
Sum2	7.27	0.79
z-scored time(continuous)		
	1.55	0.43

Table 4.1.: Linear mixed effects model of F1 of FOOT and STRUT in the South East

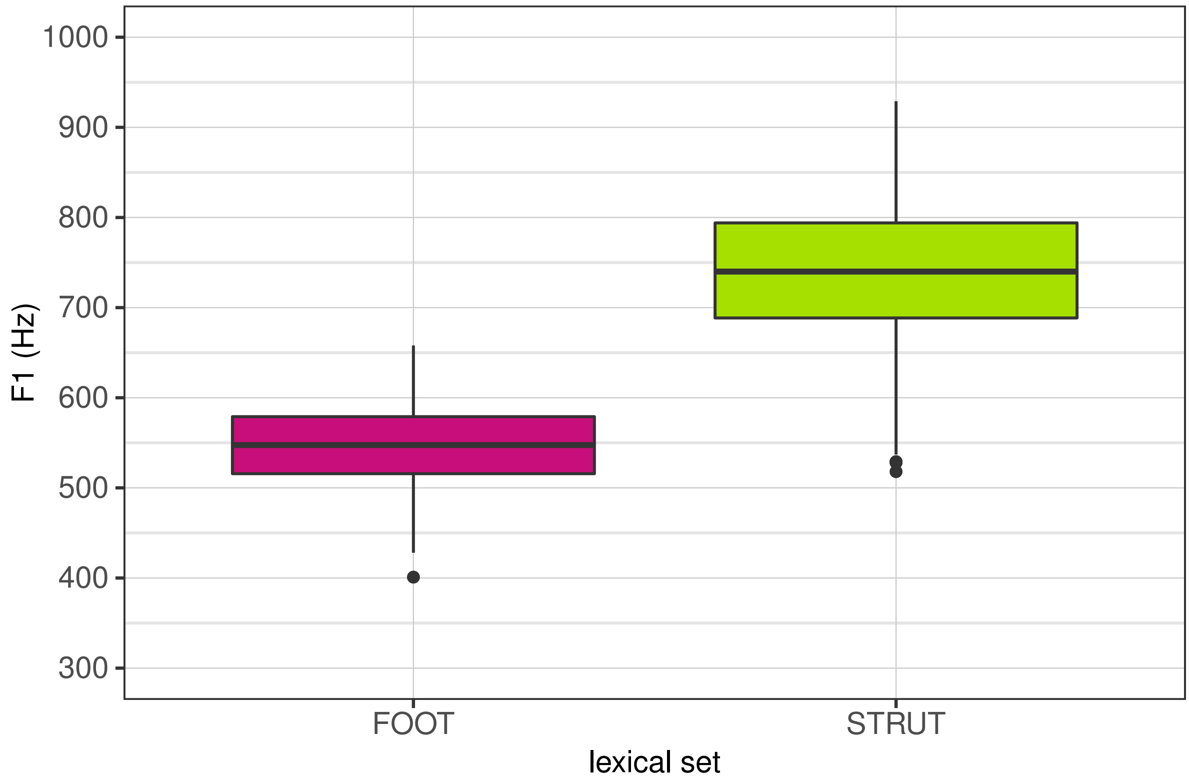

Figure 4.2.: F1 of FOOT and STRUT in CoRP-SE speakers

4.2.1.2. F2 of the CoRP-SE speakers

Modelling F2 (see table 4.2) showed an intercept of 1383Hz, the mean of the FOOT lexical set. While there is a lot less variation according to lexical set than seen in F1, there is a difference of -86Hz (t =-2.58) implying that the mean of the STRUT vowel is slightly further back than the mean of the FOOT vowel in south-eastern speakers (see figure 4.3). This difference is not large and it can be seen in figure 4.3 that the inter-quartile range is almost completely overlapping; the STRUT words sit almost completely within the range of the FOOT words. Figure 4.3 has points that represent individual speaker means and shows that some speakers have a split in F2, with STRUT words occurring further back in the acoustic vowels space, whereas for others the STRUT vowel has approximately the same frontness as the

FOOT words. This reflects literature that suggests variation in the frontness of the STRUT vowel, with some authors transcribing with ɐ and others with ʌ. There is also a small but significant effect of style.

fixed effect	estimate	t-value
(Intercept)	1386.41	29.16
lexical set (baseline FOOT*)*		
STRUT	-85.98	-2.58
speaker sex (sum-coded)		
Sum1	-13.82	-0.56
age group (sum-coded)		
Sum1	-20.93	-0.84
following manner (sum-coded)		
Sum1	10.77	0.23
Sum2	-1.30	-0.05
Sum3	-52.11	-1.37
preceding segment (sum-coded)		
Sum1	-26.20	-0.69
Sum2	-15.12	-0.60
Sum3	-62.95	-1.74
Sum4	107.84	1.81
z-scored Zipf frequency (continuous)		
	-2.61	-0.17
style (sum-coded)		
Sum1	62.19	3.24
Sum2	-43.60	-1.88
time (continuous)		
	15.37	1.67

Table 4.2.: Linear mixed effects model of F2 of FOOT and STRUT in the South East

A further model was run on the same speakers including STRUT, THOUGHT, and schwa, to check the frontness of the STRUT vowel of in comparison to other vowels that would be found at a similar height in the vowel space; the model summary can be found in appendix E.1. The comparison shows that the STRUT vowel in these speakers is significantly further forward than the THOUGHT vowel (-399Hz, t=-13.50) and also significantly further back than

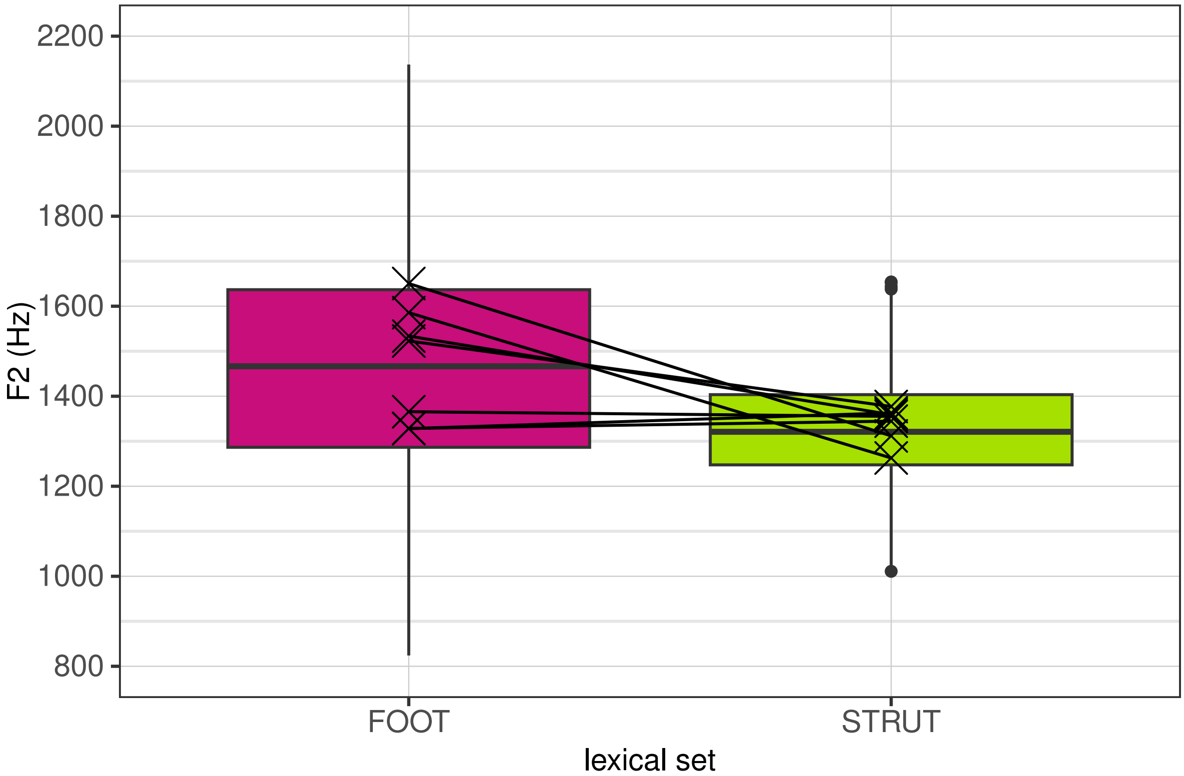

Figure 4.3.: F2 of FOOT and STRUT in CoRP-SE speakers, including speaker means

the schwa (305Hz, t=8.16), placing it between the two in the vowel space, but closer to schwa (see figure 4.4) than THOUGHT. ‘

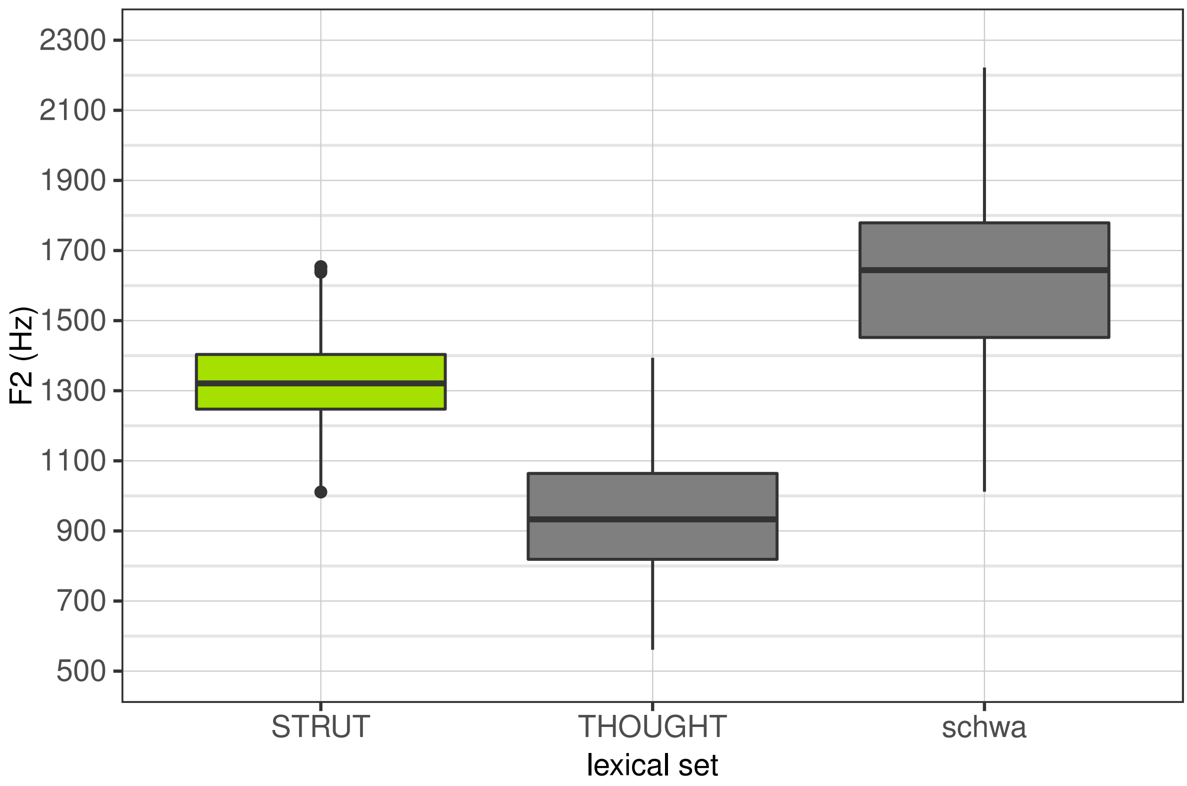

Figure 4.4.: F2 of STRUT, THOUGHT, and schwa in CoRP-SE speakers

4.2.2. The FOOT-STRUT Split in DECTE speakers

Overall, the DECTE speakers do not show a split in height, as measured by F1, and while they show some F2 differences between FOOT and STRUT in the old age group (+147Hz), the pattern is clearly different to that found in the CoRP-SE speakers. It can be concluded that according to this sample, the majority of state-educated speakers in the North East do not show a FOOT-STRUT split. Further analysis of the F1 and F2 dimensions is continued below.

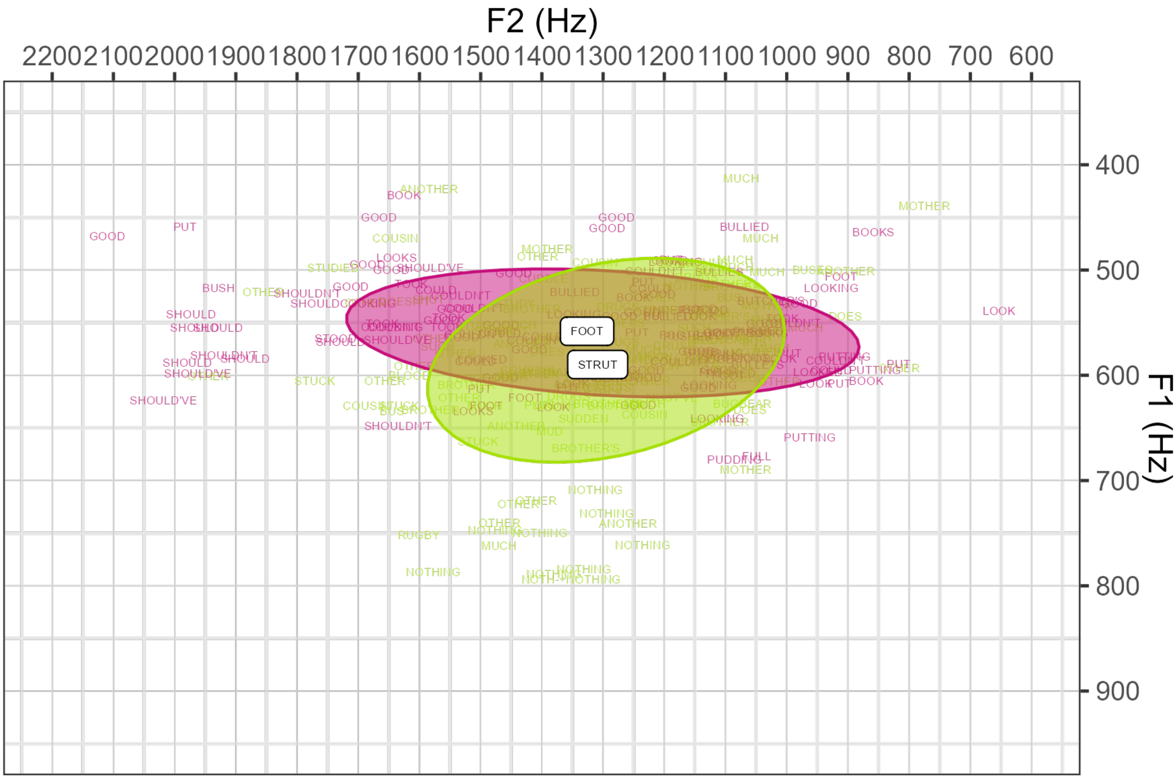

Figure 4.5.: Vowel Space plot of FOOT and STRUT in the DECTE speakers

4.2.2.1. F1 of the DECTE speakers

Modelling F1 of the FOOT and STRUT words in the DECTE speakers gives an intercept of 572Hz (table 4.3), which is the mean of the FOOT words, and no significant effect of lexical set (seen in figure 4.6). However, the points in figure 4.6 show that the individual differences are not as clear. While investigating individual speakers in detail is beyond the scope of this investigation, and the best fit model did not include speaker, it is worth noting that the lack of distinction is not homogenous and may be affected by other factors.

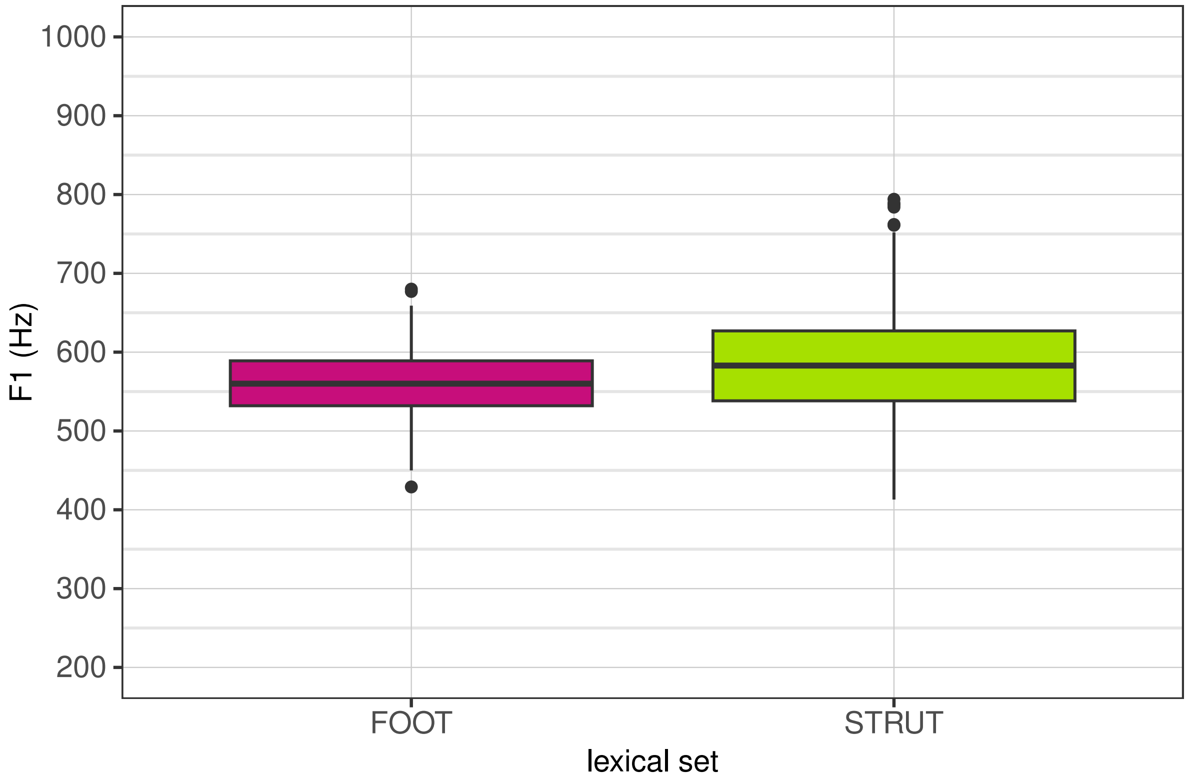

Figure 4.6.: F1 of FOOT and STRUT in DECTE speakers

fixed effect	estimate	t-value
(Intercept)	571.28	27.70
lexical set (baseline FOOT*)*		
STRUT	27.12	1.47
speaker sex (sum-coded)		
Sum1	3.53	0.72
age group (sum-coded)		
Sum1	-8.55	-1.70
following manner (sum-coded)		
Sum1	-19.81	-0.86
Sum2	5.81	0.41
Sum3	14.90	0.69
preceding segment		
Sum1	24.82	1.14
Sum2	-4.07	-0.30
Sum3	1.52	0.07
Sum4	-19.53	-0.65
z-scored Zipf frequency (continuous)		
	7.14	0.75
z-scored time (continuous)		
	11.41	1.84

Table 4.3.: Linear mixed effects model of F1 of FOOT and STRUT in DECTE speakers

4.2.2.2. F2 of the DECTE speakers

The best fit model of F2 of the FOOT and STRUT words in the DECTE speakers (table 4.4) includes an interaction of lexical set and age group. The intercept (mean of FOOT words in old speakers) is 1232Hz, the mean of FOOT words in young speakers is 1638Hz. The split in the old age group is +147Hz whereas in the young age group it is -26Hz. It is possible that the split in the younger speakers is affected by the Tyneside phenomenon wherein many FOOT words move to the GOOSE lexical set, which seems to be more prevalent in the younger speakers in this age group (figure 4.7).

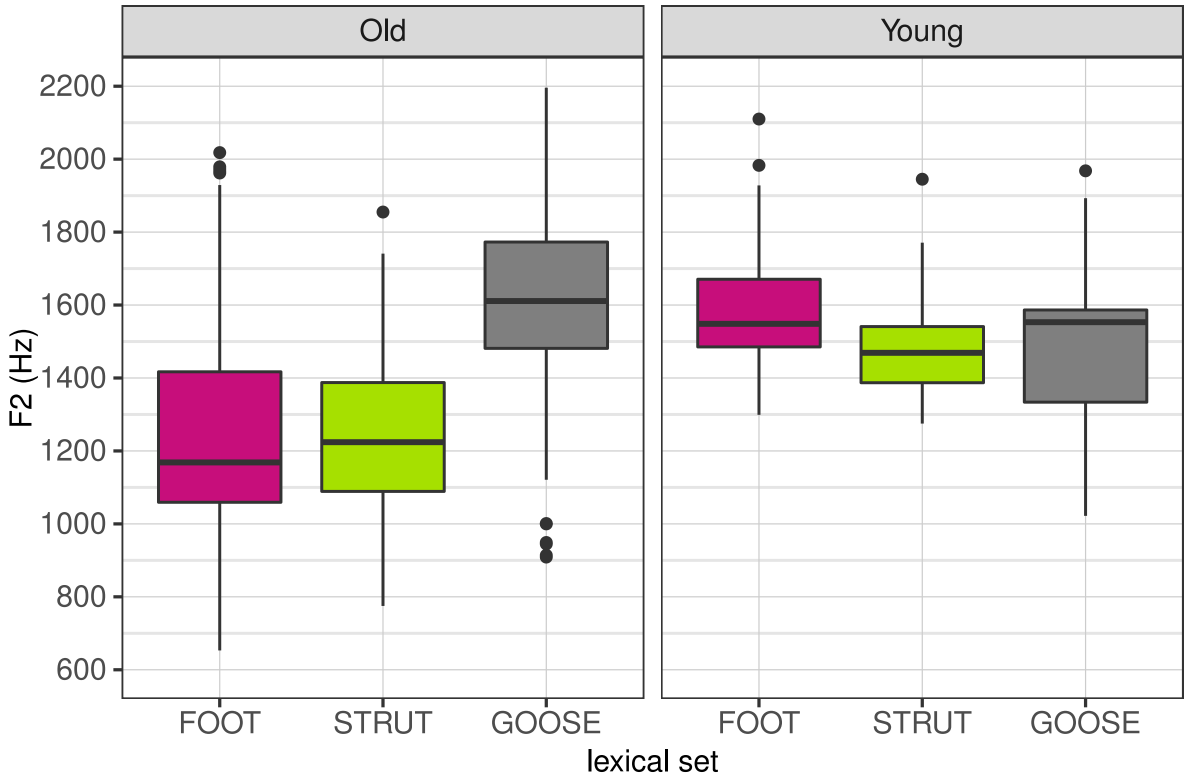

Figure 4.7.: F2 of FOOT, STRUT, and GOOSE in DECTE speakers, by age group

fixed effect	estimate	t-value
(Intercept)	1232.05	19.07
lexical set (baseline FOOT*)*		
STRUT	146.67	2.67
age group (baseline old)		
young	405.58	6.41
speaker sex (sum-coded)		
Sum1	-48.29	-1.80
following manner (sum-coded)		
Sum1	-73.52	-1.13
Sum2	-12.44	-0.31
Sum3	11.04	0.18
preceding segment (sum-coded)		
Sum1	-150.48	-2.42
Sum2	-139.84	-3.60
Sum3	-148.60	-2.36
Sum4	458.85	5.36
z-scored Zipf frequency (continuous)		
	25.87	0.95
z-scored time (continuous)		
	-37.39	-1.96
interactions		
lexical set STRUT age group young	-172.85	-2.88

Table 4.4.: Linear mixed effects model of F2 of FOOT and STRUT in DECTE speakers

4.2.3. The FOOT-STRUT Split in CoRP-NE speakers

CoRP-NE speakers show some evidence of a FOOT-STRUT split, particularly in F1 (on average 110Hz, higher in female speakers) but little evidence of an F2 split except in one speaker. Further analysis of the F1 and F2 differences can be seen below, but from the vowel space plot in figure it can be seen that the STRUT words are lower than the FOOT words (with some overlap) but have similar frontness.

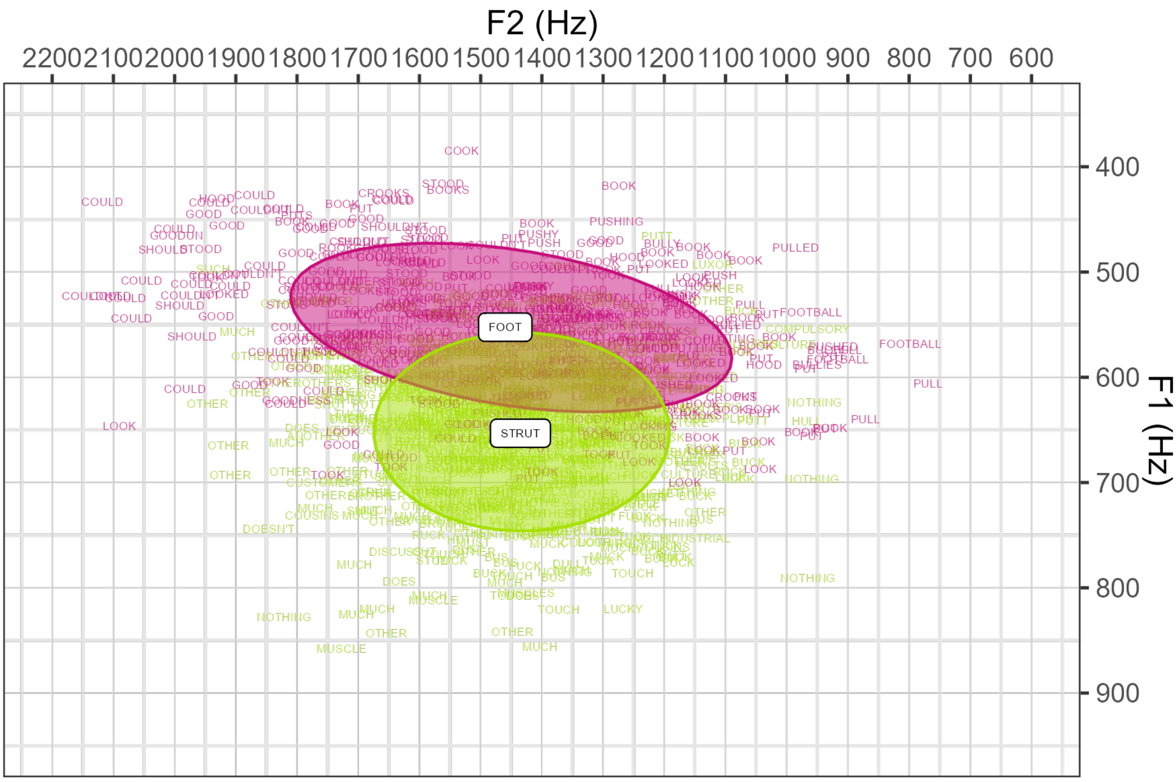

Figure 4.8.: Vowel Space plot of FOOT and STRUT, in the CoRP-NE speakers

4.2.3.1. F1 of the CoRP-NE speakers

The best fit model for F1 of the CoRP-NE speakers can be seen in table 4.5 and shows a FOOT-STRUT split in F1 that interacts with both speaker sex and age group. The mean value of FOOT is 541.53 and the mean value of STRUT is 651Hz, showing an average split of 110Hz. However, this split is overall higher for female speakers (mean=142Hz) compared to male speakers (mean=78Hz) (64 Hz difference), there is also a slightly larger split in older speakers. Overall the size of split is ranked: old female, young female, old male, young male. The difference between these splits can be seen in figure 4.9.

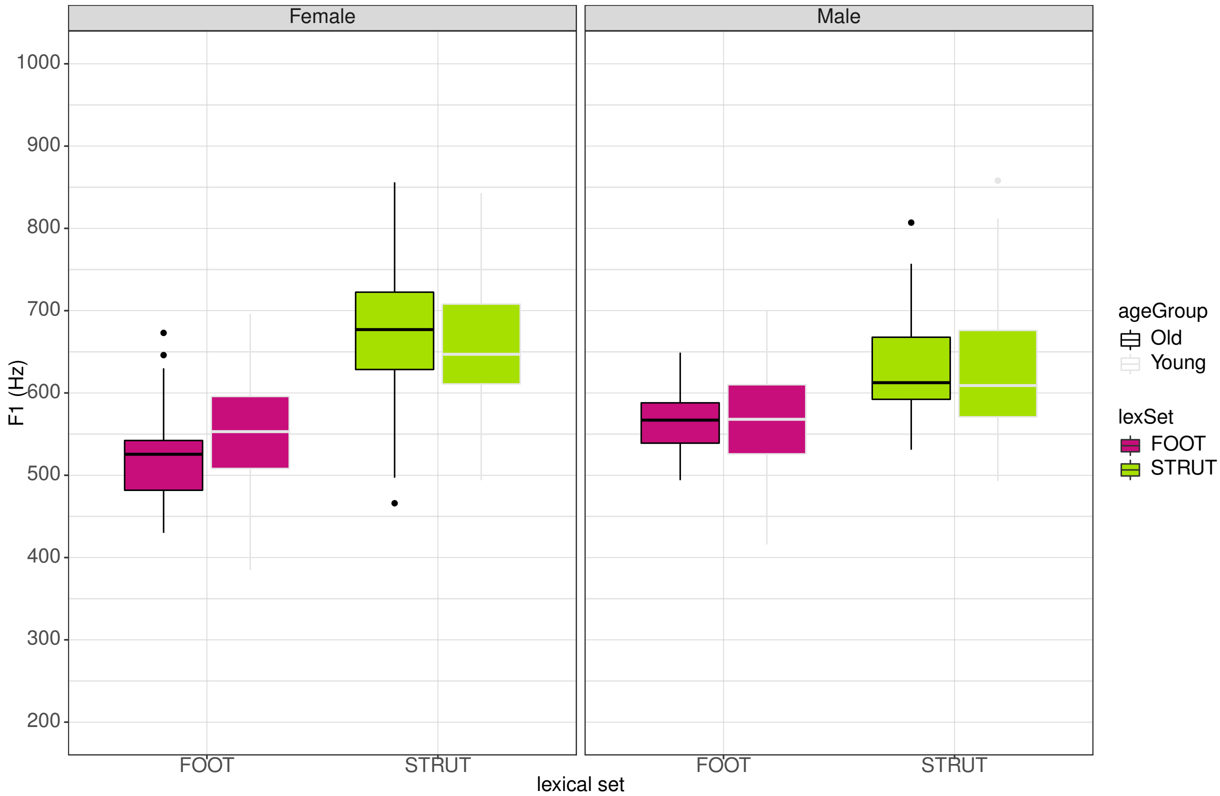

Figure 4.9.: F1 of FOOT and STRUT in CoRP-NE speakers

fixed effect	estimate	t-value
(Intercept)	516.30	27.03
lexical set (baseline FOOT*)*		
STRUT	172.92	14.45
speaker sex (baseline female)		
male	32.17	1.22
age group (baseline old)		
young	36.03	1.88
following manner (sum-coded)		
Sum1	-12.31	-1.08
Sum2	11.62	1.59
Sum3	-5.03	-0.52
preceding segment (sum-coded)		
Sum1	24.05	2.49
Sum2	12.59	1.82
Sum3	-22.43	-2.08
Sum4	-9.66	-0.56
z-scored Zipf frequency (continuous)		
	2.55	0.61
speech style (sum-coded)		
Sum1	-12.35	-2.17
Sum2	5.40	0.78
interactions		
lexical set STRUT: sex male	-89.79	-5.32
lexical set STRUT: age group young	-61.66	-5.40
sex male: age group young	-36.95	-1.18
lexical set STRUT: sex male: age group young	50.89	2.53

Table 4.5.: Linear mixed effects model of F1 of FOOT and STRUT in the North East

	FOOT	STRUT	*size of split*
Old Female	516.3	689.22	*172.92*
Old Male	548.47	631.6	*83.13*
Young Female	553.07	664.33	*111.26*
Young Male	548.29	620.65	*72.36*
Mean	541.53	651.45	*109.92*

Table 4.6.: Table showing interactions effects of speaker age group and sex on F1 of FOOT and STRUT in CoRP-NE speakers (interactions in table 4.5)

4.2.3.2. F2 of the CoRP-NE speakers

The best fit model for F2 of the CoRP-NE speakers (see table 4.7) also includes a three way interaction between lexical set, sex, and age group (summarised in table 4.8). The mean value of FOOT is 1339Hz and the mean value of STRUT is 1380Hz, showing very little distinction. When the interaction is considered, the largest distinction between FOOT and STRUTis in the old male speakers. However this distinction is not the same way round as in CoRP-SE speakers, it shows a STRUT vowel that is slightly fronter than the FOOT vowel. As can be seen in figure 4.10, similar to the CoRP-SE speakers (figure 4.3) the difference between lexical sets is not homogenous across speakers.

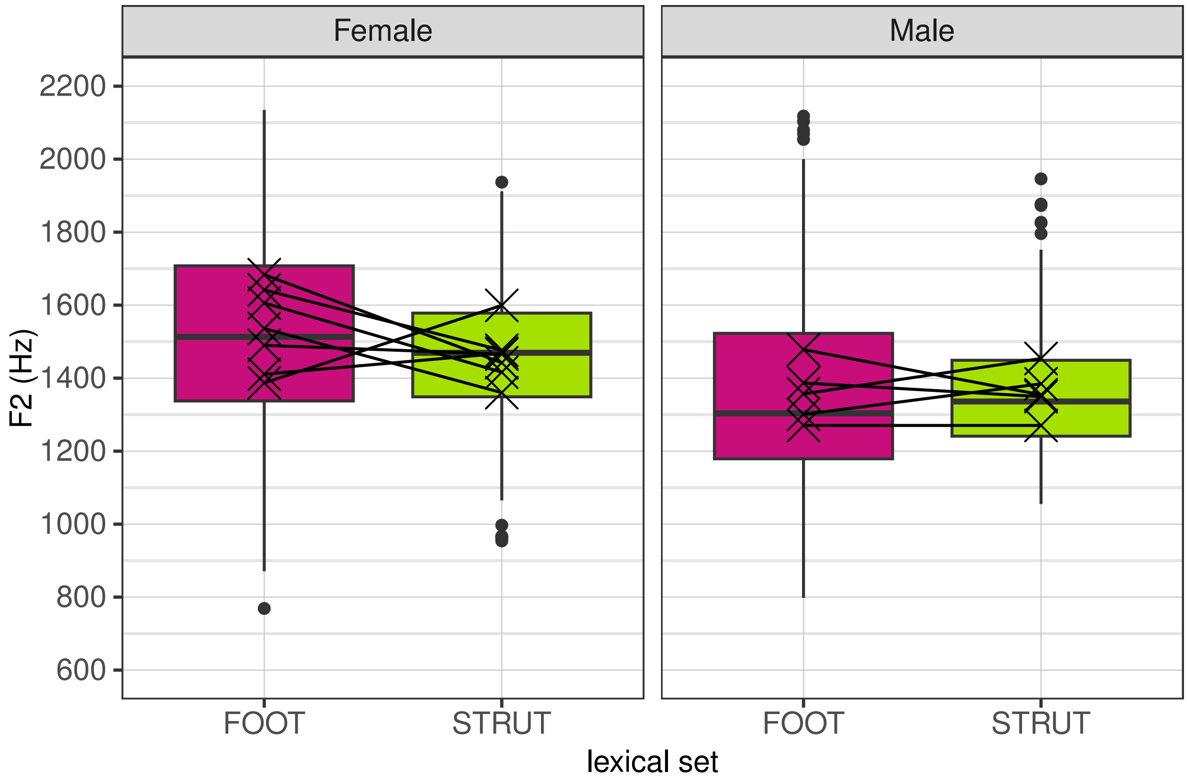

Figure 4.10.: F2 of FOOT and STRUT in CoRP-NE speakers, including speaker means

fixed effect	estimate	t-value
(Intercept)	1468.01	25.11
lexical set (baseline FOOT*)*		
STRUT	-27.34	-0.68
speaker sex (baseline female)		
male	-261.86	-3.42
age group (baseline old)		
young	-77.10	-1.45
following manner (sum-coded)		
Sum1	81.82	1.71
Sum2	52.98	1.89
Sum3	-209.68	-5.95
preceding segment (sum-coded)		
Sum1	-45.91	-1.21
Sum2	-28.81	-1.10
Sum3	-68.50	-1.71
Sum4	161.72	2.57
z-scored Zipf frequency (continuous)		
	28.41	1.74
style (sum-coded)		
Sum1	56.67	3.08
Sum2	-30.87	-1.52
time (continuous)		
	-6.79	-0.83
interactions		
lexical set STRUT: sex male	203.82	4.34
lexical set STRUT: age group young	52.55	1.67
sex male:age group young	162.76	1.84
lexical set STRUT: sex male: age group young	-240.35	-4.31

Table 4.7.: Linear mixed effects model of F2 of FOOT and STRUT in the North East

	FOOT	STRUT	*size of split*
Old Female	1468.01	1440.67	*-27.34*
Old Male	1206.15	1382.63	*176.48*
Young Female	1390.91	1416.12	*25.21*
Young Male	1291.81	1280.49	*-11.32*
Mean	1339.22	1379.98	*40.76*

Table 4.8.: Table showing effects of speaker age group and sex on F2 of FOOT and STRUT in CoRP-NE speakers (calculated from table 4.7)

4.2.4. Conclusions on the nature of the FOOT-STRUT Split in all three speaker groups

In the CoRP-SE speakers, the FOOT and STRUT vowel are distinguished in F1 (a difference of +110Hz) and F2 (-86Hz). Since the scales of F1 and F2 are different a difference of 110 is far larger within F1 than 86 is within F2. The F1 difference is clearly meaningful but while the F2 difference is present and statistically significant it needs to be considered more carefully because it is only a small difference within the acoustic vowel space and may not be phonologically meaningful. Overall the FOOT-STRUT split can be described as mostly characterised by height and slightly by frontness. It was also found that the frontness of the STRUT vowel in the South East is between the schwa and THOUGHT vowels, but closer to schwa. If the CoRP-NE speakers were behaving like the CoRP-SE speakers, we would expect them to have a similar pattern in both F1 and possibly F2.

