Variation in English /l/:
Synchronic reflections of the life cycle of phonological processes

A thesis
submitted to The University of Manchester
for the degree of
Doctor of Philosophy
in the Faculty of Humanities

2014

Danielle Turton

School of Arts, Languages and Cultures

Contents

Final word count: 70113

List of Figures

List of Tables

ABSTRACT

This thesis is an articulatory investigation into phonological variation and change in English /l/-darkening. Although syllable-based accounts of /l/-darkening state that light [l] occurs in onsets (e.g. 'leap') and a dark variant in codas (e.g. 'peel'), numerous works linking phonology with other subfields of linguistics have shown that this simplified distinction cannot fully account for the variation found. Firstly, /l/-darkening is sensitive to morphosyntactic structure, as shown through overapplication of the process in certain morphosyntactically defined positions: e.g. word-finally in phrases such as 'heal it', or stem-finally before a suffix in words such as 'healing'. In addition, analyses of /l/-darkening from several phonetic studies have led to some arguing against an allophonic distinction altogether, stating that the difference between light and dark variants is merely two extremes of one continuum. Not only does this interpretation challenge the traditional categorisation of /l/-darkening but, given the clear sensitivity to morphosyntactic boundaries that /l/-darkening displays, it also raises questions for a modular architecture of the grammar if phonetics can be morphologically conditioned.

This dissertation is an empirical analysis of /l/-darkening, presenting data from nine varieties of English. Given the difficulty in measuring liquid consonants reliably, ultrasound tongue imaging is used to provide a thorough account of the prime articulatory correlations of darkening processes. The present study provides hitherto absent instrumental evidence confirming the varying degrees of morphosyntactic sensitivity across different dialects. I demonstrate that, rather than being contradictory or chaotic, variation to morphosyntactic boundaries cross-dialectally makes complete sense under an analysis that pays due consideration to the diachronic evolution of phonological processes. Moreover, my data show that the majority of speakers display both categorical allophony of light and dark variants, and gradient phonetic effects coexisting in the same grammar. Therefore, an adequate account of English /l/-darkening presupposes both a theory of the morphosyntax-phonology interface, and the phonetics-phonology interface.

I interpret these results by assuming the modular architecture of the life cycle of phonological processes, whereby a phonological rule starts its life as a phonetically driven gradient process, over time stabilising into a phonological process at the phrase level, and advancing through the grammar. Not only does the life cycle make predictions about application at different levels of the grammar, it also predicts that stabilised phonological rules do not replace the phonetic processes from which they emerged, but typically coexist with them. Moreover, the obvious intimate link between /l/-darkening and /l/-vocalisation can be explained in terms of the life cycle, in the way of lenition trajectories. The results here show that, as predicted, the more recent stage of the lenition trajectory is harsher in terms of its phonetic effect, as well as less advanced in the grammar, applying at a lower level than darkening when the two co-occur in the same variety.

I conclude by arguing that the proposed analysis demonstrates that a full understanding of /l/-darkening in English requires an approach that considers variation under phonetic, phonological and morphosyntactic terms. The wide range of dialectal diversity, for which this thesis provides only a small subset, shows a great deal of orderliness when paying due consideration to the diachronic evolution of variable phonological processes.

Declaration

I declare that no portion of this work referred to in this thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

Copyright

ACKNOWLEDGEMENTS

I could not have produced this work without the help and guidance of the best supervisory team ever: Ricardo Bermúdez-Otero, Maciej Baranowski and Yuni Kim. Ricardo has spent literally thousands of man hours (often in the early hours of the morning) commenting on my papers and conference presentations over the last five years. Ricardo has been instrumental, not only in helping my development as a researcher during my PhD, but throughout my undergraduate and masters degrees, in helping me secure funding, and shaping me into a good researcher.

I am forever grateful to my co-supervisors Maciej Baranowski and Yuni Kim, who have provided so much support over the years. Maciej's Manchester project was one of the reasons I first became involved in linguistic research, and I'm very much looking forward to all of the exciting research we have planned for the future. I'm also grateful for the thorough and insightful commentary he has provided on this thesis at the very last minute. Yuni Kim has also been an incredible source of advice not only on the thesis, but on academic life in general. Everyone doing a PhD should have a Yuni, in order to stay focused and sane.

I'd like to thank all of my lovely informants who offered their tongues for the good of linguistics. In particular, I'd like to thank anonymous informant JWRB, the first ever test subject on the University of Manchester's ultrasound machine, who has provided numerous samples of 1920s conservative RP speech for this project. A special mention must go to the AHRC, as they have funded me through both Masters and PhD study, and this research wouldn't have been possible without their financial support.

I certainly would have been bored without fellow PhD friends, but I actually don't think I could have survived the last four months or so without Laura Arman, Mary Begley, James W.R. Brookes and James Murphy. We've eaten like kings each and every day in Ellen Wilkinson, and laughed and laughed. It's been so much fun! Lots of love to Fernanda McDougall, Farah Nazir, Michaela Hejna and all of the other PhD students at Manchester who are too numerous to thank individually.

I mustn't forget my fellow PhD-ers who have since left, Jon Morris, Michael Ramsammy and Patrycja Strycharczuk. If I didn't have such high achievers as academic big siblings, I certainly wouldn't have submitted to so many conferences, or attempted such ambitious methodologies such as ultrasound, so they've really made me the researcher I am today. Michael Ramsammy deserves a notable mention for his last minute proof-reading efforts, and stellar advice on articulatory matters of the thesis. Patrycja is responsible overall for improving the quality of my research, encouraging me to learn R and LaTeX early on. More recently, we've struggled through ultrasound methodological and statistical issues together, and I'm really indebted to her for all her advice over the years.

I'd like to thank my academic advisor Erik Schleef, who has employed me on many research projects over the course of postgraduate study. I think I might have starved without him. Erik's facilities this year have been particularly invaluable, as working on his project meant I had my own permanent desk, and I really can't imagine how I'd have finished in time without it. This is in addition to him serving as my panel member, and keeping meetings to a strict 30 minutes, which was much appreciated. He's also kept me caffeinated during our collaborative work on *like*, which has been a lot of fun.

Thanks to Maciej's connections, I was lucky enough to spend a semester of my PhD at the University of Pennsylvania, where I was temporarily supervised by the great Bill Labov. One wouldn't expect someone as eminent as Bill to give up his time so happily to look over my data and papers, and keep in touch to provide references for jobs, but he has been so helpful. It was very impressive to see Bill glance at some of my Manchester data and instantly comment on several features that I'd yet to notice. And then emerge from the back room of the Linguistics Lab with several old reels of interviews he conducted in Manchester when he was there for a bit in the 70s… I also need to thank numerous other people at Penn, such as Gillian Sankoff for making me feel immediately welcome on my arrival in Philadelphia, and Gene Buckley for the weekly phonology discussions.

The PhD community at Penn are the loveliest people. I especially have to thank Joe Fruehwald for several acts of kindness: picking me up at the airport, finding me a house to live in, inviting me over to the Fruehwalds' for Thanksgiving and teaching me how to use R and KnitR. Meredith Tamminga has also been fantastic, hosting me on my arrival and spending time teaching me FAVE. Hilary Prichard and Constantine Lignos have provided help on all sorts of issues. There are too many of this lot to mention individually, but all of them made my time at Penn very special and it's so nice to have friends all over the world that I get to see every now and again at conferences. In addition to academic colleagues, I made two friends for life during my stay at 5015 Catharine Street, West Philadelphia. Elika Bergelson and Rose Levine were the best American housemates one could hope for. We had such a fun time and they truly made my experience in Philadelphia the best it could be.

A special mention must go to someone who was originally a Penn friend, but who is now a very firm fixture in my Manchester life. Laurel MacKenzie has been an amazing addition to the Manchester community, and has become a great friend. She has supported me so much over the last two years, and I hope I can repay her in some way one day. I trust that we will continue on our collaborative projects together over the years.

There are too many fantastic staff at Manchester to mention, but a special shout out must go to George Walkden. George has given me so much great advice from the perspective of a newly employed researcher. I actually think his advice played a large role in landing me a job in the end. That brings me to Newcastle University, who I have to thank for giving me a job before I finished this thing.

I also need to mention the people who I have met along the way who have given up their time to facilitate my research. Jim Scobbie and Alan Wrench have taught me everything I know about ultrasound tongue imaging (so if there are any problems in my analysis it must be their fault). Jeff Mielke has given up his time to show me how to run numerous tests in R, and provided me with scripts to do so. Bruce Hayes has provided me with in-depth feedback on my phonological re-analysis of his /l/-darkening data. These people, alongside many more, demonstrate how generous and helpful the linguistics academic community are and I am grateful to everyone who has ever given me advice. I also have to thank the many conferences and seminar audiences over the years, in particular the Manchester Phonology Meeting.

I've been in Manchester for the most part of 29 years now, and will be so sad to leave. My family are the best that anyone could ever wish for. If it wasn't for my Mum, Gill, I never would have gone back to university in the first place, and would still be working for The Man in a bank. I'd like to, er, I'd like to, er, thank my Dad, Mike, who has also been an inspiration. My sister Holly is the best, as are my brothers Ryan and Tom. I love you all so much, thanks for supporting me over the years. And Lily, you haven't contributed much to this thesis really, but thanks anyway. I could have done with getting you on the ultrasound whilst you were acquiring language, but never mind.

Last but most importantly, I'd like to thank my boyfriend Connor, who has done everything for me over the last few months, despite having an extremely demanding job of his own. Connor has been a continuous source of support over the last four years. This past 12 months has been the best and I'm so excited to start our new life in Newcastle together.

Chapter 1

Introduction

The process of /l/-darkening in English shows a remarkable amount of cross-dialectal variation depending on its target position in the word or phrase. Thus far, the complex nature of patterns found in English /l/ have resulted in analyses which struggle to account for its variability, its morphosyntactic sensitivity, and the evidence for both categorical and gradient processes coexisting in the same grammar. This thesis aims to address the unanswered questions of the nature of /l/-darkening by uniting the fields of phonology, phonetics and variation, thus conducting an analysis which considers the interests of all three.

This thesis is an investigation into the processes of lenition which affect the lateral consonant /l/ in English. The term lenition is used to refer, for the most part, to the process of /l/-darkening. /l/-darkening shows sensitivity to morphosyntactic boundaries, displays categorical allophonic variation, but also gradient phonetic effects. Articulatorily, a dark [ɫ] is more lenited and less consonantal than a light [l], as it exhibits a lowered tongue body, retracted tongue root and reduced tongue tip gesture. General accounts of /l/-darkening state that light [l] is found in onsets and dark [ɫ] in codas (Halle and Mohanan 1985; Giegerich 1992; Roach et al. 2006). In addition, the process of /l/-vocalisation will be analysed. Vocalisation of /l/ is seen as the next stage of lenition on the trajectory whereby /l/ loses its tongue tip contact altogether, and there is no central contact between the tongue and the alveolar ridge. The two processes are often treated separately in the existing literature, but the present investigation seeks to interpret the clear link between the different stages of lenition.

Such processes have been investigated within a wide range of linguistic subdisciplines, including phonology, phonetics and sociolinguistics. However, the analyses given by these more narrow approaches often overlook the evidence from outside the particular subfield. This thesis aims to provide an account of /l/-darkening and vocalisation by uniting these fields under one analysis.

The complex structure of liquid consonants has led many researchers to advise that an empirical phonetic investigation, preferably articulatory, is required in order to conduct a reliable account of variation in English /l/-darkening (Gick et al. 2006; Lawson et al. 2010). Thus, this thesis uses ultrasound tongue imaging as the empirical basis of the investigation. Ultrasound tongue imaging is still relatively novel in linguistic study, and this thesis provides a unique perspective of /l/-darkening processes across many dialects of English. Ultrasound tongue imaging allows us to monitor the entire tongue, giving insight into tongue shape as a whole, as well as the activity in the tongue root area.

1.1 Goals of the thesis

The overarching goal of the thesis is to answer the theoretical questions in Section 1.1.1 by empirical phonetic investigation.

1.1.1 Theoretical issues

This thesis aims to utilise innovative empirical phonetic methodologies to address long-standing questions in phonology. In particular, the thesis will focus on two aspects of /l/-darkening discussed in the existing literature: its sensitivity to morphosyntactic factors, and its gradient or categorical nature. In this section these two primary research questions will be addressed in turn. Following this, several peripheral aims and investigations of the thesis will be addressed.

The first research question of the thesis concerns the interactions of phonology and morphosyntax:

1. Traditional descriptions of /l/-darkening posit a simple onset-coda distinction between light [l] and dark [ɫ]. In this view, however, some dialects show opaque overapplication when a word-final prevocalic /l/ is resyllabified into the following word, giving onset dark [ɫ] in a phrase like ***heal** it*: [hiː.ɫɪt], or preceding a stem-suffix boundary in a word like ***heal**-ing*: [hiː.ɫɪŋ]. Other dialects, on the other hand, may show transparent lack of application here. Furthermore, others may show dark [ɫ]s in additional morphosyntactically conditioned environments.

    i. Does the dialectal typology of /l/-darkening reveal implicational relationships whereby overapplication in one morphosyntactically defined environment entails overapplication in another?

    ii. If so, which theory of the morphosyntax-phonology interface best accounts for these implicational relationships, thus providing the best fit for the range of morphosyntactic conditioning effects attested synchronically?

    iii. And to what extent can this dialectal typology be illuminated by consideration of patterns of diachronic change?

The second concerns phonetics-phonology interactions:

2. /l/-darkening has traditionally been treated as a categorical phonological process resulting in two separate allophones. However, phonetic research shows that there is gradience in the signal. Indeed, some studies have argued for a purely gradient interpretation of the facts, asserting that there is no categorical difference between light and dark allophones.

   i. To what extent do we find evidence of categorical effects of English /l/-darkening, and can they fully account for the variation found?

   ii. To what extent do we find evidence of gradient effects of English /l/-darkening, and can they fully account for the variation found?

   iii. Do we find compelling evidence for both? If so, which approach to the architecture of the grammar best explains the interaction of categorical and gradient effects.

1.1.2 Empirical issues

Given the problems in accounting for the realisations of /l/ auditorily and acoustically, the best way to conduct an empirical investigation in this area is by articulatory means. Electropalatography is probably the method most frequently used for studying /l/-vocalisation (see Chapter 2), as this monitors contact between the tongue and hard palate. However, electropalatography provides no information about overall tongue shape. Considering the primary process of interest in this thesis is /l/-darkening, ultrasound tongue imaging is a better option.

Therefore, the empirical questions which this thesis seeks to address are:

1. Drawing on research question i) above, to what extent can we find articulatory corroboration of morphosyntactically conditioned sensitivity in English /l/-darkening and vocalisation?

2. For research question ii), does the ultrasound tongue imaging data support a categorical phonological interpretation of the /l/-darkening and vocalisation, or a phonetically gradient one? Can we argue for evidence of both?

1.1.3 Peripheral goals

Although the main research goals of the thesis focus on questions of morphosyntax-phonology interactions and categorical vs. gradient processes, there are several additional outcomes which will come out of this analysis and form novel contributions to the field.

The first of these involves variationist and sociolinguistic studies. Few studies in sociolinguistics have investigated /l/-darkening, and it is fair to say that, for many, it is not seen as a sociolinguistic variable of interest. The allophonic distinction is claimed not to exist for many dialects of English, such as those in Northern England or North America (see Chapter 2, Section 2.3), and this thesis seeks to determine to what extent this is true. More sociolinguistic studies have shown an interest in vocalisation, however. This is unsurprising, given this process involves more drastic lenition and is therefore a) more salient phonetically, b) a more recent innovation diachronically and c) subject to sociolinguistic conditioning.

Nonetheless, it is crucial to understand how /l/-darkening and /l/-vocalisation are intimately connected. Thus, a rigorous articulatory description of /l/-allophony across a range of English dialects will make it possible to put sociolinguistic observations about vocalisation in their proper typological and diachronic context. The result should be a fuller and more complete understanding of the (l) variable in general.

Indeed, it is not only vocalisation, but also darkening, that should attract the attention of variationists. One issue, for example, concerns the secondary articulation of the /l/. In the following chapter, this will be discussed in reference to the term *velarisation*, and how this term is not synonymous with darkening, as many acoustic or descriptive accounts would suggest. From the point of sociolinguistic variation, we may see speakers of different social groups and of different dialects displaying variation in the secondary articulation of /l/, providing new insights into socio-articulatory interactions of the variable.

Another issue which is of interest for variationists concerns the range of environments where darkening overapplies. If these turn out to be arranged along a diachronic pathway, this immediately creates the expectation that patterns of synchronic variation within speech communities will be organised along similar lines. To date, there has been no serious sociolinguistic investigation of these issues.

1.1.4 Questions which will not be addressed

A temporal account of /l/-darkening was researched early on in the experimental procedure, but it was identified as being non-crucial for the majority of dialects in the thesis. Given that the ultrasound unit used in this experiment (see Chapter 5 for details) provides only a very basic idea of the relative phasing of gestures, it was felt that an in depth temporal analysis would not be addressed here, but instead left for future work, as outlined in the conclusion.

Although the varieties in question are referred to by their geographical location, the work here makes no claim to represent that speech variety as a whole. For the most part, each analysis is based on one speaker, and there could very well be idiolectal variation which does not reflect the entire speech community of that area. Nevertheless, it is often possible to make suggestive connections between the pattern shown by a speaker and the reports for his or her dialect area in the descriptive and sociolinguistic literature. Moreover, the overall range of variation across all the experimental subjects examined in this thesis provides useful information about the range of variation in English as a whole, regardless of the extent to which this variation is itself geographically focalised or reproduced internally within each speech community.

1.1.5 Outcome of the present investigation

The thesis presents results from two separate experiments eliciting /l/ in various environments. Experiment 1 addresses the first research question, monitoring darkness at different morphosyntactically defined environments, comparing results across dialects. Here we see evidence of varieties with transparently conditioned darkening processes alongside more complicated patterns of where darkening overapplies at word and stem-suffix boundaries, and darkening feeding vocalisation.

Experiment 2 employs a much larger dataset to investigate evidence of categorical and gradient processes in English /l/ variation. Again, we see the patterns vary from speaker to speaker. Some have an obvious distinction between discrete allophones, which is shown through the articulatory and acoustic data. Others have what looks like a more gradient distinction, whilst some have no difference whatsoever.

The outcome of the present investigation shows, by articulatory means, that /l/-darkening is indeed morphosyntactically conditioned in variable categorical terms, but also has gradient phonetic effects overlaid on top of this categoricity. By assuming the model of the life cycle of phonological processes, as proposed by Bermúdez-Otero (1999) and Bermúdez-Otero and Trousdale (2012), we shall see that all aspects of complex patterning in English /l/-darkening can be accounted for.

1.1.6 Organisation of the thesis

The thesis is presented as follows. Chapter 2 provides an overview of the existing work on /l/-darkening and vocalisation. I first consider phonological studies which generally adopt a categorical account of /l/-darkening. Phonetic study makes up the largest component of Chapter 2 and acoustic and articulatory work over the years is presented in turn, alongside an overview of possible articulatory techniques used to study /l/. This section also considers other phonetic aspects to the data, such as duration effects, and secondary articulations. The third subfield focussed on in Chapter 2 is the sociolinguistic and dialectological analyses. Most of these are collated from descriptive work, alongside the occasional quantitative study into /l/-darkening or vocalisation. These descriptive studies provide an informative overview into the possibilities of variable patterns found in the processes, including the claimed lack of distinction in Northern Englishes, alongside style-shifting patterns and other sociolinguistically conditioned variation. The debate in this chapter also considers the two theoretical issues of the thesis and how existing interpretations of the data have failed to account for the patterns. These issues are dealt with in detail in Chapters 3 and 4.

Chapter 3 addresses the issue of morphosyntactically conditioned variation in English /l/-darkening, and how this varies from dialect to dialect. We shall see that different dialects vary in their morphosyntactic sensitivity: e.g. /l/ darkens before a stem-suffix boundary in a word such as *heal-ing* in some dialects, but not in others. In Chapter 3, I argue that we need a theory which can account for such variable patterns whilst not over-predicting impossible dialects. I suggest that the best results are provided by a diachronic approach based on the life cycle of phonological processes and couched in a stratal theory of the grammar. The life cycle accounts for varying morphosyntactic sensitivity by explaining the synchronic patterns through a process's diachronic trajectory. In addition to this, the life cycle also makes sense of the clear link between darkening and vocalisation in terms of lenition trajectories. The life cycle predicts that in the synchronic grammar, the older rules, affecting milder types of lenition, have narrower cyclic domains (i.e. darkening) than the younger rules, affecting more drastic types of lenition (i.e. vocalisation). On top of this, the life cycle can also account for the fact that we find evidence of categorical and gradient effects of the same process operating within the same grammar. This is fully discussed in Chapter 4.

Chapter 4 focusses on the debates in the existing literature regarding categoricity and gradience. As Chapter 3 will unveil, phonological analyses of /l/-darkening tend towards a purely categorical interpretation of the data, stating that light and dark /l/ are two discrete allophones. Many phonetic analyses dispute this, as they find gradient effects in the data which a purely categorical approach cannot account for. This chapter begins by addressing the debate in the wider fields of phonetics and phonological, before concentrating on studies of /l/. The section on /l/ in particular takes some of the most famous studies of darkening, discussing their positive aspects alongside their deficiencies. I argue that the life cycle of phonological processes is the only theory which can efficiently account for the presence of both effects, alongside variability and morphosyntactic sensitivity.