In the DECTE speakers we see what can be considered 'typical' North East FOOT-STRUT positions of FOOT and STRUT. The speakers show no significant difference between FOOT and STRUT in F1, and in F2 a small difference in mainly older speakers (+147Hz). It can be concluded that there is not a robust FOOT-STRUT split in the North East as found in the CoRP-SE speakers.

In CoRP-NE speakers we see some evidence of a split. They show an F1 difference of 110Hz (mean, male and female speakers behave differently) and some F2 difference in the opposite direction to CoRP-SE. This suggests that these speakers have a split, but it is not identical to the split found in the CoRP-SE speakers. It is not as large in height (F1 is 110Hz vs the 199Hz found in CoRP-SE speakers) and does not seem to exist in frontness.

By analysis of the FOOT and STRUT words in all three speaker groups it can be concluded that CoRP-NE speakers do not behave identically to either of the other speaker groups. However, conclusions can be drawn in answer to research question 1. The DECTE speakers (state educated in the North East) do not show any FOOT-STRUT split and CoRP-NE speakers clearly have at least some distinction between FOOT and STRUT so it can be concluded that they are not behaving consistently regionally. While the CoRP-NE speakers do not show as large a split as the CoRP-SE speakers, there is clearly a split present, demonstrating at least a tendency towards non-regional behaviour.

In order to further understand the nature of the difference between the split in different speaker groups the STRUT words were modelled separately, results of this are discussed in section 4.3.

4.3. STRUT VOWEL ONLY

4.3.1. F1 of STRUT only

Modelling the STRUT vowel alone (see table 4.9) shows that the CoRP-NE speakers are significantly different in F1 from both the CoRP-SE and the DECTE speakers. The best fit model included an interaction between corpus and speaker sex, the results of which are shown in table 4.10. These results partially support those above but add an extra dimension. The mean F1 value for STRUT in CoRP-NE female speakers is between the CoRP-SE and DECTE values, but the male speakers show a vowel that is very similar to the DECTE speakers (visualised in figure 4.11). This implies that while female speakers are not behaving regionally, male speakers are.

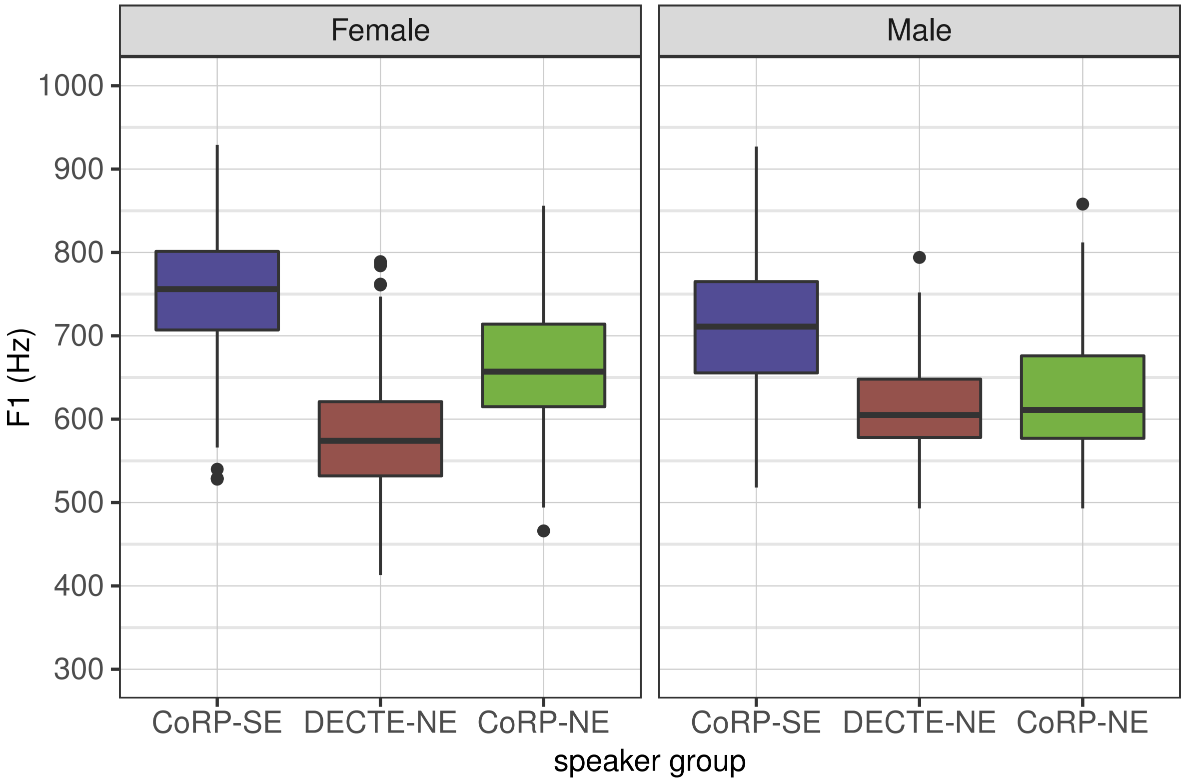

Figure 4.11.: F1 of STRUT in CoRP-SE, DECTE, and CoRP-NE speakers, by speaker sex

fixed effect	estimate	t-value
(Intercept)	671.04	43.85
corpus (baseline CoRP-NE)		
DECTE	-79.47	-5.16
CoRP-SE	93.87	6.42
speaker sex (baseline female)		
male	-47.85	-3.41
age group (sum-coded)		
Sum1	3.91	0.76
following manner (sum-coded)		
Sum1	-3.69	-0.25
Sum2	12.89	1.48
Sum3	-21.14	-1.79
preceding segment (sum-coded)		
Sum1	12.66	0.86
Sum2	15.56	1.72
Sum3	-17.46	-1.47
Sum4	-21.44	-0.93
z-scored Zipf frequency(continuous)		
	6.54	1.06
style (sum-coded)		
Sum1	-20.18	-2.28
Sum2	14.45	1.48
interactions		
corpus DECTE: sex male	77.98	3.00
coprus CoRP-SE: sex male	4.27	0.19

Table 4.9.: Linear mixed effects model of F1 of STRUT

	CoRP-SE	DECTE	CoRP-NE
Female	764.91	591.57	671.04
Male	721.78	621.7	623.19
Mean	*743.35*	*606.64*	*647.12*

Table 4.10.: Table showing effects of corpus group and speaker sex on F1 of STRUT (calculated from coefficients in table 4.9)

4.3.2. F2 of STRUT only

Modelling F2 of the STRUT words alone again shows an interaction of corpus and speaker sex (table 4.11). In each corpus the male speakers have a less front vowel, and overall the CoRP-NE speakers have the most front vowel, in both male and female speakers. Even if the CoRP-SE vowel is taken as the prototypical STRUT vowel there is variation between male and female, with male speakers producing a less front vowel (82Hz difference). The DECTE speakers show a relatively similar frontness to the CoRP-SE speakers, however the CoRP-NE speakers behave differently, both male and female speakers produce a fronter STRUT vowel than either of DECTE or CoRP-SE speakers, though the female difference is larger (see tables 4.11 and 4.12).

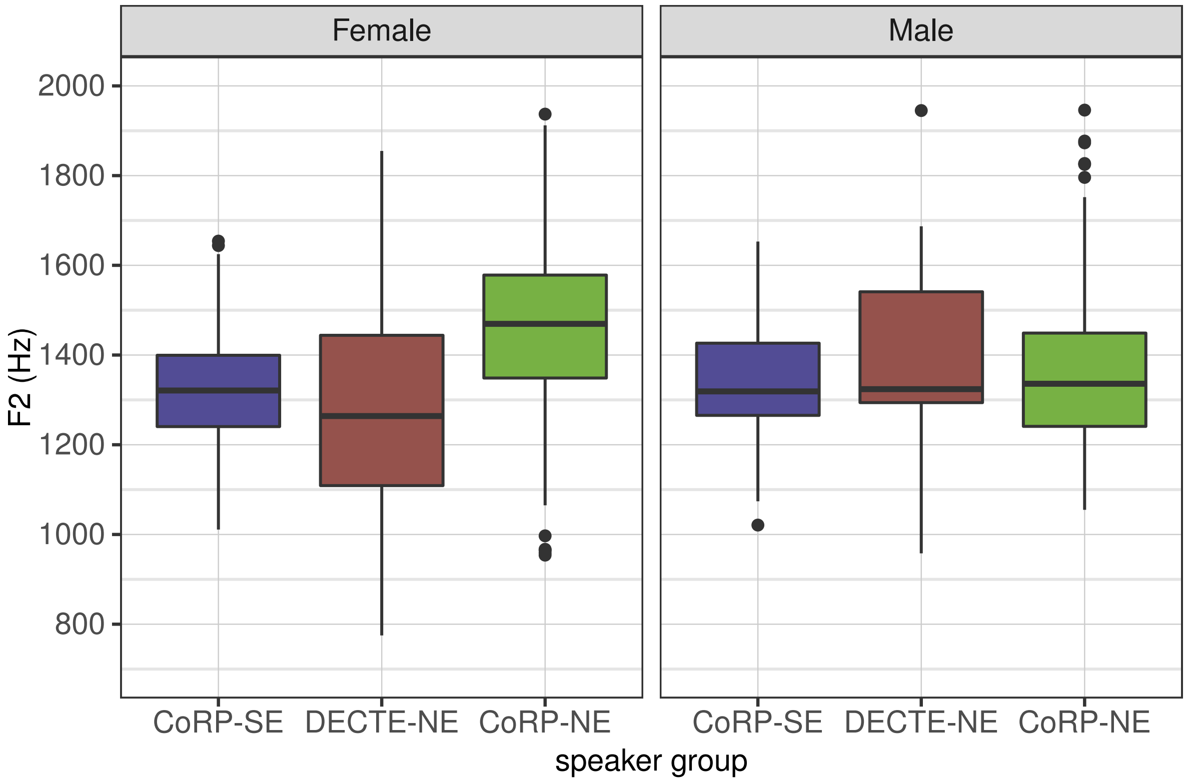

Figure 4.12.: F2 of STRUT in CoRP-SE, DECTE, and CoRP-NE speakers, by speaker sex

fixed effect	estimate	t-value
(Intercept)	1476.10	31.13
corpus (baseline CoRP-NE)		
DECTE-NE	-149.03	-2.54
CoRP-SE	-167.30	-2.77
speaker sex (baseline female)		
male	-122.73	-2.18
age group (sum-coded)		
Sum1	-14.63	-0.73
following manner (sum-coded)		
Sum1	59.85	1.64
Sum2	27.27	1.30
Sum3	-106.91	-3.83
preceding segment (sum-coded)		
Sum1	-58.13	-1.65
Sum2	-21.82	-1.01
Sum3	-52.51	-1.86
Sum4	132.08	2.43
z-scored Zipf frequency (continuous)		
	12.46	0.85
style (sum-coded)		
Sum1	10.33	0.51
Sum2	-11.97	-0.58
interactions		
corpus DECTE: sex male	208.16	2.32
corpus CoRP-SE: sex male	159.08	1.71

Table 4.11.: Linear Mixed Effects Model of F2 of STRUT

	CoRP-SE	DECTE	CoRP-NE
Female	1308.8	1327.07	1476.1
Male	1226.93	1263.47	1353.37
Mean	*1267.87*	*1295.27*	*1414.74*

Table 4.12.: Table showing effects of corpus on F2 of STRUT (calculated from interaction effects in table 4.11)

4.4. Discussion & Conclusion

The analysis above has shown the nature of the FOOT-STRUT split in each speaker group and compared the quality of the vowel in STRUT words in each.

Modelling the behaviour of both FOOT and STRUT in section 4.2 showed that the FOOT-STRUT split is mostly characterised by a difference in height, though some evidence of a frontness difference is found. The CoRP-NE speakers were found to have some evidence of a FOOT-STRUT split, the difference between the two lexical sets was larger than in the DECTE speakers but smaller than in the CoRP-SE speakers. Further analysis of the STRUT vowel alone showed that this difference is mainly due to the female speakers. The majority of male speakers seem to maintain the local pattern found in the DECTE speakers. The F2 pattern is a little more complex and likely is less informative when it comes to behaviour of the FOOT-STRUT split. As stated above, there is only some evidence of a split in frontness in the CoRP-SE speakers, who only show a difference of -86Hz, which is not a large difference within the scale of F2. In the focussed analysis of STRUT alone it can be seen that the CoRP-NE speakers have a different F2 to both the DECTE and CoRP-SE speakers. This shows that their production of the vowel, particularly in the female speakers is diverging from both groups.

4.4.1. Research Questions

There are two factors that need to be considered in relation to the results described above. First, the differences between privately educated and state-educated speakers as a type of social class variation, since while it is not occupation based, it is a form of social stratification connected to income, education and other social opportunities (see chapter 2 for further discussion of this approach). It is a known effect that male and female speakers diverge in socially stratified accent variation (Labov, 2001*b*), and this effect has been specifically attested in Newcastle (Watt, 1998). In general it has been seen that female speakers tend towards a standard, and local variants are more often found in male speakers (for example the Tyneside centring diphthongs in GOAT and FACE are found in male working class speakers by Watt (1998)). The pattern we see in the CoRP-NE speakers where male speakers have a STRUT vowel that is similar to the DECTE speakers whereas female speakers have a STRUT vowel that is different to them and more similar to the CoRP-SE speakers, clearly reflects this tendency. Male speakers are producing a local variant (though local here is likely 'northern' rather than Tyneside specific) and female speakers are tending towards a non-local variant.

Secondly, CoRP-NE speakers are creating a split where one is not found locally, but the vowel is not the same as a southern STRUT vowel. Their F1 is higher than in the DECTE speakers but lower than in the CoRP-SE speakers, and the F2 is slightly higher than the CoRP-SE speakers. This difference may be the source of the description occasionally given to STRUT realisations found in the north as 'schwa-like' (Braber and Flynn, 2015; Jansen and Braber, 2020). There are multiple possible explanations for this variation. It is possible that speakers are creating a different target for the STRUT vowel, reflecting the variable targets available in the input they receive. This input could potentially consist of both split and non-split patterns. It is also possible that they are aiming for the target as realised by CoRP-SE and other southern speakers but not hitting it. Not being able to reach the target could be be due to later acquisition of the split (Evans and Iverson, 2007). It is not possible to commit to one of these explanations without more in-depth data on the speech community and the acquisition process. While information on parent's education and region was collected, there are not enough speakers to draw meaningful conclusions. Another approach to understanding this would be to record people still at school (at various stages) to see when the FOOT-STRUT split appears in their speech. In order to answer whether the new variant is created by speakers targetting the South East form or avoiding the North East form would also need data on regional attitudes towards the variants alongside the acquisitional data discussed above.

In summary, in answer to research questions 1 and 3, both regional and non-regional patterns are seen, depending on speaker sex. The male CoRP-NE speakers behave regionally but the female speakers have a pattern that is not the local or 'RP' variation. Further discussion on the implications of the variation found here,in relation to research question 2 will be undertaken in chapter 7.

Chapter 5.

Analysis - TRAP-BATH split

5.1. Introduction

This chapter will present the analysis of the TRAP-BATH split in all three speaker groups. Each section first takes the CoRP-SE and DECTE speakers to understand the 'typical' RP and 'regional' patterns and then compares the CoRP-NE speakers to both of these groups.

The origins of the BATH lexical set are described as lengthening and backing of TRAP, and it is generally considered to be merged with the PALM (and START) sets in southern varieties of English (see chapter 2). Therefore, two approaches to modelling were taken. First, within each corpus group models included TRAP, BATH, and PALM (START words were coded as PALM since there is no difference in non-rhotic speakers) lexical sets to ascertain the pattern of BATH in relation to the other two sets of words. This analysis included F1 (to check for vowel height difference), F2 (to understand the reported difference in frontness), and duration. Regarding lengthening, for vowels of the same quality Labov and Baranowski (2006) say that a difference of 50msec is needed to form a change in vowel category (i.e. if BATH is only separated from trap by length, a 50msec difference would be needed to suggest a phonemic split). However, since BATH moving from TRAP to PALM is also a change in vowel quality, a smaller difference in duration may contribute to a change in category. The second part of the analysis will take the BATH words alone and look at the effects on their realisation by both speakers group and morpho-phonological environment. Since the BATH set is backed on phonological environment (pre-fricative, and occasionally pre-nasal), following manner was not included in the models to avoid collinearity.

It is found in the analysis below that the TRAP-BATH is as expected in the CoRP-SE speakers, being found in F2 and duration, with a small effect of F1. The DECTE speakers also behave as expected, with no significant difference between TRAP and BATH. The CoRP-NE speakers are found to broadly pattern with the DECTE speakers, with some variation observed.

Formant measurements for the below analysis were taken from the normalised FAVE (Rosenfelder et al., 2014) output, which is taken at one third duration (see chapter 3), and normalised using the Lobanov method. The syllable information used was produced using the transcriptions from MFA/FAVE and the syllabifyr package (Fruehwald, 2020).

In order to prevent over fitting of models the linguistic predictors were plotted with F1 and F2, and only included if variation was seen between the levels of the predictor. Models were compared using CAIC, including using the stepCAIC() function from the cAIC4 package (Saefken and Ruegamer, 2018) to determine the model with the best fit. Unless involved in an interaction, categorical predictors were sum coded (using contr.sum() from the stats package R Core Team 2021) in order to understand the intercept as a mean in real terms rather than at a combination of single levels of the predictors Winter (2019), those that were sum-coded are marked in the model tables as 'predictor*Sum*'. Continuous predictors were scaled to a z-score using the scale() function (R Core Team, 2021). Duration is measured by FAVE in seconds, but was converted to milliseconds (x1000) for readability and $\log_{10}$ transformed to remove positive skew (Winter, 2019).

5.2. The Split

5.2.1. The Split in CoRP-SE speakers

The TRAP-BATH split as seen in the South East speakers is confirmed as the vowel in BATH words patterning with PALM rather than TRAP. This is seen in an F1 difference between TRAP and BATH/PALM of -98Hz and an F2 difference of -363Hz. There was also a duration difference of 28msec found. The BATH (and PALM) words have vowels that are higher, further back, and longer than the vowels in TRAP words. There was also some interaction between speakers sex and lexical set in the F1 dimension, full analysis of the split is found below.

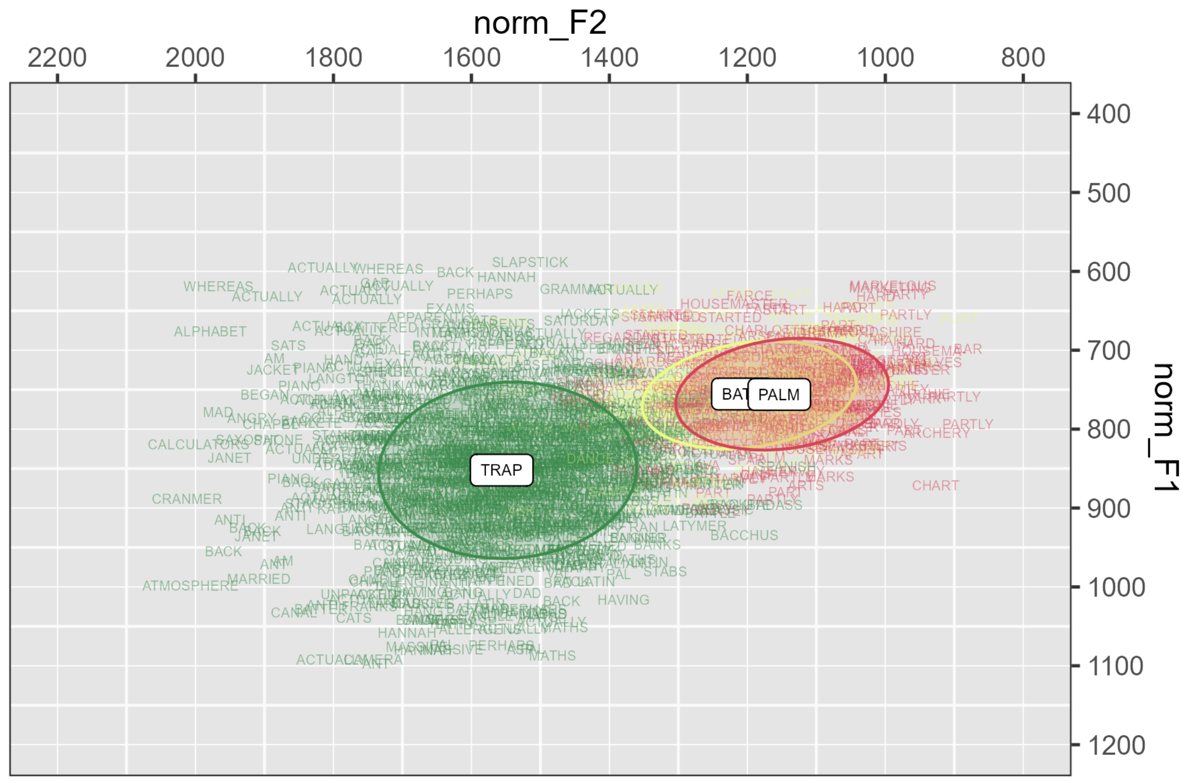

Figure 5.1.: Vowel Space plot of TRAP,BATH and PALM in the CoRP-SE speakers

5.2.1.1. F1 of the CoRP-SE speakers

The best fit model of F1 of TRAP, BATH, and PALM in the CoRP-SE speakers is shown in table 5.1, which includes an interaction effect of lexical set and speaker sex. The calculation of the interaction (from intercept values) can be seen in table 5.2 (note, the cells in grey are marked to show that none of the effects that distinguish the PALM values from the BATH values had a t-value greater than 1.4 so could be considered negligible). This demonstrates that BATH patterns with PALM and the TRAP-BATH split in the South East speakers has an overall lower vowel in TRAP than in BATH (see figure 5.2). A mean F1 difference of +98Hz (calculated between BATH and TRAP, excluding the PALM effect). This is larger in female speakers (+117Hz) due to a lower BATH and higher TRAP, and lower in the male speakers (+79Hz) due to a higher BATH and lower TRAP. Overall the mean F1 for vowels in BATH words is 740Hz, for PALM words 755Hz, and for TRAP words 857Hz.

fixed effect	estimate	t-value
(Intercept)	739.50	61.88
lexical set PALM	15.90	1.40
lexical set TRAP	117.49	12.25
speaker sex (baseline female)		
male	13.94	0.81
age group (sum-coded)		
Sum1	-0.60	-0.09
z-scored Zipf frequency (continuous)		
	0.44	0.14
style (baseline interview)		
minimal pair	42.22	2.04
word list	32.41	3.38
presence of coda (sum-coded)		
Sum1	2.42	0.71
interactions		
lexical set PALM : speaker sex male	-9.63	-0.61
lexical set TRAP : speaker sex male	-38.95	-2.93

Table 5.1.: Linear Mixed Effects Model of F1 of TRAP,BATH, and PALM in CoRP-SE speakers

	BATH	PALM	TRAP	TRAP-BATH difference
Female	764.37	780.27	881.86	117.49
Male	778.31	784.58	856.85	78.54
Mean	771.34	782.43	869.36	98.02

Table 5.2.: Interactions effects of lexical set and speaker sex on F1 of TRAP, BATH, and PALM in CoRP-SE speakers (calculated from interactions in table 5.1)

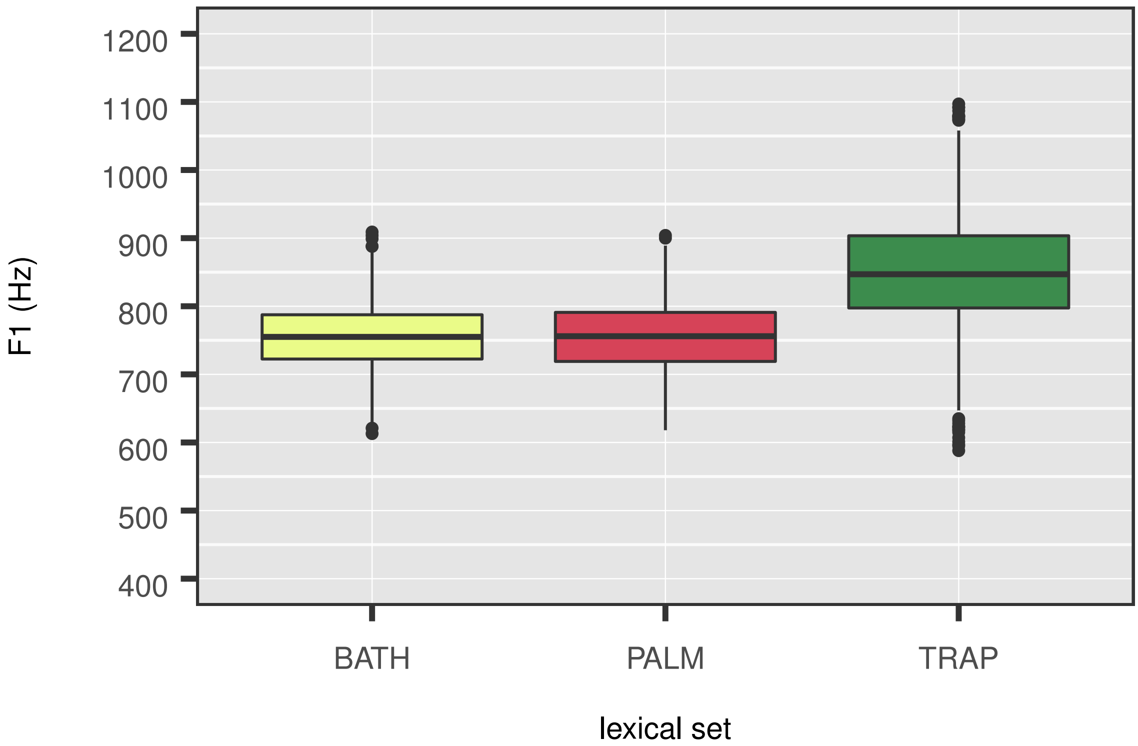

Figure 5.2.: F1 of BATH, PALM, and TRAP in CoRP-SE speakers

5.2.1.2. F2 of the CoRP-SE speakers

The best fit model for the F2 of TRAP, BATH, and PALM, can be seen in table 5.3. Similar to the model of F1 there is a difference between BATH and TRAP (+363Hz, t=18.97) but little to none between BATH and PALM (-39Hz, t = -1.81), again supporting what is stated in the literature that BATH patterns with PALM in the South East; this difference is shown in figure 5.3. The mean F2 for vowels in the BATH lexical set is 1215Hz, for PALM 1175Hz, and for TRAP 1578Hz; there is no other significant variation.

fixed effect	estimate	t-value
(Intercept)	1214.63	46.57
lexical set PALM	-39.41	-1.81
lexical set TRAP	363.19	18.97
speaker sex (sum-coded)		
Sum1	1.80	0.13
age group (sum-coded)		
Sum1	25.92	1.85
z-scored Zipf frequency (continuous)		
	3.80	0.59
style (sum-coded)		
Sum1	-25.27	-1.59
Sum2	23.53	0.90
presence of coda (sum-coded)		
Sum1	-1.73	-0.25
z-scored time (continuous)		
	-6.25	-1.46

Table 5.3.: Linear Mixed Effects Model of F2 of TRAP,BATH, and PALM in CoRP-SE speakers

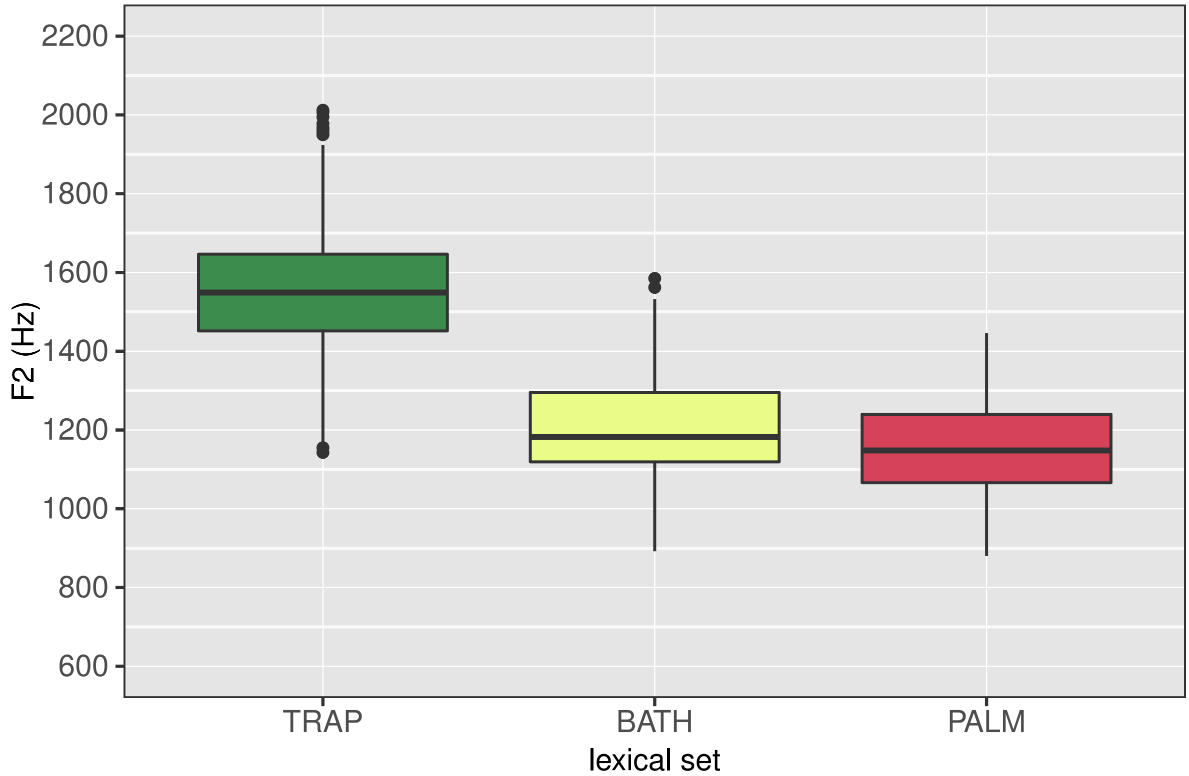

Figure 5.3.: F2 of BATH, PALM, and TRAP in CoRP-SE speakers

5.2.1.3. Duration of the CoRP-SE speakers

The model for duration of the TRAP, BATH, and PALM vowels is shown in table 5.4. The best fit model included an interaction between lexical set and presence of coda, as shown in table 5.5, but the effect is not statistically significant and is not large enough to be phonologically meaningful. The mean duration of the vowel in the BATH words is 123msec. However, the only significant effect seen is that of TRAP, which is 28msec shorter than PALM and BATH. This shows that BATH words are broadly the same length as PALM words and different to TRAP words, demonstrating that the TRAP-BATH split is also found in duration.

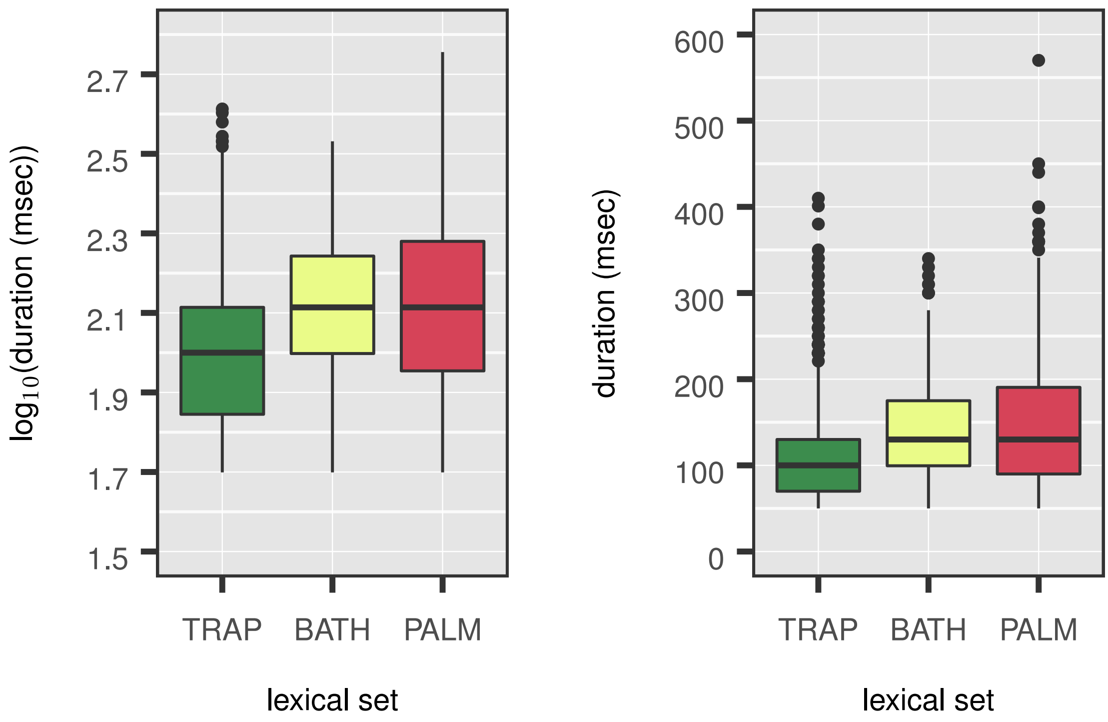

Figure 5.4.: Duration of BATH, PALM, and TRAP in CoRP-SE speakers

fixed effect	estimate	t-value
(Intercept)	2.09	30.02
lexical set PALM	0.06	0.78
lexical set TRAP	-0.12	-2.08
presence of coda in syllable (logical)		
	0.03	0.44
speaker sex (sum-coded)		
Sum1	0.02	2.48
age group (sum-coded)		
Sum1	0.00	0.33
z-scored Zipf frequency (continuous)		
	-0.00	-0.03
style (sum-coded)		
Sum1	-0.05	-1.19
Sum2	-0.04	-0.47
z-scored time (continuous)		
	0.04	1.93
following voicing (sum-coded)		
Sum1	0.01	0.82
interactions		
lexical set PALM: coda_TRUE		
	-0.01	-0.15
lexical set TRAP: coda_TRUE		
	0.05	0.74
style (sum-coded)		
Sum1: *z-scored time (continuous)*		
	-0.03	-1.36
Sum2: *z-scored time (continuous)*		
	0.02	0.52

Table 5.4.: Linear Mixed Effects Model of $\log_{10}$(duration) of TRAP, BATH, and PALM in CoRP-SE speakers

	BATH	PALM	TRAP	TRAP-BATH	PALM-BATH	TRAP-PALM
no coda	123.0269	141.2538	95.49926	-27.53	18.23	-45.75
coda	131.8257	147.9108	114.8154	-17.01	16.09	-33.10
average	127.4263	144.5823	105.1573	-22.27	17.16	-39.42

Table 5.5.: Interaction effects of lexical set and presence of coda on duration of TRAP, BATH, and PALM in CoRP-SE speakers (converted to msec)

5.2.2. The Split in DECTE speakers

In models of F1, F2, and duration, the DECTE speakers do not show significant differences between TRAP and BATH words. There are some details in variation which are outlined in the sections below. The difference between TRAP and PALM is 70Hz in F1 and 205Hz in F2. These speakers show a more spread out PALM distribution than the CoRP-SE speakers (F2 standard deviation: CoRP-SE = 117, DECTE = 198) but they show no difference between TRAP and BATH; there is no evidence of a TRAP-BATH split.