Chapter 5 outlines the methodology used in the empirical part of the thesis. This chapter begins by recapping the goals of the thesis alongside the experimental methods used in order to address the research questions. The details of ultrasound tongue imaging and the experimental procedure are given here, as well as information on the participants and dialects surveyed. Phonetic aspects of the investigation are presented, such as acoustic and articulatory measures of darkness. This section also explains the quantitative methods used in the thesis, which support the investigations of morphosyntactic sensitivity and categoricity and gradience.

Chapter 6 presents the results of Experiment 1, the investigation into the morphosyntactic conditioning of /l/ realisation. Analysis of the ultrasound data reveals a variety of patterns from both British and American dialects, differing in various respects. The first concerns phonetic differences, showing that some dialects have a much larger articulatory difference between initial and final /l/s than others. In addition, dialects that are said to display no variation at all in their /l/ realisations show a small but significant effect of darkening between initial and final position. The second concerns morphosyntactic sensitivity. We find evidence of a three-way distinction between light, dark and vocalised /l/ for one variety, and a similar distinction to a less convincingly categorical extent in another. These results are supported by the quantitative method of Smoothing Spline ANOVA, which shows the significant difference between two tongue shapes. The patterns also fall in line with the predictions made by the life cycle in terms of lenition trajectories: as a more recent stage of lenition, vocalisation applies at a lower level of the grammar than darkening.

Chapter 7 presents the results of Experiment 2, the larger investigation which seeks to address the issue of categoricity and gradience in /l/-darkening processes. We shall see that there are dialects which provide clear evidence for two allophonic categories in some dialects, but less convincing distributions for others. The articulatory distributions of eight speakers are presented here, showing tongue contours for all dialects studied. These contours are subjected to a principal components analysis, a novel way of quantifying ultrasound tongue splines. The results for this analysis form the basis of several statistical tests used in the is chapter, including tests for bimodality, significant differences between phonological contexts, and linear regression. The results of the acoustic analyses are also presented here, which are subjected to a mixed effects logistic regression. The results from all of these tests are summarised to provide pictures of several different grammars: those with a clear categorical distinction, those with a difference which may or may not be categorical, and those which show no variability at all.

The investigations also shed light on pharyngealisation vs. velarisation in secondary articulations of /l/-darkening, as well as the tongue contours for vocalised /l/s, which may or not not be largely pharyngealised, depending on dialect. We see that some vocalising dialects have evidence of a three-way distinction, whilst others do not. Nevertheless, all patterns found in the two chapters can be accounted for under the predictions of the life cycle of phonological processes.

Chapter 8 concludes and comments on the suitability of the conclusions given the statistical tests. I suggest some areas of future research for phonologists, phoneticians and sociolinguists.

Chapter 2

Previous studies of English /l/

As discussed in the introduction, /l/ in English shows variation depending on its position in the word or phrase. Within this, we find variation in the extent of lightness in different positions, the articulatory magnitude of difference between different contexts, as well as the acoustic and articulatory cues implemented to convey these differences. We also see variation cross-dialectally, with speakers of different varieties of English showing distinct patterns in terms of their phonetic articulation and phonological distribution. This chapter aims to give an overview of all of the complexities found in the realisation of English /l/. Firstly, let us define the basic aspects of the process.

/l/-darkening is the process whereby /l/ is realised with a reduced tongue tip gesture articulatorily, or a delayed tongue tip gesture temporally. Traditional descriptions of English /l/-darkening posit two discrete allophonic categories: light [l] and dark [ɫ] (Chomsky and Halle 1968; Halle and Mohanan 1985; Giegerich 1992; Roach et al. 2006). It is noted that light [l] occurs in canonical onsets (*like, love*) and dark [ɫ] in canonical codas (*pool, dull*). Darkening is reported widely for many dialects of English, although the phonetic distinction for onset and coda /l/ is much smaller in some varieties than others. For some dialects, it is claimed there is no allophonic distinction whatsoever.

Some dialects exhibit /l/-vocalisation alongside or in place of /l/-darkening, which represents a harsher form of lenition. /l/-vocalisation is characterised by the lack of contact between the active tongue and the passive alveolar ridge, alongside enhancement of the dorsal gesture. Vocalisation is typically described as occurring word-finally or pre-consonantally (Wells 1982:259). It is reported widely for dialects of Southern British English (Johnson and Britain 2007; Hardcastle and Barry 1989; Przedlacka 2001; Tollfree 1999; Wells 1982), as well as American English (Ash 1982; Hall-Lew and Fix 2012; Pederson 2001) and Southern Hemisphere Englishes (Borowsky 2001; Borowsky and Horvath 1997; Horvath and Horvath 2002). In Southern dialects of British English, /l/-vocalisation is often described as being accompanied by strong labialisation, so that /l/ sounds more like [ʊ] or [w]. For the purposes of this dissertation, vocalised /l/ is transcribed with the centralised ramshorn vowel [ɤ̈], to remain impartial about rounding or labialisation.

The difference between a dark and vocalised /l/ can not always be sharply defined.

Giles and Moll (1975) found that what is referred to as dark [ɫ] in American English can often be vocalised, by articulatory standards, (i.e. there is no tongue-tip contact) especially in faster speaking rates, and when /l/ followed a low vowel. This touches on the debates of categoricity vs. gradience. /l/-vocalisation is treated as a categorical phenomenon by most sociolinguistic and phonological analyses, and as either categorical or gradient by phonetic ones. There is evidence for both, as Section 2.2 may demonstrate, although this discussion will be saved for Chapter 4.

Vocalised /l/ occurred historically and categorically in some sets of words during the 15th century, and this pronunciation is standardised in English today e.g. *half, talk, calm* (Cruttenden 2008:218). It is probably inaccurate to describe these sounds in the same way as other kinds of vocalisation in present-day British English, as there really is no /l/-like sound there at all. The consonant was completely deleted and the vowel underwent compensatory lengthening. This is also the case for many speakers of English in rapid speech e.g. *shall we go?* as [ʃə wi ˈɡəʊ], or in highly frequent words and compounds such as *always, all right* (Jones 1966:93).

This chapter aims to give an overview of studies of /l/-darkening and vocalisation across several subfields of linguistics. Section 2.1 considers analyses in phonology, including proposals of phonological rules for the processes, and also sensitivity to morphosyntactic conditioning (Section 2.1.1). The interaction of phonology and morphosyntax forms the basis for the first major research question of the thesis. These ideas are alluded to in Section 2.1.1, but are discussed in full in Chapter 3. Section 2.2 gives an overview of some of the phonetic studies conducted to investigate /l/-darkening and vocalisation, from acoustic to articulatory variation. We shall see that some of the findings conflict with a traditional phonological approach to the phenomena. These effects will be discussed briefly here, but the main debate reserved for Chapter 4, which address the evidence for gradient and categorical processes of darkening. Finally, Section 2.3 collates the evidence from descriptive claims, dialectological findings and sociolinguistic studies to summarise the kind of variation we find cross-dialectally.

2.1 Phonological analyses

As stated in the introduction, the simple description of /l/-darknening processes in English is that we find light [l] in onsets, and dark [ɫ] in codas. Wells (1982:258) posits the rule in (1) for dark [ɫ], and the rule in (2) for vocalised [ɤ̈].

(1)

$$/\text{l}/ \longrightarrow [\text{ɫ}] \;/\; \left\{ \begin{array}{c} || \\ \#_o C \end{array} \right\}$$

(2)

$$/\text{l}/ \longrightarrow [\ddot{\text{ɤ}}] \,/ \left\{ \begin{array}{c} || \\ \#_o C \end{array} \right\}^1$$

These rules state that an underlying light /l/ becomes dark or vocalised when in word-final position, or before a consonant. Phonological analyses generally class the laterals as having both consonantal and vocalic features (Jakobson et al. 1952; Johnson and Britain 2007:19)

Johnson and Britain (2007:298) analyse vocalisation as 'the emergence of the unmarked', noting that children acquiring English tend strongly to vocalise /l/ even if that is not a feature of their dialect. They also cite cross-linguistic evidence and the rapid sociolinguistic spread of vocalisation to support their case. They provide an Optimality Theoretic analysis arguing that vocalisation is a natural phenomenon and should be expected in any language or dialect that has a clear dark dichotomy.

Tollfree (1999) discusses 'discrete vs. continuous' (i.e. categorical vs. gradient) interpretations of /l/-darkening and vocalisation for London English. She uses Government Phonology (GP) and Articulatory Phonology (AP) to represent each interpretation respectively. She explains lenition in /l/ through element loss in GP, but highlights that the theory's explanation of vocalisation would also falsely predict a vocalised [ɤ̈] in intervocalic words such as *fellow.* She also makes reference to the gradient properties of /l/ lenition processes and how a framework such as GP cannot account for such continuity in the signal. For this reason, she asserts that Articulatory Phonology is a better way of accounting for the variation. In Articulatory Phonology (Browman and Goldstein 1986; 1989; 1992) sounds are accounted for in terms of gestures and their continuous affiliations. An AP account of /l/-darkening and vocalisation has phonetic motivation from some of the studies discussed in the next section, and as this draws on the debates of categoricity and gradience, this aspect of the analysis is reserved for the discussion in Chapter 4. For the time being, Tollfree (1999) concludes that an interpretation using GP is not sufficient to account for the variation in English /l/, and that an AP account is superior. Neither framework, however, can account for the fact that /l/ seems to show varying morphosyntactic sensitivity between dialects, which is discussed in the following section.

2.1.1 Morphosyntactically conditioned phonological processes

One thing which has been viewed as a potential source of conflict in /l/-darkening studies is the variable results found with regard to morphosyntactically complex environments. For example, what if the /l/ is word final, but is resyllabified into the onset

[1]Note that Wells uses the [o] phone to transcribe vocalised /l/, whereas the present study uses its centralised unrounded equivalent [ɤ̈], as [o] may be too characteristic of Cockney.

when followed by a vowel in the next word? The phonological literature would suggest that if an /l/ is resyllabified, it is onset-like and should be light (Cruttenden 2008; Ladefoged 2001), however, existing studies suggest this may vary.

Halle and Mohanan (1985:65) account for the allophonic variation found in /l/ through Lexical Phonology, using the original strata of Mohanan (1982:48), as in Example (3). They postulate the rule that /l/ gains the feature [+back] when in the rime. They account for word-final prevocalic /l/ remaining light when resyllabified by claiming that it occurs before darkening, and then darkening applies post lexically. They claim that darkening can occur in compounds (e.g. *the seal office*) but not in other word-final prevocalic contexts (e.g. *the seal offered a doughnut*). Turning to their derivation in Example (3), they state that /l/-darkening occurs at stratum 4, to account for a light /l/ preceding a stem-suffix boundary, e.g. *wheel-ing*. As we shall see, although some dialects certainly do follow the pattern described by Halle and Mohanan, there is plenty of evidence for light [l]s occurring across word boundaries in other dialects.

(3)

Stratum 1:	Class I derivation, irregular inflection
Stratum 2:	Class II derivation
Stratum 3:	Compounding
Stratum 4:	Regular inflection

Bermúdez-Otero (2007a; 2011) summarises the findings of previous studies, accounting for the lack of concurrence by claiming that different varieties of English may show different types of morphosyntactic conditioning, as demonstrated in Table 2.1. His analysis of /l/-darkening concerns the interaction of both synchronic and diachronic processes. These analyses follow the predictions of the life cycle of phonological processes, and will be discussed fully in Chapter 3.

	***l**ight*	*he**l**ium*	*hea**l**-ing*	*hea**l** it*	*hea**l***	
RP	[l]	[l]	[l]	[l]	[ɫ]	Cruttenden (2008); Jones (1966)
Am. Eng. 1	[l]	[l]	[l]	[ɫ]	[ɫ]	Sproat and Fujimura (1993); Gick (2003)
Am. Eng. 2	[l]	[l]	[ɫ]	[ɫ]	[ɫ]	Olive et al. (1993)
Am. Eng. 3	[l]	[ɫ]	[ɫ]	[ɫ]	[ɫ]	Hayes (2000); Yuan and Liberman (2011)

Table 2.1: /l/-darkening in different morphosyntactic environments. Adapted from Bermúdez-Otero (2007a)

One of the studies mentioned by Bermúdez-Otero is Hayes's (2000) Optimality Theoretic approach to /l/-darkening in English, followed up with Stochastic Optimality Theory in Boersma and Hayes (2001). Boersma and Hayes (2001); Hayes (2000) present data demonstrating that /l/-darkening is morphosyntactically conditioned, yielding higher frequencies of dark [ɫ] in complex words such as *heal-ing* than in monomorphemic words such as *Hayley* (the contextual equivalent of *helix* in the Table 2.1). Such alternations raise challenges for the study of variation as, in addition to a model of variable processes, they require a theory of the morphosyntax-phonology interface. For this purpose, Hayes adopts Output-Output Correspondence (OOC), a theory incorporating constraints that demand identity between morphologically related surface forms (Benua 1995; 1997; Kenstowicz 1996). However, OOC has crucial drawbacks, which become apparent in his own approach to /l/-darkening. Bermúdez-Otero (2011:2043) shows that the OOC constraints used by Hayes can generate an impossible dialect in which darkening overapplies at stem-suffix boundaries (e.g. *hea*[ɫ]*-ing*) but not at word boundaries (e.g. *hea*[l] *it*). These problems, discussed further in Chapter 3, warrant the exploration of alternative approaches to morphosyntactically conditioned phonological variation.

2.1.2 A note on syllabification

It may be appropriate at this point to provide a stronger defence of the onset maximal syllabification assumptions touched on so far. In this thesis, syllabification is onset-maximal at all levels. Evidence from the Middle English period suggests that English developed phrase-level syllabification into onsets (Minkova 2003), meaning that word-final consonants, such as the /l/ in *heal it*, are in the coda at the word level, but at the phrase level when followed by a vowel-initial word, they are resyllabified into the onset. Strategies, such as ambisyllabicity (Kahn 1976), or coda maximisation (Wells 1990; 2008) fail to account for the patterns of variation found, not only in /l/-darkening, but in numerous phonological processes, in some cases creating paradoxical predictions.

Ambisyllabicity was originally developed by Kahn (1976) to explain the two environments in which American /t/-flapping occurs. Kahn noted that word-internal foot-medial intervocalic /t/ (e.g. *better*) and word-final prevocalic /t/ (e.g. *hit it*) showed the same allophony in American English as both were realised as the voiced flapped variant [ɾ] in these positions. The syllabification claim is that, on the surface, these two /t/s occupy the same position in syllable structure (the ambisyllabic position), hence flapping in both cases. Word-final prevocalic consonants are always ambisyllabic. Giegerich (1992:284) explains /l/-darkening in terms of ambisyllabicity. He claims that the /l/ in phrases such as *feel it* is light, and the /l/ would be ambisyllabic by the phrase level rule of onset capture. This is the same, he argues, as a word such as *feeling* which would be expected to have the same realisation as the /l/ is also ambisyllabic, this time by the word level rule of coda capture. Despite the initially perceived elegance for such an idea, the theory cannot account for the varying patterns of /l/-darkening found in different dialects on English, such as those displayed in Table 2.1. As will be discussed in more detail in the following chapter, ambisyllabicity cannot explain the articulatory data collected by Sproat and Fujimura (1993), whereby the /l/ is light in *Beelik* and dark in *Beel equates*. For the speakers of American English studied by Sproat and Fujimura, ambisyllabicity may work to explain one process, that of /t/-flapping, but encounters problems for others, such as /l/-darkening. Bermúdez-Otero (2007b) presents this /l/-darkening paradox, amongst others, to argue against ambisyllabicity in English and for a stratal approach (see Section 3.2.1 or Bermúdez-Otero 2007b for further details). Under his approach, /t/-flapping is foot-based, as also argued for by Kiparsky (1979) and Harris (2003).

Although it could be argued that a stratal approach with ambisyllabicity could account for such differences, Bermúdez-Otero presents another paradox which is fatal to ambisyllabicity in any framework. In what he refers to as the *látèx paradox*, Bermúdez-Otero (2007b:7), draws on the process of /t/-flapping, but this time its interaction with the process of pre-fortis clipping, whereby vowels are subject to clipping (i.e. shortening), when they are followed by a fortis consonant within the same syllable (Wells 1990). This accounts for the difference in vowel length in pairs such as *leaf* [liˑf] vs. *leave* [liːv], whereby the former vowel is slightly shorter, or clipped, preceding the voiceless consonant. Under an ambisyllabic approach, the /t/ in *latex* is not ambisyllabic, as it is foot-initial, and hence is not flapped. However, the initial vowel in the word *is* clipped, giving [leɪˑtɛks]. Bermúdez-Otero (2007b) points out that ambisyllabicity cannot accommodate this pronunciation: if the /t/ is not ambisyllabified, it cannot cause the vowel to be clipped, but if it is ambisyllabified, then the /t/ must flap. As Bermúdez-Otero (2014) more recently demonstrates, this interaction of processes and their environments is identical to that of flapping and Canadian Raising (Chambers 1973) in many dialects of North American English.

Similarly, a theory which maximises intervocalic consonants as coda-syllabified, such as Wells (1990; 2008) comes across serious issues when accounting for many observed patterns in English. The *látèx paradox* raises the same problems for coda maximisation as it did for ambisyllabicity. Wells's (2008) syllabification of *latex* keeps the /t/ in the first syllable, in line with his Main Syllabification Principle (consonants are syllabified with the more strongly stressed of the two neighbouring vowels; Wells 1990), giving [leɪt.ɛks]. This correctly predicts clipping of the diphthong. However, Wells states that a /t/ flaps in syllable final position, which incorrectly predicts flapping in this environment. Note that the transcription in the *Longman Pronunciation Dictionary* is correct, but the posited rules do not work in this case. Furthermore, Gilbert (2013) provides instrumental data which falsify Wells's syllabification predictions in the pairs in (4) and (5).

(4) *nitrate* [naɪtr.eɪt] *night rate* [naɪt.reɪt]

(5) *sucrate* [suːk.reɪt] *souk rate*[2] [suːk.reɪt]

Wells accounts for the differing phonetic realisation in the pair in (4) by assuming the /tr/ cluster is a monosegmental affricate, which is syllabified into the coda. This is in contrast with the pair in (5), where both items would be syllabified in the same way. However, by measuring the lag in voice onset after /t/ and /k, and the F3 minimum of the following /r/ (where F3 lowering indicates the presence of /r/), Gilbert found that /tr/ and

[2] As Gilbert (2013) points out, a *souk* is an Arab market.

/kr/ clusters showed the same properties. The point of voice onset and F3 minima were near-simultaneous in *nitrate* and *sucrate*, but F3 minimum lagged in relation to voice onset in *night rate* and *souk rate*. This suggests an onset-maximising strategy where the clusters are tautosyllablic in the words, but heterosyllabic in the phrases. Such patterns are easily accounted for in a stratal analysis that makes reference to the word and phrase levels.

Thus, this thesis assumes a syllabification theory of onset maximisation at all levels. In combination with a stratal approach, this is able to account for the numerous patterns which occur in empirical data, not only for /l/-darkening, but many other other phonological processes. It must be noted, however, that theories such as those mentioned in this section may not ever be able to account for issues of phrasal prosody. Resyllabification across word boundaries can be blocked by major phrasal prosodic boundaries, as demonstrated by Cho et al. (2014) in their EMA study of CV gestures. Cho et al. found that /C#V/ gestures across word boundaries would only show the same pattern as tautomorphemic /CV#/ gestures if there was no intonational phrase boundary present. As phrasal prosodification is not only conditioned by syntax, but also by speech style and speech rate, we might expect inter and intra-speaker variation with respect to the placement of such boundaries, thus resulting in potentially variable and unpredictable syllabification strategies.

2.2 Phonetic analyses

This section will give an overview of studies into /l/-darkening and vocalisation that investigate its realisation phonetically. Section 2.2.1 focusses on the acoustic studies, which use formant data to analyse /l/ variation. This will help inform the empirical component of the thesis. Section 2.2.2 considers articulatory investigations into /l/ lenition processes, firstly giving an overview of the methodologies and equipment used to conduct such studies. We will also consider other aspects of phonetic analyses such as the temporal properties over the course of the /l/, duration, coarticulation and the difference between velarising and pharyngealising dark [ɫ]s.

2.2.1 Acoustic analyses

In many languages, lateral consonants display a considerable degree of acoustic variation and English is no exception (see Proctor 2009 for a cross-linguistic overview of /l/). In English, the spectrum of acoustic investigation concerns the difference between light and dark /l/, with vocalised /l/ being very difficult to distinguish from dark variants by acoustic means (Hall-Lew 2011). The primary acoustic correlate of darkness is the difference between the first and second formants: dark [ɫ] is characterised by a close proximity between F1 and F2, whereas light [l] has a relatively high F2 and a low F1 (Carter 2002; 2003; Carter and Local 2007; Gick et al. 2006; Hawkins 2004; Ladefoged and Maddieson 1996; Recasens and Espinosa 2005:3). Many phonetic studies of /l/ realisations take a solely acoustic approach to their analyses (Carter 2002; 2003; Carter and Local 2007; Huffman 1997; van Hofwegen 2010; Lehiste 1964; Morris 2013; Nolan 1983; Yuan and Liberman 2009; 2011). Carter (2003) says that these acoustic properties are more easily observed in light [l] than for a dark [ɫ]. Carter (2002) makes a comparison of formant measurements in previous acoustic studies, taking data from RP (Nolan 1983: 90) and American English (Lehiste 1964:14). Table 2.2 shows that RP has a much clearer distinction between the first and second formants than American English. Nevertheless, Lehiste (1964) still reports a clear/dark distinction between phonological contexts in American English, with syllable-final /l/ having a smaller difference between F2 and F1. This result was also found in the British English varieties spoken Manchester and Leeds, and Carter (2002; 2003) describes these dialects as having very dark [ɫ]s everywhere, but still retaining the distinction between initial and final position.

Study	F1	F2	F3
RP: Nolan (1983)	360	1350	3050
AmE: Lehiste (1964)	295	980	2600

Table 2.2: Mean formant frequencies for /l/ from two studies (adapted from Carter 2002)

However, many acoustic studies have come across problems when attempting to measure /l/ reliably. Umeda (1977:846) discusses how, in her acoustic study of consonants, word-final /l/ was often 'totally impossible' to measure reliably. One of the biggest problems with acoustic analysis of liquids is the vast amount of coarticulation which may be found in these consonants. Studies from both British English (Bladon and Al-Bamerni 1976) and American English (Lehiste 1964) have found that the F1 and F2 of adjacent vowels affect the formant frequencies of the /l/. This effect is said to more notable in initial /l/, whereas final /l/ exerts a strong influence on the preceding vowel, whilst staying pretty much independent in itself (Lehiste 1964). Nolan's (1983) acoustic study of liquids focused purely on formant values of onset /l/ before different vowels, as previous investigation had shown this position to be the most susceptible to coarticulation. The results show that, as expected, the F2 of /l/ is highest with a following /i/ and lowest when preceding /ɒ/; the F1 of /l/ is lowest preceding /i, u/ and highest before /æ/.

Acoustic transitions in and out of the /l/ have been found to vary between speakers and dialects. When comparing different dialects of British English, Carter (2002:159) found that the transition in and out of the /l/ for his Manchester speaker was very long (particularly the transition in), with a very short steady state. In contrast, the Sunderland speaker had a much longer portion of steady state, with comparatively short transitions. Ladefoged and Maddieson (1996:361) show that, for their American English speaker with a clear acoustic difference between initial and final position, initial /l/ has a relatively short F2 transition, and in final position the low F2 is achieved before the consonant ends. Ladefoged and Maddieson (1996) use this observation to argue against the Sproat and Fujimura (1993) study discussed later in Section 2.2.2, where claims of secondary articulation are said to show mirroring of gestures.

The claim that a long transition into the /l/ is an inherent property of dark varieties does not seem to hold, however. Carter also reports results from two rhotic cities, to contrast liquids with the non-rhotic Manchester and Sunderland: the dark [ɫ] variety of Fife in Scotland, and the reportedly light [l] variety of Tyrone in Ireland. Although Fife also shows considerably longer transitions into the /l/ than Tyrone, the transitions out are shorter. Moreover, the transitions in and out overall are shorter than the clear /l/ in Sunderland.

Lehiste (1964:14) notes that the F2 of /l/ rises before a following /i/, /u/ or /ʊ/, but that a preceding /u/ and /ʊ/ cause the F2 in a following /l/ to lower. Lehiste (1964:26) did not find that /i/ and /ɪ/ had any effect on a following final /l/. Initial /l/ appears to have a greater influence on F1 and F3 of the following vowel, whereas final /l/ influences F2 of the preceding vowel. In initial position, F2 anticipates the F2 of the following vowel. F1 and F3 are influenced by the preceding vowel. On the other hand, Lehiste (1964:14) says F3 shows no discernible pattern and does not appear to be significant. Fant's (1960) work suggests that the high F3 found in /l/s may actually be an F4 with an absence of F3 indicating an /l/.

Although many studies of /l/ consist entirely of acoustic data, more recent articulatory analysis has suggested that the spectrogram does not show all relevant information. Overt gestures may not be picked up in the signal, as well as tongue shapes which show no acoustic difference, such as velarised vs. pharyngealised. In order to gain a true insight into the variable possibilities of English /l/, acoustics needs to be accompanied by articulatory data.

2.2.2 Articulatory analyses

Possibly due to the difficulties of measuring /l/ reliably in the acoustics, studies of darkening and vocalisation have been the topic of numerous articulatory analyses in both British English (Barry 2000; Hardcastle and Barry 1989; Scobbie and Wrench 2003; Scobbie 2007; Scobbie and Pouplier 2010; Wrench and Scobbie 2003; Wright 1989) and American English (Gick et al. 2006; Giles and Moll 1975; Lee-Kim et al. 2013; Lin et al. 2014; Narayanan et al. 1997; Sproat and Fujimura 1993; Stone 1990). The primary articulatory correlate of the difference between light and dark /l/ is said to be the amount of tongue retraction, with the tongue generally being less retracted for light [l] than dark [ɫ] (Giles and Moll 1975). However, as we shall see, evidence from some studies show that the variation may be more complex than just a retracted tongue body, and that /l/ consists of two lingual constrictions. These accounts of /l/-darkening suggest that investigating the allophonic distinction could involve temporal analysis, comparisons between contexts, as well as an insight to the movements of the tongue body and tongue tip.

Studies of /l/-vocalisation, on the other hand, have a much easier deal in describing what constitutes a vocalised /l/, that is, lack of contact between the tongue and the palate (Scobbie and Pouplier 2010; Wrench and Scobbie 2003; Wright 1989). However, both lenition processes show a remarkable amount of variation depending on various aspects of the speaker or speech including dialect, speech rate, syllable structure and possible coarticulation with neighbouring sounds. This section will give an overview of some of the articulatory methods that have been used to investigate /l/ darkening and vocalisation in English (Section 2.2.2.1), as well as a description of the findings of some of such studies (Section 2.2.3). These studies inform the motives behind the methodology of the present investigation.

2.2.2.1 Articulatory methodologies

Below is an overview of different technologies used in the study of articulatory phonetics, which will be discussed in the following section.

2.2.2.1.1 Ultrasound

Usually associated with scans during pregnancy, ultrasound is used by articulatory phoneticians to image the tongue in the mouth during speech. The ultrasound image is produced as a result of the reflective properties of sound waves through the omission of ultra high-frequency sound. Ultrasound works by monitoring density changes in tissues, reflecting back from the interface of change between two types of density.

In ultrasound tongue imaging, the transducer is attached under the chin and the sound wave travels towards the tongue, reflecting back from the upper tongue surface. As the tongue body has a different density to both the palate bone and the surrounding air, the echo created in the ultrasound image is fairly strong and we are able to visualise it through a bright white line on the scan. We can also observe less prominent echoes on the screen, from other tissues and tissue interfaces in the mouth.

Use of other techniques in linguistics can be limited due to safety purposes, particularly ones which expose subjects to ionising radiation. Ultrasound, in comparison, is (relatively) non-invasive and safe, so can easily be used with both adults and children alike. It is also relatively cheap and portable in comparison to other articulatory equipment, making it an increasingly popular choice in phonetic research.

As ultrasound captures most of the tongue root, it is very applicable to the study of /l/, as we can see what the whole tongue surface is doing. Its major drawback for this study is the reliability of tongue-tip contact. The palate can be traced before the recording (see the methodology chapter for full details), so we have an idea where it is, but actual tongue tip contact or a very close approximation is difficult to tell apart. Electropalatography (see below) is better suited for actual contact. An additional drawback with the machine used in this study is that the low framerate (30 frames per second, 60 deinterlaced) means that temporal analysis can only be conducted on the crudest of scales. See Section 2.2.3.1

for why this might be a weakness for studying /l/.

For further information on the technicalities of ultrasound and ultrasound tongue imaging the reader may consult Hedrick et al. (1995) and Stone (2005) respectively. For an overview of its uses in linguistics and sociophonetics the reader may visit the *Seeing Speech* website (Stuart-Smith et al. 2013).

2.2.2.1.2 Electropalatography (EPG)

In EPG, the speaker wears a palate in the mouth, which has 62 electrodes (in 8 rows) exposed to the lingual surface. The sensors are connected to a computer, which records when contact is made. EPG provides dynamic real-time visual feedback of the location and timing of tongue contacts with the hard palate.

EPG analyses are more expensive and invasive than ultrasound tongue imaging, as the palates required are speaker-specific and are fitted by a dentist. This makes EPG a more costly and time-consuming process overall than some other articulatory techniques. In addition, EPG gives no overview of general tongue body position, or no information of articulation beyond the velar region, such as tongue root retraction. However, it does monitor tongue tip contact, which ultrasound cannot do reliably. This would make it advantageous for the study of /l/-vocalisation and examining contact across the midsagittal plane.

2.2.2.1.3 X-ray microbeam

X-ray microbeam makes it possible to examine tongue movements by tracking pellets fixed to various articulators and flesh points. It has a high temporal resolution, and can locate the exact position of the tongue in the mouth, as well as precise points which are of interest (e.g. the tongue tip). However, it is expensive and is toxic for the informants (although not as toxic as the regular x-ray machines, as it has a focussed beam). X-ray microbeam is perfect for monitoring temporal movement, such as gestural phasing, but less so for giving an overall picture of the tongue body, as the pellets are only placed on a few locations on the tongue. There is also no access to the tongue root.

2.2.2.1.4 Magnetic resonance imaging (MRI)

Like MRI for medical purposes, the speaker is place in a strong magnetic field and scanned during speech. MRI images the entire vocal tract, but it is relatively invasive (although not dangerous) and very expensive.

2.2.2.1.5 Electromagnetic (mid-sagittal) articulography (EM[M]A)

The subject is placed in a device which creates a magnetic field with pellets attached to the tongue. These are tracked during speech. We cannot get an image of the entire vocal tract with EMMA, however. It is comparatively expensive and invasive in relation to ultrasound.

2.2.2.1.6 Cinefluorography

Cinefluorography is the process of making X-ray film by photographing the image from a fluorescent screen. A benefit of this method is that it allows tongue body movement to be mapped relative to the entire vocal tract.

2.2.3 Articulatory studies of /l/

Earlier investigation into the articulation of /l/, such as Giles and Moll's (1975) cinefluorographic study, tended to focus on the tongue position at a point in time, rather than the relative phasing of gestures. Although Giles and Moll took tracings of transitions from the preceding vowels and into the following ones, it was the steady state of the /l/ which formed the basis of their analysis. This provided a new insight into the realisation of American /l/, showing variation between different phonological contexts.

It was Sproat and Fujimura's (1993) X-ray Microbeam study, however, that changed the way phoneticians and lab phonologists thought about and analysed /l/. In their seminal paper, Sproat and Fujimura (1993) analysed variation in American English /l/ across nine phonological environments. They found that /l/ is made up of two gestures: a more 'consonantal' coronal (or apical) gesture, and a more 'vocalic' dorsal gesture. The relative timing of these gestural components is asymmetrical: the vocalic gesture precedes the consonantal gesture for a dark [ɫ], and lags in light [l]. With respect to environment, they found that the multiple gestures result in greater constriction of the consonantal gesture in onset position, and greater constriction of the dorsal gesture in coda position. Temporally, they found that the dorsal gesture occurs earlier relative to the coronal one in coda position. They argue that the gestures are aligned with the relevant neighbouring sounds, so that the dorsal gesture aligns itself with the vowel. They also found that duration was correlated with darkness, and that darker /l/s with greater tip delay were longer, which they argue accounts for all of the variation, rather than an allophonic distinction. Acoustically, they found similar results: the F2-F1 was smaller in the darker articulations, and the results were correlated with duration.

They also find interesting effects for phonological environment, which links in with some of the morphosyntactic conditioning discussed earlier in Section 2.1.1. As demonstrated in Table 2.1 above, word-final prevocalic /l/ in a phrase such as *heal it* or *Beel equates* has dorsal lead, i.e. is dark. That is, even though it has been resyllabified into the onset, it remains dark. This result was replicated by Gick (2003:9), who also found that, in terms of tip delay, word-final prevocalic /l/s behaved like pre-consonantal /l/s for his American speaker. Gick (2003:9) found very small, non-significant differences in tongue dorsum backing across the three contexts he looked at (initial, word-final prevocalic and word-final preconsonantal), but a little more defined differences in the predicted direction for initial /l/.

As mentioned above in Section 2.2.1, Ladefoged and Maddieson (1996:361) dispute the view of Sproat and Fujimura (1993), in that the secondary articulation of a consonant will be implemented closer to the end if it is an initial consonant, and to the beginning if it is a final consonant. Ladefoged and Maddieson (1996) provide spectrographic evidence to show that initial and final /l/s are not 'mirror-images' of one another, which they use to argue against Sproat and Fujimura's claim.

For British English, the majority of articulatory investigation has focused on the phenomenon of /l/-vocalisation, where studies have usually been conducted with EPG. Hardcastle and Barry (1989) used EPG on six speakers of English. The speakers came from the West Midlands, the South-East of England, and Australia, but the authors note their speakers' dialects are most likely dampened through the pressures of education and thus do not reflect typical instances of the varieties. Sproat and Fujimura (1993:292) cite this study as providing evidence for a gradient interpretation of /l/ lenition processes, but Hardcastle and Barry (1989:14) actually make no claims to that effect. Although their data do show partial /l/-vocalisation as being more frequent preceding sibilants to a gradient degree, it is clearly the case that these speakers do not display the kind of categorical /l/-vocalisation typical of Cockney if they are speaking relatively standard English. A gradient interpretation of their results may be correct, but it does not follow that /l/-vocalisation as a categorical phenomenon does not exist. The debate between categorical and gradient realisations of /l/, however, is reserved for Chapter 4.

Scobbie and Wrench's (2003) EPG study of /l/-vocalisation in sandhi environments compared English, Scottish and American speakers. Four of their eight speakers showed almost 100% vocalisation pre-pausally and pre-consonantally, and three of these showed vocalisation before vowels. They also found that word-initial and intervocalic /l/ was almost invariably consonantal, that is, there was contact between the tongue and the alveolar ridge. In a similar study, Wrench and Scobbie (2003) combine EPG data with EMA and ultrasound, looking at a wider variety of contexts, finding a range of possible systems of vocalisation. Their analysis did not distinguish between dark and light variants however, as this is difficult (even impossible) to monitor with EPG. This study is discussed further with respect to morphosyntactic conditioning in the following chapter. One issue which is briefly discussed in this study is the amount of lip rounding we may find in vocalising speakers. The descriptions of Cockney English indicate that vocalisation is accompanied by definite lip rounding, with the diacritic often used to represent this being [w]. However, as Wrench and Scobbie (2003) show, this is not the case for all vocalising dialects, with their Scottish speaker showing no lip rounding in the clearly vocalised /l/s.

In her EPG study of connected speech processes, Wright (1989) notes an effect of /l/ being surprisingly more consonantal in fast speech rates for her young Cambridge vocalisers. Very slow speech rates were more likely to vocalise prevocalic /l/ in phrases such as *call Andy*, than very fast rates, which produced a more consonantal /l/. This is surprising, as we might expect more lenition in fast speech rates. Wright calls this phenomenon '/l/ clarification', and suggests that the process facilitates linking, in a similar way as linking /r/ might. We know since from Sproat and Fujimura (1993) that /l/s with larger dorsal displacement take longer to articulate, which may explain why speakers do not vocalise in this position in these very rapid speech rates.

Scobbie and Pouplier (2010) look at vocalisation in their EPG comparison of Scottish Standard English and Standard British English to investigate the role of syllable structure in external sandhi. They reject that categorical resyllabification occurs in word-final prevocalic words, as they find tongue dorsum retraction in this position. However, their word-final prevocalic /l/ are in phrases such as *peel Eve*, where /l/ occurs before a stressed vowel. It might be worth investigating potential resyllabification in this environment compared with an environment such as *peel it* to test differences here.

Another consideration of vocalisation of English /l/ is that we may wish to consider the multidimensional phonetic space. That is, a combination of factors should be taken into consideration, including tongue body position, palato-alveolar contact and also lip rounding.

2.2.3.1 Temporal Analysis

As discussed at the beginning of the section, temporal analysis has been an important part of many previous studies, most notably Sproat and Fujimura (1993), who claim that the ordering of the gestures is the most important articulatory correlate of darkness. One of the issues raised by their study was the idea that gestures may be inherently linked to vowels or consonants. Gick et al. (2006) conducted a cross-linguistic overview of laterals in several languages from a temporal perspective. Data comparing [w] and [l] from a previous study (Gick 2003) challenge Sproat and Fujimura's claim that the gestures are aligned with syllables based purely on their degree of manner or constriction i.e. the claim that the vocalic dorsal gesture is aligned with a more central point in the syllable and the consonantal coronal gesture is more peripheral. Gick (2003) argues that, because [w] consists of two vocalic approximate constrictions (lip rounding and tongue dorsum retraction), Sproat and Fujimura's model should predict no lag effect. However, this is not what he finds. Instead, [w] patterns very similarly to [l] in American English in that the lip gesture is more peripheral in the syllable, much like the coronal gesture in [l]. Gick et al. (2006) attempts to address the implied universality principle behind this timing pattern across the world's languages, looking at a subset of languages which allow liquids in both initial and final position. They considered the same three environments as Gick's (2003) study. In comparison to previous studies, they found greater negative lag in prevocalic position, and a shorter lag in postvocalic position, but the general pattern was the same for American English as that discussed by Sproat and Fujimura (1993).

However, they conclude by saying that no single timing generalisation can characterise the syllable position cross-linguistically, and that perceptual and biomechanics influences cannot explain everything here. Carter (2002:83) also discusses this effect with liquid resonance, stating that many languages show no dark [ɫ] but have light [l] syllable finally (see Delattre 1965). We may see this across dialects as well as languages. No one language shows the same timing effects, so the relative phasing of gestures may also vary from dialect to dialect.

As Sproat and Fujimura (1993) were using X-ray microbeam, they were easily able to track the tongue pellets over time, which is not as easy or reliable with ultrasound. Gick et al.'s (2006) ultrasound study measured the /l/ firstly by placing intersect lines perpendicular to 'relevant' gestures and then tracking the movement of this intersection over time. They calculated 'temporal lag' from the movement trajectory for each individual token. This was done by subtracting the frame number of maximum displacement of the dorsal gesture from that of the coronal gesture. This means that, in Gick et al.'s study, 'positive' lag would be the dark variant and 'negative' lag would be the light variant. This is plotted against the average movement trajectory of the tongue, which was located in each frame for this study. As Gick et al. (2006) did not use a stabiliser such as a helmet, it turns out that their method is much more complicated than would be necessary if one had access to a stabilising device. The method used by Wrench and Scobbie (2008) is much simpler. They performed a small-scale study looking at initial /l/ and word-final prevocalic /l/ in the two phrases *pale Eva* and *pay laver*. They took measurements of two radial distances from the origin of the ultrasound pulse scanline towards the tongue root and towards the tongue tip. This is a fair measurement, but also focusses in on just two points, rather than a holistic approach which would account for the whole tongue.

2.2.3.2 Pharyngealisation vs. velarisation

Although sources tend to agree that darkness involves some kind of secondary articulation, the exact nature of this is unclear. Many studies use the term *velarisation* of /l/ as being interchangeable with /l/-darkening. Velarisation is described as the secondary articulation whereby the tongue dorsum is raised towards the velum, or soft palate, and this is the description that Cruttenden (2008:216) gives for British English. This is seemingly in contrast with other studies of /l/, where it is said that tongue body *lowering* or tongue root *backing* is the primary articulatory correlate of /l/, although it would be possible to have velum raising and tongue body lowering. However, pharyngealisation is also a strategy used by speakers to produce dark [ɫ]s. This is a different secondary articulation, where the tongue root moves towards the pharynx (Trask 1996:374; Narayanan et al. 1997:1072; Müller 2011). Pharyngealisation is more commonly associated with vowels sounds (Ladefoged and Maddieson 1996:306), and here it is said that there is very little difference between velarised and pharyngealised vowels, with no languages distinguishing between the two (Ladefoged and Johnson 2014:245). Scobbie and Pouplier

(2010:241) also list uvularisation as a possible form of darkening.

It is clear from some of the existing literature (mainly outside of phonetics) that velarisation is used as a synonym for darkening, when there is no articulatory evidence for it. Although a speaker can produce a dark [ɫ] by velarisation, pharyngealisation is another strategy in which speakers frequently employ (Recasens et al. 1996:64; Narayanan et al. 1997:1072). However, as Nolan (1995:25) advises, although dark [ɫ] is often described as velarised, it is likely to be pharyngealised in the case of many speakers. Honorof et al. (2011:27) suggest that pharyngealisation may be a more appropriate way of describing darker laterals, although this is based on just one model speaker producing EMA stimuli for their perceptual study of /l/.

Müller (2011:17) discusses the lack of work which addresses this question through a large-scale empirical study, noting that we do not yet know whether use of one strategy over another is down to language, dialect or idiolect, and just how much intra-dialect or intra-speaker variation there is. She also puts forward the possibility of a pharyngeal-velar continuum. However, other articulatory studies have touched on this, so we can piece together information from these.