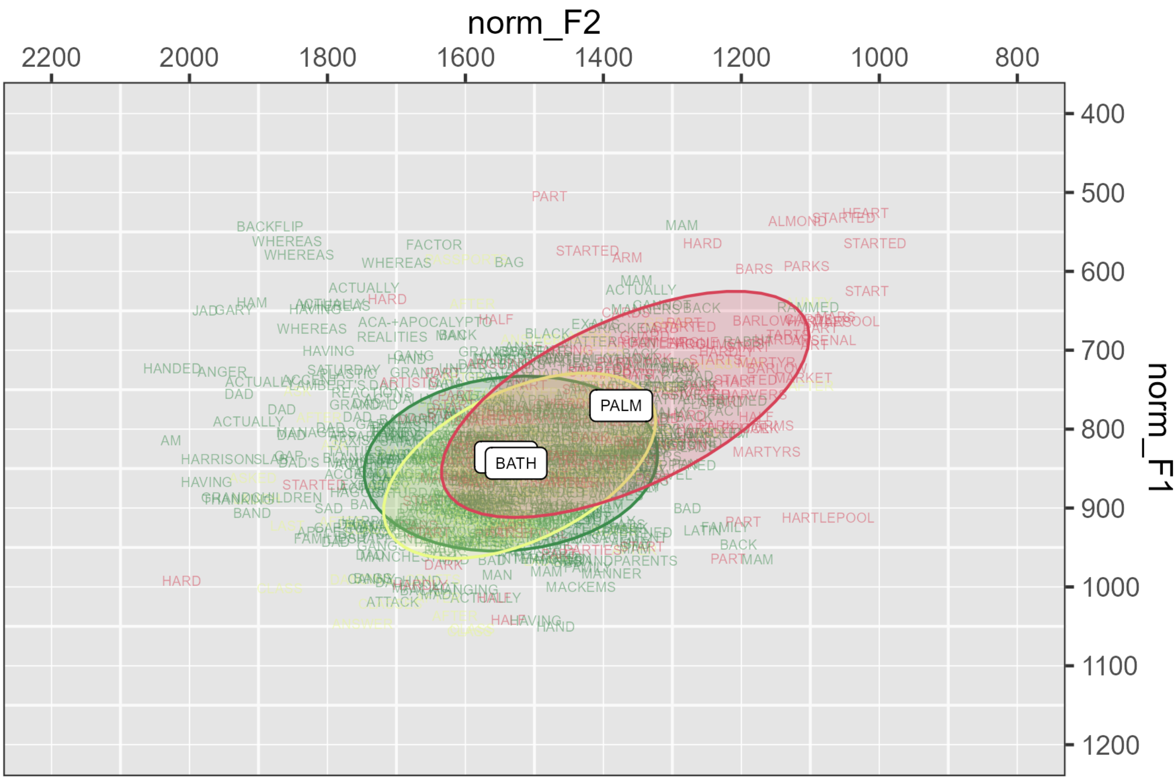

Figure 5.5.: Vowel Space plot of TRAP,BATH and PALM in the DECTE speakers

fixed effect	estimate	t-value
(Intercept)	841.89	65.90
lexical set PALM	-70.43	-4.26
lexical set TRAP	-0.90	-0.07
speaker sex (baseline female)		
male	-45.42	-2.39
age group (sum-coded)		
Sum1	23.54	6.07
z-scored Zipf frequency (continuous)		
	8.10	1.69
presence of coda (sum-coded)		
Sum1	2.41	0.48
z-scored time (continuous)		
	-4.42	-0.85
interactions		
lexical set PALM: speaker sex male	-9.28	-0.35
lexical set TRAP: speaker sex male	-1.91	-0.09

Table 5.6.: Linear Mixed Effects Model of F1 of TRAP,BATH, and PALM in DECTE speakers

5.2.2.1. F1 of the DECTE speakers

Table 5.6 shows the best fit model for F1 of the TRAP, BATH, and PALM words in the DECTE speakers. The mean F2 of the vowel in the BATH words is shown to be 842Hz, 70Hz higher (lower in the mouth) than the PALM words, which have a mean of 771 Hz, and almost identical to the TRAP words (mean = 841Hz). As would be expected from speakers in the North of England, the DECTE speakers are not showing an evidence of a TRAP-BATH split in F1.

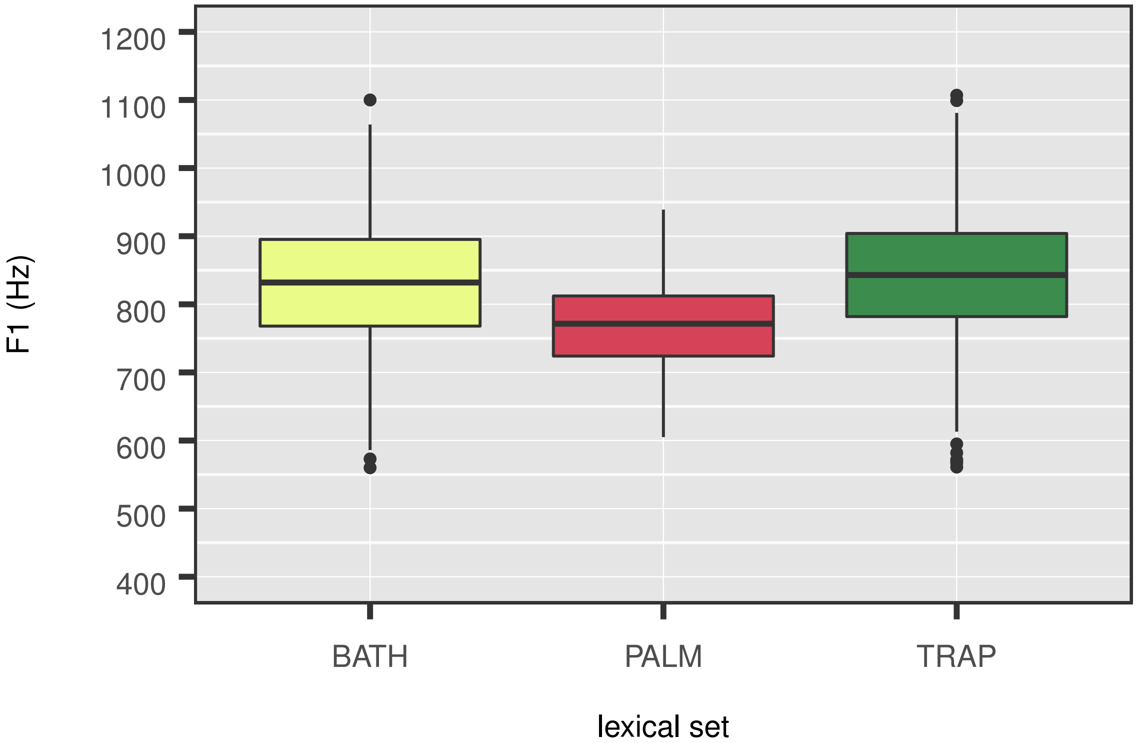

Figure 5.6.: F1 of TRAP, BATH and PALM in DECTE speakers

5.2.2.2. F2 of the DECTE speakers

The best fit model for the F2 of the TRAP, BATH, and PALM words is shown in table 5.7. The vowels in the BATH words have a mean of 1505Hz and in PALM words, 1324 Hz (difference of 181Hz), showing that BATH is further forward in the vowel space. This is supported by the negligible difference between BATH and the TRAP words, which have a mean of 1530 Hz. The difference between PALM and TRAP/BATH is also smaller than in CoRP-SE speakers (CoRP-SE = 400Hz difference). This seems to be caused by a PALM vowel that is further forwards (mean = 1324Hz) than in CoRP-SE speakers (mean = 1175); TRAP is in a similar place (CoRP-SE = 1578Hz, DECTE = 1530Hz) in both groups.

fixed effect	estimate	t-value
(Intercept)	1504.91	37.87
lexical set PALM	-180.69	-6.42
lexical set TRAP	24.61	1.06
speaker sex (sum-coded)		
Sum1	36.40	1.12
age group (sum-coded)		
Sum1	-17.55	-0.53
z-scored Zipf frequency (continuous)		
	0.74	0.08
presence of coda (sum-coded)		
Sum1	-5.21	-0.57
z-scored time (continuous)		
	-18.10	-1.99

Table 5.7.: Linear Mixed Effects Model of F2 of TRAP, BATH, and PALM in DECTE speakers

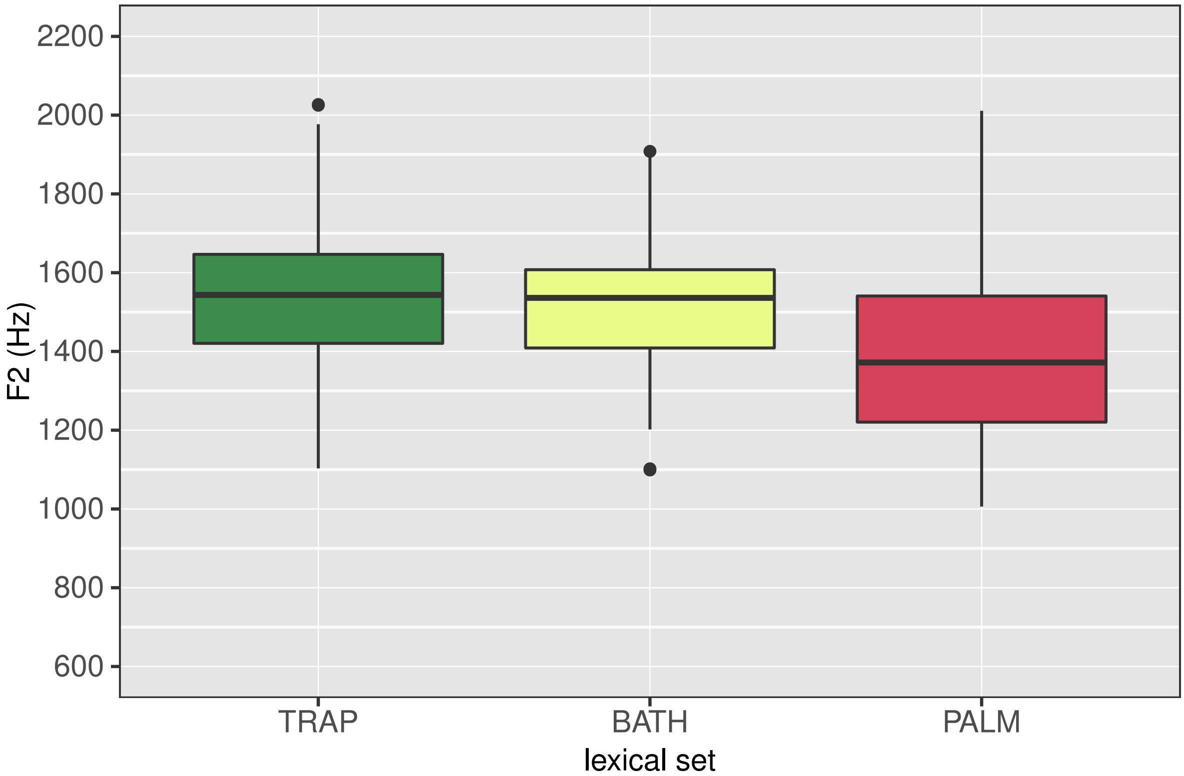

Figure 5.7.: F2 of TRAP, BATH and PALM in DECTE speakers

5.2.2.3. Duration of the DECTE speakers

fixed effect	estimate	t-value
(Intercept)	1.99	33.16
lexical set PALM	-0.04	-0.40
lexical set TRAP	-0.11	-1.86
presence of coda in syllable (logical)		
	-0.03	-0.46
speaker sex (sum-coded)		
Sum1	0.00	0.06
age group (sum-coded)		
Sum1	-0.01	-0.87
z-scored Zipf frequency (continuous)		
	-0.00	-0.21
z-scored time (continuous)		
	-0.00	-0.16
following voicing (sum-coded)		
Sum1	0.01	0.44
interactions		
lexical set PALM: coda_TRUE		
	0.14	1.26
lexical set TRAP: coda_TRUE		
	0.09	1.33

Table 5.8.: Linear Mixed Effects Model of duration of TRAP,BATH, and PALM in DECTE speakers

The best fit model of duration in the DECTE speakers can be seen in table 5.8 and the interaction calculation (converted to msec) is shown in table 5.9. The model shows that the mean vowel length in the BATH words is 94msec. Those with a coda are approximately the same length as the TRAP words with codas (-4Hz difference), though in the words without codas the difference is larger (-22msec) due to shorter TRAP vowels in open syllables . There is a difference between TRAP and PALM (-20msec), however, all of these values are closer to the value for the TRAP words in the CoRP-SE speakers than to the PALM and BATH values in the CoRP-SE speakers demonstrating that not only do the DECTE speakers not show a TRAP-BATH split in duration, they also generally do not have a long PALM vowel. This supports the conclusion drawn from the variation found in F2 that the PALM vowel is not consistently the same in DECTE speakers as in the CoRP-SE speakers.

	BATH	PALM	TRAP	TRAP-BATH	PALM-BATH	TRAP-PALM
no coda	97.72372	89.12509	75.85776	-21.87	-8.60	-13.27
coda	91.20108	114.8154	87.09636	-4.10	23.61	-27.72
mean	94.4624	101.9702	81.47706	-12.99	7.51	-20.49

Table 5.9.: Interaction effects of lexical set and presence of coda on TRAP, BATH, and PALM in DECTE speakers

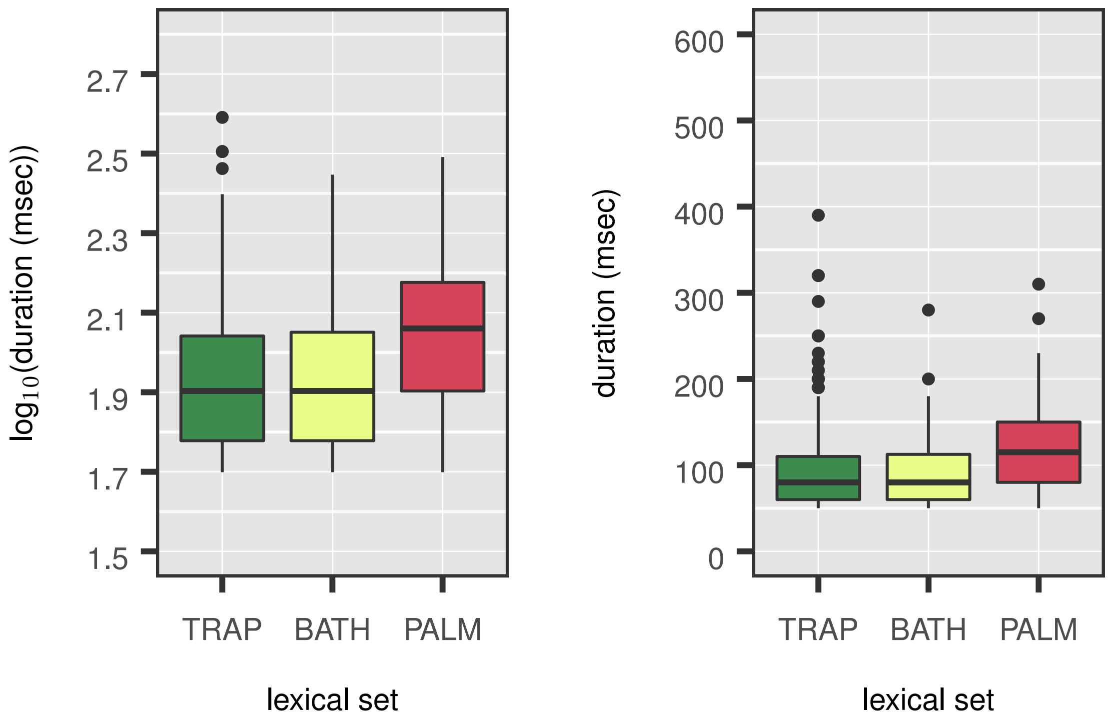

Figure 5.8.: Duration of TRAP, BATH and PALM in DECTE speakers

5.2.3. The Split in CoRP-NE speakers

The CoRP-NE speakers show little to no difference in F1 or F2 that could be evidence for a TRAP-BATH split (though F2 is a little more complex, details below). The difference between TRAP and PALM in F1 is 86Hz and in F2 is 329Hz.

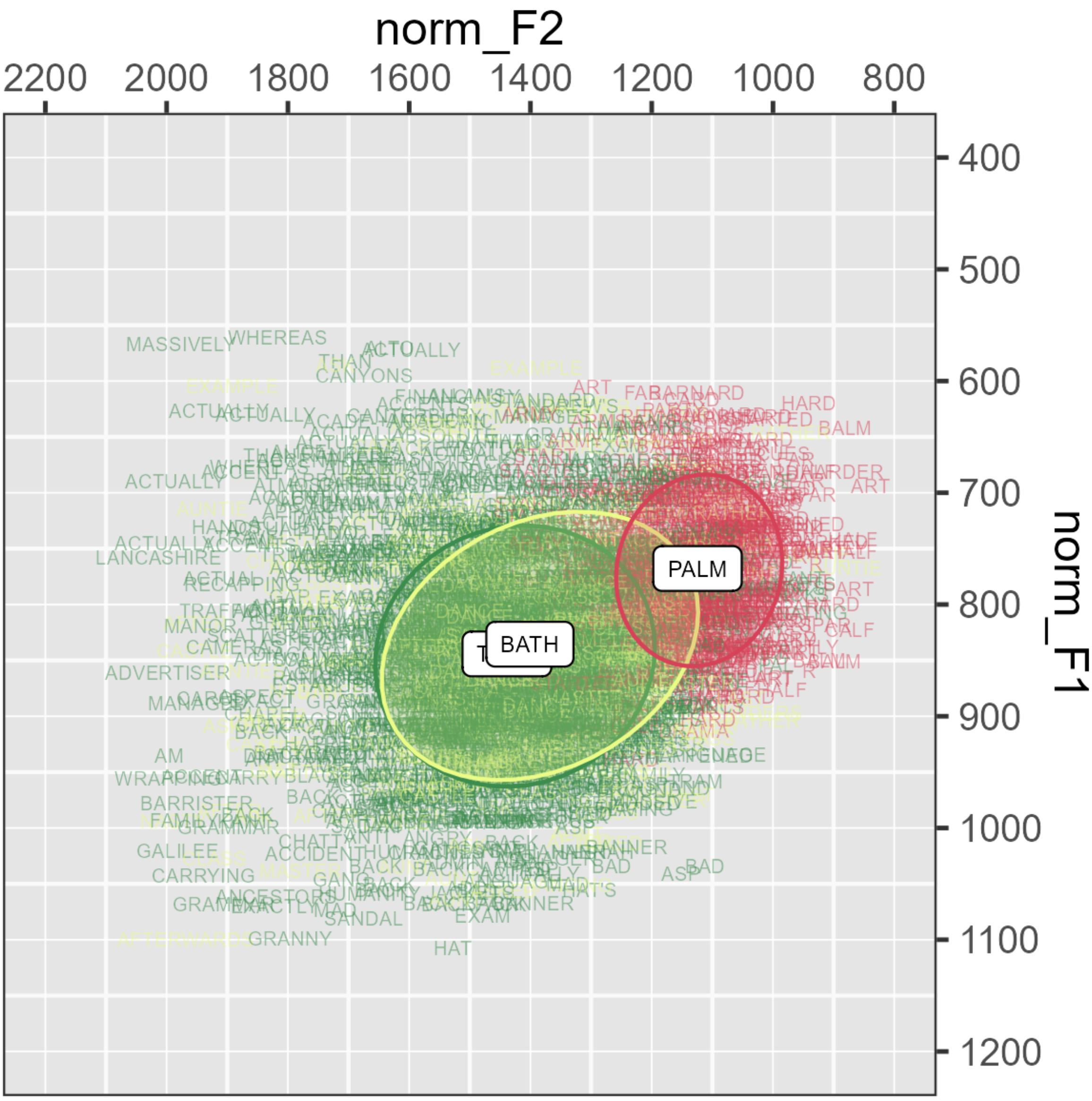

Figure 5.9.: Vowel Space plot of TRAP,BATH and PALM in the CoRP-NE speakers

5.2.3.1. F1 of the CoRP-NE speakers

Table 5.10 shows the best fit model for the CoRP-NE speakers. There is a three way interaction between lexical set, speaker sex, and speaker age group, which is summarised in full in table 5.11. However, many of the interaction terms, while they improve the fit of the model (according to cAIC), do not have a t value greater than 2.00; table 5.12 shows the interactions excluding these terms. It can be seen in table 5.12 that the height difference between TRAP and PALM is larger for the Old Female speakers than for any other speakers. This is caused by a lower TRAP vowel in that speaker group. All speakers have a BATH vowel 53Hz lower than the PALM vowel, for most speakers this is a similar position to TRAP, except for the Old Female speakers for whom it sits between the heights of TRAP and PALM.

Overall it can be concluded from this analysis that the CoRP-Ne speakers do not show a significant TRAP-BATH in acoustic vowel height.

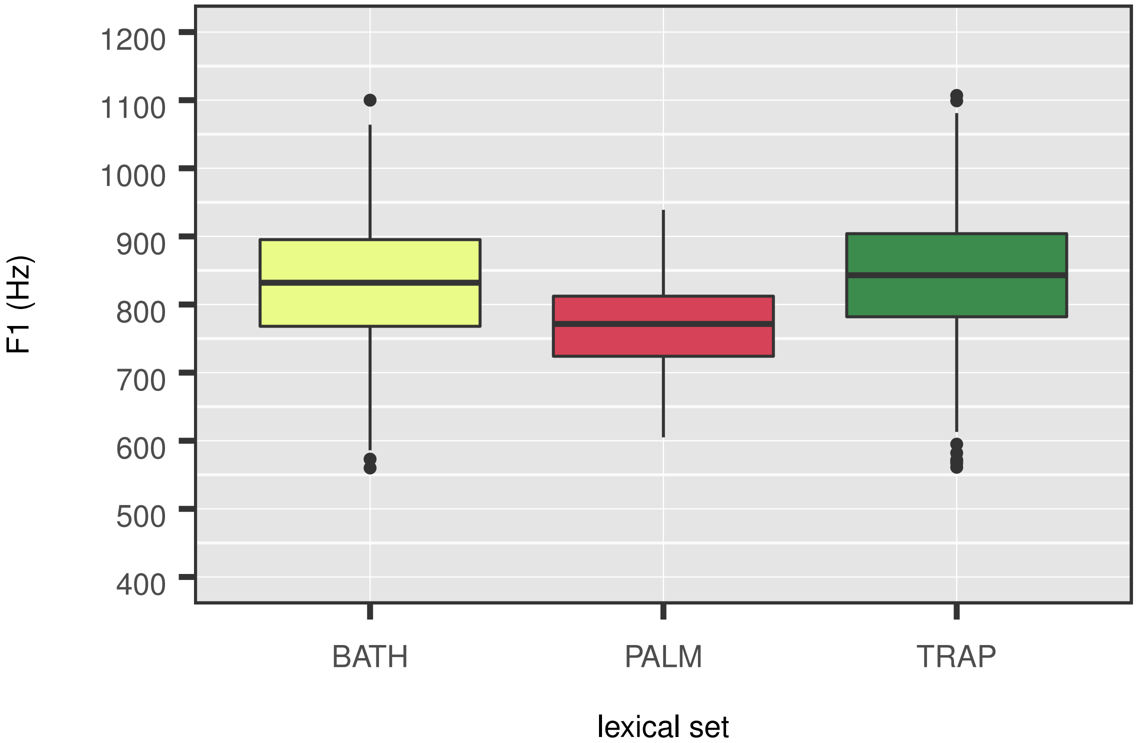

Figure 5.10.: F1 of TRAP, BATH and PALM in CoRP-NE speakers

fixed effect	estimate	t-value
(Intercept)	814.30	43.81
lexical set PALM	-52.55	-2.88
lexical set TRAP	78.88	4.97
speaker sex (baseline female)		
male	10.88	0.40
age group (baseline old)		
young	45.78	2.41
z-scored Zipf frequency (continuous)		
	1.58	0.53
style (sum-coded)		
Sum1	-16.23	-2.16
Sum2	-4.94	-0.39
presence of coda (sum-coded)		
Sum1	1.69	0.52
z-scored time (continuous)		
	1.50	0.64
interactions		
lexical set PALM: speaker sex male	15.55	0.55
lexical set TRAP: speaker sex male	-55.72	-2.34
lexical set PALM: age group young	-9.22	-0.48
lexical set TRAP: age group young	-74.15	-4.47
sex male: age group young	-33.80	-1.10
lexical set PALM: sex male: age group young	-30.44	-0.98
lexical set TRAP: sex male: age group young	59.93	2.25

Table 5.10.: Linear Mixed Effects Model of F1 of TRAP, BATH, and PALM in CoRP-NE speakers

	BATH	PALM	TRAP	TRAP-BATH	PALM-BATH	TRAP-PALM
Old Female	814.30	761.75	893.18	78.88	-52.55	131.43
Old Male	825.18	788.18	848.34	23.16	-37.00	60.16
Young Female	860.08	798.31	864.81	4.73	-61.77	66.50
Young Male	837.16	760.50	846.10	8.94	-76.66	85.60
Mean	834.18	777.19	863.11	28.93	-57.00	85.92

Table 5.11.: Interaction effects of lexical set, speaker sex and speaker age group on F1 of TRAP, BATH, and PALM in CoRP-NE speakers (calculated from interactions in table 5.10)

	BATH	PALM	TRAP	TRAP-BATH	PALM-BATH	TRAP-PALM
Old Female	814.30	761.75	893.18	78.88	-52.55	131.43
Old Male	814.30	761.75	837.46	23.16	-52.55	75.71
Young Female	814.30	761.75	819.03	4.73	-52.55	57.28
Young Male	814.30	761.75	823.24	8.94	-52.55	61.49
Mean	814.30	761.75	843.23	28.93	-52.55	81.48

Table 5.12.: Interaction effects of lexical set, speaker sex and speaker age group on F1 of TRAP, BATH, and PALM in CoRP-NE speakers (not including terms with t<2.00)

5.2.3.2. F2 of the CoRP-NE speakers

Table 5.13 shows the best fit model for F2 of the CoRP-NE speakers; the model included a three way interaction between lexical set, speaker sex, and speaker age group, the results of which are shown in table 5.14, and in table 5.15 with terms with t<2.00 removed.

Before considering the TRAP-BATH split in frontness in these speakers, it is important to note that the mean PALM value for the CoRP-NE speakers is 1119Hz, which is similar to the mean PALM vowel in CoRP-SE speakers (1175Hz) and further back than the mean PALM vowel of the DECTE speaker (1324Hz). This means that the vowels at the bottom of the vowel space are further apart in both CoRP groups of speakers than in the DECTE speakers.

Within the interaction effects taken into account (using table 5.15) it can be seen that the main variation within this speaker group is seen in the Old Male group (only one speaker), who has a TRAP vowel that is further forwards than the other speakers. Despite this, the TRAP-BATH difference is usually 60-80Hz (and approximately 190Hz for that one speakers), and the PALM-BATH difference is 215Hz demonstrating that these speakers have a BATH vowel that is far closer to TRAP than to PALM. The difference of 60-80Hz could be caused by individual BATH words being realised with a PALM vowel rather than a shift of the entire lexical set. Further analysis of the BATH vowel alone can be seen in section 5.3 below.

fixed effect	estimate	t-value
(Intercept)	1334.28	29.30
lexical set PALM	-214.93	-5.86
lexical set TRAP	80.85	2.52
speaker sex (baseline female)		
male	5.92	0.08
age group (baseline old)		
young	45.92	0.92
presence of coda (sum-coded)		
Sum1	5.25	0.73
z-scored Zipf frequency (continuous)		
	-1.75	-0.26
z-scored time (continuous)		
	-2.45	-0.58
interactions		
lexical set PALM: speaker sex male	24.34	0.46
lexical set TRAP: speaker sex male	106.53	2.36
lexical set PALM: age group young	-28.76	-0.79
lexical set TRAP :age group young	-26.68	-0.84
sexmale: age group young	39.27	0.46
lexical set PALM :sex male: age group young	-89.44	-1.52
lexical set TRAP :sex male:age group young	-124.64	-2.48

Table 5.13.: Linear Mixed Effects Model of F2 of TRAP, BATH, and PALM in CoRP-NE speakers

	BATH	PALM	TRAP	TRAP-BATH	PALM-BATH
Old Female	1334.28	1119.35	1415.13	80.85	-214.93
Old Male	1343.20	1152.61	1530.58	187.38	-190.59
Young Female	1380.20	1136.51	1434.37	54.17	-243.69
Young Male	1428.39	1119.60	1464.45	36.06	-308.79
Mean	1371.52	1132.02	1461.13	89.61	-239.50

Table 5.14.: Interaction effects of lexical set, speaker sex and speaker age group on F2 of TRAP, BATH, and PALM in CoRP-NE speakers (calculated from interactions in table 5.13)

	BATH	PALM	TRAP	TRAP-BATH	PALM-BATH	TRAP-PALM
Old Female	1334.28	1119.35	1415.13	80.85	-214.93	295.78
Old Male	1334.28	1119.35	1521.66	187.38	-214.93	402.31
Young Female	1334.28	1119.35	1415.13	80.85	-214.93	295.78
Young Male	1334.28	1119.35	1397.02	62.74	-214.93	277.67
Mean	1334.28	1119.35	1437.24	102.96	-214.93	317.89

Table 5.15.: Interaction effects of lexical set, speaker sex and speaker age group on F2 of TRAP, BATH, and PALM in CoRP-NE speakers (not including terms with t<2.00)

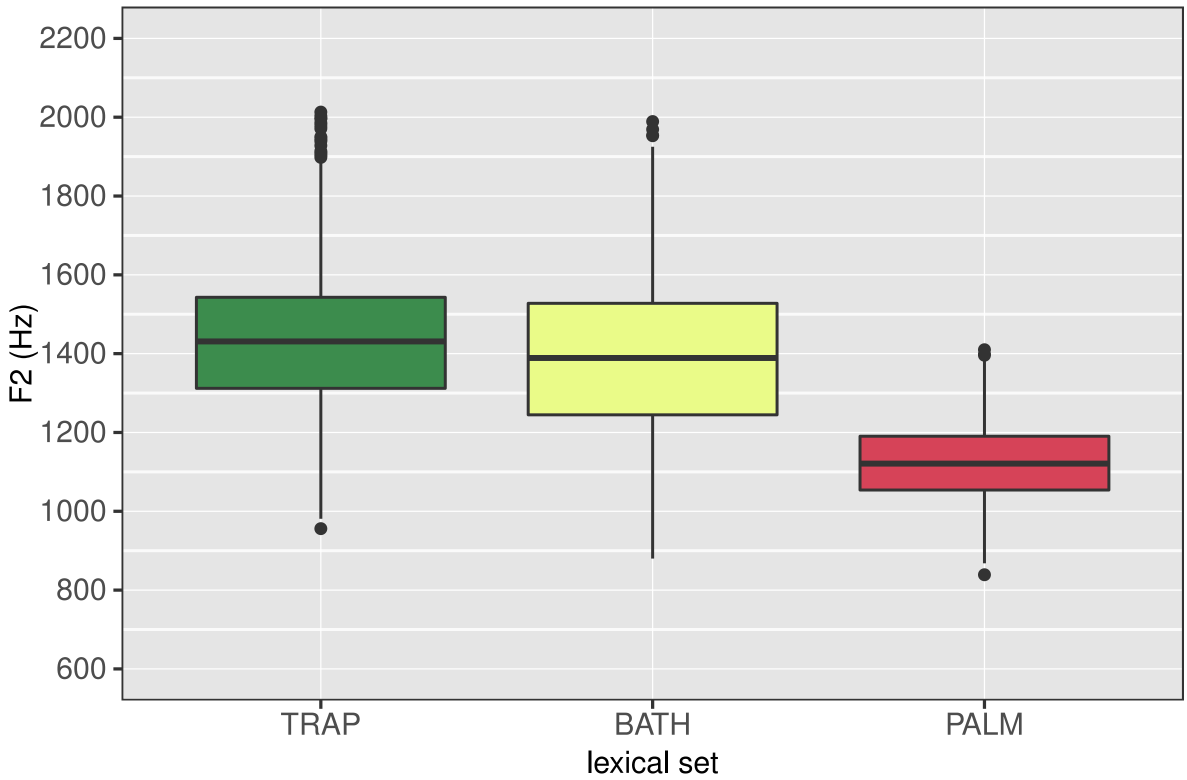

Figure 5.11.: F2 of TRAP, BATH and PALM in CoRP-NE speakers

5.2.3.3. Duration of the CoRP-NE speakers

Table 5.16 shows the best fit model for duration in the CoRP-NE speakers. An interaction effect is seen between lexical set and presence of a coda. This interaction is summarised in table 5.17. TRAP and PALM are on average 62 msec apart, and the PALM length is closer to the length found in the CoRP-SE speakers, implying that CoRP-NE speaker do not have the shorter PALM vowel found in the DECTE speakers. There is a TRAP-BATH difference present in the words with a coda (-20msec) but none in words without a coda (visualised in figure 5.12). The difference is less than the difference between TRAP-PALM, and while the mean duration of the BATH (136msec) words is longer than the TRAP words (126msec) it is not as long as the PALM words (189msec). In a form similar to the F2 results above, the duration of the BATH values could be interpreted as a target that is slightly longer than TRAP, or as some individual BATH words having a PALM target and so pulling the mean duration higher.