Narayanan et al. (1997) conducted an MRI and EPG study of /l/ in four speakers of American English. Although narrowing of the airway in the uvular and upper-pharyngeal region could be observed for all informants during dark [ɫ] production, their strategies for doing so varied. Two of their speakers produced dark [ɫ] with a significantly raised tongue body in the velar region, but the other two did so through a retracted backing of the tongue (1997:1072). Although the sample is much too small to consider any sociolinguistic effects of gender, and the speakers are not from the same area, it is nonetheless noteworthy that the two velarisers are male and the pharyngealisers are female.

Scobbie (2009) poses the question as to whether Scottish /l/ shows both pharyngealisation and velarisation. He finds that ultrasound data for Scottish speakers who have no onset/coda difference in palatality of their /l/s do show increased pharyngealisation in coda position. He suggests that, for these speakers, vocalised /l/ may be velarised but consonantal, or onset, /l/ may be pharyngealised.

This raises interesting questions for the present study. Will we find varying strategies of secondary articulation in this dataset, and will this be consistent within dialects? It also shows that future investigation from sociophoneticians may uncover all kinds of interesting speaker strategies in this regard.

2.2.3.3 Frequency effects

Lin et al. (2014) report a frequency effect in post-vocalic /l/ realisations in their ultrasound study of American English speakers, mentioned above in Section 2.2.5. They looked at complex-coda /l/s preceding /p, t, k/ in words of high frequency (e.g. *help, milk*) and low frequency (e.g. *whelp, ilk*). They predicted that the coronal (or anterior) gesture of /l/, would become more reduced in high-frequency words, so that *milk* would

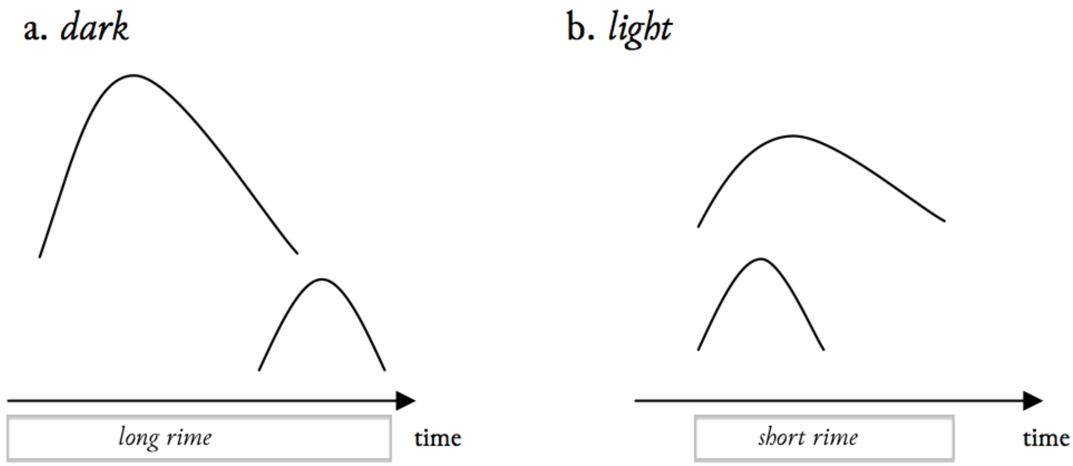

Figure 2.1: The effects of rime duration on pre-boundary /l/ (Sproat and Fujimura 1993:307). The dorsal gesture is higher up and the coronal gesture lower down. Adapted from Bermúdez-Otero (2013:17).

be more reduced, i.e. have a weaker alveolar constriction, than *elk*. They made no such prediction of the dorsal gesture. Contrary to their expectations, they did find a small but significant effect of frequency in that tongue dorsum aperture was slightly higher in high frequency words, although it did not show the interaction with place.

2.2.4 Duration

As alluded to earlier, Sproat and Fujimura (1993) claim that the conditioning of light and dark variants of /l/ can be accounted for purely by duration: the longer the pre boundary rime, the darker the /l/. They argue that this is because the dorsal gesture has a relatively large displacement, and given the time, has chance to reach its maximum peak. This is visualised in Figure 2.1. Example (a) shows a typically dark realisation of /l/, found in Sproat and Fujimura's (1993) word-final contexts, and shows the dorsal gesture preceding the coronal gesture and reaching its extremum earlier. Its displacement is also much larger. Example (b) on the right side, however, shows that the dorsal gesture is still attached to the nucleus of the /l/, but because it is much shorter, it occurs at a similar time to the coronal gesture. The displacement is also smaller because it does not have as much time to reach its target.

Although the evidence for the effects of duration seem convincing here, the authors' claims that duration solely accounts for darkness have been criticised by many other studies. Lee-Kim et al. (2013) pick apart the word-internal intervocalic contexts in Sproat and Fujimura's acoustic plot, which show varying degrees of darkness with no particular change in duration. Their argument does not hold as well for Sproat and Fujimura's articulatory data, however (see the replication of their plot in Chapter 4) which is perhaps why they chose to target the acoustics here. Barry (2000) suggests that Sproat and Fujimura's study serves better as an investigation into the morphosyntactic perturbations of /l/ articulation and that their categorical and gradient arguments are misplaced.

Nevertheless, there seems to be a tendency across many languages to lengthen final elements in an utterance (Vaissière 1983; Turk and Shattuck-Hufnagel 2007:60). In English, this has been found in word-final position, but even more so in phrase-final position (Oller 1973). There are arguments for such lengthening processes being a natural tendency which characterises all planning units. For example, such effects are found in music, birdsong and insect chirps, as well as speech (Vaissière 1983:60). However, final vowels in infant speech do not show the final lengthening processes present in adult speech, so it has been argued that this is learned behaviour (Oller and Smith 1977).

Final lengthening processes seem to be corroborated in the acoustic data of American English /l/ analysed by Yuan and Liberman (2009; 2011). However, there is an allophonic interaction. They find that duration is correlated with darkness, but for the dark [ɫ]s only. However, it is not clear how much of this could be accounted for with final lengthening, as the intervocalic /l/s are dark for their speakers, e.g. in words such as *helix*. This study is discussed further in Chapter 4 with respect to categories interacting with other gradient phonetic effects.

Lehiste (1980:7) discusses the causes behind 'pre-boundary lengthening', the observation that the final position in a word produces longer syllables. This builds on the previous observation of 'pre-pausal word lengthening' (Gaitenby 1965:38), that a following sentence boundary results in much longer syllable duration. Lehiste (1980) points out, however, that other boundaries within a syntactic unit may show phonetic effects, even if they are not pre-pausal. She analysed durations of words in sentences consisting of four different foot types, finding that the final element of a sentence was considerably longer than the other elements (potentially double the length of an equivalent initial word), regardless of foot-type. She also found speakers would lengthen segments that preceded any boundary (1980:22). Streeter (1978:1583) also found a link between duration and prosodic boundaries, in that duration was the most important cue for signalling a phrase boundary in disambiguating algebraic expressions, e.g. $(A+E)\times O$ vs. $A+(E\times O)$. Moreover, Lehiste et al. (1976:1201) showed that, when removing all other cues (such as pitch, modification of segmental sounds and insertion of pauses), duration of the preboundary segment was enough to effectively indicate the syntactic boundary. Lehiste (1980:24) concludes by saying, "In English, the controlled timing of articulatory gestures takes the rhythmic structure of speech into account." Newton's (1993) results also support this by investigating the link between darkness and duration through a perceptual study. He found that phoneticians perceived longer /l/s as being darker than shorter /l/s, even though no other differences apart from duration were present in the signal.

However, it is not clear how duration plays a role in /l/-darkening for all varieties of English. In his thesis, Carter (2002:158) notes that the prediction of darkness correlating with duration bears out for his Sunderland speaker, who shows very light /l/s in all contexts, but not for his Manchester speaker, who shows very dark [ɫ]s in all contexts. In fact, the Manchester speaker has the opposite pattern: in the greater the duration of the initial /l/, the lighter it is (looking at F2 values). This is only an observation, however, and does not reach statistical significance, potentially because of the small dataset. Huffman (1997) also argues that longer /l/s are not always darker. Other studies find no effect of duration. Van Hofwegen (2010:287) found that duration was not a significant factor in the apparent time change of /l/-darkening in AAE, nor did Barry (2000) in his EPG study of light, dark and syllabic /l/s in RP.

In summary, it seems like there is compelling evidence for duration being associated with /l/-darkness, but it is not clear how this plays out in different phonological environments, and across different dialects. Sproat and Fujimura (1993:293) attempt to take these observations further by claiming duration can solely account for any positional differences in /l/-darkening, over categorical dimensions such as lightness and darkness. This discussion, however, is reserved for Chapter 4.

2.2.5 Coarticulation

Although it is usually the position of the word which is said to have the most influence over the lightness or darkness of /l/ realisation within a particular system, it is well known that the flanking vowels can also have a significant effect on its articulation. A neighbouring front vowel, such as /iː, e/ will produce a lighter realisation than a neighbouring back /uː, ɒ/. It is usually the following vowel which is said to condition this articulation, rather than the preceding (Jones 1966:90). The acoustics overview in Section 2.2.1 discussed some studies which provide evidence for this, such as Bladon and Al-Bamerni (1976); Lehiste (1964); Nolan (1983).

If it is said that the lateral constriction towards the alveolar ridge is the most important component of /l/ articulation, the tongue body is free to change its articulation (Carter 2002:80). This implies that the following vowels can influence the rest of the tongue body however they like. This also raises interesting questions for the effect of preceding vowels on word final /l/: will they also affect the tongue body, or will the associated dorsal gesture with word-final /l/ mean that the tongue-tip will show more variation in this position? As outlined in Section 2.2.1 above, Bladon and Al-Bamerni (1976) find that light /l/ is much more susceptible to following vowels than dark /l/s in their study of American English. This could be due to the major influence of following vowels over preceding vowels, or the physiological fact that the tongue dorsum is less ‘free’ than the tip to move with its neighbouring vowels.

Neighbouring consonants have also been shown to affect articulation of /l/. Lin et al. (2014) used ultrasound to analyse the effect on /l/ of a following consonant in complex codas with labials, velars and coronals such as *help, milk and melt.* Clearly, /l/ is expected to be less anterior when preceding a velar or labial than when preceding a coronal, and this has also been confirmed by Giles and Moll (1975), Hardcastle and Barry (1989), Scobbie and Wrench (2003) and Wrench and Scobbie (2003) for English and by Recasens and Espinosa (2009) for Catalan. However, Ash (1982) found no such effect of place in her auditory Philadelphia study. Although it could be argued that articulatory data is needed to monitor such patterns, other studies have successfully accounted for this auditorily. In their study of Australian and New Zealand English, Horvath and Horvath (2002) found the effect of consonant place of vocalisation to be velar > labial > coronal in order of strength, and this analysis was conducted auditorily. The same consonantal effects have accounted for by auditory coding in both in Britain (Johnson and Britain 2007), and in the USA (Dodsworth 2005).

This effect has also been observed across word boundaries depending on the following consonant. Scobbie and Pouplier (2010) found that word-final /l/ was more likely to have contact with the palate when the following consonant was /h/, rather than /b/.

Wells (1982:259) states that /l/-vocalisation has 'massive implications' for the reorganisation of the vowel system. He discusses the breaking effect that dark or vocalised /l/ can have on the preceding vowel which will possibly give rise to eventual phonemes such as /ɪʊ/ and /ɛʊ/. Vowel mergers before /l/ usually occur when the /l/ has undergone full vocalisation, or when /l/ is particularly dark word-finally (Jones 1966:92). In the South-East, this may happen in the back vowels, so that speakers have a merger between *pull, pool* and *Paul*, or some combination of the three (reportedly all three are merged in Southampton; Hughes et al. 2012:90). This may be sensitive to morphosyntax, with mergers reported before some boundaries but not others (Jones 1966). In Blackburn in the North-West, we find a merger between *bowl* and *ball*. As Shorrocks (1999) notes, the /l/-vocalisation found in the South is widely reported, but little is known about the kind found in Lancashire.

2.2.5.1 Vowel off-glide

There is some discussion in the literature of the possible schwa-like off-glide which occurs between the vowel and the dark [ɫ] word-finally (Gick and Wilson 2006). Jones (1966:91) mentions that, although this may often be perceived, it is purely an 'incidental transitory sound which need not be symbolised in phonetic transcripts.' This claim is interesting from the perspective of categoricity and gradience. Jones's comments imply that this schwa-like sound is simply some kind of result of gestural overlap, and that it is no way cognitively controlled by the speaker. However, auditory observation in some areas of Greater Manchester suggest that this may have become phonologised to some extent, that is, the former schwa-like glide seems longer in duration than a mere transitory sound, and results in the monosyllabic word sounding more disyllabic. The Manchester speakers in the present study do not produce these kind of sounds (I associate this more with Ashton, Bolton and the surrounding areas further out) but it is certainly a topic ripe for future study. Mees and Collins (1999:193) report obvious breaking for Cardiff English before a final dark [ɫ], giving [miːəɫ] for *meal*. There is seemingly also a strong effect in American English, given Hayes's (2000) study of /l/-darkening in which he uses a schwa off-glide as a proxy for darkness, rather than measuring the /l/ itself (see Chapter 4 for more details of this study).

2.3 /l/ in varieties of English

Perhaps due to its lack of salience or striking sociolinguistic variation, studies of /l/-darkening from a sociolinguistic or variationist perspective are scarce. For the most part, we are relying on descriptions of /l/ from studies of dialects in general which happen to mention the quality of /l/ in passing. Alternatively, we may find details from pronunciation guides about English, which typically focus on RP but may divulge one or two facts about /l/ regionally. This dearth of information on dialectological accounts of /l/-darkening is less true for /l/-vocalisation, and so the two are treated separately in this section.

The different realisations of /l/ across varieties of English are one of the motivating research goals behind the present thesis, so it is necessary to get some insight into what we might expect to find. Section 2.3.1 gives an overview of what the dialectal and sociolinguistic literature has to say about the allophonic realisations of light and dark /l/, as well as the minority of phonetic studies that have looked at /l/ from the perspective of different varieties of English. Section 2.3.2 focusses on those varieties which show complete vocalisation. Although this process has been described clearly in articulatory terms above, the studies in this section primarily focus on auditory analysis of categorical /l/-vocalisation, which can be coded fairly reliably by auditory means.

Figure 2.2 shows the description in *The Linguistic Atlas of England.* The description of vocalised /l/ seems to be fairly accurate, but the all light [l] in the North, as we shall see, does not tally with the situation today. Whether this is evidence of rapid change, or perhaps an inaccuracy as a result of the auditory analysis is not clear.

2.3.1 /l/-darkening

Although American dialects of English tend to be tarred with the 'all dark' brush (Jones 1966:92), the descriptions of British dialects vary widely from region to region. Wells (1982:370) says that dialects of the North of England often lack the light/dark distinction found in RP and the varieties in the South. From the descriptive literature, this certainly seems to be the case, although the phonetic extent to which we find similarities across phonological contexts is not clear. It is said that Manchester speakers have dark [ɫ]s in all contexts (Cruttenden 2008:218; Hughes et al. 2012:149; Kelly and Local 1986), or more generally speakers from Lancashire (Beal 2008:130).[3] We have acoustic

[3]Manchester was historically part of the county of Lancashire until the Local Government Act in 1972 which led to the creation of Greater Manchester 1974. Therefore we might expect Mancunians to exhibit the same /l/ allophony (or lack of) as their Lancastrian neighbours. The study of the effect of such borders

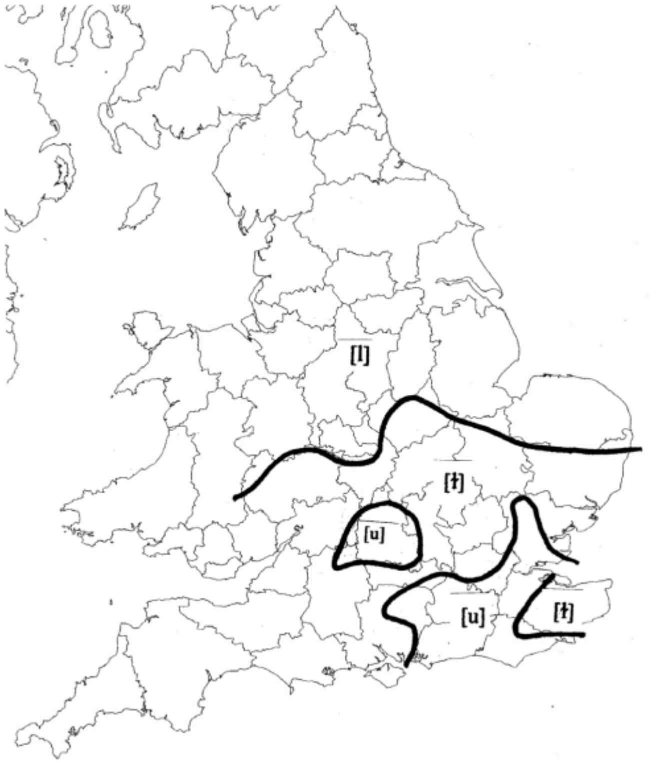

Figure 2.2: The realisation of coda /l/ in the mid 20th century in the Linguistic Atlas of England Orton et al. (1978)

corroboration of these claims to an extent. Carter (2002:151) included Manchester in his acoustic study of /l/-darkening. Although he found that the /l/s both initial and final position were phonetically very dark, there was a small but significant difference acoustically. This is also noted in other acoustic studies of English /l/: varieties for which descriptive observations report no difference between initial and final /l/, the acoustics find a small but significant one (Newton 1994:168).

In the North-East, the reports usually tally with the description of a lack of distinction between phonological contexts in Northern dialects, but inversely to the North-West. Wells (1982:371) summarises /l/ variation in the North by differentiating between the 'far North', which he says roughly corresponds to the historical borders of Northumberland and Durham, and the rest of the North-East such as Yorkshire. The 'far North' has a light realisation in all environments, as opposed to the rest of the North. Light /l/ in all positions is reported for Newcastle (Watt and Milroy 1999) and Northumberland (Hughes et al. 2012:155). Again, this is to some extent confirmed acoustically by Carter (2002; 2003) for both Newcastle and Sunderland. However, there is still a small difference between initial and final realisations, just like in Manchester. Orton (1933:7) describes South Durham /l/ as being 'thin' i.e. light, but not as 'thin' as the Northumbrian vernacular. The claims of light /l/ seem only to reach Tyneside and Teeside, however, with Middlesbrough showing an intermediate allophonic distinction (Llamas, pc) which

on accent variation is beyond the scope of this study, but may provide interesting results for sociolinguists.

seems to be toward the darker end of the spectrum (Hughes et al. 2012:120).

In Yorkshire, we find descriptions closer to North-West varieties, with Hughes et al. (2012:109) describing /l/ as generally quite dark in both onset and coda position. Hull is described as 'lightly velarised in all positions' (Williams and Kerswill 1999:148)[4], and Leeds shows a similar acoustic distribution to Manchester (Carter 2003; Carter and Local 2007). There are conflicting reports for Sheffield, with Stoddart et al. (1999:76) describing the variety as having light [l]s in all positions but Kirkham (2014) finding very dark laterals in his thorough acoustic analysis. It is possible that this contradiction is the result of a change in progress. However, given the results of nearby Leeds, and the fact that Stoddart et al.'s (1999) speakers are aged as young as 12 in 1999, and Kirkham's are aged 14 in 2012, a change in progress seems unlikely. The discrepancy is more likely to reflect how difficult it is to reliably code /l/ variation auditorily, and the importance of phonetic analysis with this liquid consonant. Interestingly, we find that Carlisle in the North-West, and very close to the Scottish border, is described as having a 'moderately clear' /l/ Hughes et al. (2012:125). However, this is not necessarily surprising given that this area of the country has many linguistic aspects in common with the North-East, rather than the North-West (Jansen 2012).

In addition to his general results, Kirkham (2014) found a noticeable difference in ethnicity for his Sheffield speakers. Anglo speakers were the ones producing the typical dark variants, whereas Pakistani speakers had very light [l]s in all contexts. This is unsurprising, as very light [l]s have been reported for other varieties of British Asian English, such as Bradford Panjabi (Heselwood and McChrystal 2000), and Glasgow Asian (Stuart-Smith et al. 2011:45), as well as Indian English generally (Sharma 2012; Wells 1982). Stuart-Smith et al. (2011) suggest that these clearer variants of /l/ found in British Asian speakers could ultimately go back to Punjabi or Urdu, which also has very light [l]s (Naseem 2002). However, light /l/ does not seem restricted to just speakers of Asian background in Britain, but could occur more as a general multicultural influence, as it is also reported for Caribbean Englishes (Wells 1982:570).

The situation in the Midlands is less well-reported, although we find some descriptions. Mathisen (1999:111) claims Sandwell /l/ is dark [ɫ] in most positions, but that there is a gender distinction, with females producing lighter variants prevocalically. Many dialects in the Midlands report descriptive sociolinguistic variation for ongoing changes in /l/, which usually make reference to /l/-vocalisation and are discussed below.

Jones (1966:91) describes some speakers in the South of England to have lighter than expected /l/s in final position, although he does not mention a specific location. This might be Norwich, as Trudgill (1999:133) notes that there was originally a light [l] syllable-finally in Norwich, which nowadays has an RP-like light/dark distinction. He poses the question as to whether we will see vocalisation in Norwich in years to come. This aside, the South of England does not seem to show any particularly interesting pat-

[4]Kerswill (pc) confirms that this description does not necessarily reflect a velar vs. pharyngeal contrast, but in this case *velarised* is used as a synonym for darkening in general.

terns with regard to darkening, but really comes into to its own with /l/-vocalisation, as discussed in the next section.

Outside of England, Irish /l/ is described as 'strikingly light' in all environments (Wells 1982:431) and there are numerous other reports of this for both Northern and Southern Ireland (Jones 1966:92; Hickey 2005:272; Hughes et al. 2012:141). However, there is sociolinguistic evidence that this may change, with claims of Derry /l/ showing possibly dark realisations syllable finally (McCafferty 1999:250).

In North Wales, Morris (2013) looked at liquid variation in the system of Welsh-English bilinguals, taking a sociophonetic approach. Although both varieties of English and Welsh were though to be 'velarised' in this area, Morris found that English [ɫ] is lighter than Welsh [ɫ] word-initially for females, and for everyone intervocalically. South Wales is reported as having light [l] in all positions (Hughes et al. 2012:94), but Cardiff contrasts to the rest of the region by exhibiting an allophonic light/dark distinction (Mees and Collins 1999:187). This kind of pattern is found for many other aspects of Cardiff English, such as t-glottalling, where upwardly mobile speakers seem to have Southern British English as more of a standard than their own Welsh varieties (Mees and Collins 1999).

Scottish /l/ is described as being dark in all positions (Jones 1966:92), and we have a plethora of acoustic and articulatory evidence to back this general description up thanks to the number of articulatory studies done on Scottish varieties. Many of the studies listed in the articulatory section include Scottish speakers. Although darker initial variants are long-established in Scottish English, Speitel's (1983) study of Edinburgh shows that the variable is subject to sociolinguistic variation. WC speakers tend to have darker variants in both initial and final position, and MC speakers showed the RP-like light/dark dichotomy. This class pattern is also reflected in Glasgow speech (Stuart-Smith 1999:210). However, young MC males in Edinburgh were commonly found to be using dark [ɫ] initially. Light [l] in initial position was used slightly more by females and in more formal styles in Edinburgh, although these effects were small.

Given the existing descriptions of dialects such as Irish English and Lancashire, it poses the question as to the quality of /l/ in Liverpool. Although Liverpool is geographically surrounded by the dark [ɫ] varieties of Manchester and Lancashire, we know the accent is heavily influenced by Irish English (Honeybone 2007), which has very light [l]s (see below). In contrast to Jones (1966:92) who says that Liverpool /l/ may be light in all positions from the Irish influence, Knowles (1973:256) claims that 'Scouse' (the Liverpool accent) /l/ is velarised and does not vary much according to syllable position. He states that before a vowel, it is darker than RP, but after a vowel it is clearer than the Southern English pharyngealised /l/. This is an interesting description from several perspectives. Firstly, Knowles uses the term *velarisation*, but for him it does not serve as a synonym for darkening, as he also comments on the standard pharyngealised variants found post-vocalically in RP. Secondly, it would seem from this description that we might expect Liverpool /l/ to behave more like Manchester English, if it is always velarised (hence darkened) in all positions, with a small difference between initial and final position. This description provides further support for Wells's claim that Northern English /l/s fall in between the possible extremes of Southern ones for British English.

Aside from these passing descriptions, there is little sociolinguistic study on the realisation of /l/-darkening and quantitative analysis is rare, although there are a few exceptions to this, mainly from American English. van Hofwegen (2010) looked at the changing speech of African American English (AAE) speakers in North Carolina, comparing sociolinguistic interviews with ex-slave recordings from the 1930s and 40s. She found that /l/-darkening was changing in apparent time, and found that AAE /l/s are moving away from their original lighter target (which she tentatively links to the dialects of former slaves), and getting darker, towards the speech of General American English.

2.3.2 Vocalised /l/

In the UK, the process of /l/-vocalisation is generally associated with the Cockney dialect spoken in London. Wells (1982:259) dates /l/-vocalisation as a relatively new phenomenon, estimating it as 'less than a century old' in 1982. He notes there is no orthographic marking of vocalisation in Pygmalion's Eliza Doolittle, who has *l* present in *gel* and *spoil.* Wright (1905:59) states that /l/ has remained the same in all positions. In London, vocalised /l/ does not just lose its alveolar contact, but is accompanied by rounding (Wells 1982:95), giving it a distinctive sound, and explaining why many transcribe it with a [w]. In their sociolinguistic study of London high schools, Hudson and Holloway (1977) show that /l/-vocalisation is led by young WC females. Although Wells states in 1982 that word-final prevocalic /l/s in contexts such as *legal info* cannot vocalise in London, by 1999 Tollfree documents vocalisation in this position for the younger speakers only. This is evidence for domain narrowing of the phonological rule, which is discussed in detail in the next chapter. However, there is no evidence of further progression, with vocalisation never appearing in intervocalic words such as *shallow, Eleanor* in her data (Tollfree 1999:175).

RP is said to be showing various changes towards more urban variants. This has been shown for t-glottalling (Fabricius 2000), but /l/-vocalisation is also listed as something which modern RP speakers may display (Wells 1982:106, 259) . It will be interesting to see if we have any evidence of this in the current study. It is found in many other areas of the South, and often shows evidence of a change in progress. Younger speakers lead in Cambridge (Wright 1987; 1989), showing little effect of speech rate on their vocalised /l/s, in contrast to the older speakers. It is attested in Southampton by Hughes et al. (2012:90), who note that Wallace (2007:218) reports /l/-vocalisation in the dialects of Hampshire from the Survey of English Dialects incidental material. It is a typical feature of Essex English (Johnson and Britain 2007; Gibb 2014).

Although /l/ vocalisation is usually associated with the South, it is not restricted to this area. Lancashire /l/ may be vocalised (Hughes et al. 2012:149), particularly in areas like Blackburn, resulting in the vowel mergers discussed above in Section 2.2.5. Vocalised /l/ seems to be creeping into the Midlands, with reports for many dialects showing change in apparent time. These are usually descriptions, but span a wide area including Leicester in the East Midlands (Hughes et al. 2012:102), and Sandwell in the West Midlands (Mathisen 1999:111), In Derby, vocalisation is variable and socially stratified, being preferred by males, WC speakers and younger speakers (Docherty and Foulkes 1999:52). The highest vocalisers are young WC males, who vocalise 77% of the time in word-list style, and the authors note that it is even higher in sociolinguistic interview, but do not provide figures for this. Note that many of the EPG studies listed above including Scottish speakers showed evidence of advanced vocalisation, with informants vocalising prevocalically as well as prepausally and pre consonantally (Scobbie and Wrench 2003; Scobbie and Pouplier 2010; Wrench and Scobbie 2003). Scottish English does not seem to be associated with the typical vocalised /l/ realisation, perhaps because it is different to the stereotypical Cockney realisation.

/l/-vocalisation is not only found on this side of the Atlantic, with many descriptions of American English showing that /l/ is completely vocalised. Ash (1982) did a comprehensive study of /l/-vocalisation in Philadelphia, looking at both final and intervocalic /l/. Philadelphia is unusual in having /l/-vocalisation occurring intervocalically, so that words such as *bounce* and *balance* can become near-homophones.

It is not clear whether vocalisers have a three-way distinction between light, dark and vocalised /l/, or just a two-way distinction between light and vocalised. It is most likely that this varies from dialect to dialect. There is evidence of a three-way distinction in Cambridge, as Wright (1989) also reports her results in terms of light, dark and vocalised /l/. She states that pre-vocalically, speakers were either light or vocalised, rarely dark [ɫ] (Wright 1989:363), which indicates that dark [ɫ]s were most likely found word-internally, perhaps before some kind of stem-suffix boundary.

2.3.2.1 Style-shifting

Style-shifting does not crop up in any investigations of /l/-darkening, even those which have trouble avoiding the Observers Paradox. However, we do find numerous descriptions of speakers style-shifting in /l/-vocalisation studies. Hughes et al. (2012:81) note that in their recording of a speaker of London West-Indian English, a variety of dark and vocalised /l/ variants appear in her sociolinguistic interview. However, when then get to the reading and wordlist styles, she produces ‘oddly exaggerated’ light [l]s in final position. They hypothesise she associated this pronunciation as clearer, more precise articulation appropriate for reading list style. Hughes et al. (2012:125) also report ‘conspicuously clear’ realisations for their Carlisle speaker’s wordlist, in both complex coda position (e.g. *felt, guilty*) and word-finally (*until, people*). Wright (1989) also finds such an effect in her EPG study of Cambridge speakers, stating that they produce ‘aberrant’ clear /l/s in final position, which she takes as a formal effect of the laboratory setting.

2.3.2.2 Intrusive /l/

To the bemusement of many speakers which do not have heavy vocalisation, some varieties of English exhibit intrusive /l/. In the UK, this is a feature of Bristol (Hughes et al. 2012:87; Jones 1966:94; Wells 1982:344) so much so that the original name of the city was Bristow. Intrusive /l/ results in the phoneme being added to words which end in schwa, so that *America* may be realised /əˈmeɪɪkəl/. It is not a sandhi phenomenon like intrusive /r/ in the UK, so can apply word-finally to a word in isolation, leading to the joke about the three sisters from Bristol named Idle, Evil and Normal (quoted in Wells 1982: 344). The explanation of intrusive /l/ is similar to that of intrusive /r/: speakers who are heavy vocalisers are unaware of where an /l/ usually appears, and so ‘hypercorrect’ by inserting them everywhere.

Intrusive /l/ is not restricted to the UK, and in fact is well documented in Philadelphia in the USA, even producing them word-internally before a stem-suffix boundary, giving *drawling* for *drawing*. The likelihood of capturing these rare occurrences is low, given the low rate at which they occur and the heavy stigmatisation. However, Gick (1999; 2002) provides a through phonological and sociolinguistic summary of the phenomenon for US dialects. The chance of evidencing them on the ultrasound, sadly, is slim to none.

2.4 Summary

This chapter has given an overview of the phonological, phonetic and sociolinguistic investigations into /l/-darkening and vocalisation. The goal of the present study is to investigate variation in /l/-darkening from the perspective of phonological theory. However, as the phonetic evidence shows, it is important to have an empirical insight into these processes, in order to best understand them. The empirical evidence thus far points towards articulatory methodology being the most robust in terms of reliability. The existing literature also show different findings from study to study, which have been taken as contradictory by some so far. However, it is clear from cross-dialectal investigation that what is true for one variety may not be of another. Because of this, it is important to collect data from speakers of many different varieties, to gain an insight into the possible typologies of /l/-darkening systems. This is why it is important to investigate claims from the sociolinguistic literature.

However, like many of the auditory claims above, these dialectal descriptions are very general and do not represent the more fine-grained phonological or phonetic patterns that we see through phonetic analysis. As shown in the acoustic evidence from previous studies such as Carter (2002), even accents which are described as being light in all positions, or dark in all positions, still show different phonetic resonances, i.e. final

/l/ in these dialects is usually darker than initial /l/. The claim being made by descriptive accounts is usually in reference to the standard RP. By conducting an articulatory study of different varieties of English, we gain added insight into the possible variation across and within dialects, which sociolinguists could then use to conduct larger analyses across the speech community.

Overall, the existing literature poses many questions which will form the basis of the present research to greater or lesser extents. The first is the morphosyntactic conditioning of /l/-darkening. We saw in the first section in Table 2.1 that different studies report differences in which phonological environments darkening can appear in, and we also see this from studies of vocalisation. Chapter 3 will discuss the diachronic trajectory of this kind of morphosyntactic sensitivity, and how and why this may differ from dialect to dialect.

We have also seen that some phonetic studies of /l/ find evidence which conflicts with the traditional phonological distinction reported, stating that the two allophones are merely two ends of a continuum. Chapter 4 will address the possibilities between a categorical phonological approach to /l/-darkening, alongside a gradient phonetic approach. There are numerous other interesting factors concerning /l/ lenition processes which have been described in this chapter, such as the link between darkening and vocalisation, the effect of duration, secondary articulations and dialectal variation. Although these factors will not form the basis of the main research questions of the thesis, they will be accounted for along the way, and will help inform our current understanding of these issues within the phonological, phonetic and variationist paradigms.

Chapter 3

The Life Cycle of Phonological Processes

As explored briefly in the previous chapters, the process of /l/-darkening in English shows sensitivity to morphosyntax, but the level at which this applies can vary from dialect to dialect. We need a theory which can account for the fact that, for example, in some varieties of English, /l/ darkens before a stem-suffix boundary, and in others it does not. The life cycle of phonological processes can account for such differences, alongside many other complexities found in the distribution of English /l/.

This chapter provides an overview of the theory of the life cycle of phonological process, as proposed by Bermúdez-Otero (1999) and Bermúdez-Otero and Trousdale (2012), as well as related ideas and observations (see also Bermúdez-Otero 2007b; 2010b; Ramsammy forthcoming). The main idea behind the life cycle concerns the synchronic reflections of diachronic processes within patterns of sound change, showing the progression of the change through it becoming increasingly embedded with morphosyntactic structure over time. The ability of the life cycle to account for differing patterns of morphosyntactic sensitivity within a language make it an attractive theory for variable phonological processes which show varying application in different morphosyntactic domains from dialect to dialect.

The chapter will proceed as follows: Section 3.1 presents the ideas behind the life cycle, as well as the different stages, followed by discussion of /l/-darkening in relation to the cycle in Section 3.2. Sections 3.3 to 3.5 give an overview of the related ideas of the life cycle, including rule generalisation, lenition trajectories and rule scattering.

3.1 The life cycle

The idea of a phonological process proceeding through the grammar in an orderly fashion dates back to the early discussions of sound change in generative grammar (Schuchardt 1885), and for a while, these ideas were not revisited. More recent work has returned to this idea with the view to approaching phonology from an amphichronic perspective (Bermúdez-Otero 2013; Kiparsky 2006:222). Bermúdez-Otero (2013:1) states that, in amphichronic phonology, ‘synchronic and diachronic explanation feed each other.' The life cycle of phonological processes seeks to explain the path in which a change has trod, from its initiation to its current state in the synchronic grammar. The idea, described step-by-step below, and visualised in Figure 3.1, is that phonetically-driven gradient processes phonologise and become cognitively controlled, before stabilising as phonological rules at the phrase level of the grammar. Over time, a rule may advance through the grammar, advancing to the word level, and finally the stem level. In time, a phonological rule may advance to the lexicon, through morphologisation or lexicalisation.

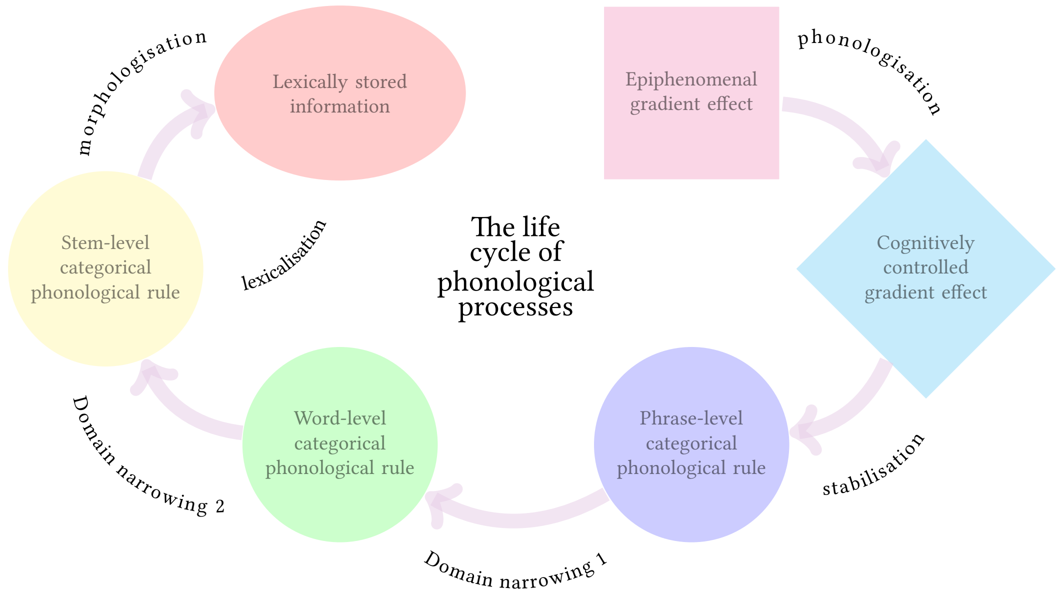

Figure 3.1: The life cycle of phonological processes (adapted from Ramsammy forthcoming). Circles represent the phonology; shapes 2-6 are under the control of the grammar.

Bermúdez-Otero (2006; 2011) uses the example of post-nasal stop deletion in English to illustrate the workings of the life cycle over time. In Early Modern English, the stop /g/ was retained after velar nasal /ŋ/,[1] so words such as *sing* were pronounced /sɪŋg/, and *singer* and *finger* were exact rhymes.[2] Over time, this stop was deleted when the consonant cluster was in the coda, as in (6).

(6) /ŋg/ → [ŋ] / __ σ]

In present day Received Pronunciation (i.e. the British English standard; henceforth RP), this is found not only at the end of a word but also when the /ŋg/ precedes a stem suffix boundary, e.g. *singer* is now /sɪŋə/ in Standard British English, not /sɪŋgə/. This

[1] In fact, Bermúdez-Otero (2011:3) discusses postnasal stop deletion in bilabial homorganic cluster /mb/. In present-day English, /mb/ clusters in the coda result in a deleted /b/, e.g. *thumb, bomb.* It is fair to class this as a stem-level process as the /b/ does not resurface when a stem-suffix boundary is attached, e.g. *thumbing, bombing*. However, the /b/ is retained when a Latinate suffix is added, in words such as *thimble, bombard.* Note that the same happens for /ŋg/ clusters in words such as *longer.* Bermúdez-Otero explains this with reference to suffixation to roots and stems.

[2] The post-nasal stop is still found in many dialects in the North (West) of England, including Manchester (Baranowski and Turton Forthcoming) and the West Midlands (Mathisen 1999) and is variable in Liverpool (Knowles 1973:293).

may be surprising given the the /g/ could be resyllabified into the following word and because, phonotactically the /ŋg/ in *finger* is the same as in *singer*. As we shall see, the difference between these two words lies in the morphosyntax, and the reason they are not exact rhymes in RP today is because of this, and will be explained in terms of the life cycle in the next section.

We have historical details of the trajectory of this change, from James Elphinston's descriptions of his own language (Elphinston 1765; as discussed by Bermúdez-Otero 2011; Garrett and Blevins 2009). In the next subsections, I will refer to this sound change to explain each stage of the life cycle, as laid out in Table 3.1. As the life cycle can predict, it is the importance of of the morphosyntactic boundaries which help explain these different effects.

	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 3.1: Adapted from Bermúdez-Otero (2011:2025)

3.1.1 Phonologisation

Under the theory of the life cycle, a phonological process is initiated as a result of phonetically motivated effects. These could be physiological restrictions of the articulators, perhaps resulting in gestural overlap, as detailed in the framework of Articulatory Phonology (Browman and Goldstein 1986; 1989; 1992). Another possibility could be a listener as the source of sound change model (Ohala 1981), where change is driven through hypocorrection and hypercorrection. The source of these are not the concern of the life cycle, which focusses more on the general cause rather than the specific reasons, which could arise from a multitude of factors. With phonologisation, some kind of epiphenomenal detail in the phonetics is reinterpreted as a new cognitively controlled phonetic process which is still gradient.

In Table 3.1, phonologisation would reflect an intermediate, variable point between Stage 0 and Stage 1 of where there has been gradient lenition of the final stop. Although it is more difficult to attempt a description of this intermediate stage of the life cycle than a seemingly categorical replacement of /g/ with [Ø], one could imagine a in-between scenario, perhaps in rapid speech or pre-consonantally, where a complex coda was reduced it some way that was not identical to full stop deletion. This would be the first circle in Figure 3.1, and this kind of gradient production is what is then phonologised by the learner to become the second circle. At this stage, the process in question will most likely subject to variation, depending on linguistic and sociolinguistic factors.

The process of phonologisation may be better understood with a phonetic shift which is perceived as more typically gradient, i.e. a vowel shift rather than a consonantal change. Take /u/-fronting in British English an example, the process whereby the /u/ vowel is produced at the front of the mouth with a raised F2. This may be transcribed as a centralised [ʉ] (Harrington et al. 2008), but in some dialects this vowel approaches that of /i/ on the F2 dimension (Baranowski 2013). However, it has not resulted in a vowel merger thus far (Chládková and Hamann 2011), even though younger speakers are much more advanced than the older generation in some dialects. Hamann (2014) argues that these younger or advanced speakers are not displaying different phonological category, but instead are showing phonetic change i.e. effects of phonologisation. She gives the example of homorganic glide insertion to make this point (see also Uffmann 2010). For the example below in (7), the hiatus breaker is based on the high front vowel and results in jod insertion. The similar process in (8) follows as high back vowel in RP, so here instead of [j] being inserted, we get [w]. This is the case in varieties where /u/ is fronted almost as far as /i/, or /ʏ/, as found in Manchester (Turton and Baranowski 2014). There are no reported dialects on English whereby the inserted glide is a [j] for the examples in (8). Hamann argues that this is evidence that the underlying representation of the vowel has not changed across generations (i.e. [+back]). This is an example of a gradient but cognitively controlled phonological process which arises as a result of phonologisation. Another feature standing in the way of a merger here would be [± round], which would presumably have to change to complete a potential merger between /u/ and /i/.