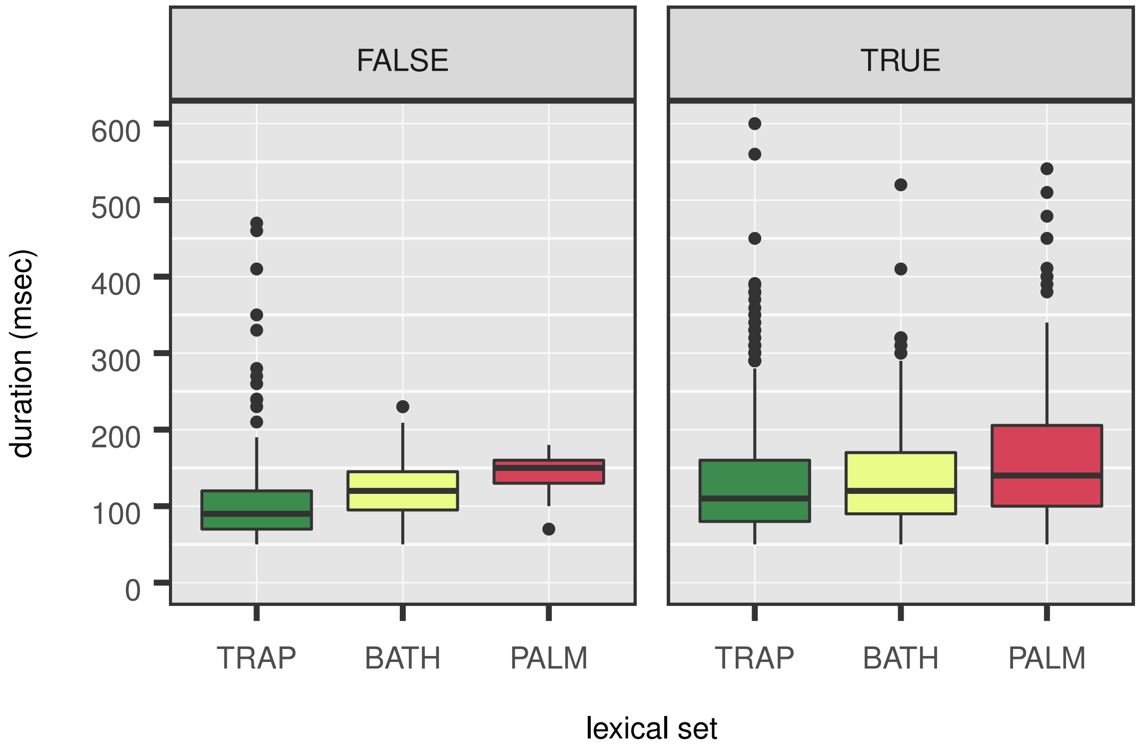

Figure 5.12.: Duration of TRAP, BATH and PALM in CoRP-NE speakers

fixed effect	estimate	t-value
(Intercept)	2.14	35.52
lexical set PALM	0.16	1.49
lexical set TRAP	-0.07	-1.25
presence of coda in syllable (logical)		
	-0.01	-0.22
speaker sex (sum-coded)		
Sum1	0.02	1.68
age group (sum-coded)		
Sum1	0.00	0.28
z-scored Zipf frequency (continuous)		
	0.01	1.27
style (sum-coded)		
Sum1	-0.11	-5.35
Sum2	0.06	1.74
z-scored time (continuous)		
	0.00	0.38
following voicing (sum-coded)		
Sum1	-0.01	-0.85
interactions		
lexical set PALM: coda_TRUE		
	-0.04	-0.34
lexical set TRAP: coda_TRUE		
	0.07	1.14

Table 5.16.: Linear Mixed Effects Model of duration of TRAP,BATH, and PALM in CoRP-NE speakers

	BATH	PALM	TRAP	TRAP-BATH	PALM-BATH	TRAP-PALM
no coda	138.0384	199.5262	117.4898	-20.55	61.49	-82.04
coda	134.8963	177.8279	134.8963	0.00	42.93	-42.93
mean	136.4674	188.6771	126.193	-10.27	52.21	-62.48

Table 5.17.: Interaction effects of lexical set and presence of coda on duration of TRAP, BATH, and PALM in CoRP-NE speakers.

5.2.4. Conclusions on the nature of the TRAP-BATH split in all three speaker groups

As attested in the literature, the analysis above has confirmed that the major feature of the TRAP-BATH split is a distinction in vowel frontness, as shown by an F2 difference, and duration. The TRAP-BATH split is the BATH words being found in a PALM-like position as opposed to a TRAP-like position (this is a change, as outlined in chapter 2 , but as a phonemic split is stable in the South of England). For CoRP-SE speakers it was found that the difference between TRAP and BATH is characterised by a difference of 98Hz in F1, 363 Hz in F2 and -28msec in duration.

There are two points of variation that impact analysis of the split but may not be required characteristics of it. First, in the CoRP-SE speakers who have a clear TRAP-BATH split in F2 and duration, there is the potential for variation in the height of TRAP. In this speaker group an F1 difference is found between TRAP and BATH/PALM, which may be an indicator of the split but is not focussed on as a primary characteristic. Figure 5.1 shows that the F1 variation in TRAP is larger than in either of BATH or PALM. Secondly, in the DECTE speakers PALM is further forward than than in the CoRP-SE speakers. Even without considering BATH, the difference in PALM between DECTE and CoRP-SE means that there are two different low vowel systems that CoRP-NE speakers are exposed to, as well as different distribution of BATH words between the two possible positions. If the CoRP-NE speakers showed a similar position to DECTE speakers it would impact analysis of the split as a movement from a TRAP-like position to a PALM-like position.

The DECTE speakers do not show a split, there is a statistically significant but not linguistically meaningful meaningful F1 difference of +1HZ and F2 difference of +24Hz, and a duration difference of +13msec.

Without considering BATH the CoRP-NE speakers seem to show a low vowel system like CoRP-SE. However, with respect to BATH, they broadly pattern with the DECTE speakers, though with some important differences. They show an F1 difference of +28Hz, an F2 difference of 60-80Hz, and a duration difference of -20Hz (if the vowel is followed by a coda). However, this speaker group is not identical to the DECTE speakers, as mentioned, the PALM vowel is closer to the CoRP-SE speakers, and there is some variation within the BATH vowel, as seen by a higher F2 and duration difference than in the DECTE speakers. This difference will be discussed in the analysis of the BATH vowel below.

5.3. BATH vowel alone

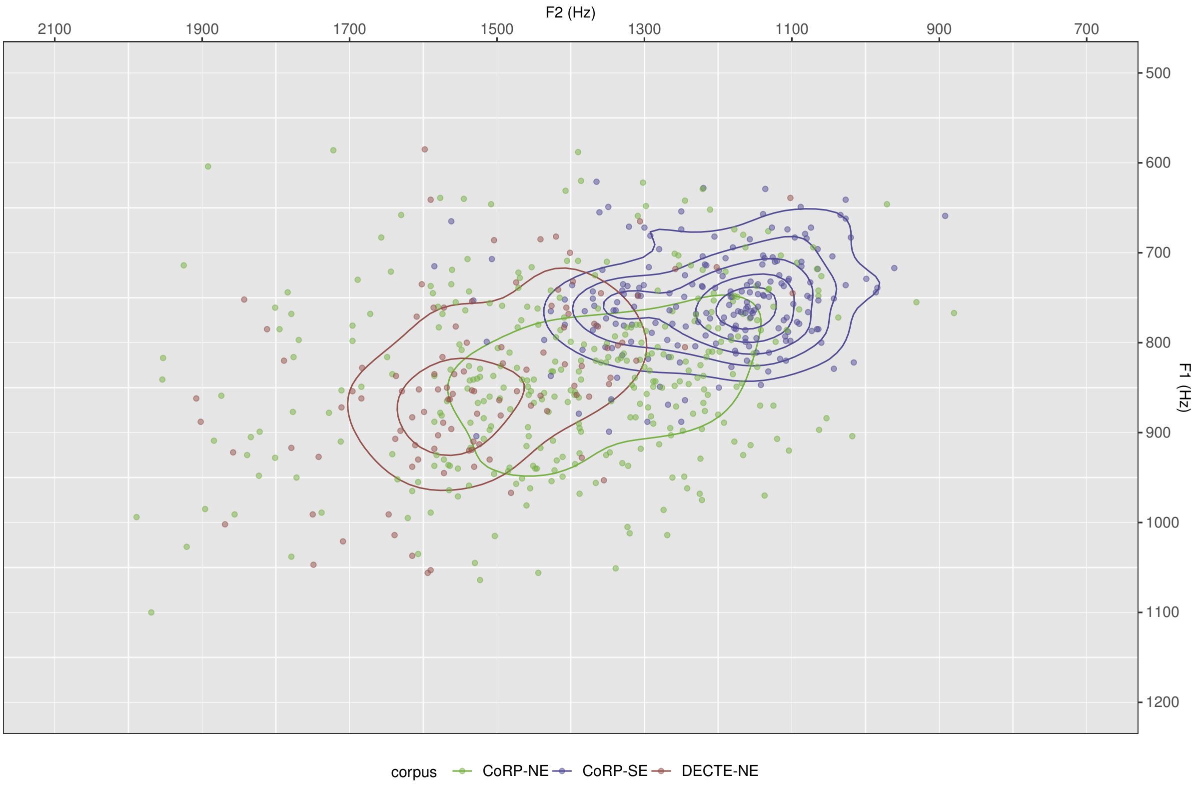

Figure 5.13.: BATH in all three speaker groups

5.3.1. F1

The best fit model of BATH in all speaker groups is shown in table 5.18. There is a three way interaction between sex, age group and corpus, shown in table 5.19. It can be seen that overall the CoRP-NE speakers have a BATH vowel at 828Hz, very similar to the height of the DECTE speakers (837Hz), whereas the CoRP-SE speakers have a higher vowel, at 762Hz. While there is variation between male and female, and old and young, none of this variation reaches overlap between the lower vowels (CoRP-NE and DECTE), and the higher vowels (CoRP-SE). Despite height not acting as the primary indicator of historical BATH movement, these values align with the predicted TRAP and PALM positions in section 5.2 respectively. The difference further demonstrates that in all north-eastern speakers (DECTE and CoRP-NE) the BATH vowel has not moved and remains in the same place as TRAP whereas in the CoRP-SE speakers it has moved to the PALM position.

fixed effect	estimate	t-value
(Intercept)	815.06	38.48
DECTE	61.18	2.53
CoRP-SE	-68.19	-2.82
age group (baseline old)		
young	40.96	1.96
speaker sex (baseline female)		
male	2.08	0.07
z-scored Zipf frequency (continuous)		
	0.78	0.12
style (sum-coded)		
Sum1	-23.64	-2.19
Sum2	1.15	0.06
presence of coda (sum-coded)		
Sum1	-12.15	-1.69
z-scored time (continuous)		
	-4.63	-1.05
interactions		
corpus DECTE: age group young	-78.54	-2.30
corpus CoRP-SE: age group young	-21.00	-0.70
corpus DECTE: speaker sex male	-40.70	-1.05
corpus CoRP-SE: speaker sex male	19.85	0.52
age group young: speaker sex male	-25.99	-0.78
corpus DECTE: age group young: speaker sex male	-8.42	-0.15
corpus CoRP-SE: age group young: speaker sex male	-7.77	-0.15

Table 5.18.: Linear Mixed Effects Model of F1 of BATH

	CoRP-NE	DECTE-NE	CoRP-SE	Mean
Old Female	815.06	876.24	746.87	812.72
Old Male	817.14	837.62	768.80	807.85
Young Female	856.02	838.66	766.83	820.50
Young Male	832.11	791.62	755.00	792.91
Mean	830.08	836.04	759.38	808.50

Table 5.19.: F1 of BATH vowel for all three speaker groups

5.3.2. F2

Table 5.20 shows the best fit model for F2 of the vowel in BATH words in all three speaker groups. Modelling the BATH words alone shows that the CoRP-NE speakers have a vowel with F2 between the CoRP-SE speakers (-144Hz lower) and DECTE speakers (173Hz higher) (see figure). From this model it is difficult to tell if this is truly a vowel with a mean in between or the effect of both positions existing with the set of tokens.

fixed effect	estimate	t-value
(Intercept)	1346.66	38.65
DECTE	173.11	3.62
CoRP-SE	-143.53	-3.12
speaker sex (baseline female)		
male	40.09	0.92
age group (sum-coded)		
Sum1	-5.17	-0.32
z-scored Zipf frequency (continuous)		
	26.29	1.58
presence of coda (sum-coded)		
Sum1	-24.64	-1.27
z-scored time (continuous)		
	-5.45	-0.72
interactions		
corpus DECTE: speaker sex male	-148.92	-2.07
corpus CoRP-SE: speaker sex male	-50.06	-0.69

Table 5.20.: Linear Mixed Effects Model of F2 of BATH

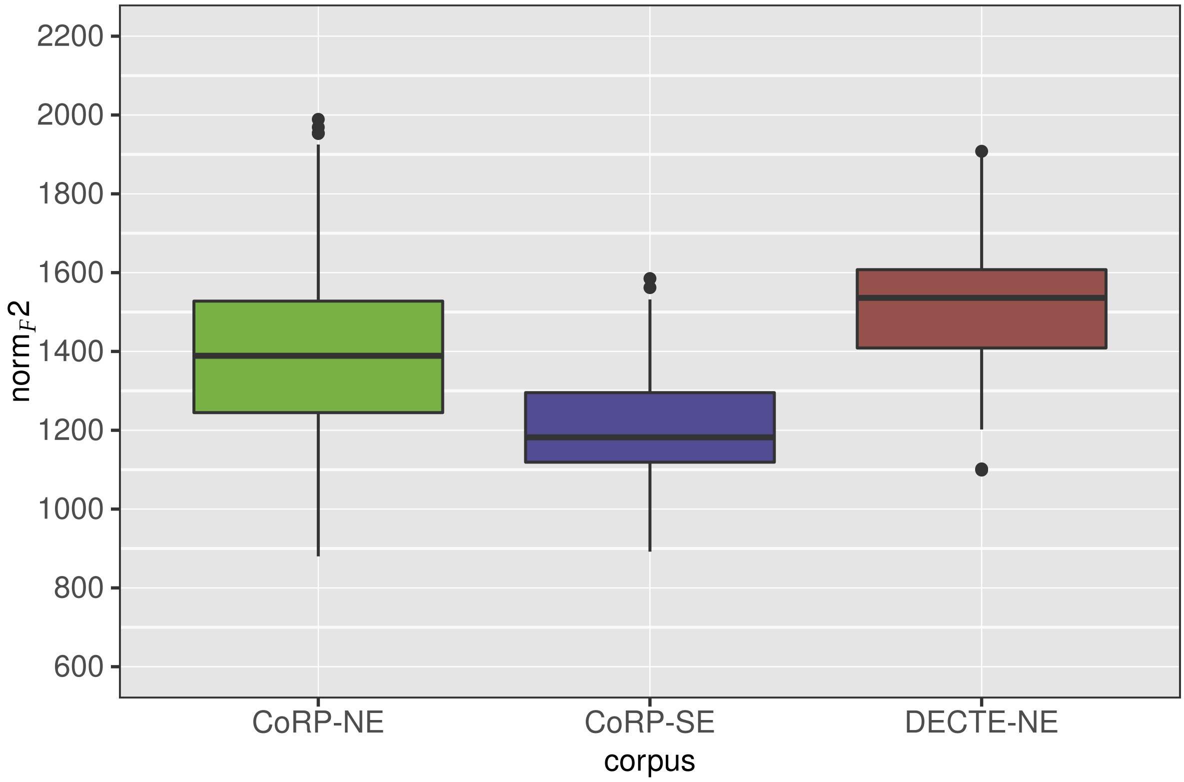

Figure 5.14.: BATH vowel only in all three speaker groups

5.3.3. Duration

Table 5.21 shows the best fit model for the duration of the BATH vowel, and the interaction effect is summarised in table 5.22. These show that the BATH vowel duration is not patterning in the same way as F2, instead we see that the CoRP-NE speakers have a similar (though not identical in value or coda interaction) duration to CoRP-SE speakers, whereas DECTE speakers have a shorter duration. These differences are not large but considering the history of the TRAP-BATH split, which began with a change in length, it is possible that this difference is reflective of the beginning of a split in the CoRP-NE speaker group.

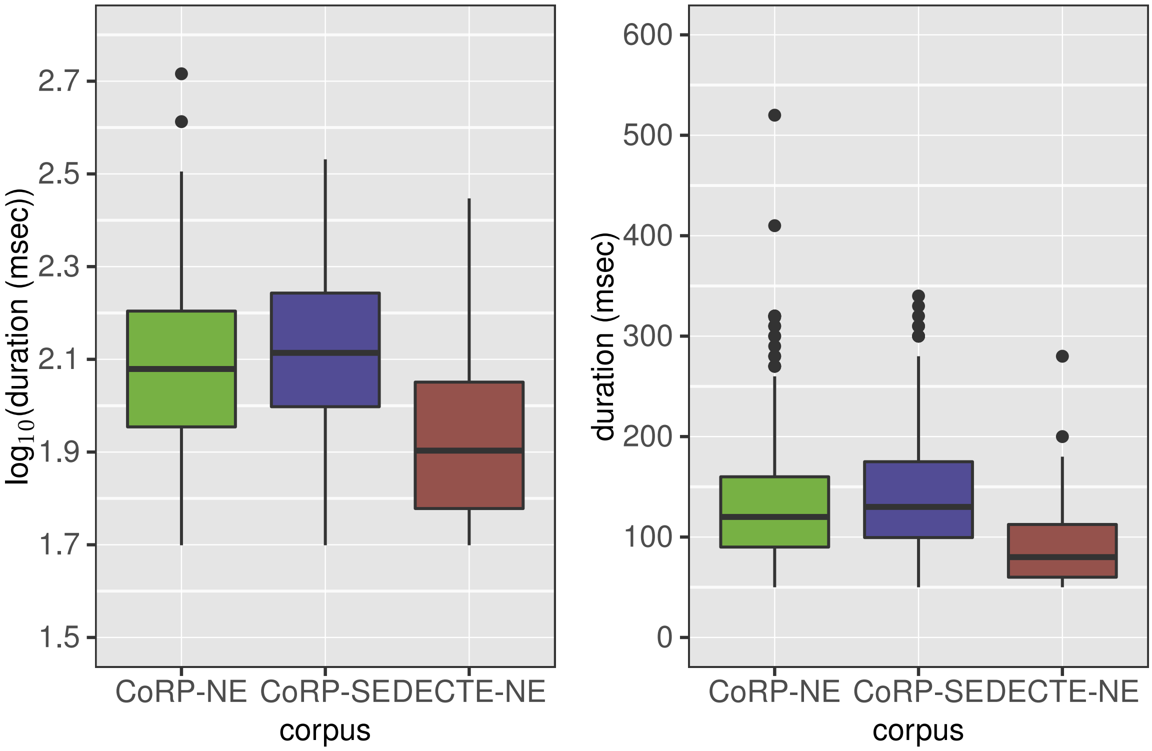

Figure 5.15.: BATH

fixed effect	estimate	t-value
(Intercept)	2.15	38.53
DECTE	-0.12	-1.88
CoRP-SE	0.00	0.01
presence of coda in syllable (logical)		
	-0.02	-0.42
speaker sex (sum-coded)		
Sum1	0.03	2.13
age group (sum-coded)		
Sum1	0.01	0.43
z-scored Zipf frequency (continuous)		
	0.03	1.56
style (sum-coded)		
Sum1	-0.12	-3.99
Sum2	0.07	1.32
z-scored time (continuous)		
	-0.00	-0.35
following voicing (sum-coded)		
Sum1	-0.05	-3.10
interactions		
corpus DECTE: coda_TRUE		
	0.03	0.52
corpus CoRP-SE: coda_TRUE		
	0.05	1.05

Table 5.21.: Linear Mixed Effects Model of $\log_{10}$(duration) of BATH, in all three speaker groups

	CoRP-NE	CoRP-SE	DECTE
has_codaFALSE	141.2538	141.2538	107.1519
has_codaTRUE	134.8963	151.3561	109.6478
Mean	138.075	146.3049	108.3999

Table 5.22.: Interaction effects on the duration of BATH in all three speaker groups (calculated from $\log_{10}$(duration) in table 5.21)

5.4. Discussion and Conclusion

The parameters of the TRAP-BATH split in the CoRP-SE speakers are clearly canonical, supporting defining the TRAP-BATH split as found in F2 and duration (with some impact of F1). DECTE speakers clearly do not show a TRAP-BATH split, with no significant difference in F2 or duration between TRAP and BATH.

In analysis of F2 and duration of the BATH vowels that it is difficult to tell if the variation seen in the CoRP-NE speakers is the result of a different BATH target or individual speakers, or s behaving differently. This question is discussed further below. At this point it is important to consider the linguistic status of the variables being considered, not just the modelled values. Two relevant points about the BATH vowel are important to consider to understand this analysis. First, a BATH realisation between TRAP and PALM is not attested in the literature. Secondly, TRAP-BATH split is a move of a subset of the TRAP lexical set (in pre-fricative and pre-nasal environment) from one existing target (a low front vowel) to another existing target (*long* low back vowel, i.e. the PALM position). Despite the CoRP-NE speakers showing a PALM target more like the CoRP-SE speakers, the BATH vowel would still not be a new vowel in their inventory so there is no reason for an intermediate vowel quality to appear. There is the remote possibility that the CoRP-NE speakers are producing BATH in the position where DECTE speakers would have PALM but this would require combining targets from two different systems. Therefore, unlike what is seen in the STRUT vowel in chapter 4, it is unlikely that CoRP-NE speakers are creating a new split.

In addition to the model of F2 of BATH in section 5.3.2, it is possible to understand this data further by visualising the vowel positions and including individual speaker means. Figure 5.11 (reproduced below as figure 5.16), split by speaker sex to aid in seeing individual speakers, not to show any particular pattern) shows the CoRP-NE speakers alone and looking at the speaker means and connections between them it can be seen that BATH consistently patterns with TRAP, there is no evidence of a TRAP-BATH split in these speakers.

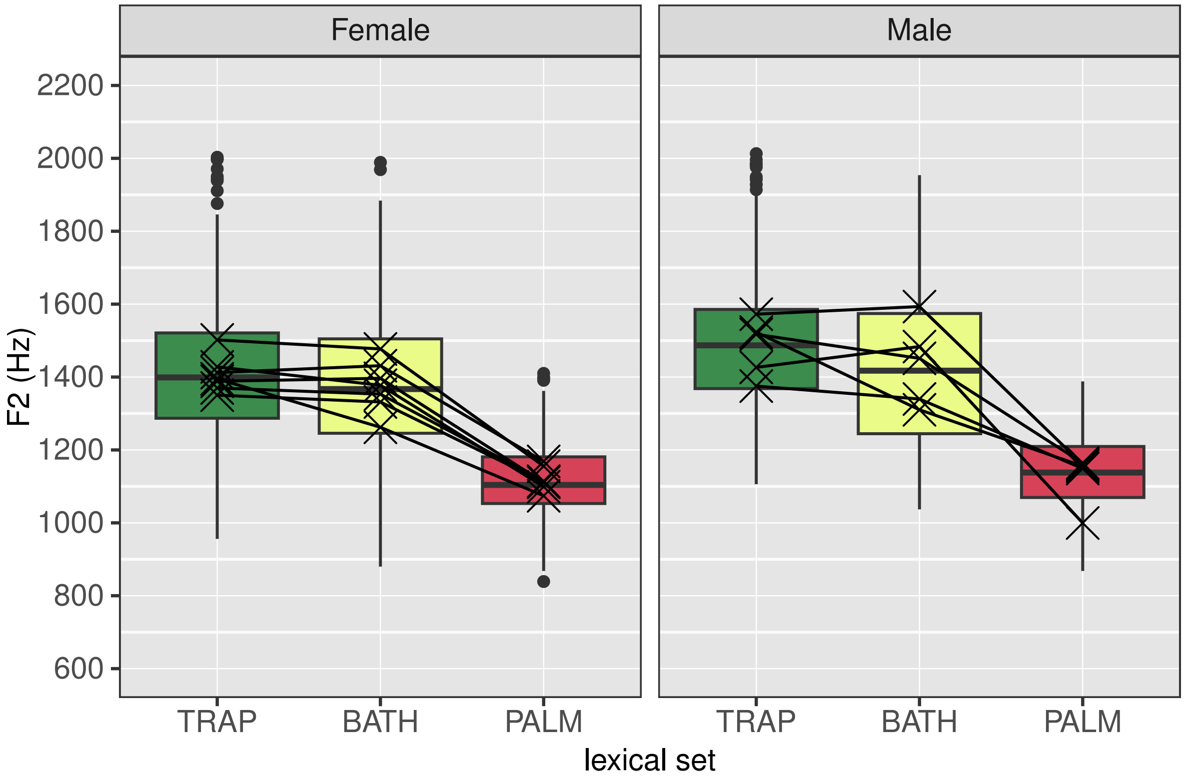

Figure 5.16.: Figure 5.11 including speaker means

Considering figure 5.14 (reproduced below), it is clear why the model shows a BATH value for CoRP-NE that is in between those for CoRP-SE and DECTE. The CoRP-NE BATH vowels are wider in distribution than either of the other two speaker groups, with some overlap in both direction, implying that some speakers in CoRP-NE produce BATH words with a PALM-like realisation.

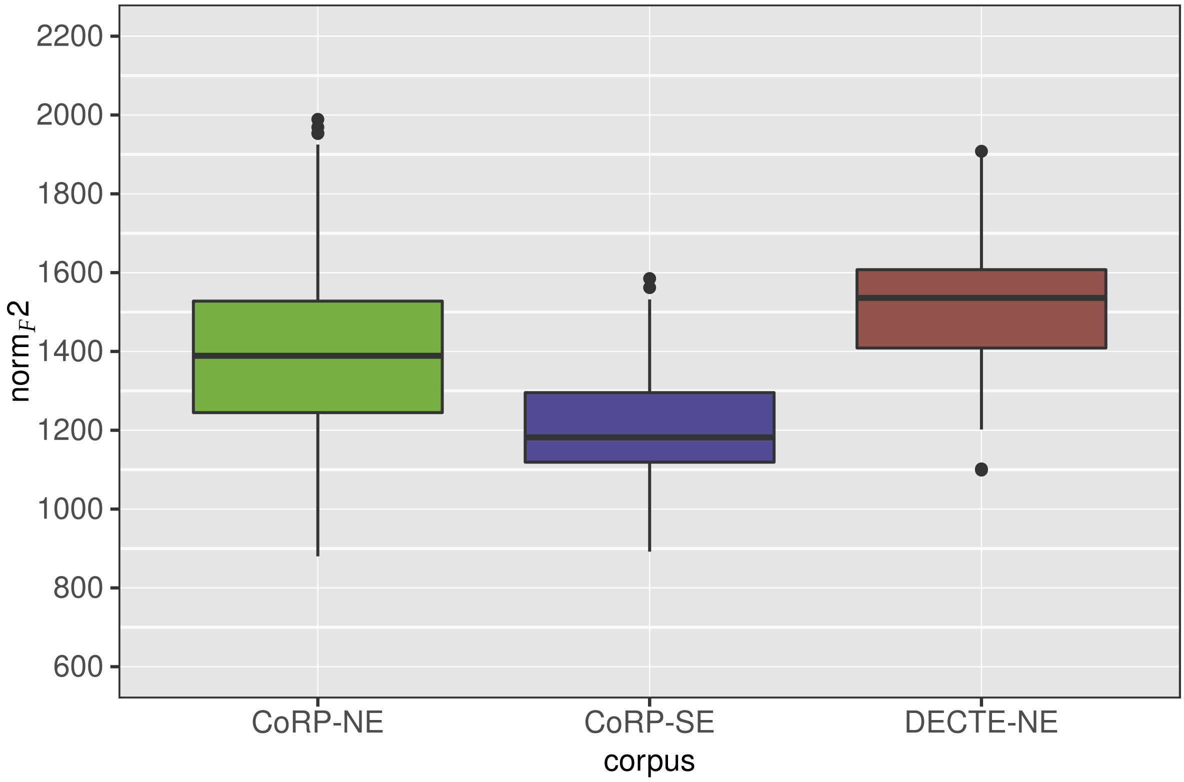

Figure 5.17.: Figure 5.14

However, a figure that looks at all the data together provides a slightly different perspective. Figure 5.18 shows that within the CoRP-NE speaker group there exist BATH vowels that are in overlap with the CoRP-SE BATH and PALM position. For example see the lowest F2 value for BATH from CoRP-NE. However, if that speaker is taken as an example, they show a PALM vowel with a lower F2 again, demonstrating that their BATH is not produced in their PALM position.

It is possible that the spread of F2 values for CoRP-NE speakers is due to individual words. It is known that the TRAP-BATH split is not consistent within the pre-fricative context, even in southern speakers who maintain the split (see chapter 2), e.g. *gas* vs *grass*. In CoRP-NE speakers we may be seeing the effect of some "BATH' words remaining in the TRAP position, but others moving to the PALM position. As discussed in chapter 2, it is very

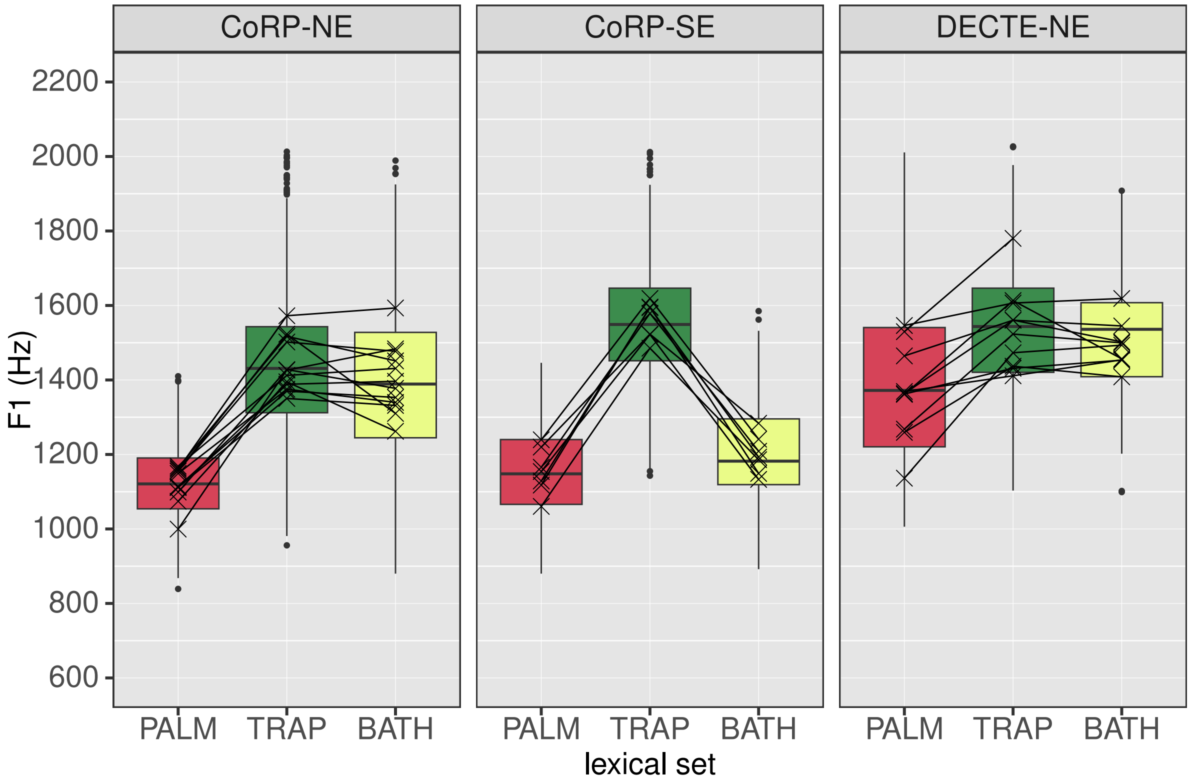

Figure 5.18.: TRAP, BATH, and PALM of all three speaker groups, including speaker means

possible that TRAP-BATH is controlled by a complex system similar to Philadelphia or New York TRAP tensing, and there is evidence for those systems varying and changing within speech communities (Labov et al., 2013, 2016). However, this data set is not large enough to be able to take reliable word means that can explain the pattern of this variation.

The variation in the CoRP-NE group is further seen in vowel duration where the mean of BATH in CoRP-NE is closer to CoRP-SE than DECTE suggesting that pre-fricative lengthening may be occurring in the CoRP-NE speakers, which historically was the precursor to the split in the South of England, and hence could be understood as the beginning of a split happening in this speech community. The duration changing before the vowel quality would suggest that this is a phonetic process not the speakers aiming towards the South East variant.

In conclusion, broadly CoRP-NE speakers behave in a regional manner with respect to the TRAP-BATH split. However, there are some indicators of variation due to the mixed input these speakers receive. There is indication of a PALM-like quality in some words, and also of a mean longer vowel length than found in DECTE. These would suggest that either a more detailed phonetic process is taking place or some speakers (in some words) are targetting the South East variant. I would predict the former because the change is so slight and lacking the split is a strong northern shibboleth.

Chapter 6.

Analysis - GOAT allophony

6.1. Introduction

As discussed in chapter 2, the GOAT vowel is predictably a diphthong in the South East (usually [oʊ] or [əʊ]) but highly variable in the North East, including (in addition to the nationally common diphthong) two monophthongal variants ([oː] and occasionally [ɵ]) and a centring diphthong ([ʊə]).

In order to determine which variants were present in each group of speakers, the particular variation caused by the following /l/ environment, and hence the GOAT split, multiple levels of analysis were undertaken.

First, the monosyllabic tokens in non pre-/l/ condition was modelled separately, to understand the general variation of the GOAT vowel, then modelled together with the monosyllabic pre-/l/ condition to consider the behaviour of the simple split (GOAT-GOAL).

Next the three pre-/l/ contexts (*hole, holy, holey*) were modelled together to understand the interactions between pre-/l/ variation and morphological conditioning.

6.2. Modelling of the GOAT vowel

This section will expand on the modelling procedure described in chapter ch:Methodology and describe the specific modelling decisions made throughout the analysis.

The GOAT vowel analysis was conducted using generalised additive mixed models in order to capture the variation within the vowel trajectory and adequately compare diphthongal and monophthongal variation between speakers and context. All GAMMs included demographic and phonological environment predictors.

Models were built according to procedures laid out by Sóskuthy (2017), Coretta , and Stanley et al. (2021) (one of the most recent sociolinguistic publications to use dynamic modelling). Parametric terms were used to capture overall differences in height of trajectories (Sóskuthy, 2017), a single smooth was fit at the reference levels (term = s(measurement.no)) and difference smooths (term = s(x, by=y)) to quantify the difference between the reference level and other levels of the predictor. Since duration can have an effect on the overall trajectory shape a tensor product interaction between duration and measurement number, this effectively controls for the effect of duration. The models were built stepwise (manually) by adding predictors and using the compareML() function from the itsadug package (van Rij et al., 2020) to test whether the predictor improved the model fit. gam.check(), from the mgcv package Wood (2017) was used to check the appropriate number of knots was used in each smooth (the number of knots determines how much 'wiggliness' is allowed in the spline and too high a value can lead to over smoothing). Full code of the model selection process can be found in the github repositories listed in Appendix A. least fit model included lexical set (or corpus for the *hope* context only models) , and the other potential terms were:

- age group (parametric term and difference smooth)
- sex (parametric term and difference smooth)
- preceding segment (difference smooth, because parametric terms cannot be controlled for when extracting model predictions)
- an interaction including any 2-3 terms of lexical set, age group, and sex (parametric term and difference smooth)
- duration (tensor product interaction with measurement no.)
- individual speaker (random smooth)

- word (random smooth)

Plots were made from model predictions, exported using the predict_gam() function from the tidymv package (Coretta, 2020), which excludes terms that are not wanted for plotting by setting them to 0 and calculating predictions from the model (as stated above, the predict_gam() function cannot exclude parametric terms, therefore since preceding segment was included as a control rather than as expected phonological variation it was only included in a smooth term). The confidence intervals are plotted using the geom_smooth_ci() function from the same package. The default package settings were used, a z-value of 1.96 for 95% confidence intervals.

Stanley et al. (2021) notes that the output of GAMMs are long and complex and hence interpretation of results are often dependent on visualisation of the predicted values from the models. Throughout this chapter, graphs created as described above are used to interpret the results, primarily focussing on the predictors relevant to the research questions but also assessing statistically significant predictors within those groups.

6.3. GOAT vowel in *hope* context only

To understand only the *hope* variation models were built of the GOAT vowel tokens without any following lateral, including corpus as a predictor. It was found that in F1 CoRP-NE behaved like CoRP-SE, with a downward trajectory, while DECTE speakers had a flatter trajectory, and in F2 there was less variation between the three speaker groups apart from the older DECTE speakers having a less front vowels (lower F2). Details on the analysis below.

6.3.1. *hope* F1

The best fit model of F1 is shown in tables 6.2 and 6.4; it included an interaction between corpus and speaker sex but as can be seen in figure 6.1 the difference between male and female speakers is not significant. Table 6.2 shows that the midpoint measurements vary from 574Hz (Female, DECTE) to 609Hz (Male SE), a range of 35Hz. The mean midpoint (table 6.1) does not vary in any meaningful way between the speaker groups.

corpus	mean F1 at midpoint
CoRP-SE	601 Hz
DECTE	578 Hz
CoRP-NE	582 Hz

Table 6.1.