(7) a. *see* [j] *it*
    b. *be* [j] *on*

(8) a. *do* [w] *it*
    b. *Sue* [w] *is*

3.1.2 Stabilisation

Stabilisation is the stage of the life cycle when a gradient phonologised process develops into a categorical phonological process, reflected in Stage 1 in Table 3.1. Taking the example in Table 3.1, the learner hears *sing* being pronounced with some kind of gradient reduction of the stop, and posits the bare velar-nasal form /sɪŋ/ as her underlying category. As Table 3.1 and Figure 3.1 imply, this stage of the life cycle makes a crucial prediction in that the stabilisation of a gradient phonetic process into a categorical phonological one initially applies at the phrase level.

When a phonological rule applies at the phrase level, it can see across word boundaries. This means that the next word is 'visible' to the segment to which the rule is applying. To clarify this point, let us consider the domain structure of the words and phrases *singer, sing it* and *sing* in Table 3.2. The middle rows shows the phrase *sing it* at each level. When the process applies at the phrase level, the following word, and therefore following vowel, is visible to the consonant cluster, and so the /g/ is resyllabified into the next word.

After stabilisation, the rule of post-nasal stop deletion can still see across word boundaries, and the whole phrase is taken into consideration for syllabification *before* the rule applies. This explains why we do not find application of the rule in *sing it* just yet: the /g/ is resyllabified into the onset of the following syllable before the rule applies, and is saved from deletion as it no longer forms part of a coda cluster.

Environment	**Stem level**	**Word level**	**Phrase level**	**/ŋg/ cluster in coda?**
sing	[sɪŋg]	[sɪŋg]	[sɪŋg]	SL, WL, PL
sing it	[sɪŋg]	[sɪŋg]	[sɪŋ.g ɪt]	SL, WL
singer	[sɪŋg]	[sɪŋ.gə]	[sɪŋ.gə]	SL

Table 3.2: Cycles meeting the conditions for post-nasal stop deletion.

In terms of the /u/-fronting example above, this will mean that the category [+back] associated with the vowel becomes [-back]. This may happen in future generations when /u/ becomes so front in all phonological contexts (i.e. after velars and coronals alike) that the learner reanalyses it as a front vowel. Only then may we expect to witness speakers saying *do* [j] *it*, rather than *do* [w] *it*.

3.1.3 Domain narrowing

The next stage of the cycle involves domain narrowing. The process which starts off as phrase level, applying across word boundaries, over time will narrow its domain of application. The process, initially showing no sensitivity to the morphosyntax, now climbs up to the word level and can no longer see across word boundaries (Stage 2 in Table 3.1, or the word level column in Table 3.2). At this stage, a phrase like *sing it* contains no post-nasal stop. Although resyllabification would have broken up the complex coda cluster anyway, it occurs *after* the rule applies, counterbleeding the process.

This stage happens as a result of input restructuring, meaning that the learner has reinterpreted the pattern in Stage 1 into the pattern in Stage 2, due to exposure of phrase-level tokens. That is, the learner hears a Stage 1 token of *sing* with no final /g/, and reinterprets the phrase-level deletion as a rule which applies at the word level. Alternatively, if this token is in the minority, the learner may posit the historical pattern. However, over time, Stage 2 may dominate, as the learner will exhibit this domain narrowed pattern.

Domain narrowing continues to the stem level in Stage 3, which is the situation we have in most varieties of present-day English. At this stage, the process has climbed to the stem level meaning that, even though a vowel follows the /ŋg/ cluster in the same word, the /g/ deletes at the stem level, before the suffix is added. Lignos (2012) provides support for domain narrowing in this process by simulation, which suggests that this particular stage would have transitioned fairly rapidly in the history of English.

Stem-level deletion is what breaks the minimal pair in *singer* ∼ *finger*, due to the presence of the stem-suffix boundary in the /sɪŋg-er/. *Finger*, on the other hand, consists of one morpheme and is not affected by such patterns.[3]

When a rule ceases to be gradient and enters the phonology, it initially applies at the widest domain of the phrase level and can see across word boundaries. This means that a rule which currently applies at the stem level progressed by previously applying at the word and phrase levels. That is, a rule which applies at the stem level is diachronically advanced.

3.2 Evidence for the life cycle

Ramsammy (forthcoming) gives an overview of phonetic and phonological evidence demonstrating the explanatory strength of the life cycle. He uses the example of nasal velarisation in unconnected dialects of Spanish to demonstrate the processes of emergence, phonologisation and stabilisation. Speakers of velarising dialects assign a [DORSAL] place feature when neutralising nasal place contrasts word-finally, which is [CORONAL] in more conservative dialects of Spanish (see also Harris 1984; Ramsammy 2011). Ramsammy shows that the innovative form applies at the word level, showing a velarised nasal both word-finally before a pause and a following vowel, but it has yet to advance to the stem level. Ramsammy (forthcoming) gives examples from other languages which provide further evidence for the effects of the cycle, including /o/-lowering and umlaut in Swiss German, and glide hardening and continuancy dissimilation in Cypriot Greek.

3.2.1 English /l/

Some studies, such as Yuan and Liberman (2009) assume a different method of syllabification to the current paper, classifying intervocalic segments as ambisyllabic. Note that this approach accounts for their data only and not the results of other /l/-darkening studies such as Sproat and Fujimura (1993). As discussed in the previous chapter, ambisyllabicity has been posited previously by studies which need to account for phonological patterns found in all intervocalic positions, both within a word *and* across word-boundaries, such as American English /t/-flapping (Kahn 1976). However, ambisyllabicity has been shown to provide an inconsistent account of allophony in English (see Kiparsky 1979; Jensen 2000; Harris 2003; Bermúdez-Otero 2007b; 2011). As discussed in Chapter 2, Bermúdez-Otero (2007b) provides an ambisyllabicity paradox specific to /l/-darkening, amongst others. Using Sproat and Fujimura's data, he points out a categorical discrepancy between the two supposed ambisyllabic positions. As we shall see in more detail in the following chapter, for a word-final prevocalic /l/ (such as in *Beel equates*) the dorsal gesture precedes the coronal, but for a word-medial intervocalic /l/

[3]However, there are said to be varieties which do have a bare velar nasal in *finger*, such as those in the west of Scotland and Northern Ireland (Wells 1982:63). This could be an entirely different process altogether, as it is said to have been this way for centuries. On the other hand, it could be to do with Rule Generalisation, which is discussed below in Section 3.3.

(e.g. *Beelik*) the coronal gesture precedes the dorsal gesture. Furthermore, an ambisyllabic analysis is problematic for the other varieties discussed below where /l/ darkens in different domains.

As mentioned briefly in the previous chapter, Bermúdez-Otero (2007a; 2011) provides an overview of /l/-darkening typologies in English by collating results from other studies and descriptions of the phenomena, as Table 3.3 shows. This kind of typological structure is evidence of synchronic reflections of the life cycle. RP represents a conservative /l/-darkening dialect, whereby darkening applies to /l/s in the coda at the phrase level. This would represent the stage of the life cycle after the cognitively controlled gradient effect has been stabilised into a categorical phonological rule at the phrase level. One could imagine how such a process could have come about in the first place, given evidence from the phonetic literature in the previous chapter. If Sproat and Fujimura (1993) are right, then the dorsal gesture of the /l/ is aligned with the preceding vowel. If they are right about the durational effects, this phrase final position allows the dorsal gesture of the /l/ to reach its maximum extremum, which in turn results in a darker /l/. This could have originated purely as an epiphenomenal effect, which was phonologised, and later stabilised, giving us varieties such as present-day RP.

	***l**ight*	*he**l**ix*	*hea**l**-ing*	*hea**l** it*	*hea**l***	
RP	[l]	[l]	[l]	[l]	[ɫ]	Cruttenden (2008); Jones (1966)
Am. Eng. 1	[l]	[l]	[l]	[ɫ]	[ɫ]	Sproat and Fujimura (1993); Gick (2003)
Am. Eng. 2	[l]	[l]	[ɫ]	[ɫ]	[ɫ]	Olive et al. (1993)
Am. Eng. 3	[l]	[ɫ]	[ɫ]	[ɫ]	[ɫ]	Hayes (2000); Yuan and Liberman (2011)

Table 3.3: /l/-darkening in different morphosyntactic environments. Adapted from Bermúdez-Otero (2007a)

American English dialects have advanced further, so it would seem from the synchronic patterns. In the Sproat and Fujimura (1993) study discussed in the previous chapter, speakers produced dark [ɫ] in phrases such as *Beel equates the actors.* Although the authors do not subscribe to the categorical distinction, these tokens had significant tip delay and an advanced dorsal gesture typical of a dark [ɫ] in prevocalic position. These speakers show that domain narrowing has occurred, and that /l/-darkening has advanced to the word level in this variety. The /l/ darkens at the word level, and cannot see across word boundaries. This results in opaque application of a resyllabified /l/ which darkens before a following vowel. American English 2, the acoustic study by Olive et al. (1993) shows an interesting distribution where two phonotactically identical /l/s (represented by *he**l**ix* and *hea**l**-ing* in the table) show different realisations. This is due to their morphosyntactic sensitivity. The /l/ in *heal-ing* precedes a stem-suffix boundary whereas the /l/ in *helix* is part of a monomorpheme. This is the exact situation we saw above with the difference between *sing-er* and *finger* in present day English. The process has advanced to the stem level and cannot see across stem-suffix boundaries: the /l/ darkens in the coda at the stem level, and the /g/ deletes in a complex-coda at the stem level. This shows the second stage of domain narrowing in the life cycle.

American English 3 exhibits a foot-based process of darkening, whereby the /l/ darkens not only in the coda, but also anywhere outside of foot-initial onset (see also Jensen 1993; Carter 2003; Carter and Local 2007). This is not surprising, given that many lenition processes in English target consonants not only in codas, but also in foot-medial onsets. This advancement through the prosodic scale is discussed in detail below in Section 3.3.

We can observe evidence supporting the predictions of the life cycle in other phonetic studies of /l/ lenition processes. /l/-vocalisation is typically described as occurring pre-pausally and pre-consonantally, i.e. to phrase-level codas, and this is tends to be the majority finding in phonetic or sociolinguistic studies. However, we do have articulatory evidence of the process operating at the word level, with /l/ vocalising prevocalically.

Wright's (1989) EPG study of young Cambridge speakers shows this, albeit variably, with 22% completely vocalised /l/s in word-final prevocalic position. This is supported by Tollfree (1999:174) in London, who points out that Wells (1982:321) suggests word-final prevocalic /l/s do not vocalise in London, but she finds it for the youngest age group. This shows that vocalisation was applying at the phrase level in London in 1982, but by the early 1990s the youngest generation had undergone domain narrowing, and vocalisation applied in the coda at the word level for these speakers.[4]

Further EPG evidence of /l/-vocalisation advancing through the grammar even further comes from (Scobbie and Wrench 2003:1872). Firstly, one speaker of American English does show consistent vocalisation across all tokens in the phrase *peel apples*, showing that vocalisation has moved up from the phrase level to the word level. Moreover, in a follow-up study including eleven phonological contexts across eight speakers, they find evidence for four different vocalising systems, which are summarised in Table 3.4 (Wrench and Scobbie 2003:316: Fig 2). In this study, Wrench and Scobbie (2003) take any kind of alveolar contact as absence of vocalisation (symbolised by the capital [ʟ] in the table), so this is not necessarily as advanced as the kind of stereotypical vocalisation we find in Cockney speech, but purely an articulatory definition involving the presence or absence of tongue-tip contact. They found that speakers were very consistent and there was little intra-speaker variation, with the exception of the speaker named USA1 in Table 3.4. This speaker is the most advanced of them all with respect to /l/-vocalisation, where vocalisation is variably present in intervocalic monomorphemes, such as *yellow* and *bungalow*. This is a more advanced pattern than the Hayes (2000) speakers reported above, who accept dark [ɫ] in *helix*-type words, as vocalisation is the next step on the lenition trajectory.

Perhaps surprisingly, the 'Northern English' speakers are the most conservative in this dataset, even more so than RP. However, there a couple of points to address with respect to this. Firstly, these are very standard Northerners, with one being described as near-RP. Secondly, if vocalisation is a further stage of darkening and if Northern En-

[4]Tollfree (1999:163) collected her data between 1990 and 1994.

glishes do lack the sharp clear/dark allophony of the South, it could be hypothesised that Northern accents (even more RP-like ones) don't have enough of a distinction between initial and final contexts for /l/s to fully vocalise (see the discussion of lenition trajectories in 3.4). RP follows the pattern predicted by Cruttenden (2008) and discussed previously, although it is interesting to note that they tend to vocalise the /l/ in unstressed final segments, such as *apple*. However, Barry (2000) does not report loss of contact in syllabic /l/ in his EPG study of two RP speakers, so it is likely that this environment shows inter-speaker variation, particularly for a variety so broadly defined as RP.

Variety	*leap*	*believe*	*bungalow*	*yellow*	*apple*_V	*peel*_V	*elm*	*apple*	*peel*	*apple*_C	*peel*_C
Northern England	[L]	[L]	[L]	[L]	[L]	[L]	[L]	[L]	[L]	voc	voc
RP	[L]	[L]	[L]	[L]	[L]	[L]	[L]	[L]	voc	voc	voc
USA2/SE/Glasgow	[L]	[L]	[L]	[L]	[L]	[L]	voc	voc	voc	voc	voc
USA1	[L]	[L]	50/50	50/50	voc	voc	voc	voc	voc	voc	voc

Table 3.4: Summarising the findings of Wrench and Scobbie (2003:316) in their study of /l/-vocalisation. N.B. the capital [L] notation refers to the fact that the study found some form of contact with the palate, and does not distinguish between typical light or typical dark realisations. 'Voc' means the /l/ was completely vocalised, and 50/50 that some speakers were variable. SE refers to a speaker from the South-East of England.

One additional prediction made by accounts of vocalisation concerns the first stage of the life cycle. In her Articulatory Phonology account of darkening and vocalisation, Tollfree (1999:182) notes that the interpretation by this framework would crucially rely on /l/-vocalisation originating in preconsonantal position, due to the gestural overlap account. However, in her London data, vocalisation has proceeded too far and applies equally in all phrase-final coda positions (i.e. prepausally and preconsonantally) so it is impossible to disentangle the two. We could monitor potential effects in a variety where vocalisation was more recent, however. This would mean an epiphenomenal element of speech whereby the coronal gesture of /l/ is cut short by a following consonant is phonologised, applying to all coda /l/s at the phrase level. This, combined with extra phrase-level duration, eventually results in a phrase-level stabilised phonological rule of vocalisation.

3.3 Rule generalisation

Rule generalisation, in short, concerns the 'internal expansion of the sound laws' (Venneman 1972:53 via Schuchardt 1885). Although it is linked with domain narrowing in the sense that both observe the tendency of a phonological process to become more inclusive over time, the term rule generalisation is usually reserved for phonetic analogy, rather than morphosyntactic advancement through the grammar. However, when we consider phonological context alone, the expansion of the sound laws predicts a different kind of advancement through the grammar. As discussed by Bermúdez-Otero

(2007a; 2010a:2, 9) and Kiparsky (1988:14.3.1) a sound change first applies in a phonetically favourable phonological context, before encompassing more environments over time. The most favourable phonological environment for lenition is the syllable coda and, unsurprisingly, we find all kinds of variable sound changes of lenition originating in coda position. Over time, however, these changes may advance not only by domain narrowing to the coda of the word level and then stem level, but also to higher prosodic environments. That is, a process may advance from applying to codas, to applying to any position outside of a foot-initial onset. It is fair to say that 'any position outside a foot-initial onset' is a more general prosodic environment that 'coda position'. For example, with /l/-darkening, we may expect coda /l/ in a word such as *heal* to be more likely to darken over an intervocalic /l/ such as *helix*. Be that as it may, darkening in *helix*-type words is reported for many dialects of English, and adhered to above with the American English 3 dialect in Table 3.3 English (from Hayes 2000 but see also Carter and Local 2007 for evidence in British English dialects).

We have observed evidence for a prosodic target other than the syllable with respect to /l/-darkening, but such effects can be found in many other phonological processes. It is demonstrated for /r/ in Harris's (2006) description of rhoticity in varieties of English. Harris compares rhoticity in three dialects of English as shown in Table 3.5: the typical rhotic variety found in Scotland, Ireland and America (R1); the typical non-rhotic variety with no post-vocalic /r/, but /r/ resurfacing before a following vowel, found in RP and most dialects in England (R2); and a less well documented broad non-rhotic dialect, attested for the American South (R3).

			R1	R2	R3
(a)	[*r*v́	**red, rack, rude**	+	+	+
(b)	[*r*v̌	**ravine, revolt, resort**	+	+	+
(c)	C*r*	**tray, agree, petrol**	+	+	+
(d)	V[*r*v́	**Corinne, terrain, carouse**	+	+	+
(e)	*r*v̌	**very, parent, sheriff**	+	+	−
(f)	*r*]v̌	**bear a, before a, poor again**	+	+	−
(g)	*r*]v́	**bear up, before eight, poor Eva**	+	+	−
(h)	*r*C	**board, cart, source**	+	−	−
(i)	*r*]C	**bear to, before nine, poor man**	+	−	−
(j)	*r* ‖	**bear, before, poor**	+	−	−

Table 3.5: Rhoticity in three varieties of English, across ten phonological environments. Adapted from Harris (2006:359)

Harris (2006:359) argues that, although non-rhoticity is often used as a prime example of the coda's inability to licence contrasts, a syllabic target may be too narrow for all phonological rules and that some may need a broader prosodic scope. By this, he means that the coda cannot account for all /r/-vocalisation, and we might need a higher prosodic unit such as the foot. For example, Dialect R3 in Table 3.5 is a problem for syllable based analyses when we get to the monomorphemic intervocalic examples in (d) and (e). These behave differently with respect to rhoticity in that the foot-initial syllables retain /r/ in (d), but the non-foot-initial examples in (e) have /r/-vocalisation. In this dialect, the rule of /r/-vocalisation has not only moved up the hierarchy of stem, word and phrase levels, but has also moved up the prosodic domain from the coda to the foot.[5]

Rule generalisation can be observed in other variable phonological processes in English, such as t-glottalling, whereby /t/ is replaced by a glottal stop. This was typically reported for coda /t/s in words such as *cat*, but is reported at higher rates more recently in intervocalic position e.g. *better*, that is anywhere outside the foot-initial onset. Although t-glottalling is now accomplished in intervocalic position for many varieties of British English, the diachronic prosodic advancement is reflected in its rate of application. For example, in Manchester, in the social group where glottalling initiated (now middle-aged, working class males), word-final t-glottalling occurs almost 100% of the time, but just over 50% of the time intervocalically (Baranowski and Turton Forthcoming). Further evidence for rule generalisation in Manchester comes from glottalling in the so-called *-ee/-oo* environments such as *tattoo, canteen, eighteen, cartoon* (Harris and Kaye 1990:271). As such tokens bear primary stress, we would not expect them to lenite. However, we do not see lenition of stressed /t/ in any variety British English when the syllable is preceded by a weak syllable e.g. *attack*.[6] This shows that, for some speakers, glottalling has advanced to encompass the *-ee/-oo* set, whereby /t/ is part of a stressed syllable, but this syllable is not the only stressed syllable of the word.[7]

The tapped realisation of /r/ in conservative RP is also sensitive to prosodic factors.[8] The tap is considered a weak allophone of the alveolar approximant because of its reduced duration. Although the tap makes contact with the alveolar ridge, unlike the approximant, this contact is rapid, much shorter than a typical alveolar approximant and can be as short as 20ms (Ladefoged and Maddieson 1996:231). Rubach (1996:220) outlines the environment for tapped /r/ in RP as intervocalically before a non-stressed vowel, or as linking /r/. Therefore we expect a tap in (9) and (11) but not in (10) .

(9) [ɾ]: *courage, very, sorry, baron, laurel, story, period*

(10) [ɹ]: *courageous, reduce, red, bright, Henry, walrus*

(11) [ɾ]: *for example, for instance, the other end*

[5] Or arguably just moved on the prosodic domain and not the syllabic domain. For an analysis of rates of application of both domains simultaneously, see Turton (2012).

[6] However, the word-initial /t/ in *to* and its variants *today, tonight* may glottal when cliticised to the preceding word.

[7] It is worth noting that the traditional stress pattern in *tattoo* seems to have shifted for these speakers. A conservative pronunciation may or may not involve the first vowel reduced to schwa, but the first syllable would not be marked for stress. The OED lists /tæˈtu/, but the speakers who show advanced glottalling in this stressed position have /ˌtæˈtu/ as their more standard variant and /ˌtæˈʔu/ as their glottalled variant.

[8] Tapping is an extremely conservative variable of RP, and Wells (1982:282) reports that very few speakers of modern RP have the realisation in the early 80s, so it is likely not many remain today. However, it does persist in dialects in Scotland and Wales (Hughes et al. 2012:90), and cities in the North of England, such as Leeds (Wells 1982:368) and Manchester (Turton 2010). In Liverpool, tapped /r/ is reported as being the most common realisation, occurring more often than the alveolar approximant (Honeybone 2004:7, Watson 2007:352). The phonological distribution of /r/-tapping outside of RP is not clear, however.

From Rubach (1996:220)

Rubach argues that this pattern is evidence of ambisyllabicity, and as the /r/s in (10) are part of a stressed syllable, they are not ambisyllabic. Jensen (2000:195) argues against ambisyllabicity and for a prosodic analysis of /r/-tapping; where /r/ taps when non-initial in the foot. He points out that that the ambisyllabic analysis is somewhat ironic given that RP has no true (i.e. phrase level) coda /r/s anyway. Jensen provides a prosodic account to show that ambisyllabicity is not required to account for /r/-tapping, however, he does not provide any evidence for why Rubach's is problematic. What Jensen doesn't point out is that Rubach is wrong in placing *bright* in (10), and that there is a fourth condition of tapping, as in (12):

(12) [ɾ]: *three, pride, freeze*

These /r/s are not ambisyllabic under any analysis. Taking this, it would follow that the /r/ in words such as *three, pride, freeze* are susceptible to tapping as they are not foot-initial and are weakened as part of the branching onset. These examples provide possible evidence for the argument that the life cycle is governed by both a grammatical and a prosodic force. However, it could also be plausible that this branching onset realisation could just be due to gestural co-articulation.

3.4 Lenition trajectories

Entangled with the ideas of the life cycle is evidence of lenition trajectories in different varieties of a language. A lenition trajectory refers to the output typology of a sound change which result in several allophonic variants of a phoneme at varying stages of consonantality. In English, dialects show evidence of lenition trajectories in all kinds of consonantal phonological processes. The trajectory of /r/ in present day Standard British English serves as an example of this kind of distribution. As discussed in the previous section, the majority of dialects in England (with the exception of some varieties in the South-West and the North-West) display full /r/-vocalisation word-finally or before a consonant, as in (13). What was not pointed out in the previous section, however, is that although the /r/ is present in word-final prevocalic linking position, we find an in-between realisation phonetically. That is, the /r/ is retained, but crucially it is lenited, as in (14), in that it has shorter duration (Cruttenden 2008:223), higher intensity (McCarthy 1993:179) and higher F3 (Hay and MacLagan 2010; see Bermúdez-Otero 2011:18 for a thorough overview of the articulatory and acoustic correlates of word-final prevocalic /r/ lenition, but also cf. for articulatory factorsMullooly 2004).

(13) *four pears* [fɔː pɛəz]

(14) *four apples* [fɔːɹ̞ æpl̩z]

This results in a three-way realisation of /r/ in present-day Standard British English, as in (15): word-initial consonantal /ɹ/, the word-final prevocalic lenited variant in *four apples*, and the vocalised variant in *four pears*.

(15) ɹ $>_{reduction}$ ɹ̞ $>_{deletion}$ Ø

Diachronically, successive steps in a lenition trajectory give rise to a series of separate phonological rules entering the grammar one after the other. Synchronically, older rules effect milder types of lenition and have narrower cyclic domains, that is, they apply at higher levels e.g. /r/ lenition applying at the word level. Conversely, the younger rules affect more drastic types of lenition, building further on the previous forms, e.g. /r/-vocalisation applying at the phrase level. In RP, /r/-lenition is word level, which exists at this higher level due to it being an older phonological process. Therefore, the process of /r/-lenition in RP targets /r/ in the coda at the word level. The younger harsher process of complete vocalisation occurs at a lower level, and targets /r/s in the coda at the phrase level. An /r/ in a word such as *four* undergoes lenition at the word level first. In the next round, the /r/ in *four pears* is also in the coda at the phrase level, and is vocalised completely. The /r/ in *four apples*, however, is resyllabified into the onset and avoids complete vocalisation, although lenition has already occurred in the previous cycle. Compare this to reports of South African English, where linking /r/ is said to occur rarely or not to occur at all (Hartmann and Zerbian 2010). This shows that vocalisation of /r/ has advanced to the word level, albeit variably, in South African English. Whether this supplanted a previous realisation of lenited /r/ in this position, or whether there was no in-between stage would be difficult to tell, but it poses interesting questions for the interaction of lenition trajectories and the advancing phonological changes.

Adapting Harris's table in Table 3.5 to the phonological environments in Table 3.1, with the added consideration of lenition trajectories, we might expect a typology of /r/-leniting dialects, as in Table 3.6. Given the possible permutations between [ɹ], [ɹ̞] and [Ø], this could predict numerous possible dialects, many more than those listed in Table 3.6. However, note that crucially the life cycle would only predict dialects which proceeded in an orderly fashion with respect to morphosyntax and prosodic domains, as well as and lenition trajectories. For example, we would not predict a dialect with lenition in a higher morphosyntactic domain when the full /r/ was realised in a lower one, e.g. *pou*[ɹ̞] *a drink*, but *pou*[ɹ]. Bermúdez-Otero (2011) states that such dialects are not only unattested, but they are impossible, by the Russian Doll Theorem:

(16) **The Russian Doll Theorem**
Let there be the nested cyclic domains [γ ...[β ...[α ...] ...] ...]. If a phonological process *p* is opaque in β because its domain is α, then *p* is opaque in γ.

(Bermúdez-Otero 2011: 2026)

It is clear to make the link from the rhotic patterns discussed above to the same patterns in the laterals. We have already discussed the phenomenon of /l/-vocalisation

	terrain	*very*	*pouring*	*pou**r** a*	*pou**r***
American English	[ɹ]	[ɹ]	[ɹ]	[ɹ]	[ɹ]
RP	[ɹ]	[ɹ]	[ɹ]	[ɹ̞]	Ø
South African Eng	[ɹ]	[ɹ]	[ɹ]	Ø	Ø
Broad non-rhotic (R3)	[ɹ]	Ø	Ø	Ø	Ø

Table 3.6: Typology of /r/ systems

both in this chapter and in detail in the literature review. It is not clear whether all vocalisers have a three-way distinction, or just a two-way distinction between light and vocalised /l/. As mentioned in the previous chapter, Wright (1989) reports a three-way distinction between light, dark and vocalised /l/ for her Cambridge speakers.

Liquids are a fruitful type of consonant in which to observe trajectories of this kind, as they have both consonantal and vowel-like qualities. However, they are not the only sounds to display these kinds of orderly lenition. We find this in other lenition processes in British English, such as pre-glottalisation and full glottal replacement in Newcastle (Milroy et al. 1994), or voiceless plosive lenition leading to full frication in Liverpool (Honeybone 2001; Watson 2007). Unlike the neat example of /r/-lenition and vocalisation, however, these kinds of trajectories are not stable and complete. In addition, these changes in progress are often wrapped up in all kinds of variation within the speech community. The beauty of the liquid trajectories is that, for the most part, their changes are complete and the diachronic pathway can be observed through the synchronic present-day patterns, which are no longer tied up in sensitivity to sociolinguistic factors.

3.5 Rule scattering

The life cycle does not merely involve phonologisation, stabilisation and domain narrowing, but also includes rule scattering (Bermúdez-Otero 2013). With rule scattering, the life cycle also makes predictions about the relative position of diachronically related rules coexisting in the same synchronic system. When a new rule enters a higher component of the grammar, it does not stop applying at the lower level. That is, rather than replacing the phonetic rules in which they emerged from, a new phonological process coexists alongside the former gradient phonetic process. Bermúdez-Otero (2013; 2007a): discusses evidence for rule scattering demonstrated in several studies of present-day English.

The first example concerns the results from Zsiga's (1995) electropalatography (EPG) study of present-day English palatalisation. In this study, Zsiga looked at palatalisation of /s/ plus /j/ clusters, both across word boundaries e.g., *press you*) and within a word e.g., *confession.*[9] Zsiga's results show that, although palatalisation occurs in both contexts, the articulators are doing different things. Word-internally, the EPG shows that the palate trace of the /s+j/ cluster is identical to a typical /ʃ/ in word such as *shoe*. However,

[9]Note that Chaucer's *confessioune* would have contained an /s + j/ cluster in Middle English.

when palatalisation occurs across word boundaries, the palate traces look identical to an /s/ and a /j/ articulated simultaneously. This is an example of rule scattering. Across word-boundaries, the /s/ + /j/ clusters produce a /ʃ/-like sound, even though the articulators are still moving towards their separate /s/ and /j/ targets, resulting in gradient palatalisation by gestural overlap. Word-internally, however, the process has been phonologised, giving categorical palatalisation by featural change. Crucially, the idea behind rule scattering is that the two exist in the same grammar.

Secondly, is the famous case of /æ/-tensing in Philadelphia (Labov 1994), whereby /æ/ becomes tense mainly before /m, n, f, θ, s/, but also under additional complex conditions (see Fruehwald 2013 for detailed chart). Bermúdez-Otero (2013:16) argues that we need two grammatical components to account for tense vs. lax /æ/: a stem-level rule capturing the default distribution, overlaid with a gradient phonetic rule controlling the formant location in the vowel space. Again, this calls for a rule scattering explanation, where the categorical phonological rule and gradient phonetic rule coexist in the same grammar.

Similarly, the process of /l/-darkening may show the cumulative effects of several cognate processes simultaneously overlaid in the grammar. Although Sproat and Fujimura (1993) argue for a purely gradient interpretation of /l/-darkening in their X-ray microbeam study, their data provide evidence for rule scattering, as will be discussed in the following chapter. This supports the idea that the process of darkening originated as a gradient phonetic process but over time has been reanalysed by learners and phonologised as a categorical process. This means that the original gradient process of phonetic implementation coexists in the grammar on top of the newer morphosyntactically conditioned categorical process.

3.6 Summary

This chapter has given an overview of the ideas and stages of the life cycle of phonological processes, as well as related processes and effects. We have seen that the life cycle can easily account for variable phonological processes which are morphosyntactically conditioned, such as /l/-darkening. What remains to be investigated is whether we can find true synchronic variation across different dialects of English which are reliably accounted for with robust articulatory methodology. The questions of morphosyntactic conditioning in /l/-darkening are addressed with empirical data in Chapter 6. For now, we will take a closer look at the related ideas of rule scattering, whereby categorical and gradient processes can coexist in the same grammar. These ideas are explored in detail in Chapter 4.

Chapter 4

Categorical vs. Gradient Processes

The previous chapter focussed on the ideas behind the life cycle of phonological processes, concluding with some ideas on how and why we may find both categorical and gradient effects of related phonological processes operating within the same grammar by rule scattering. This brings us neatly into the debate of categoricity and gradience in phonetics and phonology in general. Although evidence for both categorical and gradient processes is fairly non-controversial in some aspects of the field, in other subfields the debate rages on, typically by displaying non-compromising views. This chapter attempts to address the current debate in the literature by firstly considering the evidence from the existing literature in phonetics–phonology interactions (Section 4.1). In Section 4.2 arguments for the purely categorical or purely gradient interpretations of /l/-darkening data are presented in turn, focussing on the data from Sproat and Fujimura's (1993) seminal paper on the phenomenon. Using both ideas from previous phonological studies, and data from phonetic investigations, I will argue that a mixture of the two is required in order to fully account for the variation found.

4.1 Gradience and Categoricity

In this thesis, following the ideas behind the life cycle in the previous chapter, *phonological* is used to mean a cognitively controlled categorical distinction, whereas *phonetic* variation, which is borne out of the phonologisation of constraints on the tongue, is cognitively controlled, though gradient and continuous. In the previous chapter, we saw examples of phonetic and phonological processes overlaid over one another in terms of rule scattering. The next section gives an overview of several studies showing the effects of categorical phonological and gradient phonetic, in phonetics, phonology and sociolinguistics. As we shall see, it is fairly evident that categorical and gradient effects can coexist alongside one another in the grammar.

Variation and change in the phonological grammar can be ambiguous to the extent where it is not clear whether a process is phonological and categorical, or phonetic and gradient. This works on the assumption that phonetics and phonology are distinct modules in the grammar, a view that is not shared by all. Ohala (1990) takes umbrage with the idea of a phonetics-phonology interface in his conspicuously titled paper 'There is no interface between phonology and phonetics,' his general problem being that the two are far too entwined to be ever consider distinct modules. Scobbie (2005:2) airs similar views about the interface, stating non-controversially that empirical phonetic work can uncover gradient and variable patterns on top of previously transcribed categorical distinctions, which can be explained through the large overlap between phonetics and phonology. Although such views seek to challenge the kind of modular feedforward architecture implied in the life cycle, they still acknowledge the evidence of categorical phonological and gradient effects at play in the grammar.

In Articulatory Phonology (Browman and Goldstein 1986; 1989; 1992), a particular gesture specifies the formation of a linguistically significant constriction, and coarticulation and gradience are the result of gestural overlap. In turn, the field of Laboratory Phonology has created a wealth of empirical evidence for all kinds of studies which use phonetics to answer questions in phonology, as well as factors such as gradience and overlap.

Studies linking phonetics and phonology in this way have discovered that processes can be purely phonetic, phonological, or maybe somewhere in between. Ellis and Hardcastle (2002) analysed external sandhi realisations of /n/ using EMA and EPG to observe coarticulation patterns in English. They compared phrases with /n/ and /ŋ/ preceding velar consonants to see what effect the following consonant had on the coronal nasal's tongue tip gesture, and how it compared to the velar nasal generally. Interestingly, they found that the result was speaker-specific, although all were variable. Some speakers uttered the phrase *ban cuts* with categorical reduction of the tongue tip gesture in the [n], giving a realisation tantamount to the phrase *bang cuts*. Others showed next to no reduction at all, producing a regular [n]. What is interesting is that there is a third group, who have gradient reduction of the tongue tip gesture. Bermúdez-Otero and Trousdale (2012:7) discuss this is terms of the early stages of the life cycle, pointing out that speakers who vary between complete reduction and none at all provide evidence that stabilisation has taken place, whereas those in between speakers are showing phonologisation. Other phenomena showing evidence of both categorical and gradient processes operating in the same grammar include assimilatory speech processes (Nolan and Kerswill 1990), alveolar to velar assimilation (Kerswill 1985; Wright and Kerswill 1989), /s/ to /ʃ/ assimilation (Holst and Nolan 1995; Nolan et al. 1996) and nasalisation (Cohn 1993), as well as West-Flemish obstruent voicing (Strycharczuk and Simon 2013) and Polish voicing (Strycharczuk 2012). What these studies all have in common is that they provide empirical evidence for phonetic effects which, directly or indirectly, can be interpreted as having categorical and gradient counterparts. This suggests that, in order to gain reliable evidence for the full spectrum of phonetic and phonological effects, we need evidence from acoustic and articulatory measures. All of these studies have a strong empirical basis to the investigation, which is needed to look at potential interactions between categoricity and gradience.

4.1.1 Phonetics-phonology interactions and the speech community

We have seen above that a categorical phonological process is one which has a cognitively controlled difference, e.g. two distinct allophones. Some studies have attempted to equate this distinction with whether speakers are overtly aware of a process. The insight of social factors in phonetic realisation raises interesting questions for sociophonology, particularly the advancement through a change with the help of a speech community.

Kerswill (1985) uses EPG to investigate connected speech processes, and highlights the interaction between phonetic gradualness, phonological discreteness and sociolinguistic salience. Salience is defined under Labov's (1972) widely used continuum of indicators, markers and stereotypes (from most to least salient), referring to listeners' social evaluation or overt commentary of a particular feature i.e. how aware the speech community are of a particular feature. Kerswill (1985) discusses the difference between phonetically motivated gradual connected speech processes becoming 'fossilised', i.e. stabilisation in terms of the life cycle, and how sociolinguistically salient processes are fossilised and categorical. This is an interesting point, but works in one direction only: sociolinguistically salient phonetic/phonological features provide evidence for processes becoming categorical, but it is not the case that features which are not sociolinguistically salient are gradient.

Studies of phonetic dialectology and sociophonetics have given us an insight into how speakers show control over gradient phonetic effects, and how this may vary between speech communities. Erker (2012) conducted a sociophonetic study of Spanish coda /s/, considering the data from both a categorical and gradient approach. Erker analysed data from Spanish speakers in New York City (NYC) who were either from the regional groups of the Caribbean or Latin America, then those who had been brought up in NYC and those who had recently moved there, resulting in four possible social groups. He measured /s/ frication categorically by presence or absence, and also took measurements of the centre of gravity, for an approach accounting for gradience. The categorical results showed a simple situation: Caribbeans are more likely to /s/ delete, particularly those who had recently arrived in NYC; Mainlanders show no difference between deletion rates regardless of how long they have been there. However, the gradient results show a more complicated situation. Regardless of regional group, NYC raised speakers show significant differences from their recently arrived counterparts in the spectrotemporal properties of /s/ production. Speakers raised in NYC showed longer duration and higher centre of gravity in their fricative realisations i.e. more canonical /s/-like sounds. These results are compatible to a feedforward model where production of /s/ occurs categorically with an additional phonetic effect. This shows that, although the Latin American speakers delete less overall, their /s/ sounds are also more /s/-like if they were raised in NYC. The same goes for the Caribbean speakers; they had less /s/-like sounds overall, but those raised in NYC had fricative movements which were longer in duration and higher in centre of gravity.

4.2 Approaches to /l/-darkening

Thus far, we have considered the evidence for phonological categoricity and phonetic gradience in general, across various linguistic studies. This section will narrow in on the interpretation of /l/-darkening as told by phonological and phonetic studies. As we shall see, studies tend to opt for either categoricity or gradience in their interpretation of the data, rejecting viable evidence from the other side. By reanalysing the methods and data used in these studies, we will see that both sides of the debate have compelling arguments behind their reasoning, but overestimate the facts for their camp. Later in the section, we consider whether it is possible to unite these previous analyses in a non-contradictory manner, by retaining the positive aspects of both under one model.

4.2.1 A categorical approach

Traditional accounts of /l/-darkening in English tend to assume an allophonic distinction between light [l] and dark [ɫ] (Chomsky and Halle 1968; Halle and Mohanan 1985). This is the approach taken by Hayes (2000) when accounting for the varying rates of morphosyntactically conditioned /l/-darkening in his Optimality Theoretic analysis of judgement data. Hayes gathered well-formedness ratings of light and dark /l/ in a variety of environments. In a follow-up study, Boersma and Hayes (2001) used a sigmoid transformation to convert these well-formedness ratings into the conjectured frequencies illustrated in Table 4.1. As we can see from this table, the likelihood of a speaker producing a dark /l/ in word-initial position is almost at zero, and the inverse is true of the word final category. What is most interesting about these results is that the intermediate contexts reflect their relative morphosyntactic-phonological strengths: e.g., the /l/ in *mailer* is intervocalic, and we might expect it to be light, but because it precedes a stem-suffix boundary and is stem-final, it is actually dark the majority of time. Contrast this with the phonotactically similar (in terms of intervocalic) word *yellow*, which is monomorphemic and mostly light.

In Hayes's (2000) original study, 10 speakers of American English were asked to rank the acceptability of pronunciations of light and dark /l/s in the representative forms listed in Table 4.1. Hayes did not directly control the phonetic quality of /l/ in the stimuli, but rather that of the preceding vowel, justifying this procedure on the grounds that that front or high vowels and 'true diphthongs' preceding dark [ɫ] have a schwa off-glide. Therefore, the presence of breaking in the preceding vowel was taken as a proxy for darkening. For example, in a word such as *pool*, dark tokens would be realised as [puːəɫ]

Environment type	Conjectured frequency of light [l] %
light	99.96
free-ly	94.53
yellow	76.69
mail-er	16.67
mail it	0.49
bell	0.001

Table 4.1: Table showing the conjectured frequencies of light [l] in several phonological contexts. Adapted from Boersma and Hayes (2001:74)

and light tokens as [puːl]. The acceptability of a broken vowel was taken to equal the acceptability of a dark [ɫ]. Hayes argues that acceptability judgements are determined by a variable grammar, and that any variation in /l/-darkening is entirely down to category mixture of discrete allophones.

Although Hayes acknowledges there may be low-level duration-driven effects, he argues that the main distributional facts reflect probabilistic application of variable, morphosyntactically sensitive, categorical phonological processes. However, the experiment design inherently elicits scalar responses in judgement, making it very difficult to distinguish between variability at the community level and variability at the individual level, resulting in a potential overestimation of categorical variation in the individual. Hayes's account is also unable to control for the fact that duration variance is greater in some environments than others (e.g. phrase-finally). The additional problems with a purely categorical approach to /l/-darkening may not be of concern to some phonologists, but such an interpretation does not enable us to account for the small gradient effects in the data which may become much larger under phonetic inspection.

Thus, there are problems with a purely categorical interpretation of the data. Furthermore, the categorical nature of a light vs. dark allophonic distinction in /l/-darkening has been questioned, particularly by studies which focus on articulatory realisation. However, as we shall we in the next section, a purely gradient interpretation of the data is not without its own array of problems.

4.2.2 A gradient approach

Many phonetic studies of /l/-darkening have argued that their data are gradient, not categorical. Under this interpretation, it is argued that the so-called allophonic difference between a light and dark is merely two ends of one continuum. The most famous of these claims is made by Sproat and Fujimura's (1993) in their X-ray microbeam study of /l/-darkening. This investigation looked at /l/s in 9 contexts, ranging from initial position, to pre-consonantal position, and other prosodically conditioned environments in between. The technology allowed them to monitor gestural phasing, which led them to uncover a distinction which would be cited as the primary articulatory correlate of light and dark

/l/ from then on. They found that in light [l]s the coronal gesture precedes the dorsal gesture, and in dark [ɫ]s the dorsal gesture precedes the coronal, as represented by a subset of environments in Table 4.2.