The random smooths and tensor product interaction between measurement number and duration are significant. Preceding segment had improved the model fit but has very little impact on the trajectory shape. The effect of corpus can be most clearly seen in figure 6.1. CoRP-SE speakers show a downward change in F1, demonstrating a diphthong that moves upwards in the acoustic vowel space, as would be expected in a change from [ə] to [ʉ] or [u]. DECTE speakers show less change, implying little to no change in vowel height, potentially implying a monophthongal vowel. The CoRP-NE speakers show a similar trajectory shape to the CoRP-SE speakers, though with perhaps a slightly lower start point and slightly higher end point leading to a marginally flatter trajectory (see figure 6.1).

	Estimate	Std.Error	t-value	Pr (>|t|)
(Intercept)	576.42	5.94	96.99	<2e-16
speaker sex & corpus (baseline female - CoRP-NE)				
male - CoRP-NE	11.30	8.14	1.39	0.16
female - CoRP-SE	16.81	8.70	1.93	0.05
male - CoRP-SE	33.01	9.66	3.42	0.00
female - DECTE	-2.31	9.59	-0.24	0.81
male - DECTE	5.4	11.51	0.47	0.64

Table 6.2.: Parametric terms in the model of F1 in *hope* words

	edf	p-value
measurement.no	6.51e+00	<2e-16
speaker sex & corpus (baseline female - CoRP-NE)		
male - CoRP-NE	1.00e+00	0.06
female - CoRP-SE	4.13e+00	4.60e-05
male - CoRP-SE	1.00e+00	0.25
female - DECTE	4.33e+00	0.00
male - DECTE	3.21e+00	0.01
preceding segment		
liquid	1.00e+00	0.71
nasal_apical	1.00e+00	0.45
nasal_labial	1.00e+00	0.77
none	1.00e+00	0.39
obstruent_liquid	2.34e+00	0.14
oral_apical	1.00e+00	0.87
oral_labial	1.407e+00	0.82
palatal	4.31e-01	0.89
velar	1.001e+00	0.61
following manner		
affricate	1.00e+00	0.48
fricative	2.09e+00	0.61
lateral	2.12e-04	0.50
nasal	1.00e+00	0.92
none	1.00e+00	0.70
stop	3.90e+00	0.00
following place		
apical	1.43e+00	0.73
interdental	1.00e+00	0.80
labial	1.00e+00	0.84
labiodental	1.00e+00	0.75
none	1.00e+00	0.70
palatal	4.42e-01	0.89
velar	1.00e+00	0.78
tensor product interaction - measurement.no & duration (continuous)		
	1.18e+01	<2e-16
random smooth (by speaker)		
	5.17e+01	<2e-16
random smooth (by word)		
	3.10e+02	<2e-16

Table 6.4.: Table showing smooth terms (by measurement number) of the model of F1 in *hope* words

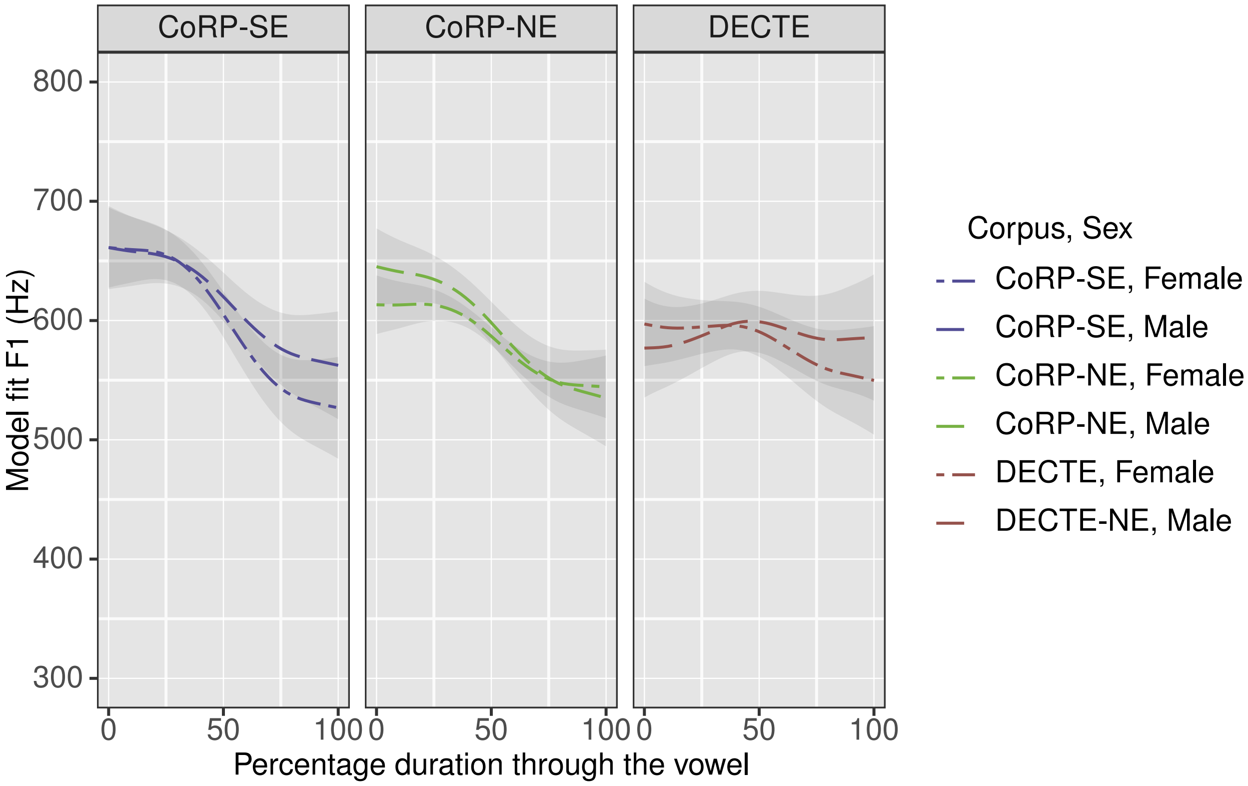

Figure 6.1.: GAMM plot of F1 of hope in CoRP-SE, DECTE, and CoRP-NE

6.3.2. *hope* F2

Tables 6.6 and 6.8 show the best fit model for F2 of the GOAT vowel in *hope* contexts only. The model shows a three-way interaction between age group, sex, and corpus, however the differences do not show a clear pattern. The mean midpoints for each corpus can be seen in table 6.5.

corpus	mean F2 at midpoint
CoRP-SE	1647.50 Hz
DECTE	1268.50 Hz
CoRP-NE	1666.09 Hz

Table 6.5.

CoRP-SE speakers have the most front vowel (by midpoint) as supported by the literature, though the difference between them and the CoRP-NE speakers is not significant and may not be meaningful ($p>0.05$). DECTE speakers have a far further back midpoint, consistent with a monophthong such as [o].

	Estimate	Std.Error	t-value	Pr (>|t|)
(Intercept)	1508.72	57.22	26.37	<2e-16
age & sex & corpus (baseline Old - Female - CoRP-NE)				
Young - Female - CoRP-NE	43.95	67.18	0.65	0.51
Old - Male - CoRP-NE	67.09	98.30	0.68	0.49
Young - Male - CoRP-NE	78.42	69.54	1.13	0.26
Old - Female - CoRP-SE	52.76	80.28	0.66	0.51
Young - Female - CoRP-SE	108.26	80.27	1.35	0.18
Old - Male - CoRP-SE	170.70	80.37	2.12	0.03
Young - Male - CoRP-SE	167.91	98.67	1.70	0.09
Old - Female - DECTE	-401.67	73.29	-5.48	4.28e-08
Old - Male - DECTE	-307.73	98.95	-3.11	0.00
Young - Male - DECTE	-11.25	99.24	-0.11	0.91

Table 6.6.: Parametric terms of the model of F2 in *hope* words

As can be seen in figure 6.2 The F2 value decreases through the vowel in CoRP-SE speakers, suggesting a diphthong that moves back in the mouth, as would be expected from [/əʊ/]. The DECTE speakers have a slightly flatter but not completely straight trajectory shape (the DECTE facet only shows male speakers, there was not enough data to look at young female speakers in DECTE). This could mean that speakers are producing overall a less front vowel than the CoRP-SE speakers (the literature supports RP speakers having a fronter diphthong than other diphthongal variants), and that it is a mixture of a monophthong and a diphthong, but the difference is not capture by any of the predictors in the model. The parametric and smooth terms both suggest that CoRP-NE speakers are producing a non-regional variant of GOAT.

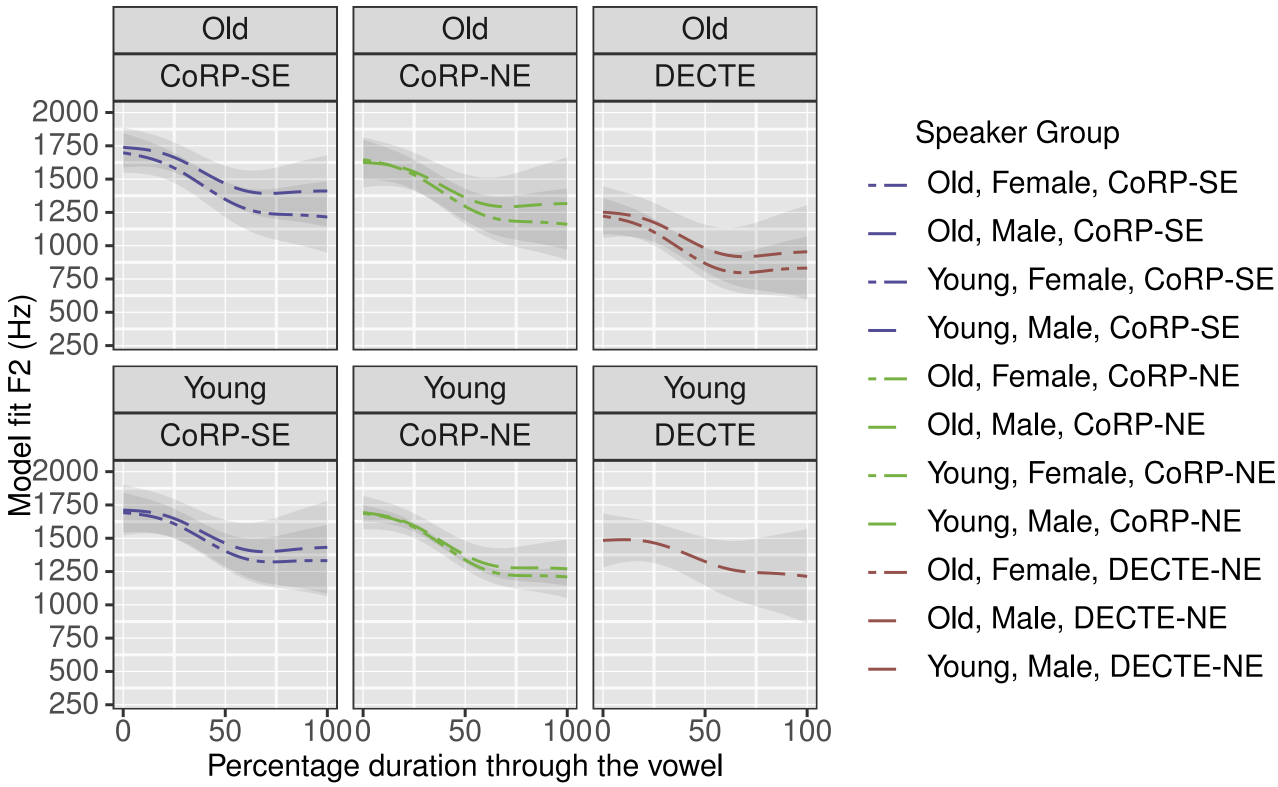

Figure 6.2.: GAMM plot of F2 of hope in CoRP-SE, DECTE, and CoRP-NE

	edf	p-value
s(measurement.no)	1.00e+00	0.05
age & sex & corpus (baseline Old - Female - CoRP-NE)		
Young - Female - CoRP-NE	2.75e+00	0.39
Old - Male - CoRP-NE	1.00e+00	0.46
Young - Male - CoRP-NE	1.00e+00	0.71
Old - Female - CoRP-SE	1.00e+00	1.00
Young - Female - CoRP-SE	1.00e+00	0.52
Old - Male - CoRP-SE	1.00e+00	0.42
Young - Male - CoRP-SE	1.00e+00	0.39
Old - Female - DECTE	3.30e+00	0.07
Old - Male - DECTE	1.29e+0	0.56
Young - Male - DECTE	2.35e+00	0.19
preceding segment		
liquid	1.00e+00	1.75e-05
nasal_apical	4.21e+00	0.00
nasal_labial	3.10e+00	<2e-16
none	2.378e+00	0.10
obstruent_liquid	1.00e+00	8.25e-06
oral_apical	1.799e+00	0.13
oral_labial	1.002e+00	2.05e-06
palatal	4.540e-04	0.50
velar	3.56e+00	0.00
following manner		
affricate	1.00e+00	0.25
fricative	1.00e+00	0.12
lateral	3.32e-04	1.00
nasal	1.00e+00	0.17
none	4.81e+00	0.01
stop	1.00e+00	0.11
following place		
apical	1.35e+00	0.45
interdental	1.001e+00	0.15
labial	2.01e+00	0.33
labiodental	1.75e+00	0.38
none	1.01e+00	0.96
palatal	1.07e-03	1.00
velar	1.00e+00	
tensor product interaction - measurement.no & duration (continuous)		
	1.09e+01	0.09
random smooth(by speaker)	5.015e+01	<2e-16
random smooth (by word)	2.821e+02	<2e-16

Table 6.8.: Table showing smooth terms of the model of F2 in *hope* words

6.3.3. *hope* context summary

In summary, CoRP-SE and CoRP-NE show evidence of a diphthong in both the F1 and F2 results, with little to no difference between the speaker groups. The DECTE results show little to no variation in F1 or F2, lending evidence to at least some monophthongal productions. In conclusion, CoRP-NE are showing non-regional behaviour with respect to the GOAT vowel in *hope*-like (no pre-/l/) contexts.

6.4. The GOAT split in monosyllabic environments (*hope* vs *hole*)

6.4.1. The GOAT-GOAL split in CoRP-SE speakers

The GOAT-GOAL split as taken from the CoRP-SE speakers is characterised by little to no difference in F1 and a clear difference in F2. *Hole* is less front than hope by more than 550Hz (parametric term, midpoint) and has a significantly different trajectory shape, with more movement seen in an F2 decrease through the vowel. Models and more detailed explanation below are given below.

6.4.1.1. F1 of the CoRP-SE speakers

The best fit model for F1 of the GOAT-GOAL split in the CoRP-SE speakers can be seen in tables 6.9 and 6.10. The best fit model included an interaction between lexical set and speaker sex but there is no significant difference in the trajectory shape between these groups. The only significant term, aside from the random smooths, is the tensor product interaction between measurement.no and duration.

Figure 6.3 shows that despite the interactions leading to a better fit model, the confidence interval for all four combinations of predictors is overlapping, leading to the conclusion that neither the height nor trajectory shape of F1 vary significantly between *hope* and *hole* in the CoRP-SE speakers. This is taken as the proto-typical pattern of the GOAT-GOAL split, so no

	Estimate	Std.Error	t-value	Pr (>|t|)
(Intercept)	599.72	6.58	91.18	<2e-16
speaker sex & lexical set (ordered)				
Linear	0.91	8.40	0.11	0.91
Quadratic	-21.30	1.73	-12.29	<2e-16
Cubic	-9.13	10.27	-0.89	0.37

Table 6.9.: Parametric terms of the model of F1 in CoRP-SE speakers

	edf	p-value
s(measurement.no)	6.631	<2e-16
speaker sex & lexical Set		
Male - hope	1.000	0.2726
Female - hole	1.000	0.0138
Male - hole	1.000	0.5894
preceding segment		
liquid	1.953	0.2082
nasal_apical	1.001	0.8053
nasal_labial	1.000	0.5248
none	1.964	0.3398
obstruent_liquid	0.686	0.7812
oral_apical	1.000	0.2531
oral_labial	1.000	0.4681
palatal	1.000	0.9612
velar	1.000	0.2953

Table 6.10.: Table showing smooth terms of the model of F1 in CoRP-SE speakers

difference would be expected in the CoRP-NE speakers even if they do show the southern version of the split.

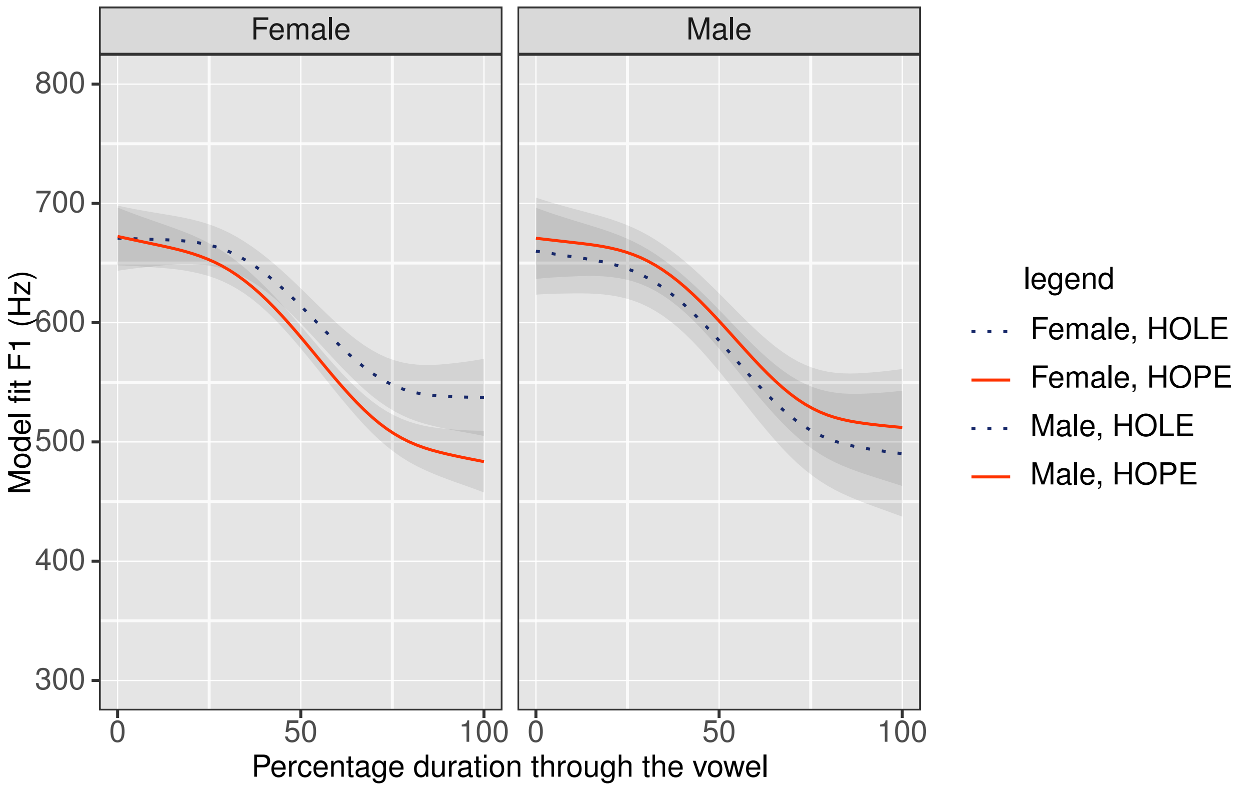

Figure 6.3.: GAMM plot of F1 in CoRP-SE speakers

6.4.1.2. F2 of the CoRP-SE speakers

The best fit model for F2 of the GOAT-GOAL split in the CoRP-SE speakers is shown in tables 6.11 and 6.12. The best fit model included an interaction term between age and sex of speaker but the variation between the levels of this interaction is not large and is only significant for one of the combinations (young male). The difference in between *hope* and *hole* words is -568.45Hz and there is also a significant difference in shape between the *hope* and *hole* words.

	Estimate	Std.Error	t-value	Pr (>|t|)
(Intercept)	1626.72	17.56	92.61	<2e-16
lexical set (baseline hope)				
hole	-568.4517	25.2506	-22.512	<2e-16
age & sex (ordered)				
Linear	58.41	28.31	2.06	0.039
Quadratic	-32.55	26.26	-1.24	0.22
Cubic	-0.8431	24.0124	-0.035	0.9720

Table 6.11.: Parametric terms of the model of F2 in CoRP-SE speakers

	edf	p-value
s(measurement.no)	1.00e+00	0.85
lexical set (baseline hope)		
hole	4.66e+00	<2e-16
age & sex (baseline Old - Female)		
Young - Female	1.00e+00	0.18
Old - Male	1.000e+00	0.20
Young - Male	2.78e+00	0.04
preceding segment		
liquid	1.00e+00	0.69
nasal_apical	4.15e+00	6.67e-05
nasal_labial	1.00e+00	0.18
none	1.00e+00	0.35
obstruent_liquid	1.88e-04	1.00
oral_apical	1.43e+00	0.08
oral_labial	1.00e+00	0.48
palatal	1.00e+00	0.02
velar	3.97e+00	2.81e-05
tensor product interaction measurement number & duration (continuous)		
	5.84e+00	0.00
random smooth by speaker		
	5.59e+00	<2e-16
random smooth by word		
	1.79e+02	<2e-16

Table 6.12.: Table showing smooth terms of the model of F2 in CoRP-SE speakers

The difference in shape and frontness between the *hope* and *hole* words can be seen in figure 6.4. Despite the age-sex interaction improving the model it is clear that the important effect is the distinction between the *hope* and *hole* words. In the CoRP-SE speakers there is a frontness difference of -568Hz (*hole* is further back). The *hope* and *hole* words also have a significantly different shapes. The vowel in *hope* words does not change much in F2 (as would be expected in a movement from a schwa position to a [/ʊ/] but the vowel in the *hole* words shows movement in F2, particularly decreases from the 10% to 70% points demonstrating the backing caused by the following /l/ segment.

If we take the CoRP-SE speakers as again having the prototypical split, the GOAT-GOAL split in F2 can be described as being found in difference in both frontness and shape of the vowel trajectory, with most of the set at around 1600Hz and the pre-/l/ context at around 1000Hz (parametric terms, table 6.11).

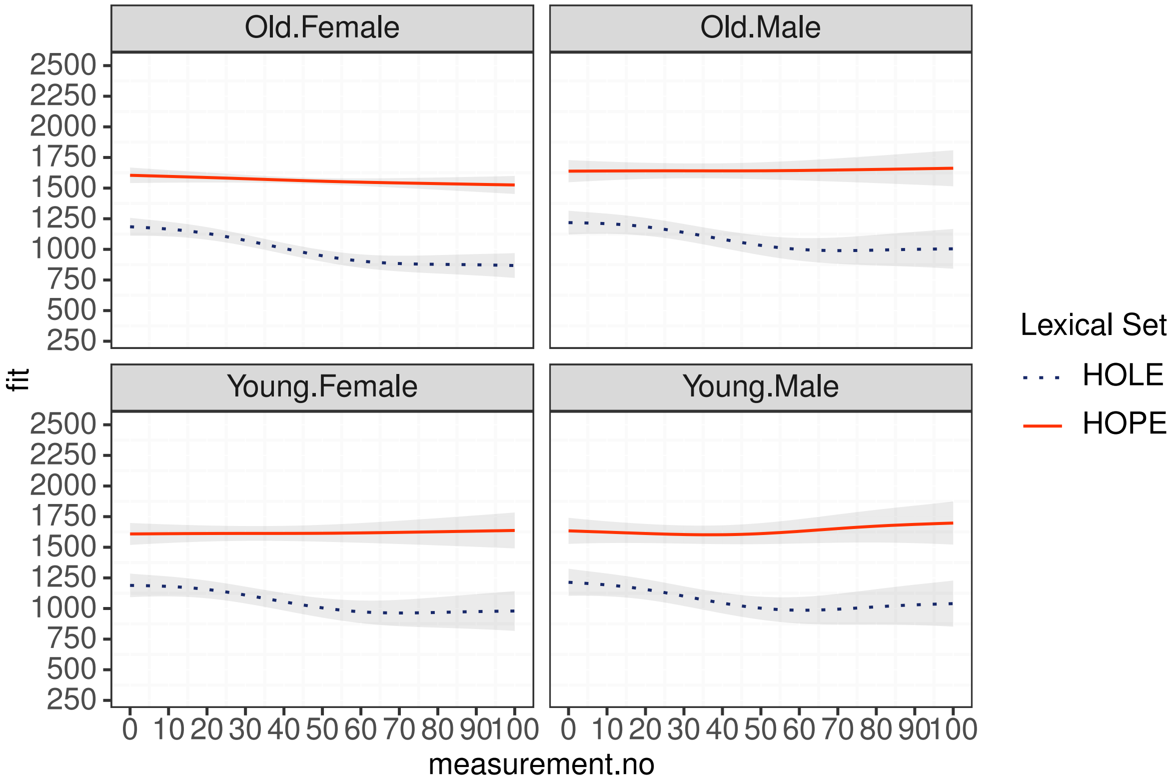

Figure 6.4.: GAMM plot of F2 in CoRP-SE speakers

6.4.2. DECTE

The patterns exhibited by DECTE speakers show very little evidence of a GOAT-GOAL split. The only potential variation is in young vs old. The young (male) speakers show potential evidences of a higher and less front *hole* vowel but the observed differences are not conclusive. Overall the DECTE vowel in nearer [o] and has similar frontness to the CoRP-SE pre-/l/ context.

6.4.2.1. F1 of the DECTE speakers

The best fit model for F1 of the DECTE speakers is shown in tables 6.13 and 6.14. The model includes an interaction between age and lexical set, implying that old and young speakers show a different relationship between the *hope* and *hole* sets of words. The intercept (old - *hope*)is 585Hz and there's a significant difference between that and *hope* from the young speakers (23Hz), which is unlikely to be linguistically meaningful. However, there's a larger effect from the young speakers' *hole* words, which have an estimate of -102Hz, showing a higher midpoint to the vowel. In summary, with regard to midpoint, Old & Young *hope* and Old *hole* are similar whereas Young *hole* is higher in the acoustic vowel space.

In smooth terms the two random smooths are significant, and there is some effect of preceding segment, but the only difference between speaker groups is the young speakers *hope* words in comparison to the base line (Old *hope*).

	Estimate	Std.Error	t-value	Pr (>|t|)
(Intercept)	584.76	5.94	98.47	<2e-16
age & lexical set (baseline Old - hope)				
Young - hope	23.47	9.60	2.44	0.01
Old - hole	-0.18	13.94	-0.01	0.99
Young - hole	-102.94	43.30	-2.38	0.018

Table 6.13.: Parametric terms of the model of F1 in DECTE speakers

The parametric and smooth effects can be seen in figure 6.5; the plotted smooth for the young *hole* words is overall lower than the other smooths, though the confidence intervals overlap through most of the trajectory. This may be evidence for an acoustically higher *hole* vowel in younger speakers.The literature on the GOAT vowel in pre-/l/ context does not attest an effect on F1 (vowel height), and it is unclear what we are seeing here. It can be determined from the confidence interval that there is a large variation in the height within the acoustic vowel space of younger speakers' productions of the *hole* context.

	edf	p-value
s(measurement.no)	4.021e+00	0.000175
age & lexical set (baseline Old - hope)		
Young.hope	1.00e+00	0.00
Old.hole	1.001e+00	0.90
Young.hole	1.00e+00	0.90
preceding segment		
liquid	1.00e+00	0.09
nasal_apical	1.00e+00	0.01
nasal_labial	7.96e-05	1.00
none	1.00e+00	0.02
obstruent_liqui	1.84e+00	0.10
oral_apical	1.00e+00	0.04
oral_labial	1.00e+00	0.06
palatal	1.00e+00	0.06
velar	1.59e+00	0.66
random smooth by speaker		
	5.70e+00	<2e-16
random smooth by word		
	1.01e+02	<2e-16

Table 6.14.: Table showing smooth terms of the model of F1 in DECTE speakers

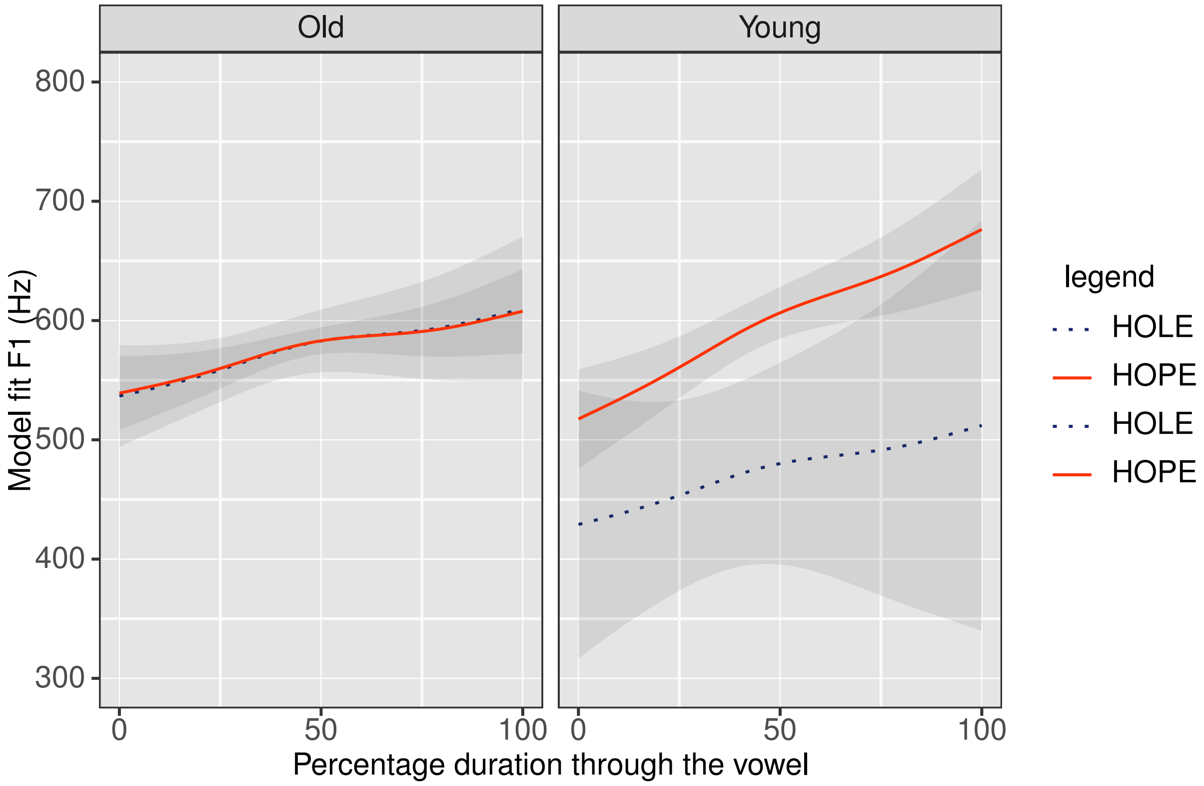

Figure 6.5.: GAMM plot of F1 in DECTE speakers

6.4.2.2. F2 of the DECTE speakers

The best fit model for F2 of the DECTE speakers can be seen in tables 6.15 and 6.16. The intercept is 1095Hz and the model includes a three way interaction between lexical set, age, and sex. The baseline group (Old Female, *hope*), Old Male *hole*, and Young Male *hole*, are not significantly different to each other in frontness. There is a 118Hz difference between the baseline and Old Male *hope*, a 381Hz difference between the baseline and Young Male *hope*, and a -101Hz difference between the baseline and Old Female *hole*. The conclusion that can be drawn from this is that at least in terms of midpoint intercept there is some variation. Specifically, Old Female speakers have a -101Hz difference between *hope* (more front) and *hole* (less front), Old Male speakers have a -117Hz difference between *hope* and *hole*, and Young Male speakers a -283Hz difference between *hope* and *hole* (not enough data is present in the DECTE group to look at young female speakers). There is little to no significant effect of trajectory shape in the smooth terms (table 6.14); from the confidence intervals in figure 6.5 there is overlap within the older speaker groups. However younger male speakers show some evidence of a split apart, with overlap t the beginning and end of the vowel but distinction the middle of the vowel. The is the same pattern as seen in the F1 of this speaker group and implies that a GOAT-GOAL split is beginning to emerge but there is still variation in the production of the *hole* context.

	Estimate	Std.Error	t-value	Pr (>|t|)
(Intercept)	1094.81	23.89	45.82	<2e-16
age & sex & lexical set (baseline Old Female hope)				
Old - Male - hope	118.30	44.57	2.65	0.01
Young - Male - hope	381.00	45.14	8.44	<2e-16
Old - Female - hole	-101.45	30.95	-3.28	0.00
Old - Male - hole	1.64	78.11	0.02	0.98
Young - Male - hole	97.46	105.54	0.92	0.36

Table 6.15.: Parametric terms of the model of F2 in DECTE speakers

	edf	p-value
s(measurement.no)	5.032e+00	<2e-16
age & sex & lexical set (baseline Old Female hope)		
Old - Male - hope	1.00e+00	0.23
Young - Male - hope	3.437e+00	0.00
Old - Female - hole	1.000e+00	0.98
Old - Male - hole	1.000e+0	0.24
Young - Male - hole	1.000e+00	0.08
preceding segment		
liquid	1.000e+00	0.44
nasal_apical	3.164e+00	0.01
nasal_labial	1.000e+00	0.08
none	1.296e+00	0.89
obstruent_liquid	1.000e+00	0.43
oral_apical	1.556e+00	0.72
oral_labial	1.000e+00	0.05
palatal	1.216e-04	1.00
velar	1.000e+00	0.13
tensor product intearaction measurement.no & duration (continuous)		
	7.496e+00	<2e-16
random smooth by speaker		
	4.199e+00	<2e-16
random smooth by word		
	1.105e+02	<2e-16

Table 6.16.: Table showing smooth terms of the model of F2 in DECTE speakers

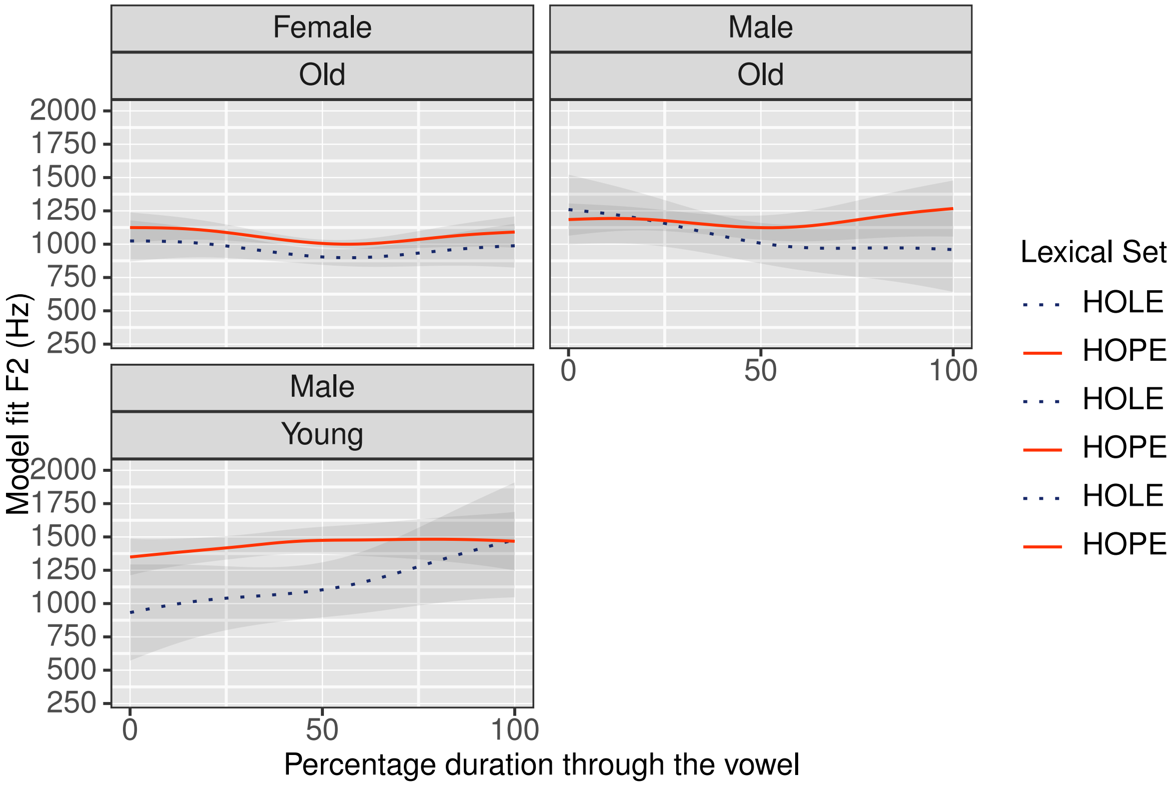

Figure 6.6.: GAMM plot of F2 in DECTE speakers

6.4.3. CoRP-NE

CoRP-NE speakers show little to no distinction in F1 between *hope* and *hole*, in F2 they show a distinction between the two contexts. More details and models are discussed below.

6.4.3.1. F1 of the CoRP-NE speakers

CoRP-NE speakers show very little variation in F1 of *hope* and *hole* There is a small difference of 32Hz at the midpoint of the vowel, that is unlikely to be linguistically meaningful (see table 6.17), and the only significant variation in the smooth terms is in the random smooths and the interaction between duration and shape. The shape of both trajectories can be seen in figure 6.7. The F1 midpoint (580Hz) is similar to both CoRP-SE (599Hz) and DECTE (584Hz), the trajectory moves in a similar direction to the CoRP-SE speakers, but not as much, which is more similar to DECTE speakers.

	Estimate	Std.Error	t-value	Pr (>|t|)
(Intercept)	580.01	4.36	132.95	<2e-16
lexical set (baseline hope)				
hole	31.79	6.99	4.55	5.44e-06

Table 6.17.: Parametric terms of the model of F1 in CoRP-NE speakers

	edf	p-value
s(measurement.no)	6.46	<2e-16
lexical set (baseline hope)		
hole	2.27	0.15
preceding segment		
liquid	1.00	0.29
nasal_apical	1.00	0.29
nasal_labial	1.858	0.48
none	2.184	0.18
obstruent_liquid	0.88	0.53
oral_apical	1.00	0.91
oral_labial	1.00	0.63
palatal	1.52	0.62
velar	1.001	0.80-
tensor product interaction measurement.no & duration (continuous)		
	11.079	<2e-16
random smooth by speaker		
	36.97	<2e-16
random smooth by word		
	288.48	<2e-16

Table 6.18.: Table showing smooth terms of the model of F1 in CoRP-NE speakers

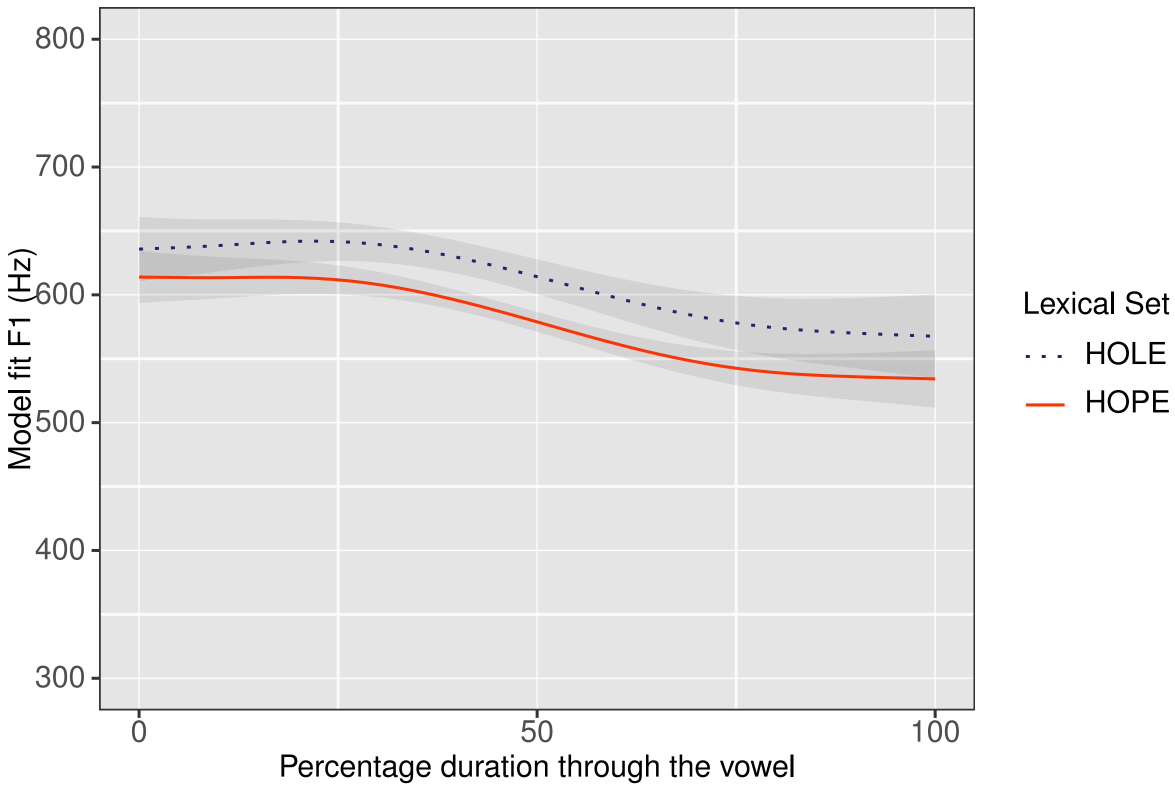

Figure 6.7.: GAMM plot of F1 in CoRP-NE speakers

6.4.3.2. F2 of the CoRP-NE speakers

The best fit model of F2 of the CoRP-NE speakers included speaker age group and sex, but in the parametric terms of the model neither of these have a significance or a meaningful effect size. The effect of the lexical set parametric term is a difference between *hope* and *hole* of -522Hz, similar to that of the CoRP-SE speakers. The smooth terms show some effects of preceding segments but due to small token counts in the categories these are difficult to interpret clearly. There is also an effect of lexical set but not of age group or speaker sex. As can be seen in figure 6.8, all four speaker groups in the model show a difference in acoustic frontness and some difference in shape (statistically significant, as seen in table 6.20 but not clearly visible in the graphs) between the *hope* and *hole* words.

Section 6.4.4 below will directly compare the trajectories of all three groups.

	Estimate	Std.Error	t-value	Pr (>|t|)
(Intercept)	1542.39	44.99	34.28	<2e-16
lexical set (baseline hope)				
hole	-521.97	22.57	-23.13	2e-16
age group (baseline Old)				
Young	6.18	48.34	0.13	0.90
speaker sex (baseline Female)				
Male	36.33	42.46	0.86	0.392

Table 6.19.: Parametric terms of the model of F2 in CoRP-NE speakers

	edf	p-value
s(measurement.no)	1.00e+00	0.04
lexical set (baseline hope)		
hole	4.07e+00	2.16e-06
age group		
Old	2.46e-04	1.00
Young	1.00e+00	0.72
speaker sex		
Female	2.72e+00	0.11640
Male	1.00e+00	0.94
preceding segment		
liquid	1.00e+00	6.05e-06
nasal_apical	4.99e+00	<2e-16
nasal_labial	2.48e+00	8.67e-06
none	1.00e+00	0.00
obstruent_liquid	1.00e+00	2.81e-06
oral_apical	2.29e+00	0.014
oral_labial	1.78e+00	1.47e-06
palatal	9.47e-04	1.00
velar	4.39e+00	<2e-16
tensor product interaction measurement.no & duration (continuous)		
	1.15e+01	<2e-16
random smooth by speaker		
	3.629e+01	<2e-16
random smooth by word		
	2.81e+02	<2e-16

Table 6.20.: Table showing smooth terms of the model of F2 in CoRP-NE speakers

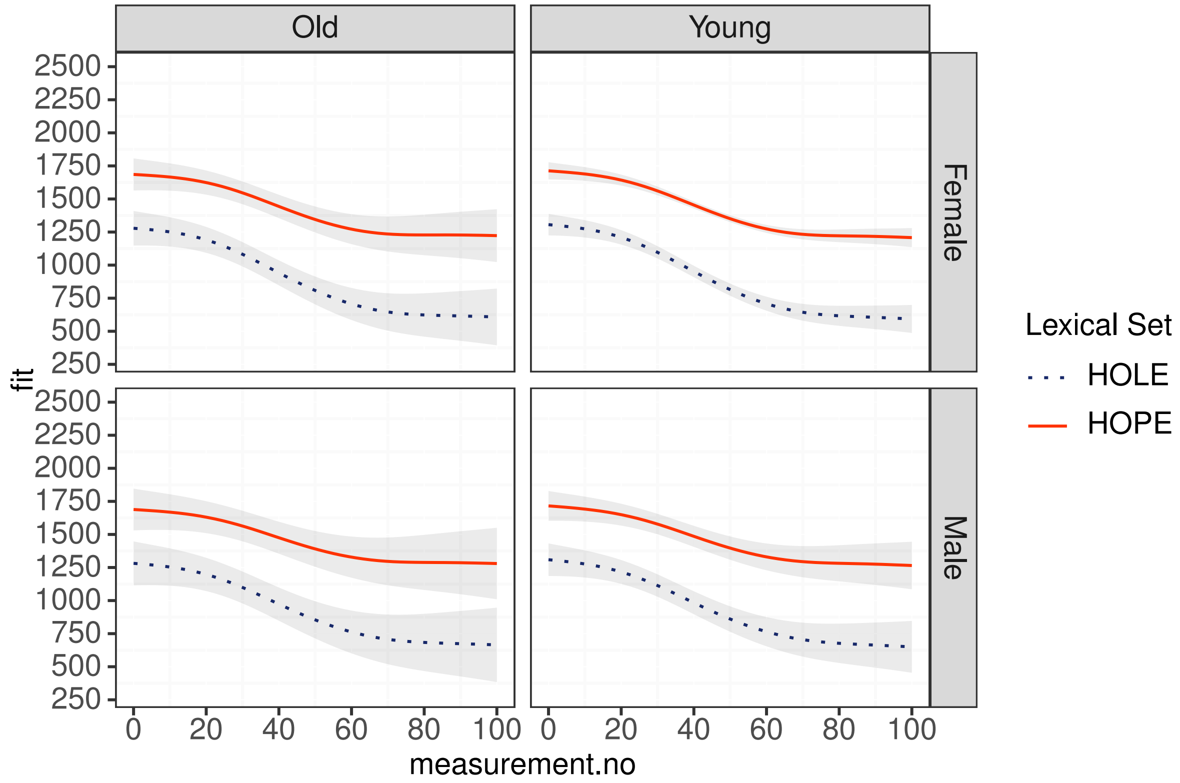

Figure 6.8.: GAMM plot of F2 in CoRP-NE speakers

6.4.4. Direct Comparison of the GOAT-GOAL split in all three speaker groups

6.4.4.1. GOAT-GOAL F1 comparison

Figure 6.9 compares all three speaker groups in F1. It is clear that while the F1 trajectories are broadly in a similar position, and have similar midpoint measurements, the shape and directions are very different.

The DECTE speakers are also the only ones that show any distinction between the *hope* and *hole* words, with younger speakers showing a lower F1 in *hole* words.

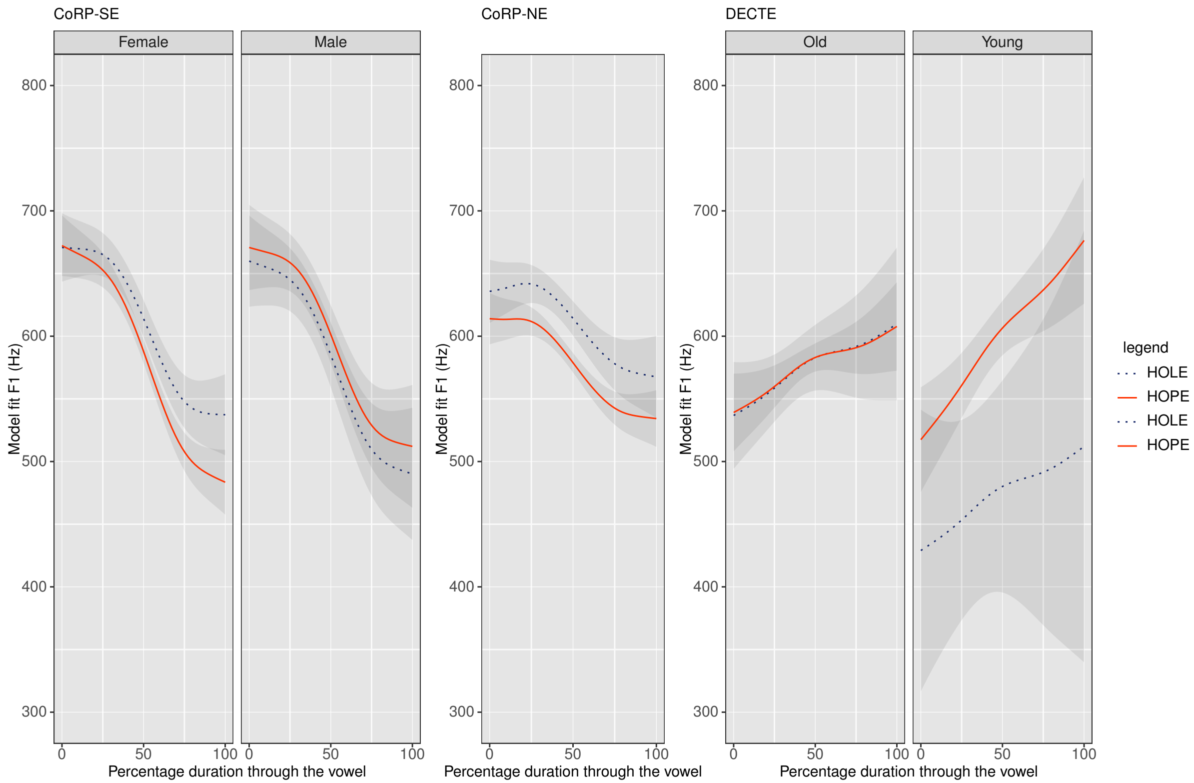

Figure 6.9.: GAMM plot of F1 in all three speaker groups

6.4.4.2. GOAT-GOAL F2 comparison

Figure 6.10 compares all three speaker groups in F2 by plotting all three models discussed in the sections above together. It can be clearly seen that CoRP-SE and CoRP-NE speakers have a GOAT-GOAL split in F2 whereas the old DECTE speakers do not and the young (male) DECTE speakers have indication of a split emerging but it is not across the whole vowel trajectory as found in the CoRP groups. The midpoint differences (as seen in the parametric terms) are similar (CoRP-SE: 568Hz, CoRP-NE: 522Hz) but there is more difference in the trajectory shape in the CoRP-SE speakers, who show a flatter F2 trajectory. It is possible to say that CoRP-NE speakers are behaving broadly non-regionally with respect to the GOAT-GOAL split.

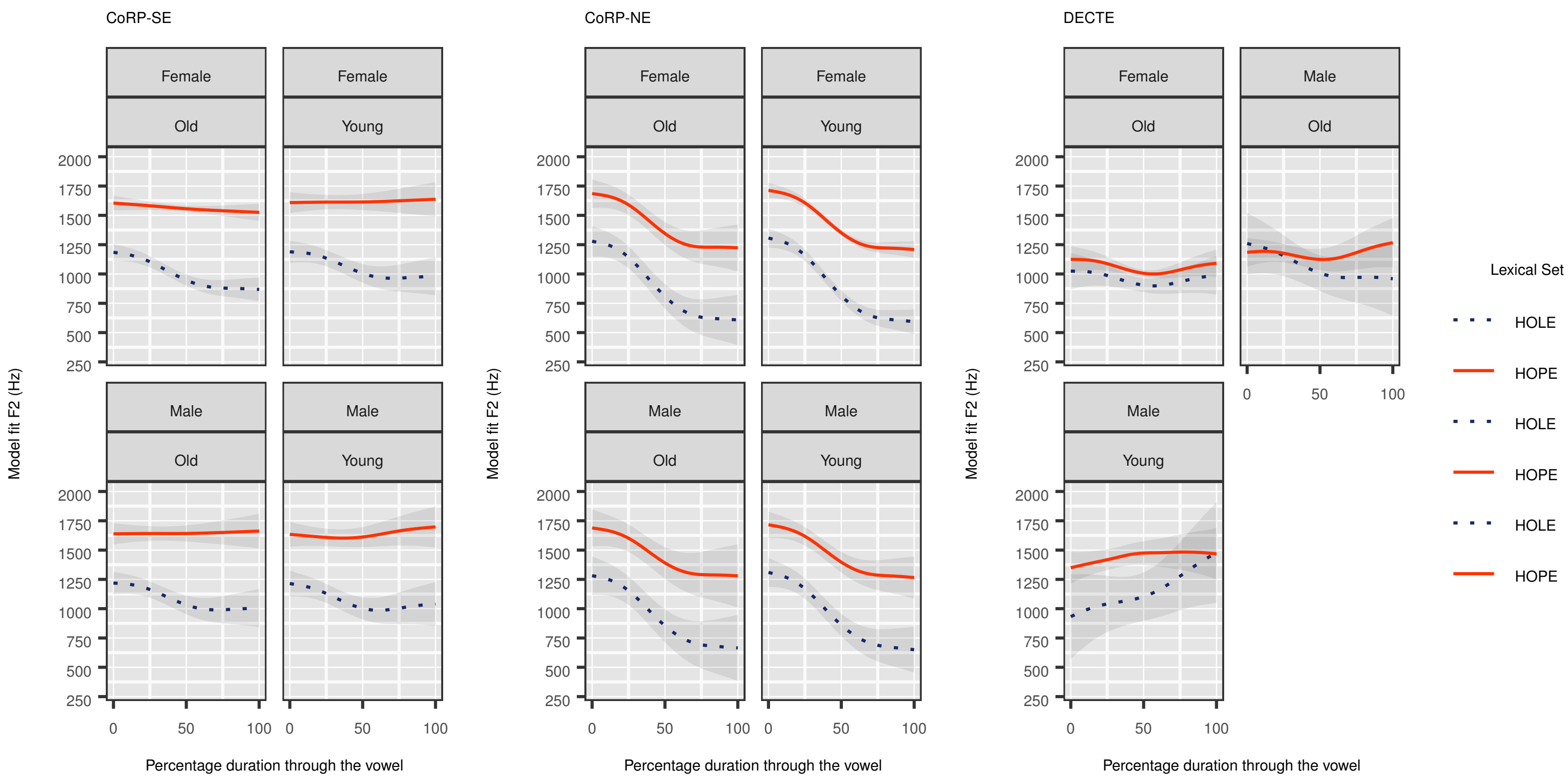

Figure 6.10.: GAMM plot of F2 in all three speaker groups

6.5. Morpho-phonological Conditioning of the GOAL context (*hole, holy, holey*)

This section will consider the GOAL context (that is GOAT vowels before an ɫ segment) in 3 different morpho-phonological conditions. Monosyllabic, monomorphemic (*hole*), disyllabic, bimorphemic (*holey*), and disyllabic, monomorphemic (*holy*). The analysis aims both understand the morphological variation in general but also to apply the theory of the lifecycle of phonological processes (chapter 2) compare the relationship and potential direction of change between the CoRP-SE and CoRP-NE communities.

6.5.1. Morpho-phonological Conditioning in CoRP-SE

The CoRP-SE speakers show no distinction between the pre-/l/ contexts in F1, and show a distinction between all three contexts in F2.

6.5.1.1. CoRP-SE F1

In F1 of the CoRP-SE speakers, the best fit model included a three way interaction between age group, speaker set and morpho-phonological context. However, very few of these differences were large or significant, as can be seen in figure 6.11. F1 is variable between these contexts, as would be expected because the potential variation for *hole*,*holey*, and *holy* is within the GOAT-GOAL difference.

	Estimate	Std.Error	t-value	Pr (>|t|)
(Intercept)	599.66	13.85	43.290	2e-16
age & speaker sex & lexical set(baseline Old Female hole)				
Young - Female - hole	30.10	18.87	1.6	0.11
Old - Male - hole	-0.60	18.93	-0.03	0.97
Young - Male - hole	-37.09	23.11	-1.61	0.11
Old - Female - holey	-30.41	7.62	-3.99	6.76e-05
Young - Female - holey	-27.41	20.17	-1.36	0.17
Old - Male - holey	-49.76	20.27	-2.46	0.01
Young - Male - holey	-58.16	24.32	-2.39	0.017
Old - Female - holy	-26.50	9.03	-2.93	0.00
Young - Female - holy	-2.82	20.88	-0.16	0.89
Old - Male - holy	-59.95	20.84	-2.88	0.00
Young - Male - holy	-56.06	26.24	-2.14	0.03

Table 6.21.: Parametric terms in the model of F1 in the pre-/l/ contexts in CoRP-SE speakers

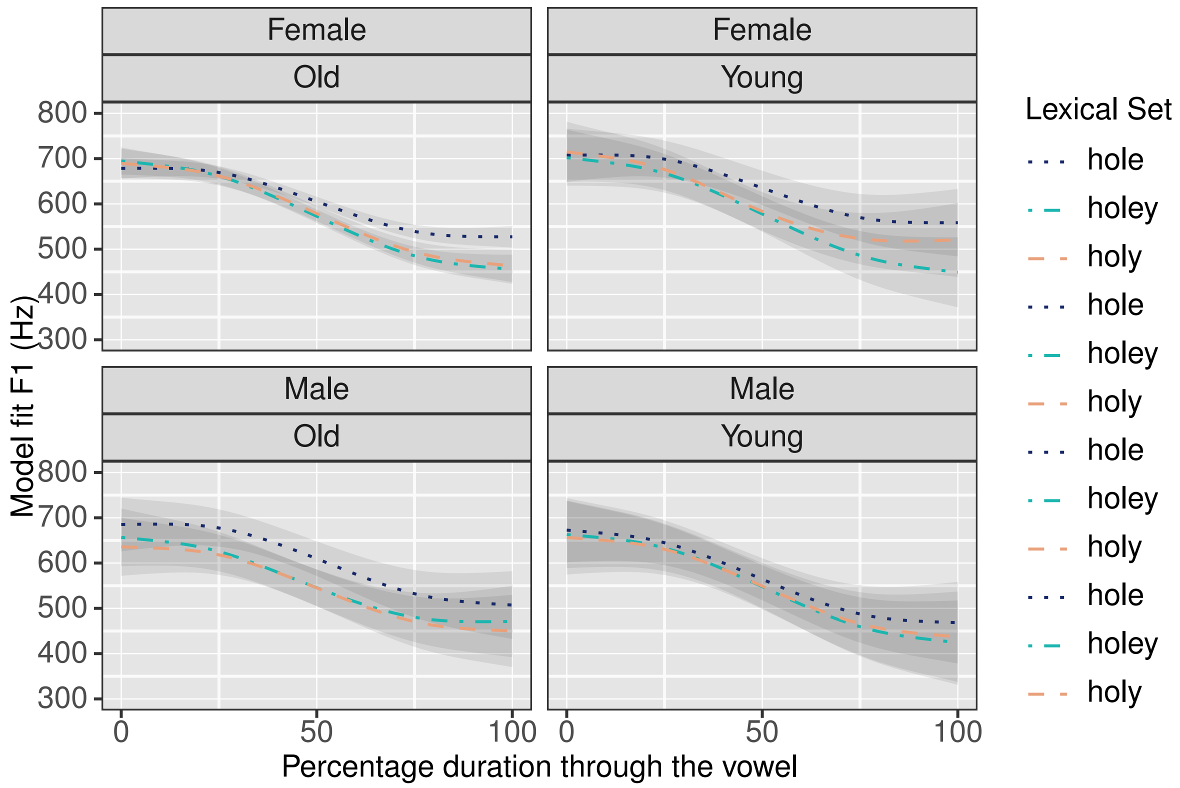

Figure 6.11.: GAMM plot of F1 of the GOAL context in CoRP-SE speakers

	edf	p-value
s(measurement.no)	6.16	<2e-16
age & speaker sex & lexical set (baseline Old Female hole)		
Young - Female - hole	1.0	0.97
Old - Male - hole	2.08	0.26
Young - Male - hole	1.76	0.35
Old - Female - holey	1.53	0.00
Young - Female - holey	1.00	0.08
Old - Male - holey	2.22	0.25
Young - Male - holey	1.16	0.26
Old - Female - holy	1.00	0.01
Young - Female - holy	2.72	0.06
Old - Male - holy	1.00	0.56
Young - Male - holy	1.00	0.37
preceding segment		
liquid	1.00	0.60
nasal_labial	1.00	0.33
none	1.51	0.30
obstruent_liquid	1.00	0.35
oral_apical	1.78	0.47
oral_labial	1.00	0.48
palatal	0.51	0.97
velar	1.26	0.70
tensor product interaction measurement.no duration (continuous)		
	11.14	<2e-16
random smooth by speaker		
	5.87	<2e-16
random smooth by word		
	62.90	<2e-16

Table 6.22.: Table showing smooth terms the model of F1 in the pre-/l/ contexts in CoRP-SE speakers

6.5.1.2. CoRP-SE F2

The best fit model of F2 for the CoRP-SE speakers shows that there is a difference between the midpoints of the *hole*, *holey*, and *holy* contexts (see table 6.23 but very little difference in the shape of the smooths). The *holy* midpoint (1520Hz) is similiar to that of *hope* in 6.15 (=1627Hz), implying the un-backed diphthong.

As discussed in chapter 2 there is discussion over whether an interim target is possible or if speakers only have a front or backed diphthong. In order to consider this a check of the raw (not modelled) data was undertaken, which can be seen in figure 6.13 (with *hope* also included for comparison). This shows that potentially the in-between diphthong shown in the *holey* context in the model is actually caused by tokens existing in the two positions but the model not capturing what is causing this variation.

	Estimate	Std.Error	t-value	Pr (>|t|)
(Intercept)	1045.37	29.40	35.56	<2e-16
lexical set (baseline hole)				
holey	234.50	34.82	6.74	1.89e-11
holy	474.28	42.25	11.23	<2e-16
age group (baseline Old)				
Young	-24.52	27.06	-0.91	0.37
speaker sex (baseline Female)				
Male	40.27	27.10	1.49	0.137

Table 6.23.: Parametric terms of the model of F2 in the pre-/l/ contexts in CoRP-SE speakers

	edf	p-value
s(measurement.no)	5.06e+00	8.33e-06
lexical set (baseline hole)		
holey	1.00e+00	0.20
holy	1.95	0.16
age group		
Old	1.01e-04	1.00
Young	2.95e+00	0.07
speaker sex		
Female	1.10e-04	1.00
Male	1.85e+00	0.46
preceding segment		
reSegliquid	1.00e+00	0.76
nasal_labial	1.47e+00	0.65
none	1.73e+00	0.57
obstruent_liquid	1.97e-04	1.00
oral_apical	1.00e+00	0.83
oral_labial	1.68	0.52
palatal	1.00e+00	0.63
velar	1.00e+00	0.81
tensor product interaction measurement.no & duration (continuous)		
	8.28	<2e-16
random smooth by speaker		
	9.33e+00	<2e-16
random smooth by word		
	7.70e+01	<2e-16

Table 6.24.: Table showing smooth terms the model of F2 in the pre-/l/ contexts in CoRP-SE speakers

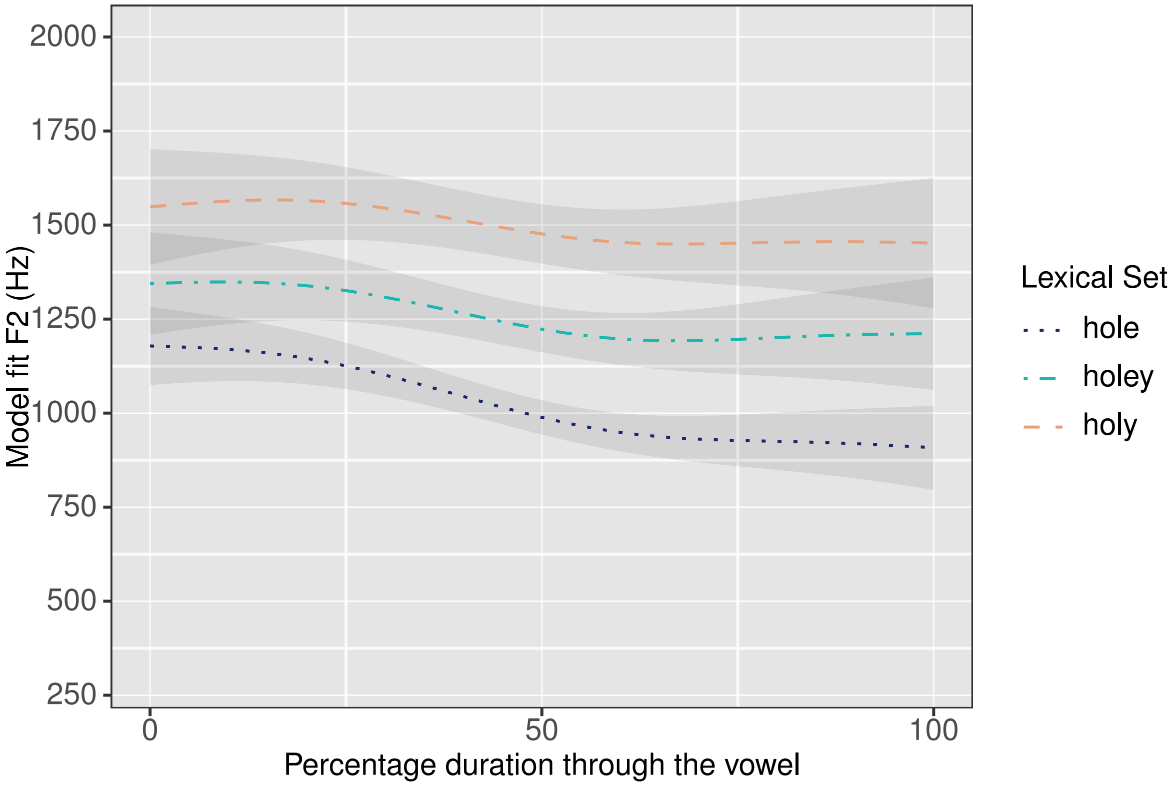

Figure 6.12.: GAMM plot of F2 of the GOAL context in CoRP-SE speakers

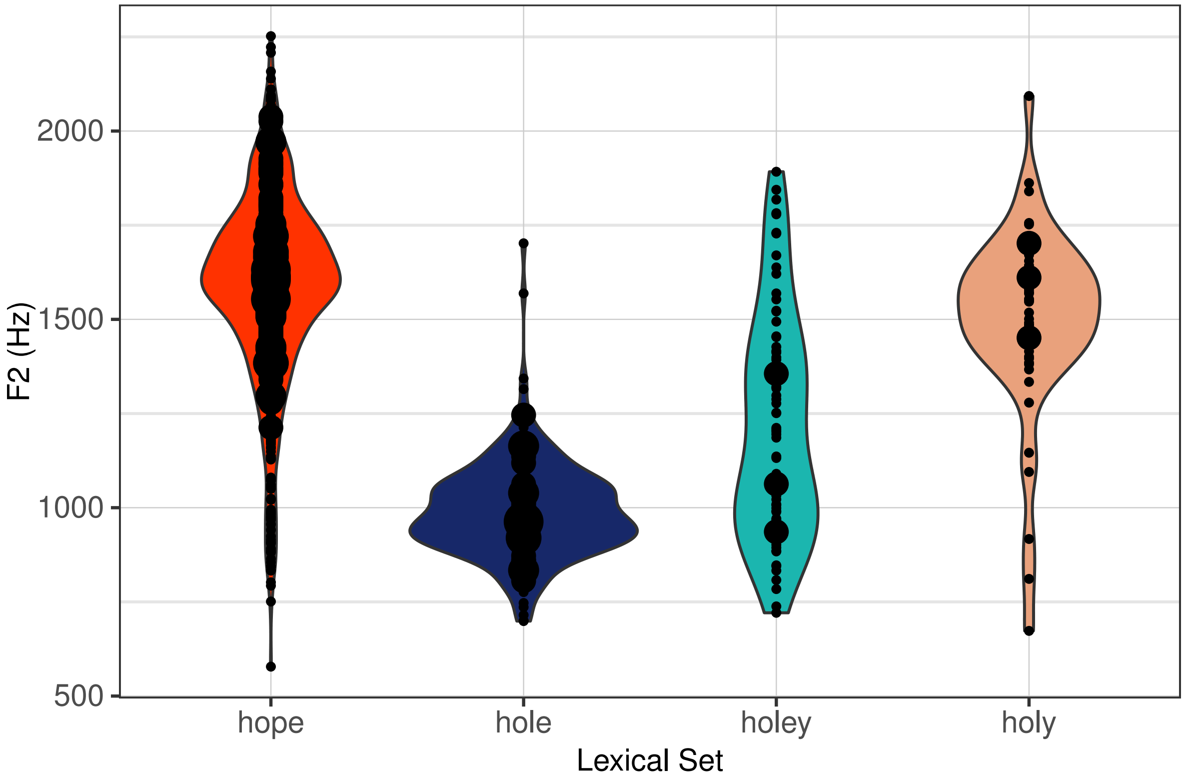

Figure 6.13.: Violin plot of the F2 midpoints (measurement.no = 50) of the GOAL context in CoRP-SE speakers (raw data)

6.5.2. Morpho-phonological Conditioning in DECTE

The pre-/l/ contexts for the DECTE speakers is a small data set, but the data available shows no distinction in either F1 or F2, models are discussed below.

6.5.2.1. DECTE F1

Tables 6.25 and 6.26 show the best fit model for F1 in the DECTE speakers (NB no bi-morphemic words occurred in the DECTE interviews so that category is missing from the model). The best fit model included age group and speaker sex but neither had a meaningful effect size or significance so are left out of the plotting of the model. The parametric terms show a marginal difference between the midpoint of *hope* and *hole* words (58Hz) but no significant smooth term variation. The intercept (Old Female, *hole*) is at 554Hz, approximately 40Hz lower (higher in the vowel space) than the CoRP-SE speakers. It is unlikely that this is a meaningful difference. From looking at the trajectory shape, it can be seen that there is little to no F1 movement in either the *hole* or *holy* words implying that the vowels do not change height, which is similar to the *hope* trajectory (section 6.3.1) but different to the CoRP-SE trajectory which shows evidence of upward moving in the vowel space. The lack of F1 movement in the DECTE vowel contributes evidence towards the conclusion that the speakers are producing a predominantly monophthongal realisation (not a centering diphthong).

	Estimate	Std.Error	t-value	Pr (>|t|)
(Intercept)	553.96	30.10	18.41	<2e-16
lexSet_ordholy	57.92	26.04	2.23	0.03
ageGroupYoung	13.13	37.74	0.35	0.73
sexMale	-19.04	37.75	-0.50	0.61

Table 6.25.: Parametric terms of the model of F1 in the pre-/l/ contexts in DECTE speakers

	edf	p-value
s(measurement.no)	1.00e+00	0.61
s(measurement.no):lexSet_ordholy	1.00e+00	0.53
s(measurement.no):ageGroupOld	1.00e+00	0.61
s(measurement.no):ageGroupYoung	1.15e+00	0.24
s(measurement.no):sexFemale	1.39e-05	1.00
s(measurement.no):sexMale	1.77e+00	0.47
s(measurement.no):preSegnone	1.00e+00	0.71
s(measurement.no):preSegobstruent_liquid	1.00e+00	0.62
s(measurement.no):preSegoral_apical	1.00e+00	0.60
s(measurement.no):preSegoral_labial	1.32e-05	0.50
s(measurement.no):preSegpalatal	2.09e+00	0.32
s(measurement.no):preSegvelar	1.39e+00	0.57
ti(measurement.no,dur)	3.02e+00	0.00
random smooth by speaker		
	1.05e+01	<2e-16
random smooth by word		
	1.45e+01	<2e-16

Table 6.26.: Table showing smooth terms the model of F1 in the pre-/l/ contexts in DECTE speakers

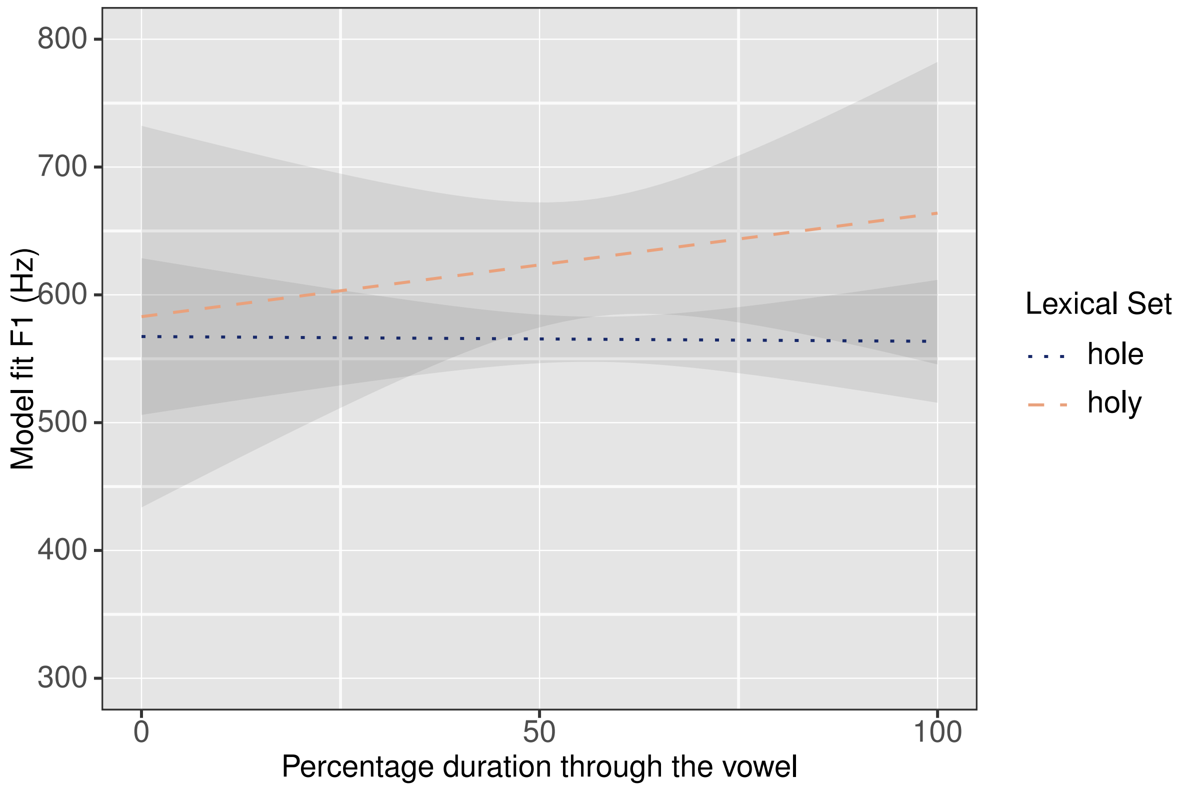

Figure 6.14.: GAMM plot of F1 of the GOAL context in DECTE speakers

6.5.2.2. DECTE F2

The best fit model for the pre-/l/ contexts in the DECTE speakers is shown in tables 6.28 and 6.29. In the parametric terms there is a significant effect of age group, and speaker sex, but not lexical set. The calculated mean midpoints can be seen in table 6.27

The female speakers show an overall less front vowel, potentially showing more backing effect than the male speakers. The male speakers show more movement in the diphthong (see fiture 6.15). However, overall, none of the DECTE speakers show any syllable structure conditioning of the GOAT vowel in pre-/l/ position.

speaker group	mean F2 at midpoint
Old Female	985.93
Young Female	1161.24
Old Male	1152.73
Young Male	1328.04

Table 6.27.

	Estimate	Std.Error	t-value	Pr (>|t|)
(Intercept)	985.93	37.48	26.31	<2e-16
lexical set (baseline hole)				
holy	-15.86	75.44	-0.21	0.83
age group (baseline Old)				
Young	175.31	36.75	4.77	2.30e-06
speaker sex (baseline Female)				
Male	166.08	37.48	4.43	1.11e-05

Table 6.28.: Parametric terms in the model of F2 in the pre-/l/ contexts in DECTE speakers

	edf	p-value
s(measurement.no)	4.31e+00	4.6e-05
s(measurement.no):lexSet_ordholy	1.00e+00	0.37
s(measurement.no):ageGroupOld	1.0e+00	0.68
s(measurement.no):ageGroupYoung	2.28e-05	0.50
s(measurement.no):sexFemale	2.67e-05	1.00
s(measurement.no):sexMale	1.00e+00	0.00
s(measurement.no):preSegnone	1.00e+00	0.478
s(measurement.no):preSegobstruent_liquid	1.00e+00	0.91
s(measurement.no):preSegoral_apical	1.00e+00	0.34
s(measurement.no):preSegoral_labial	7.45e-04	1.00
s(measurement.no):preSegpalatal	1.00e+00	0.24
s(measurement.no):preSegvelar	1.00e+00	0.95
tensor product interaction measurement.no & duration (continuous)		
	3.68e+00	0.19
random smooth by speaker		
	8.15+00	<2e-16
random smooth by word		
	1.65e+01	<2e-16

Table 6.29.: Table showing smooth terms the model of F2 in the pre-/l/ contexts in DECTE speakers

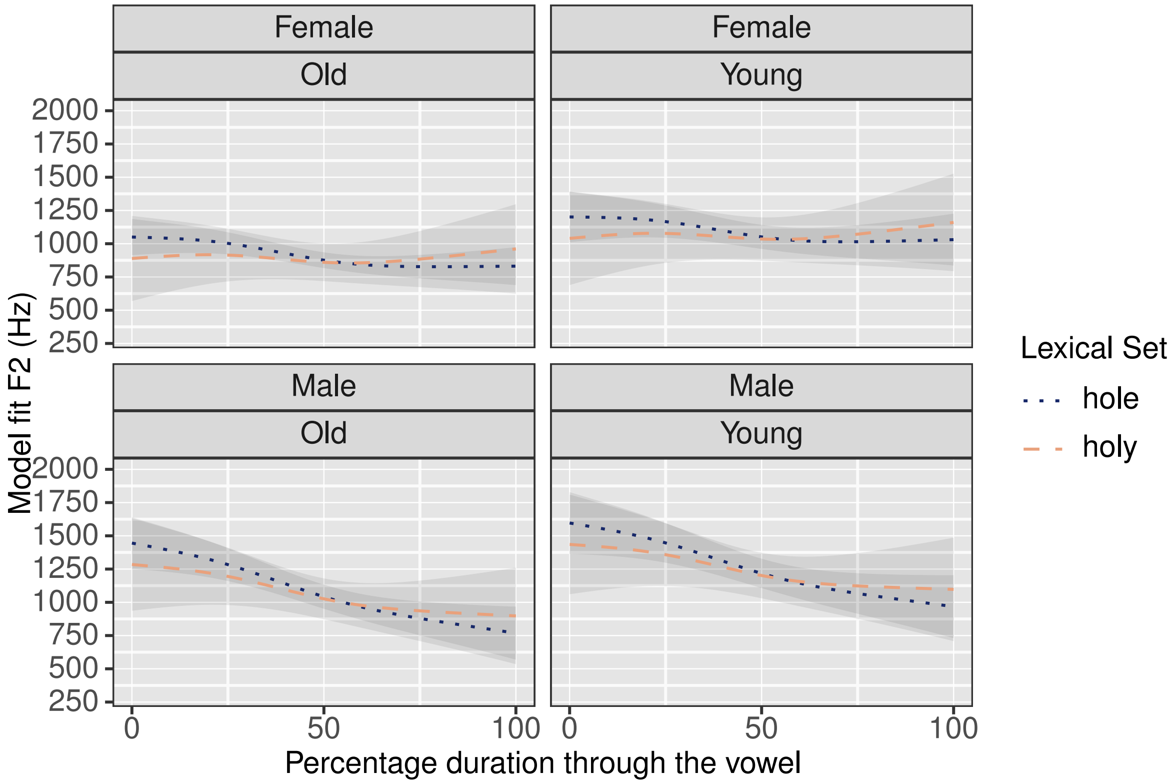

Figure 6.15.: GAMM plot of F2 of the GOAL context in DECTE speakers

6.5.3. Morpho-phonological Conditioning in CoRP-NE

6.5.3.1. CoRP-NE F1

Tables 6.30 and 6.31 show the best fit model of F1 of the CoRP-NE speakers. The interaction improved the model fit but there aren't large differences between the speaker groups. In general, the *hole* words have a higher F1 (lower in the vowel space), the *holy* words a lower F1 (higher in the vowel space), and the female speakers have a lower F1. However the confidence intervals for these differences for these differences overlap for most of the trajectory so the broad conclusion that can be drawn is that, similar to the CoRP-SE speakers, the CoRP-NE speakers show evidence of a raising diphthong in pre-/l/ position and no difference in F1 between the *hole*, *holey*, and *holy* contexts.

	Estimate	Std.Error	t-value	Pr (>|t|)
(Intercept)	591.10	11.31	52.26	<2e-16
age & sex & lexical set (baseline Old - Female - hole)				
Young - Female - hole	13.83	12.60	1.10	0.27
Old - Male - hole	83.24	19.12	4.35	1.36e-05
Young - Male - hole	38.37	13.06	2.94	0.00
Old - Female - holey	-33.76	8.02	-4.21	2.62e-05
Young - Female - holey	-4.47	14.53	-0.31	0.76
Young - Male - holey	-8.05	14.90	-0.54	0.60
Old - Female - holy	-39.25	9.23	-4.25	2.13e-05
Young - Female - holy	-9.44	15.22	-0.62	0.54
Young - Male - holy	-0.66	15.62	-0.04	0.97

Table 6.30.: Parametric terms in the model of F1 in the pre-/l/ contexts in CoRP-NE speakers

	edf	p-value
s(measurement.no)	5.95	<2e-16
age & sex & lexical set (baseline Old - Female - hole)		
Young - Female - hole	1.20	0.90
Old - Male - hole	1.28	0.87
Young - Male - hole	1.00	0.21
Old - Female - holey	1.00	0.55
Young - Female - holey	2.44	0.21
Young - Male - holey	3.11	0.01
Old - Female - holy	1.36	0.64
Young - Female - holy	1.00	0.61
Young - Male - holy	2.64	0.02
preceding segment		
liquid	1.00	0.96
nasal_labial	0.66	0.98
none	2.89	0.04
oral_apical	2.97	0.08
oral_labial	1.00	0.46
palatal	1.42	0.71
velar	1.00	0.99
tensor product interaction measurement.no & duration (continuous)		
	12.16	<2e-16
random smooth by speaker		
	28.81	<2e-16
random smooth by word		
	103.30	<2e-16

Table 6.31.: Table showing smooth terms the model of F1 in the pre-/l/ contexts in CoRP-NE speakers

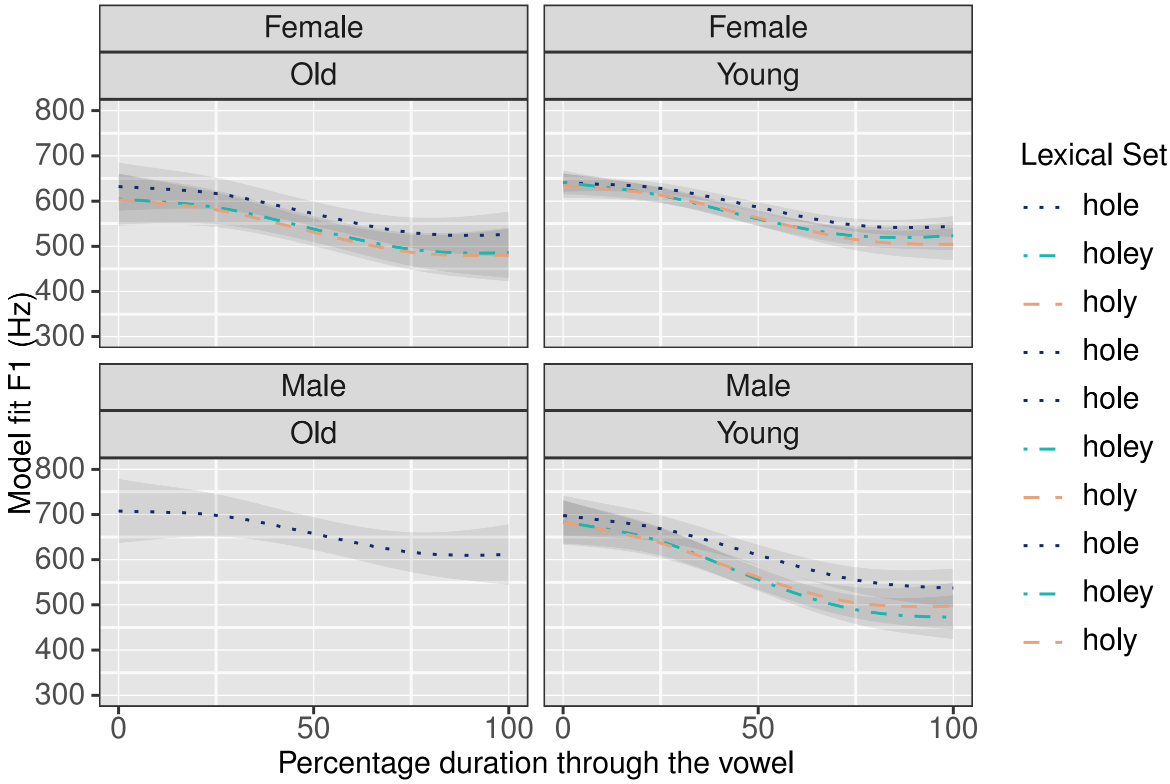

Figure 6.16.: GAMM plot of F1 of the GOAL context in CoRP-NE speakers

6.5.3.2. CoRP-NE F2

The best fit model for F2 of the pre-/l/ contexts in CoRP-NE speakers is shown in tables 6.33 and 6.34. An interaction between lexical set, age group, and speaker sex improved the model fit, the mean F2 value (calculated from the interaction values) for each of the lexical sets can be seen in table 6.32.

However, as can be seen in figure 6.17, the confidence intervals for these intervals overlap for the majority of groups (the gap seen in young female speakers is the only distinction, = 200Hz). The *holy* set is not as front as the *hope* words in table 6.19, and the overall

lexical set	mean F2
hole	1044Hz
holey	1170Hz
holy	1217Hz

Table 6.32.

position and combining that information with the overlap in confidence intervals implies that all the pre-/l/ context words have the backed diphthong.

	Estimate	Std.Error	t-value	Pr (>|t|)
(Intercept)	1116.10	56.88	19.62	<2e-16
age & sex & lexical set (baseline Old - Female - hole)				
Young - Female - hole	-99.22	62.39	-1.59	0.11
Old - Male - hole	-116.63	92.99	-1.25	0.210
Young - Male - hole	-72.33	64.63	-1.12	0.26
Old - Female - holey	126.72	41.54	3.05	0.00
Young - Female - holey	90.04	74.06	1.22	0.22
Young - Male - holey	-55.26	75.87	-0.73	0.47
Old - Female - holy	148.38	49.01	3.03	0.00
Young - Female - holy	118.36	78.47	1.51	0.13
Young - Male - holy	36.06	80.27	0.45	0.65

Table 6.33.: Parametric terms in the model of F2 in the pre-/l/ contexts in CoRP-NE speakers

	edf	p-value
s(measurement.no)	5.91e+00	<2e-16
age & sex & lexical set (baseline Old - Female - hole)		
Young - Female - hole	1.00	0.69
Old - Male - hole	1.00e+00	0.86
Young - Male - hole	1.00	0.96
Old - Female - holey	1.00	0.59
Young - Female - holey	1.00	0.93
Young - Male - holey	1.23e+00	0.83
Old - Female - holy	1.00e+00	0.64
Young - Female - holy	1.00e+00	0.91
Young - Male - holy	1.00e+00	0.92
preceding segment		
liquid	2.53e+00	0.03
nasal_labial	2.52e-04	1.00
none	1.99e+00	0.24
oral_apical	2.69e+00	0.01
oral_labial	1.63e+00	0.34
palatal	1.01e+00	0.00
velar	1.00e+00	0.03
tensor product interaction measurement.no & duration (continuous)		
	5.88e+00	<2e-16
random smooth by speaker		
	2.15e+01	<2e-16
random smooth by word		
	8.67e+01	<2e-16

Table 6.34.: Table showing smooth terms the model of F2 in the pre-/l/ contexts in CoRP-NE speakers

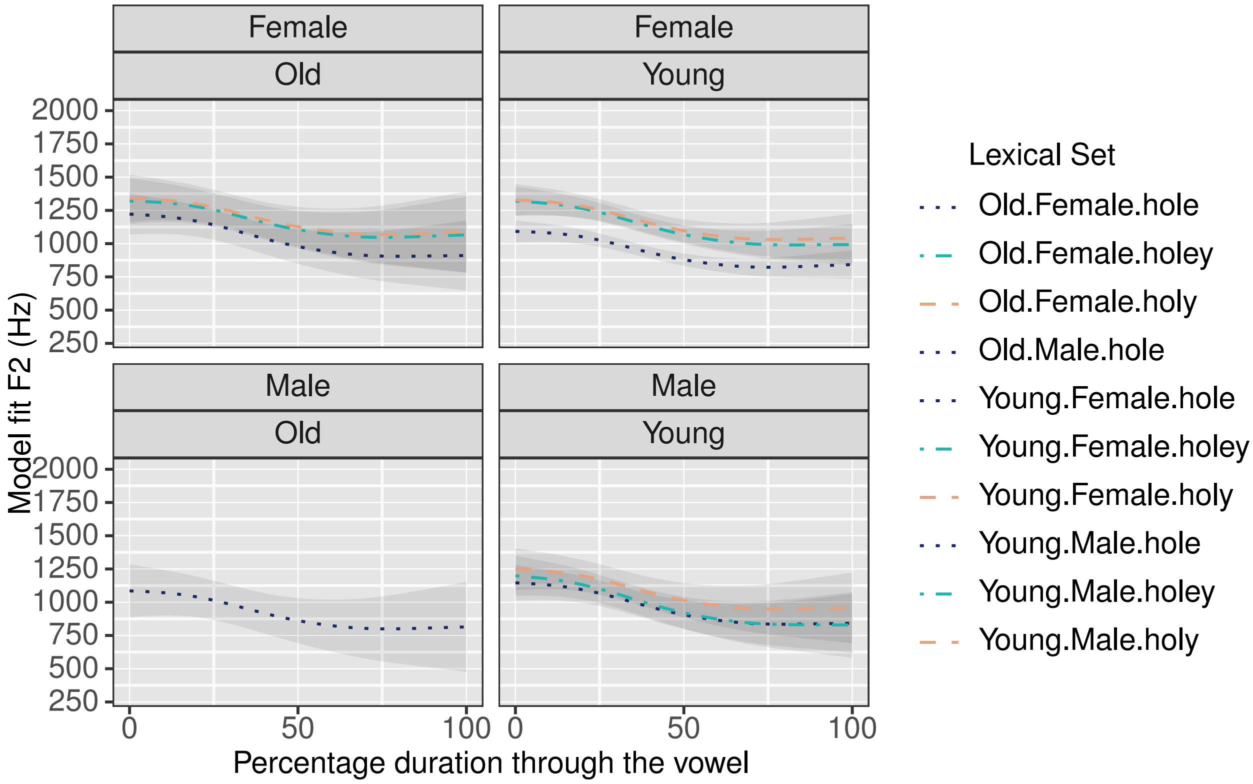

Figure 6.17.: GAMM plot of F2 of the GOAL context in CoRP-NE speakers

6.5.4. Comparison of the pre-/l/ contexts in all three speaker groups

6.5.4.1. pre-/l/ F1 comparison

Figure 6.18 shows the F1 trajectories for the pre/l/ context in all three speaker groups. The DECTE speakers have the least movement in the trajectory, implying a monophthongal variant. CoRP-SE has the most movement through the trajectory, but CoRP-NE has a similar trajectory shape. None of the groups show distinction between the different morpho-phonological environments.

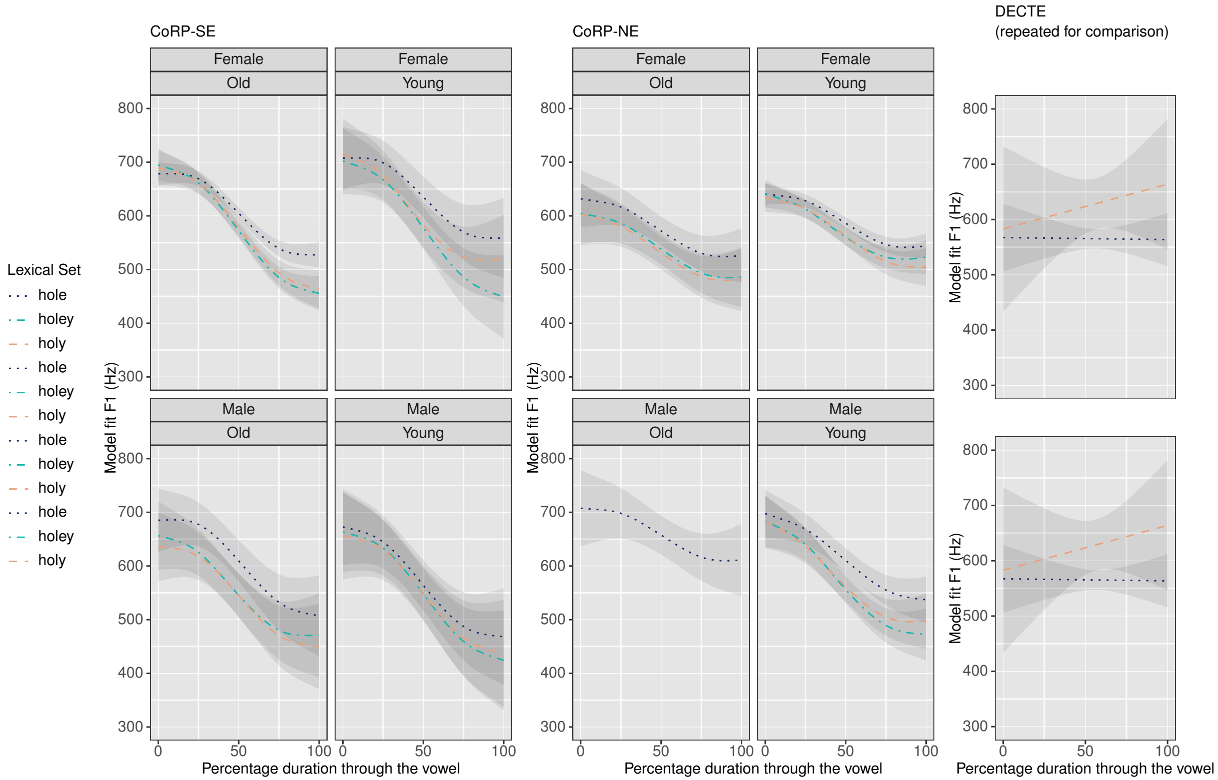

Figure 6.18.: GAMM plot of F1 of pre-/l/ position in all three speaker groups

6.5.4.2. pre-/l/ F2 comparison

Figure 6.19 shows the F2 trajectories for the pre/l/ context in all three speaker groups. There is a high level of variation and large confidence intervals in the DECTE speakers, but they clearly do not show a distinction between the contexts. The CoRP-SE speakers show distinction between all three contexts. With the *holey* words showing as between the *hole* and *holy* words. As discussed above (section 6.5.1.2) this may be showing an in-between diphthong or a situation where some words are front and some back, but none of the predictors capture the distinction. CoRP-NE speakers do not show a distinction between the contexts; the implications of this, combined with the F1 variation, will be discussed in the conclusion (section 6.6).

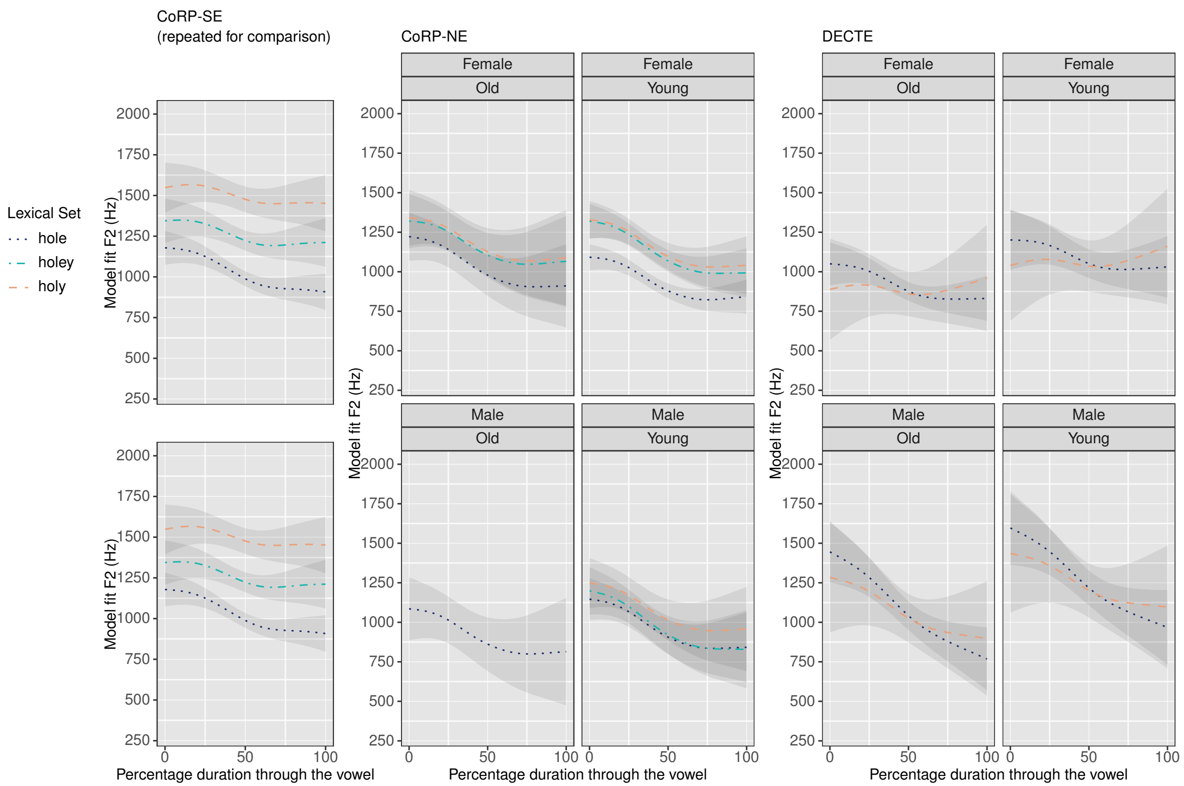

Figure 6.19.: GAMM plot of F2 of pre-/l/ position in all three speaker groups

6.6. Conclusion

WIith respect to the GOAT vowel in *hope* context, and the GOAT-GOAL split. The CoRP-NE speakers behave in a non regional manner. In the *hope* words they show a vowel that is highly similar to the CoRP-SE vowel, a diphthong that shows change in F1 and F2. In the split we see a difference in F2 but not in F1, again, similar or identical to the CoRP-SE speakers.

The situation in the pre/l/ context is far less clear. The CoRP-SE speakers show a trajectory in the *holey* words that is between the *holy* (+*hope*) words and the *hole* words. However, looking at the raw data shows that potentially this is caused by a mixture of words in the two positions rather than a third trajectory position. Unfortunately, whatever predictor is conditioning the variation between the two places is not captured in the model. It may be a combination of word and individual speaker (both modelled as random intercepts) interacting with how speakers process the morphology. for example, solar and polar could be processed as bimorphemic or monomorphemic.

	hole	***hole-y***	***holy***	***hope***
Stage 0	[həʊl]	[həʊliː]	[həʊliː]	[həʊp]
Stage 1	[hɔul]	[həʊliː]	[həʊliː]	[həʊp]
Stage 2	[hɔul]	[hɔuliː]	[həʊliː]	[həʊp]
Stage 3	[hɔul]	[hɔuliː]	[hɔuliː]	[həʊp]

Table 6.35.: Life cycle of Phonological Processes (Bermúdez-Otero, 2007; Bermúdez-Otero and Trousdale, 2012)

The theory of the life cycle of phonological processes (Bermúdez-Otero, 2015) suggests a change moving through levels of the grammar. In this context that would be the back of the GOAT vowel in pre-/l/ position. This would lead to potential degress of the changed that can be described as level, as shown in table 6.35. The CoRP-SE speakers could be considered as sitting at stage 1 or 2 of the life cycle.

The CoRP-NE speakers have the same diphthong in all three pre-/l/ positions, indicating stage 3 of the life cycle. In this they differ from both the DECTE and CoRP-SE speakers, leading to the question as to where their pattern has come from. I propose two potential explanations.

1. Stage 3 is a further stage in the change. It is possible that the North East were experiencing the same phonetically triggered change as the South East but the split has now progressed further after the South East had stopped changing, or the change has simply progressed faster than in the South East.

2. After the split had reached stage 2 in the South East, North East speakers targetted that pattern but the morphological complexity was reanalysed and simplified to a phonological rule that applies pre-/l/ whether dark or not and whatever the position in the syllable.

Both of these potential explanations can only be possible if the RP speech community in the North East is varying and changing in at least some way separately to the RP speech community in the South East, rather than acquiring and following exactly the changes from the South East as suggested by Trudgill (2008).

In answer to research questions 1 and 3 the CoRP-NE speakers behave regionally in terms of their GOAT vowel and monosyllabic GOAT-GOAL split but in terms of the morphological conditioning of the split are not behaving in a regional or non-regional manner.

In answer to research question 2, the results suggest that the proposed non-regional speech community is not homogeneously non-regional, but instead has regional variation, to at least some extent.

Chapter 7.

Discussion

7.1. Introduction

In this chapter I will discuss the results of chapters 4,5, and 6, bringing together the results of the analysis of all three variable to answer the research questions as outlined in chapter 1 (repeated below). I will then consider further potential explanations of the variation found, before outlining the limitations of the research and potential future work.

7.2. Research Questions

1. In terms of pronunciation, do British north-eastern 'RP' speakers behave regionally or non-regionally? Variables to be considered:
    - FOOT-STRUT split
    - TRAP-BATH split
    - GOAT allophony
2. What insights do these differences give us into the nature of a non-regional sociolect comprising diffuse speech communities?
3. If the results of research question 1 vary depending on the variable in question, what phonological or social factors control this?

7.2.1. Research Question 1

In terms of pronunciation, do British north-eastern ‘RP’ speakers behave regionally or non-regionally? In discussion of regional behaviour it is important here to note that in the context of the discussion of the three groups of speakers “regional" behaviour is defined as the patterns found in the DECTE speakers, and “non-regional" behaviour is defined as the patterns found in the CoRP-SE speakers. These definitions, alongside potential other causes for the variation found will be discussed in section 7.3.

7.2.1.1. FOOT-STRUT Split

In the analysis of the FOOT-STRUT split it was found that CoRP-SE and DECTE speakers behaved as expected. The CoRP-SE speakers show a split, which is mostly characterised by an F1 difference in, and some difference in F2. The DECTE speakers do not overall show evidence of a split.

The CoRP-NE speakers are variable with respect to the FOOT-STRUT split. Male speakers behave regionally, their FOOT-STRUT split patterns with the DECTE speakers. Female speakers do not behave regionally or non-regionally but instead have a different form of split.

In conclusion, the male CoRP-NE speakers behave regionally, the female CoRP-NE speakers behave neither regionally nor non-regionally. I discuss the status of this unique behaviour in section 7.3

7.2.1.2. TRAP-BATH Split

In the analysis of the TRAP-BATH split, it was again found that the CoRP-SE and DECTE speakers behave as expected. The CoRP-SE speakers show a TRAP-BATH split, characterised by a difference in F2 and duration (and a small amount of F2 difference). BATH is found to pattern with PALM. The DECTE speakers do not show any evidence of a split, again, as would be expected from speakers in the North of England. The BATH vowel in these speakers patterns with TRAP in F2 and duration. The DECTE speakers also show a

PALM vowel that is further forward and shorter than in the CoRP-SE speakers .

The CoRP-NE speakers can be described as behaving broadly regionally, with no clear TRAP-BATH split. However, they do overall show a broader range of F2 values in the BATH words and also a longer duration. It is possible that some individual BATH words are being produced with a PALM-like vowel, causing the variation seen in F2, and while the absolute position of the PALM vowel wasn't directly under study here it was found that they generally have a lower F2, more simlar to the CoRP-SE speakers. Despite showing TRAP-like F2 in their BATH vowel the duration that CoRP-NE speakers have in BATH vowels is not as short as TRAP (though also not as long as PALM). Since pre-fricative lengthening was the phonetic effect that originally started the TRAP-BATH split in the South (Wells, 1982*a*; Barber et al., 2010) the lengthening may be evidence that these speakers are moving towards a split.

In conclusion, the CoRP-NE speakers can be described as behaving broadly regionally but with some indications non-regional behaviour or a change in progress, that could lead towards non-regional behaviour.

7.2.1.3. GOAT allophony

In analysis of the GOAT vowel, CoRP-SE speakers were found to have a diphthong in all contexts, and a GOAT-GOAL split found in F2. DECTE speakers had a more monophthongal vowel and no split. With respect to the morphological conditioning of the GOAT-GOAL split, the CoRP-SE speakers seem to sit at a mixture of stage 1 and stage 2 of the life cycle of phonological processes. The DECTE speakers show a monophthongal GOAT vowel and no evidence of a GOAT-GOAL split in any morphological context.

CoRP-NE speakers show non-regional behaviour with respect to the monosyllabic GOAT vowel and the GOAT-GOAL split, but have different morphological conditioning, showing stage 3 of the life cycle, implying either a different acquisition of the morphological conditioning of the split or a more advanced progression of the change, either of which are non-regional behaviour.

In conclusion, the CoRP-NE speakers show both regional and non-regional behaviour, as well has instances of unique patterns not seen in either the CoRP-SE or DECTE speakers.

The status of these unique patterns is discussed below in section 7.3.

7.2.2. Research Question 2

What insights do these differences give us into the nature of a non-regional sociolect comprising diffuse speech communities?

Research question 2 brings together the detailed results and leads us to consider the nature of RP, as discussed in sociolinguistic literature and as understood from the speakers analysed in this project.

Observing the FOOT-STRUT split in the CoRP-NE group shows a similar pattern to Labovian change from below the level of consciousness. The female speakers are showing the new variant in STRUT as described by Labov (2001*a*) and others explaining changes from below the level of consciousness lead by female speakers. While the FOOT-STRUT split is standard elsewhere in the country, if it is not socially salient in the North East it is still possible that the change is moving in this way.

The TRAP-BATH split is not found in vowel quality of the CoRP-NE speakers. Since the split has long been complete in the South East, the length distinction found in the CORP-NE speakers is more likely to be a phonetic effect (which may lead to a change locally) than the beginning of the split spreading from the South East.

The GOAT-GOAL split is more recent (2), and while there is no effect of age, the morphological conditioning can indicate the progress through a change via the life cycle of phonological processes. The CoRP-NE speakers show a stage of the split that is either a simplified version of the southern split, or further ahead of the change. Either of these possibilities is not the innovation process suggested by Trudgill.

There are two particularly prevalent theories in the literature regarding the status RP as a non-regional accent. First, as indicated by the diagram below, that regional variation decreases up the socioeconomic spectrum (Honey, 1985; Kerswill, 2007; Wells, 1982*a*). In comparing speakers privately educated in the North East to those of a similar background in the South East regional, variation was found. Therefore, two possible conclusions are available:

1. that a 'non-regional' accent can only be found in the North East by going further up the socio-economic spectrum than is needed in the South East

   OR

2. that a regional accent does not exist, instead there is high social class/private education accent found in each area, which shows less regional variation but not none

This project has compared two equivalent groups, privately educated, in the two areas. CoRP-SE are from the canonical background of RP speakers (Fabricius, 2000*b*; Wells, 1997). Therefore, since the CoRP-NE speakers do not match the CoRP-SE speakers, either the non-regional accent does exist in the North East but has not been accessed (conclusion 1), or neither group is speaking a non-regional accent (conclusion 2). Conclusion 1 could support the view of Trudgill (2008) and others, that RP is not disappearing and there are no fewer native speakers, instead change is occurring but has not been captured by the speakers recorded and comparisons made in this research. Conclusion 2 would support the view of Lindsey (2019), Macaulay (1988), and others, that RP has indeed disappeared.

Secondly, that change in RP happens from innovations in working class South East accents moving up the socio-economic spectrum and then spreading across the country (Trudgill, 2008). This hypothesis is not supported by the results of this project. The FOOT-STRUT and TRAP-BATH splits are not recent innovations, the changes have been established long enough in the South East and RP to expect them to exist across the country. However, the FOOT-STRUT split is not found at all in male CoRP-NE speakers and is found in a different form in the female speakers. The morphological conditioning of the GOAT-GOAL split shows that the CoRP-NE speakers have either progressed further in the change than the CoRP-SE speakers, or have acquired a simplified version of the split. Neither of these possibilities hold with the suggestion of change diffusing across the country to make one accent variety.

7.2.3. Research Question 3

If the results of research question 1 vary depending on the variable in question, what phonological or social factors control this?

There is not a clear pattern of overall regional or non-regional behaviour in the CoRP-NE speakers, instead we see broadly regional behaviour in the TRAP-BATH split (with some caveats, as described above), regional behaviour in the FOOT-STRUT split in the male speakers. We see non-regional behaviour in the simple GOAT-GOAL split. However, we also find patterns that don't fit into either of those categories. The female speakers show a STRUT vowel that does not appear in either the CoRP-SE or DECTE speakers and the pattern of the morphological conditioning of the GOAT-GOAL split in disyllabic contexts seen in the CoRP-NE speakers is not the same as in either CoRP-SE or DECTE. However, it is also not clear that there is a particular social or linguistic context that conditions regional or non-regional behaviour. The data collected here cannot answer what pattern is governing this variation but some suggestions are made in the further work section below.

7.3. Further Discussion of Variation

It is clear from the results of the research questions that there are more factors conditioning the variation than were analysed here. Due to limitations of speaker numbers and likelihood of collinearity it was not possible to include a more detailed analysis of social class including level of education, occupation, and parents' occupation and education. It is clear from the results of all three variables that the CoRP-NE speakers are showing patterns that are not 'regional' in the sense that they are not the same as the DECTE speakers. However, if regional accents are defined by the speakers who live in the area, these patterns are part of the variation found in the accents of speakers in the North East. I suggest that the differences seen between DECTE and CoRP-NE are an example of socially conditioned variation. From the current results it is clear that private education is a factor that is contributing to the variation, however, it is likely that a more complex combination of social class factors are combining and conditioning the patterns seen. This proposal is supported by the distinction between male and female found in the FOOT-STRUT split results. It has been seen in past studies of Tyneside English that speaker sex interacts with social class, for example Watt's (1998) work on the GOAT vowel found that working class male speakers were the most likely to use the centring diphthong [ʊə] whereas middle class female speakers were the most likely to use the closing diphthong [əʊ/oʊ]. Further work on understanding these patterns of variation is suggested below in section 7.6. Since this is the pattern observed in the North East it is likely that a similar process is in progress in the South East, that is that there is not a non-regional variety but that there is high levels of social conditioning, including from private education, but also from other social class factors (work of this nature has been begun by other researchers, including for example Cole and Strycharczuk 2022).

7.4. Implications for RP in Linguistics

The results discussed have shown that it is possibel that a fully non-regional variety of English is either rare and difficult to access or does not exist at all. It is then in important to now consider that basing research in English accents on a model that includes a non-regional variety is no longer sufficient. As such, it is perhaps time to consider moving away from both the concept of a non-regional variety (even as a standard), and from the term RP altogether. As considered in chapter 2, RP is used in various sub-fields of Linguistics. Before focussing on sociolinguistics I wish to briefly consider the uses of RP in other fields. First, English language teaching (EAL, TESOL, etc.) still often use RP as a term for a model variety of English. Lindsey (2019) is already clear that the sociological context for RP no longer exists, and a new standard is needed. He defines this as 'Standard Southern British (SSB)'. Secondly, Phonetics studies often use RP as an exemplifier of English (for example, when comparing to other languages). It is not helpful to suggest a full sociolinguistic study needs to be done for each speaker in a phonetics study, but it is important to stop defining RP by features and it could be possible for phonetics research to include some regional and/or social information for speakers used. It would also be possible to further consider SSB (as defined by Lindsey 2019) as an exemplifying variety of English.

Regarding sociolinguistics, in order to move away from the idea of a non-regional variety, I suggest that we move away from using RP as a standard or as a description of a sociolect. With respect to standard language ideology research is needed on the interactions between people's perception of a standard and their understanding of regional variation. For example, anecdotally, people in Newcastle perceive a southern accent as *posh*, even if it has features that a southerner would perceive as working class. With respect to sociolinguistic modelling of accent variation in England, more work is needed on understanding how social factors interact with regional variation, including working to include as much of the social spectrum as possible. This work will also need a greater understanding of social class, or at least what is defined as social class. As discussed in chapter 2, class is a socially constructed stratification of society and further understanding of how English speakers (and English society) bring together factors such as profession, economic situation, family background, and education (type and level) is needed. It is only in bringing together all these types of research that we will be able to move towards a robust understanding of the sociolinguistics landscape of English spoken in England.

7.5. Limitations

There were multiple limitations in the design and analysis of this study; chapter 3 explained the rationale for speaker selection but there were still issues surrounding the speaker selection and categorisation. When defining South East versus North East there is the question of whether to use the area the person lived with their family or where they were educated. I chose to use the region of education, and only one speaker (008_MO) had a discrepancy between the two. This issue leads into the question of boarding. Four of the CoRP-SE speakers boarded, two of the CoRP-NE speakers boarded, one in the North East and one in the South East before moving to a day school in the North East. There are also far few boarding schools in the North East overall. It is unquestionable that boarding had a significant influence on the development of RP but the mobility of these speakers (through both education and their adult lives) is not addressed when the term 'non-regional' is used. Further investigation into the impact of mobility is needed but due to constraints of time and sample size was not possible in this study.

7.6. Future Work

This section will suggest a set of potential future directions that this research could take, briefly towards answering some of the phonetic questions remaining from the analysis of the GOAT-GOAL split in chapter 6, before particularly focussing on ways to move further towards answers to research questions 2 and 3.

The GOAT-GOAL split is related to the process of /l/ darkening, with the backing triggered by the velar /l/ gesture. This project only analysed the vowel trajectory and at least in the CoRP-SE speakers, was not able to capture predictors of all the variation in the morphologically complex environments. Further work could consider the acoustic trajectory of both the vowel and /l/ segment together (as considered in the GOOSE vowel by Strycharczuk and Scobbie 2016), and also conduct ultrasound measurements of the articulatory position of the /l/ (Turton, 2014*b*). This data would indicate the relationship between /l/ darkening and vowel retraction, and alongside perceptual data regarding the processing of low frequency bimorphemic words such as *Polish* and *solar* Sampson (1985), could lead to better understanding of the morphological conditioning of the split and hence to further understanding of the differences between the CoRP-SE and CoRP-NE positions in the change.

The research covered in this thesis was a high level sociolinguistic study and did not consider detailed social network influences. Further work could include mapping of social networks (Milroy and Milroy, 1992) in both childhood and present (from the participants presented here and future recordings) to further understand the nature of the speech community at a local level and discover the influences on individual's accents. Alongside this, collecting data from teen speakers, particularly from the CoRP-NE speaker group, while they are at the private schools will help to understand if the variation seen is caregiver or school effect, or change through the lifespan.

The variation in regional versus non-regional behaviour does not follow any clear pattern based on the data that has been analysed. However there are various avenues for understanding the variation. The first is social saliency and attitudes. The TRAP-BATH split is very salient and generally marks northern identity, the FOOT-STRUT is less socially salient but is differently marked, Wells (1982*b*) describes not having the split as “vulgar". Comparing these two data sets it could be suggested that the more marking of a regional identity a variant is, the more likely speakers are to retain it, even when they don’t have other regional variants. Anecdotally a front GOAT diphthong is marked in the south of England but there is very little discussion of the social saliency of the original vowel or the split across the country. More work on understanding the saliency and social meaning of these three variables needs to be undertaken to develop this possibility further. One way that this could be done could be using matched guise tests that manipulate the BATH, STRUT, and GOAT vowels and test people’s perceptions and associations with the variants.

This project did not consider individual behaviour, but a related possibility, with respect to the lack of clear pattern that predicts regional behaviour, is that individuals are working on a hierarchy of variables. For example, if we rank the regionality of the variables as diphthongal GOAT (and GOAT-GOAL split) > FOOT-STRUT split > TRAP-BATH split, any speaker that has a TRAP-BATH split will also have a FOOT-STRUT split and diphthongal GOAT, a speaker could not have a TRAP-BATH split but have a FOOT-STRUT split and monophthongal GOAT, or a speaker could not have a TRAP-BATH split or a FOOT-STRUT split but still have a diphthongal GOAT. Further research would work on analysis of each individual and their behaviour with respect to each variable. If an order can be deduced, the next question would be what controls that order, results from the social saliency analysis mentioned above could give insight into this.

Britain (2017) calls for a new study of the accents of the elite. Since this study has shown that studying RP as one community is no longer appropriate I can echo this call, and in further work I plan to continue to investigate social conditioning of language variation in the North East and in general as a trend in English spoken in England. Social class in both complex and changing in British Society and understanding how this affects and is affected by language will be an ongoing challenge for sociolinguistics. Work will involve considering more social factors and how they relate to each other, for example by combining social information as done by Trudgill (1974).

Appendix A.

Github Repositories

- FOOT-STRUT analysis https://github.com/caitlin91/foot-strut
- TRAP-BATH analysis https://github.com/caitlin91/trap-bath
- GOAT-GOAL analysis https://github.com/caitlin91/goat

Appendix B.

Information Letter for Participants

I am doing an MA and PhD at Newcastle University, funded by the ESRC, under the supervision of Dr Danielle Turton. I am studying the interactions between class, upbringing, language and culture. I'm interviewing people from around the country and talking to them about their experiences, memories and stories.

What I do is simply sit down and talk to people for about an hour, or as long as you feel comfortable talking to me. We'll talk together about topics such as your school, your community, your neighbourhood, the games you used to play as a child, your hobbies, your work life and general personal experiences.

Although the findings of this study will not benefit you directly, by participating in this study you will be contributing to the production of new and potentially important knowledge about social and linguistic interactions and the changes going on around the country. I would be happy to send you a summary of my findings at the end of the project if you like to know how the research turns out.

Because I won't be able to remember everything everyone says exactly, I ask your permission to record our conversation. However, everything that is said is kept entirely confidential and you will remain entirely anonymous. No one will have access to the data except academic researchers and no one will have access to the recordings unless they follow the same procedures as I do for keeping it confidential and anonymous.

By signing below you agree to the above, but participation in the project is entirely voluntary and you may withdraw at any time if you choose. You can tell me not to record part of what you say, or to erase part of it. It is entirely up to you.

Many thanks,

Caitlin Halfacre
c.a.e.halfacre1@newcastle.ac.uk/caitlin.halfacre@gmail.com

Name:____

Signed:____

Date:____

Appendix C.

Participant Information Form

Date of recording: 2018

Name	
Contact email	
Date of Birth	
Sex	
Ethnicity	
Place of Birth (+partial postcode if known)	
Postcode/region for majority of childhood	
Current postcode (first half)	
Number of moves outside your current region	
Junior Education (state /private/boarding etc.)	
Secondary Education (state /private/boarding etc.)	
Level of education completed	
Current Occupation	
Region where parents were raised	
Parents' schooling (state /private/boarding etc.)	
Parents' level of education completed	
Parents' occupation	
Do you have any memory of actively changing your accent?	

Appendix D.

Summary of Variables

D.1. Linguistic Variables

Variable	Explanation
Coda	Whether the vowel is in a syllable with a coda (TRUE/FALSE)
Duration	duration in msec
Following Manner	Manner of following segment, as categorised by FAVE
Following Place	Place of following segment, as categorised by FAVE
Following Voicing	Voicing of following segment, as categorised by FAVE
Frequency Zipf	Zipf Frequency according to SUBTLEX-UK (van Heuven et al., 2014)
Lexical Set	Lexical Set Wells (1982*a*). The GOAT vowel words are sub-categorised, see chapter 6
Measurement Number	Percentage through the vowel trajectory (by time)
Preceding Segment	Preceding segment, as categorised by FAVE
Word	Word vowel is found in

Table D.1.: Explanation of linguistic variables in models

D.2. Social Variables

Variable	Explanation
Age Group	Over or under 40 years old (at time of interview)
Corpus	As explained above
ID	Unique identifier by speaker, format:Number_SexAge (DECTE didactic interviews are preceded with a D and categorised A or B for the speaker)
Region	
Sex: Male or Female	As filled in by participant in the information form
Time	Row number when every measurement from the interview is listed in order, a proxy for time to approximate the possibility of any effects of accommodation through the interview.

Table D.2.: Explanation of social variables in models

Appendix E.

Extra Models

E.1.

fixedeffect	estimate	tvalue
(Intercept)	1310.39	38.15
lexSetTHOUGHT	-377.17	-12.76
lexSetschwa	319.19	8.28
sexSum1	-21.36	-1.22
ageGroupSum1	-1.32	-0.08
folManSum1	59.86	2.04
folManSum2	20.54	1.33
folManSum3	-131.94	-7.97
preSeg_smallSum1	-45.21	-1.99
preSeg_smallSum2	-35.43	-2.57
preSeg_smallSum3	-61.31	-2.42
preSeg_smallSum4	100.69	3.16
freq.zipf_z	0.21	0.02
styleSum1	54.94	2.95
styleSum2	-45.86	-1.98

Table E.1.: Linear mixed effects model of F2 of STRUT, THOUGHT, and schwa in the South East

List of References

Alderton, R. (2019), Salience and Social Meaning in Speech Production and Perception, PhD thesis, Lancaster University.

Altendorf, U. and Watt, D. (2008), 'The dialects in the south of England: Phonology', *Varieties of English* **1**, 194–222.

Anderson, S. R. (1981), 'Why Phonology isn' t "Natural"', *Linguistic Inquiry* **12**(4), 493–539.

Badia Barrera, B. (2015), A Sociolinguistic Study of T-glottalling in Young RP : Accent , Class and Education, PhD thesis, University of Essex.

Baranowski, M. (2013), 'The sociolinguistics of GOOSE and GOAT in Manchester English'.

Baranowski, M. (2017), 'Class matters: The sociolinguistics of GOOSE and GOAT in Manchester English', *Language Variation and Change* **29**(3), 301–339.

Baranowski, M. and Turton, D. (2018), 'The FOOT-STRUT vowels in Manchester: Evidence for the diachronic precursor to the split?', *University of Pennsylvania Working Papers in Linguistics: Selected Papers from NWAV 46* **24**(2), 1–8.

Barber, C., Beal, J. C. and Shaw, P. A. (2010), *The English Language: A Historical Introduction*, 2nd edn, Cambridge University Press, Cambridge.

Bates, D., Maechler, M., Bolker, B., Walker, S., Bojesen Christensen, R. H., Singmann, H., Dai, B., Scheipl, F., Grothendieck, G. and Green, P. (2018), 'Linear Mixed-Effects Models using 'Eigen' and S4'.

Baudouin de Courtenay, J. (1972), 'An attempt at a theory of phonetic alternations', *Edward Stankiewicz (ed. and trans.) A Baudouin de Courtenay anthology: The beginnings of structural linguistics* pp. 144–212.

Beal, J. C. (1999), *English Pronunciation in the Eighteenth Century*, Oxford University Press, Oxford.

Beal, J. C. (2012), 'By Those Provincials Mispronounced: The strut Vowel in Eighteenth-Century Pronouncing Dictionaries', *Language and History* **55**(1), 5–17.

Bermúdez-Otero, R. (2006), Phonological Change in Optimality Theory, *in* K. Brown, ed., 'Encyclopedia of Language & Linguistics', 2 edn, Vol. 9, Elsevier, Oxford, pp. 497–505.

Bermúdez-Otero, R. (2007), Diachronic Phonology, *in* P. de Lacy, ed., 'The Cambridge Handbook of Phonology', Cambridge University Press, Cambridge, pp. 497–517.

Bermúdez-Otero, R. (2011), Cyclicity, *in* M. Van Oostendorp, ed., 'The Blackwell Companion to Phonology', Malden, Massachusetts: Wiley-Blackwell, pp. 2019–2048.

Bermúdez-Otero, R. (2015), Amphichronic explanation and the life cycle of phonological processes, *in* P. Honeybone and J. C. Salmons, eds, 'The Oxford Handbook of Historical Phonology', Oxford University Press, Oxford.

Bermúdez-Otero, R. and Trousdale, G. (2012), Cycles and continua: On unidirectionality and gradualness in language change, *in* T. Nevalainen and E. C. Traugott, eds, 'The {Oxford} Handbook of the History of {English}', number January 2019, Oxford University Press, Oxford, pp. 691–720.

Bjelaković, A. (2016), 'The vowels of contemporary RP: Vowel formant measurements for BBC newsreaders', *English Language and Linguistics* pp. 1–32.

Braber, N. and Flynn, N. (2015), The East Midlands, *in* R. Hickey, ed., 'Researching Northern English', John Benjamins B.V.

Britain, D. (2017), 'Beyond the "gentry aesthetic": Elites, Received Pronunciation and the dialectological gaze', *Social Semiotics* **27**(3), 288–298.

Britain, D., Leeman, A., Blaxter, T., Kolly, M.-J., Grossenbacher, S., Calame, M. and Wanitsch, D. (2016), 'Big bad dialectological data? Lessons from the English Dialect App'.

Bybee, J. (2001), *Phonology and Language Use*, Cambridge University Press, Cambridge.

Cole, A. and Strycharczuk, P. (2019), The PRICE-MOUTH crossover in the "Cockney diaspora", *in* 'Proceedings of the 19th International Congress of Phonetic Sciences (ICPhS 2019)', pp. 602–606.

Cole, A. and Strycharczuk, P. (2022), 'The search for linguistically coherent accents: Unsupervised clustering of diphthong variation in Southeast England'.

Conference, T. H. (2023), 'About HMC', https://www.hmc.org.uk/about-us/about-hmc/.

Coretta, S. (2020), 'Tidymv: Tidy Model Visualisation for Generalised Additive Models'.

Corrigan, K. P., Buchstaller, I., Mearns, A. and Moisl, H. (2012), 'The Diachronic Electronic Corpus of Tyneside English'.

Cruttenden, A., ed. (2001), *Gimson's Pronunciation of {English}*, 6th edn, Arnold, London.

Davenport, M. and Hannahs, S. (2010), *Introducing Phonetics and Phonology*, Hodder Education, London.

Deshaies-Lafontaine, D. (1974), A Socio-Phonetic Study of a Québec French Community: Trois-Rivières., PhD thesis, University College London.

Di Paolo, M. and Yaeger-Dror, M. (2011), *Sociophonetics: A Student's Guide*, Routledge, London.

Dobson, E. (1968), *English Pronunciation, 1500-1700*, 2nd edn, Clarendon P., Oxford.

Ellis, A. (1889), *On Early English Pronunciation*, Asher & co., London.

Evans, B. G. and Iverson, P. (2007), 'Plasticity in vowel perception and production: A study of accent change in young adults', *The Journal of the Acoustical Society of America* **121**(6), 3814–3826.

Fabricius, A. H. (2000*a*), 'Between Stigma and Prestige : A Sociolinguistic Study of Modern RP', *Environment* (February).

Fabricius, A. H. (2000*b*), T-Glottalling between Stigma and Prestige: A Sociolinguistic Study of Modern RP, PhD thesis, Copenhagen Business School.

Fabricius, A. H. (2002*a*), 'Ongoing change in modern RP: Evidence for the disappearing stigma of t-glottalling', *English World-Wide* **23**(1), 115–136.

Fabricius, A. H. (2002*b*), 'RP as sociolinguistic object', *Nordic Journal of English Studies* **1**(2), 355–372.

Fabricius, A. H. (2002*c*), 'Weak vowels in modern RP: An acoustic study of happY-tensing and KIT/schwa shift', *Language Variation and Change* **14**(02), 211–237.

Fabricius, A. H. (2006), 'The 'vivid sociolinguistic profiling' of received pronunciation: Responses to gendered dialect-in-discourse', *Journal of Sociolinguistics* **10**(1), 111–122.

Fabricius, A. H. (2007), 'Variation and change in the TRAP and STRUT vowels of RP: A real time comparison of five acoustic data sets', *Journal of the International Phonetic Association* **37**(03), 293–320.

Fabricius, A. H. (2019), The anticlockwise checked vowel chain shift in modern rp in the twentieth century : Incrementations and diagonal shifts, *in* 'ICPhS XVIII: Proceedings of the 19th International Congress of Phonetic Sciences'.

Fallows, D. (1981), 'Experimental Evidence for English Syllabification and Syllable Structure', *Journal of Linguistics* **17**(2), 309–317.

Foulkes, P. and Docherty, G. J. (1999), Derby and Newcastle: Instrumental phonetics and variationist studies, *in* P. Foulkes and G. J. Doherty, eds, 'Urban Voices: Accent Studies in the British Isles', online edn, Arnold, Abingdon, Oxon, pp. 47–71.

Foulkes, P., Scobbie, J. M. and Watt, D. (2010), Sociophonetics, *in* 'Handbook of Phonetic Sciences', 2nd edn, Blackwell, Oxford, pp. 703–754.

Fruehwald, J. (2020), *Syllabifyr: Syllabifier for CMU Dictionary Transcriptions*.

Fudge, E. (1976), 'Long and short [æ] in one Southern British speaker's English', *Journal of the International Phonetic Association* **7**(2), 55–65.

Gimson, A. and Cruttenden, A. (1994), *Gimson's Pronunciation of English*, fifth edn, Arnold, London.

Halfacre, C. (2017), The Life Cycle of Poshness - the Progress of GOAT Backing in Received Pronunciation English., Unpublished BA dissertation, Newcastle University, Newcastle.

Halfacre, C. (2018), North / South Dividers in Privately Educated Speakers : A Sociolinguistic Study of the \FS{} and \TB{} Distinctions in the North-East and the South-East, Unpublished MA dissertation, Newcastle University.

Halfacre, C. (2020), 'GEOFF LINDSEY, English after RP: Standard British pronunciation today. Cham: Springer, 2019. Pp. xvi + 153. ISBN: 978-3-030-4356-8', *Journal of the International Phonetic Association* pp. 1–5.

Halfacre, C. and Khattab, G. (2019), North-South dividers in privately educated speakers: A Sociolinguistic study of Received Pronunciation using the FOOT-STRUT and TRAP-BATH distinctions in the North East and South East of England, *in* 'Proceedings of the 19th International Congress of Phonetic Sciences (ICPhS 2019)', pp. 2665–2669.

Hannisdal, B. R. (2006), Variability and Change in Received Pronunciation A Study of Six Phonological Variables in the Speech of Television Newsreaders, PhD thesis, University of Bergen.

Harris, J. (1989), 'Towards a lexical analysis of sound change in progress', *Journal of Linguistics* **25**(1), 35–56.

HM Government (1868), 'Public Schools Act. CAP. CXVIII.'.

HM Government (1870), 'Elementary Education Act c.75'.

Honey, J. (1985), 'Acrolect and hyperlect - the redefinition of English RP', *English Studies* **66**(3), 241–257.

Hughes, A., Trudgill, P. and Watt, D. (2012), *English Accents and Dialects: An Introduction to Social and Regional Varieties of English in the British Isles*, 4th edn, Routledge, London.

Jansen, S. (2017), 'Change and stability in GOOSE, GOATand FOOT: Back vowel dynamics in Carlisle English', *English Language and Linguistics* **23**(1), 1–29.

Jansen, S. and Braber, N. (2020), 'FOOT-fronting and FOOT-STRUT splitting: Vowel variation in the East Midlands', *English Language & Linguistics* pp. 1–31.

Jones, D. (1917), *An English Pronouncing Dictionary*, Cambridge University Press, Cambridge.

Jones, D. (1972), *An Outline of English Phonetics*, ninth edn, Cambridge University Press, Cambridge.

Kahn, D. (1976), Syllable-Based Generalizations in English Phonology, PhD thesis, University of Massachusetts.

Kerswill, P. (2007), Standard and non-standard English, *in* D. Britain, ed., 'Language in the British Isles', number May, Cambridge University Press, Cambridge, pp. 34–51.

Kettig, T. (2016), 'The BAD-LAD split: Secondary /æ/-lengthening in Southern Standard British English', *Proceedings of the Linguistic Society of America* **1**, 1–14.

Kiparsky, P. (1979), 'Metrical Structure Assignment is Cyclic', *Linguistic Inquiry* **10**(3), 421–441.

Kiparsky, P. (1982), Lexical Morphology and Phonology, *in* I.-S. Y. f. t. L. S. of Korea, ed., 'Linguistics in the Morning Calm: Selected Papers from SICOL-1981', Vol. 1, Hanshin Publishing Company, Seoul, pp. 3–91.

Kroch, A. (1995), Dialect and Style in the Speech of Upper Class Philadelphia, *in* G. Guy, J. Baugh, D. Schriffrin and C. Feagin, eds, 'Towards a Social Sceince of Language: Papers n Honor of William Labov', John Benjamins, Philadelphia.

Krug, M. and Sell, K. (2013), Designing and conducting interviews and questionnaires, *in* M. Krug and J. Schlüter, eds, 'Research Methods in Language Variation and Change', Cambridge University Press, Cambridge.

Labov, W. (1966), 'The Effect of Social Mobility on Linguistic Behavior', *Sociological Enquiry* **36**(2), 186–203.

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

Labov, W. (2001*a*), *Principles of Linguistic Change: Internal Factors*, Oxford: Blackwell.

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

Labov, W. (n.d.), 'Plotnik'.

Labov, W. and Baranowski, M. (2006), '50 {msec}', *Language Variation and Change* **18**(3), 223–240.

Labov, W., Fisher, S., Gylfadottír, D., Henderson, A. and Sneller, B. (2016), 'Competing systems in Philadelphia phonology', *Language Variation and Change* **28**(3), 273–305.

Labov, W., Rosenfelder, I. and Fruehwald, J. (2013), 'One hundred years of sound change in Philadelphia: Linear incrementation, reversal, and reanalysis', *Language* **89**(1), 30–65.

Labov, W., Yaeger, M. and Steiner, R. (1972), *A Quantitative Study of Sound Change in Progress*, Vol. 1, US Regional Survey.

Lass, R. (1976), *English Phonology and Phonological Theory*, Cambridge University Press, Cambridge.

Lawson, E., Scobbie, J. M. and Stuart-Smith, J. (2011), 'The social stratification of tongue shape for postvocalic /r/ in {Scottish English}', *Journal of Sociolinguistics* **15**(2), 256–268.

Lindsey, G. (2019), *English after RP: Standard British Pronunciation Today*, Springer.

Luick, K., Wiessner, E. and Katz, T. (1921), *Historische Grammatik Der Englischen Sprache*, Ripol Klassik.

Macaulay, R. (1988), 'RP R.I.P.', *Applied Linguistics* **9**(2), 115–124.

MacMahon, M. K. C. (1998), Phonology, *in* S. Romaine, ed., 'The Cambridge History of the English Language', Vol. 4, Cambridge University Press, Cambridge, pp. 373–535.

Macy, M. (2001), Social Class, *in* R. Mesthrie, ed., 'Concise Encyclopedia of Sociolinguistics', Elsevier Science Ltd, Oxford, pp. 362–369.

Max Planck Institute for Psycholinguistics (2008), 'ELAN'.

McAuliffe, M., Socolof, M., Mihuc, S., Wagner, M. and Sonderegger, M. (2017), 'Montreal Forced Aligner'.

Meyerhoff, M. (2006), *Introducing Sociolinguistics*, Routledge, New York.

Milroy, L. and Milroy, J. (1992), 'Social network and social class: Toward an integrated sociolinguistic model1', *Language in Society* **21**(1), 1–26.

Minkova, D. (2003), *Alliteration and Sound Change in Early English*, Cambridge University Press.

Mugglestone, L. (2007), *'Talking Proper' the Rise of Accent as Social Symbol*, updated 2nd ed.. edn, Oxford University Press, Oxford.

Nespor, M. and Vogel, I. (1984), The Syllable and the Foot, *in* 'Prosodic Phonology', Walter de Gruyter.

Nycz, J. and Hall-Lew, L. (2014), Best practices in measuring vowel merger, *in* 'Proceedings of Meetings on Acoustics: The Acoustical Society of America through the American Institute of Physics', Vol. 20, pp. 1–19.

Ohala, J. J. (1981), The listener as a source of sound change, *in* C. Masek, R. Hendrick and M. F. Miller, eds, 'Papers from the Parasession on Language and Behavior', Chicago Linguistics Society, Chicago, pp. 178–203.

Parveen, N. (2020), 'Students from northern England facing 'toxic attitude' at Durham University', *The Guardian* .

Payne, A. (1980), Factors controlling the acquisition of the {Philadelphia} dialect by out-of-state children, *in* W. Labov, ed., 'Locating Language in Time and Space', Vol. 1, Academic Pr, New York, pp. 143–178.

Pullum, G. K. (1976), 'The Duke of York gambit', *Journal of Linguistics* **12**.

R Core Team (2021), *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria.

Ramsammy, M. (2015), 'The life cycle of phonological processes: Accounting for dialectal microtypologies', *Language and Linguistics Compass* **9**(1), 33–54.

Reid, I. (1981), *Social Class Differences in Britain*, 2nd edn, Grant McIntyre, London.

Rosenfelder, I., Fruehwald, J., Evanini, K., Seyfarth, S., Gorman, K., Prichard, H. and Yuan, J. (2014), 'FAVE (Forced Alignment and Vowel Extraction) Program Suite'.

RStudio Team (2016), 'RStudio: Integrated Development Environment for R'.

Saefken, B. and Ruegamer, D. (2018), 'cAIC4: Conditional Akaike information criterion for lme4'.

Sampson, R. (1985), 'The '*GOAT* split': A phonological puzzle in one variety of English.', *Anglia-Zeitschrift für englische Philologie* **103**, 282–296.

Schneider, E. W. (1999), 'Sociolinguistic Theory: Linguistic Variation and Its Social Significance, By J. K. Chambers, Oxford, UK: Balckwell, 1995, xxi + 284.', *Journal of English Linguistics* **27**(1), 49–56.

Sóskuthy, M. (2017), 'Generalised additive mixed models for dynamic analysis in linguistics: A practical introduction'.

Sóskuthy, M., Hay, J. and Brand, J. (2019), HORIZONTAL DIPHTHONG SHIFT IN NEW ZEALAND ENGLISH, *in* 'Proceedings of the 19th International Congress of Phonetic Sciences (ICPhS 2019)', pp. 597–601.

Stanley, J. A., Renwick, M. E. L., Kuiper, K. I. and Olsen, R. M. (2021), 'Back Vowel Dynamics and Distinctions in Southern American English', *Journal of English Linguistics* **49**(4), 389–418.

Stoddart, J., Upton, C. and Widdowson, J. D. A. (1999), Sheffield dialect in the 1990s: Revisiting the concept of NORMs, *in* P. Foulkes and G. J. Docherty, eds, 'Urban Voices: Accent Studies in the British Isles', Arnold, London, pp. 72–89.

Strang, B. M. H. (1970), *A History of English*, Methuen, London.

Strycharczuk, P. and Scobbie, J. (2017), 'Whence the fuzziness? Morphological effects in interacting sound changes in {Southern British English}', *Laboratory Phonology: Journal of the Association for Laboratory Phonology* **8**(1).

Strycharczuk, P. and Scobbie, J. M. (2016), 'Gradual or abrupt? The phonetic path to morphologisation', *Journal of Phonetics* **59**, 76–91.

Tagliamonte, S. A. and Mesthrie, R. (2006), *Analysing Sociolinguistic Variation*, Cambridge University Press, Cambridge New York.

Trudgill, P. (1974), *The Social Differentiation of English in Norwich*, Cambridge University Press, Cambridge.

Trudgill, P. (2002), *Sociolinguistic Variation and Change*, Edinburgh University Press, Edinburgh.

Trudgill, P. (2008), 'The historical sociolinguistics of elite accent change: On why RP is not disappearing', *Studia Anglica Posnaniensia* **44**.

Turton, D. (2014*a*), 'Some /l/s are darker than others : Accounting for variation in English /l/ with ultrasound tongue imaging', *Penn Working Papers in Linguistics* **20**(2), 189–198.

Turton, D. (2014*b*), Variation in English /l/: Synchronic Reflections of the Life Cycle of Phonological Processes, PhD thesis, University of Manchester, Manchester.

Turton, D. and Baranowski, M. (2020), 'Not quite the same: The social stratification and phonetic conditioning of the FOOT-STRUT vowels in Manchester', *Journal of Linguistics* pp. 1–61.

van Heuven, W. J. B., Mandera, P., Keuleers, E. and Brysbaert, M. (2014), 'Subtlex-UK: A New and Improved Word Frequency Database for British English', *Quarterly Journal of Experimental Psychology* **67**(6), 1176–1190.

van Rij, J., Wieling, M., Baayen, R. H. and van Rijn, H. (2020), 'Itsadug: Interpreting Time Series and Autocorrelated Data Using GAMMs'.

Warburton, J. (2020), The Merging of the GOAT and THOUGHT Vowels in Tyneside English., PhD thesis, Newcastle University.

Ward, I. (1929), *The Phonetics of English*, Heffer, Cambridge.

Wassink, A., Squizzero, R., Fellin, C. and Nichols, D. (2018), 'Client Libraries Oxford (CLOx) [computer software]'.

Watt, D. (1998), Variation and Change in the Vowel System of Tyneside English, PhD thesis, Newcastle University.

Watt, D. (2000), 'Phonetic parallels between the close-mid vowels of Tyneside English: Are they internally or externally motivated?', *Language Variation and Change* **12**(1), 69–101.

Watt, D. and Allen, W. (2003), 'Tyneside English', *Journal of the International Phonetic Association* **33**(2), 267–271.

Watt, D. and Milroy, L. (1999), Patterns of variation and change in three {Newcastle} vowels: Is this dialect levelling?, *in* P. Foulkes and G. Docherty, eds, 'Urban Voices: Accent Studies in the British Isles', Arnold, London, pp. 25–46.

Weber, M. (1978), *Economy and Society*, University of California Press, London.

Weber, M. (2012), *The Theory of Social and Economic Organization*, The Free Press & The Falcon's Wing Press, Glencoe, Illinois.

Weide, R. (1998), 'The Carnegie Mellon pronouncing dictionary', *release 0.6, www. cs. cmu. edu* .

Weinreich, U., Labov, W. and Herzog, M. (1968), Empirical foundations for a theory of language change, *in* W. P. Lehmann and Y. Malkiel, eds, 'Directions for Historical Linguistics', University of Texas Press, Austin, pp. 95–188.

Wells, J. C. (1970), 'Local Accents in England and Wales', *Journal of Linguistics* **6**(2), 231–252.

Wells, J. C. (1982*a*), *Accents of English: 1 An Introduction*, Cambridge University Press, Cambridge.

Wells, J. C. (1982*b*), *Accents of English 2: The British Isles*, Cambridge University Press, Cambridge.

Wells, J. C. (1997), 'What's happening to Received Pronunciation?', *English Phonetics (English Phonetic Society of Japan)* **1**, 11–24.

Wickham, H., François, R., Henry, L. and Müller, K. (2022), *Dplyr: A Grammar of Data Manipulation*.

Winter, B. (2019), *Statistics for Linguists: An Introduction Using R*, Routledge.

Wood, S. N. (2017), *Generalized Additive Models: An Introduction with R*, 2 edn, Chapman and Hall/CRC.

Zsiga, E. (2020), *The Phonetics/Phonology Interface*, Edinburgh University Press, Edinburgh.