Environment	**Example**	**Gesture**	**Realisation**
Initial	*Likkovsky*		
Intervocalic	*Mr Beelik*	Coronal gesture precedes dorsal	[l]
Intervocalic pre-boundary	*beel-ing*		
Word-final prevocalic	*Beel equates*	Dorsal gesture precedes coronal	[ɫ]
Final	*Neal*		

Table 4.2: Adapted from the results of Sproat and Fujimura (1993)

However, they do not accept that such categories are the best way of characterising the variation found, and argue that darkening is gradient and is dependent on duration. Although the relative phasing of the coronal and dorsal gestures seems categorical in nature, they argue it is simply a result of their alignment with adjacent segments, that is, the dorsal gesture is aligned with the vowel. For example, in a word such as *leap* the dorsal gesture comes second, as it is aligned with the following vowel. In *peel*, the dorsal gestures comes first, as it is aligned with the preceding vowel. They argue that, as the dorsal gesture can take more time to articulate because of its larger displacement, in longer rimes it has more chance to reach its extremum. This is why the darkness is correlated with duration: the longer the rime, the darker the /l/.

Lee-Kim et al. (2013) also conclude that /l/ darkness is on a continuum, but reject the claim that duration solely accounts for /l/ realisation. They used ultrasound to study /l/ in three phonological contexts: /l/ as a part of a suffix (e.g. *flaw-less*), before a stem-suffix boundary (e.g. *cool-est*), and pre-consonantally (e.g. *cool headphones*). Using tongue-body lowering as the articulatory correlate of darkness, they found that /l/ is darkest in the pre-consonantal context, intermediately realised before the pre-boundary /l/, and lightest when part of a suffix. They assume that three categories cannot be a possibility, and thus conclude that darkness must be gradient, not categorical. Instead they opt for an explanation of the morphological boundaries directly affecting phonetic implementation, which may result in an indirect side-effect of duration in some contexts. Note that this explanation violates the principle of morphology-free phonetics (Bermúdez-Otero 2010b; 2013), the long-held fundamental assumption that morphology and phonetics do not share an interface. Moreover, it is not clear that we can make judgements of gradience or categoricity, based on so few phonological contexts; rather we need an overview of the entire spectrum of realisation possibilities.

4.2.3 Evidence for both categoricity and gradience

Yuan and Liberman (2009; 2011) also consider duration against darkness in their investigation into the darkness of /l/ in the SCOTUS corpus (Supreme Court Justice of the

United States corpus). This data is taken from 50 years of spoken data from the Supreme Court tapes, giving a total of 21,706 tokens of /l/. Yuan and Liberman (2009) looked at /l/ in three contexts: intervocalically before a stressed syllable (e.g. *believe*, intervocalically before an unstressed syllable (e.g. *helix*, *peel-ing*) and word-finally (e.g. *peel*).[1] They found a correlation between /l/ darkness and duration, but for the latter two contexts only. In this dataset, that was the phonological contexts which were realised as dark [ɫ]s. Their follow-up paper (Yuan and Liberman 2011) considers different kinds of word-final /l/ (as their 2009 grouped in pre-consonantal and prevocalic into the same word-final category) as well as initial /l/ (e.g. *leap*). The results reinforce their original findings: initial and foot-initial tokens (*leap* and *believe*) receive negative D-scores, i.e. lighter /l/s, while the intervocalic and word-final tokens (*helix*, *peel* and *peel bananas*) received positive D-scores, and were therefore dark. Crucially, the correlation between how dark an /l/ was and its duration was found only in the *helix*, *peel* and *peel bananas*-type tokens, i.e. the dark ones.

Note that, for Yuan and Liberman's dataset, *helix* and *peel-ing*-like tokens are generally dark, although the experiment does not distinguish between the two. Intervocalic /l/-darkening is also reflected in Hayes (2000), whose informants accepted a dark /l/ in intervocalic monomorphemes (albeit less than 25% of the time; *yellow* in Table 4.1) and much more so in intervocalic /l/s preceding a stem-suffix boundary (73% of the time; *mail-er* in Table 4.1), showing that darkening has advanced not only from the word level to the stem level, but also up the prosodic hierarchy to anywhere non-initial in the foot in these varieties. Yuan and Liberman's data do not make it possible to distinguish between gradience and variance, however. Although we can see the D-score for the intervocalic tokens is not as high as the pre-consonantal ones (2011:41), whether this is phonetic gradience or category mixture is not clear.

As discussed by Bermúdez-Otero and Trousdale (2012), the most interesting aspect to Yuan and Liberman's work on /l/-darkening is this finding of duration combined with category explaining darkening i.e. a combination of categorical and gradient effects. In fact, as pointed out by Bermúdez-Otero and Trousdale (2012), when inspecting Sproat and Fujimura's (1993:303) own plot of darkening vs. duration, the same seems to be true of their data also. Figure 4.1[2] is a re-construction of the graph taken from the original 1993 paper, with an added ellipsis[3] and correlation line. The plot shows tip delay against duration. In this study, tip delay is taken as the primary articulatory correlate of darkening, due to the relative phasing of the coronal and dorsal gestures, and is plotted on the

[1]Note that Yuan and Liberman (2009) use different terminology to the present study. They refer to *helix*-type tokens as 'intervocalic coda /l/' and *believe*-type tokens as 'intervocalic onset /l/', following a syllabification of intervocalic consonants in the coda of the preceding syllable unless stressed. This is rephrased in the 2011 paper.

[2]The graph was automatically replicated using PlotDigitizer (Huwaldt 2010) by extracting the data points and replotting in `ggplot2` (Wickham 2009).

[3]Plotted with 68% confidence intervals (i.e. two standard deviations from the mean) with the stat-ellipse function in R, available on Josef Fruehwald's GitHub page: https://github.com/JoFrhwld/FAAV/blob/master/r/stat-ellipse.R

y axis. /l/s which have a negative tip delay, i.e. where the coronal gesture precedes the dorsal one, reflect lighter /l/s and are toward the bottom of the plot. Tokens where the dorsal gesture precedes the coronal reflect darker variants and are at the top of the plot. The x axis shows the duration of the rime. Different symbols reflect different phonological contexts, as shown in the legend. The point which Bermúdez-Otero and Trousdale (2012) make is that, although the top half of the panel does seem to show a correlation between darkness and duration, the bottom half is more clustered together and shows no such pattern.[4] Once again, this suggests an in-between situation for the debate around categoricity and gradience: duration does correlate with darkness, but for the dark [ɫ]s only.

Indeed, fitting a linear regression model to the replicated Sproat and Fujimura data gives the suggestion that a model which accounts for both the categorical and gradient aspects of the data gives the best fit. Table 4.3 gives the adjusted r^2 for three models with the dependent variable of tip delay, the articulatory correlate of darkness (the closer to 1 the r^2, the better the fit of the predictors; these values are discussed thoroughly in the following methodology chapter). Model 1 is the Sproat and Fujimura explanation: tip delay can be explained purely by duration, which has a reasonable fit of 0.3418. Model 2 is the Hayes explanation, that category accounts for the main distributional facts, giving a much stronger fit of 0.838. Model 3 is the Lee-Kim et al. explanation: that the morphophonological context is the best way of explaining the variation in tip delay. This provides a better fit than Model 1, with a respectable value of 0.5936, but cannot compete with a simpler model which accounts for category alone. However, the best fit is a model which considers the gradient effect of duration alongside an added predictor of **category**. Model 4 assigns separate categories for tokens with a negative tip delay (i.e. light [l]s) and a positive tip delay (dark [ɫ]) and has an adjusted r^2 of 0.8609. Model 3 is a significant improvement on Model 2, as the ANOVA comparison in Table 4.4 confirms (p = 0.001).

Model	**adjusted** r^2
1. $TipDelay \sim Duration$	0.3418
2. $TipDelay \sim Category$	0.838
3. $TipDelay \sim Context$	0.5936
4. $TipDelay \sim Category + Duration$	0.8609
5. $TipDelay \sim Context + Duration$	0.5985

Table 4.3: Comparison of adjusted r-squared from four linear models

The plot in Figure 4.1 also highlights a deficiency in the Lee-Kim et al. (2013) study mentioned in the previous section. Three separate ellipses with dashed lines are fit to the nearest equivalent phonological contexts from their study using the Sproat and Fujimura data. We can see that, indeed, these do seem to form three separate ranges, but only

[4]The correlation line is fit to all values where the tip delay is positive, and the ellipse is fit to all values where the tip delay is negative.

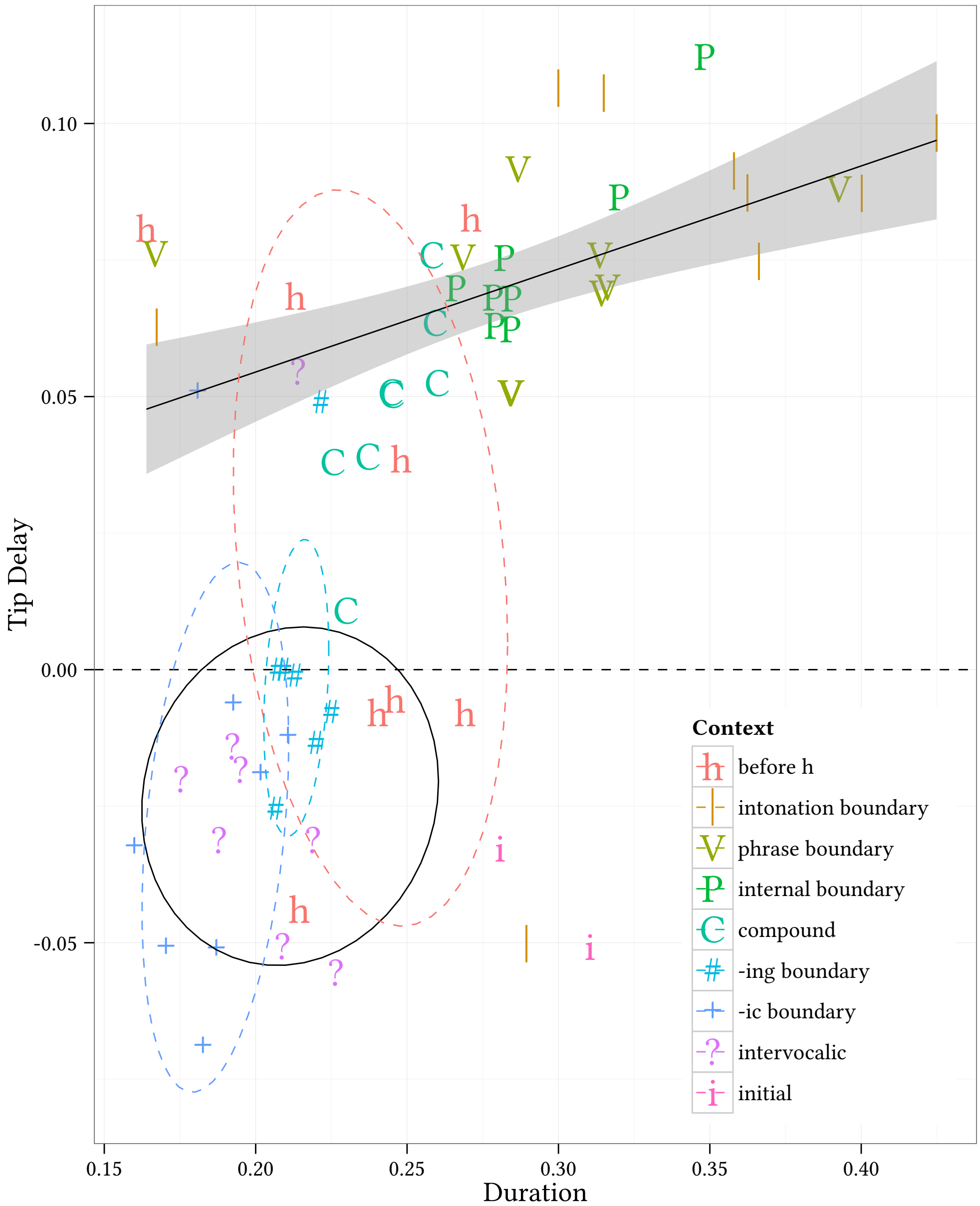

Figure 4.1: Graph adapted from Sproat and Fujimura (1993:303) showing tip delay against rime duration across nine phonological contexts. The longer the tip delay, the darker the /l/ (correlation lines and ellipses not in original plot).

	Res.Df	RSS	Df	Sum of Sq	F	Pr(>F)
1	65.000	0.028				
2	64.000	0.024	1.000	0.004	11.730	0.001

Table 4.4: ANOVA comparison of models 2 and 4 indicating that adding duration gives a better fit to a purely categorical model.

account for a subset of the data, ignoring the rest of the picture. By considering only three phonological contexts, Lee-Kim et al. (2013) have sampled the distribution too coarsely to make claims about categoricity or gradience in their dataset.

4.2.4 Further arguments for categoricity and gradience

The aforementioned studies may indicate that the categorical/gradient argument is a split between a phonological and phonetic approach. However, this is simply not the case, and many phonetic studies of /l/ have either directly called for an allophonic distinction, or such an approach can be inferred from their descriptions and results. In their seminal study, (described in Chapter 2) Giles and Moll (1975) state that the two allophones are distinct, arguing that they should be considered physiologically separate, as they consist of different motor patterns.

It is clear that this idea of having gradient and categorical factors of the same phonological process is not new, and for some phonological processes these combinatory factor are discussed in a non-controversial manner. In fact, although both effects have not widely been accepted in studies of /l/-darkening, such an approach to the facts seems to be well documented in studies in discussions of /l/-vocalisation (Kerswill and Wright 1990; Scobbie and Wrench 2003; Scobbie and Pouplier 2010; Wrench and Scobbie 2003; Wright 1989).

Kerswill and Wright (1990:261) discuss ideas of categoricity and gradience, drawing on data from other studies of connected speech processes (Kerswill 1985; Nolan and Kerswill 1990; Wright 1989; Wright and Kerswill 1989). This includes /l/-vocalisation, which they describe as tending towards articulatory discreteness. They also note that Wright's (1989) EPG work shows that vocalisation is also sensitive to speech rate. However, Wright (1989:358) discusses its relative resilience to speech rate for the younger speakers only, suggesting that for this social group /l/-vocalisation is a fully categorical process, whereas for the older speakers it may not have completely stabilised, still acting as the phonologised predecessor which is more susceptible to rapid rates of speech.

Vocalisation seems to be less robust and convincingly categorical not only in older speakers, but also more standard speakers. Sproat and Fujimura (1993:292) cite Hardcastle and Barry's (1989) study of vocalisation as providing evidence for a gradient interpretation of /l/ lenition processes. However, as mentioned in Chapter 2, the authors make no such claims and admit that their speakers have educated accents and have lost any significant local features. Although their data do show evidence of gradient vocalisation, this is most likely due to the fact that they speakers are not 'proper' vocalisers as they are speaking standard English. Thus, a gradient interpretation of Hardcastle and Barry's (1989) may be true of this study, but it does not follow that /l/-vocalisation as a categorical phonological process does not exist.

Wrench and Scobbie (2003) find evidence for both gradience and categoricity in their combined EPG, EMA and ultrasound study of /l/-vocalisation. They interpret their EPG results in a variable categorical manner, defining lack of tongue tip contact as a vocalised /l/, and any tongue tip contact at all is categorised as some kind of /l/. They discuss the difference between gestural targets for speakers, and explain that, within the general /l/-like category, the difference can be as large as those between the /l/-like and vocalised category. However, they do discuss gradient effects with the results from their EMA data, through the mean tongue tip height in speakers (a higher tongue tip indicates a lighter /l/). They find a combination of effects between EMA and EPG but overall, their results show that a speaker either has tongue tip contact, or does not, and on top of that there is gradience with respect to the strength and consistency of contact. They suggest that more extreme types of vocalisation may benefit from a categorical analysis, whereas less extreme cases of vocalisation can be analysed from the perspective of gradient weakening.

Scobbie and Pouplier (2010) also find evidence for categorical vocalisation in their EPG study of vocalisation, and support this even for varieties which have a very small distinction between onset and coda /l/. Although they do not subscribe to the traditional allophonic interpretation favoured by many phonologists, they reject Sproat and Fujimura's (1993) purely gradient interpretation, instead settling on a gestural account. Their approach is consistent with the difference between the phonologisation and stabilisation stages of the life cycle. Vocalisation starts off as a gradient process, whereby the dark [ɫ] loses its tongue tip contact due to rapid speech, following low vowels, or due to the potential motor influences of being phrase final. Whatever the reason, this results in a further lenited /l/, which may or may not become stabilised as a phonological rule at the phrase level. Whether it does is down to sociolinguistic effects, invoking the actuation problem (Weinreich et al. 1968), and this is why we see differences between dialects like Cockney and RP. Nevertheless, it gives us the distinction between those who have a cognitively controlled /l/-vocalisation process, such as the extreme vocalisers in Wrench and Scobbie (2003), and those who merely have some form of gradient weakening in the /l/, which the majority of us may display to differing extents, such as those described by Hardcastle and Barry (1989). In terms of the life cycle, the former would be after stabilisation, whereas the latter would be after phonologisation, with stabilisation yet to occur.

To summarise thus far, the existing literature provides examples of analyses which tend toward one of two ways of looking at /l/-darkening: one approach underplays the role of categorical phonological variation and the other overestimates it. However, there are compelling arguments for both sides. Additional data from phonetic analyses indicates that a combination of both effects are at play in the same grammar. Is there is a way of keeping the positive aspects of these analyses whilst overcoming their deficiencies i.e. is there an approach which can account for both categorical and gradient effects?

4.3 A modular approach

In a modular approach, categorical phonology feeds gradient phonetics. The phonology computes generalisations over discrete features, whereas the phonetics assigns targets to surface feature configurations. The phonology itself is stratified with three levels: stem level, word level, and phrase level. This follows from the ideas behind the life cycle of phonological processes as discussed in full in the previous chapter (Bermúdez-Otero 1999; 2011; Bermúdez-Otero and Trousdale 2012). To recap, phonetically driven innovations enter the grammar as gradient phonetic rules. Later, they may become stabilised as categorical phonological processes at the phrase level. Analogical change results in the new phonological process climbing up to higher levels, which narrows the domain of application of the process. This explains why word-final prevocalic /l/s darken in Sproat and Fujimura's data in phrases such as *Beel equates the actors.* One might expect /l/ to resyllabify into the following syllable and remain light, going on to conclude that /l/-darkening shows opaque overapplication in this environment because of the dark [ɫ]. However, a modular approach can explain this opacity through word-level darkening. For these speakers of American English, darkening has moved up to the word level, and /l/ darkens prior to resyllabification. We see this rising to the stem-level for Yuan and Liberman's speakers, who have dark [ɫ] in *pee**l**-ing*-type words. There is also evidence of rule generalisation on the prosodic level in some cases for both Yuan and Liberman's speakers and Hayes's (2000) informants, who show some form of darkening in intervocalic monomorphemes such as *he**l**ix*, demonstrating that darkening has advanced from the coda and applies to any position outside foot-initial onsets.

The life cycle also makes predictions about the relative position of diachronically related rules coexisting in the same synchronic system. As demonstrated in the previous chapter, the life cycle does not just involve the stages of phonologisation, stabilisation and domain narrowing, but also includes rule scattering (Bermúdez-Otero 2013). As discussed, when a new rule enters a higher component of the grammar, it does not completely stop its application at the lower level. In other words, innovative phonological rules do not replace the phonetic rules from which they emerge, but typically coexist with them. The example given in the previous chapter regarding /s+j/ palatalisation demonstrated that present-day speakers of English can show both categorical and gradient effects of the same phonological process and, crucially, the idea behind rule scattering is that the two exist in the same grammar.

Similarly, the process of /l/-darkening may show the cumulative effects of several cognate processes simultaneously overlaid in the grammar. Sproat and Fujimura's data provide evidence for rule scattering, as there is a categorical effect of positive and negative tip delay, and within that the dark [ɫ]s show a strong gradual effect of duration and the light [l]s do not. This effect in the data supports the idea that the process of darkening originated as a gradient phonetic process which was sensitive to duration, but over time has been reanalysed by learners and phonologised as a duration-insensitive categorical process. This means that the original duration-sensitive gradient process of phonetic implementation coexists in the grammar on top of the newer morphosyntactically conditioned categorical process.

4.4 Empirical diagnostics for categoricity

We have seen thus far that existing studies which find evidence for gradience in the results are quick to dismiss the possibility of categoricity. Although in some cases we may not be able to tell for absolute certain whether a phonological process is categorically conditioned in the speaker's brain, we can outline some diagnostics which we would hope to find in the empirical data from which categoricity can be inferred. As Scobbie (2005:2) deliberates, "What counts as phonological data? What gets into surface structure in the first place?" It is such questions we need to consider when judging the available evidence.

Previous work has shown a variety of ways of distinguishing between processes under phonetic or phonological control, whether they directly use it for that purpose or not. As we have seen above, assimilatory processes are a good example of this. In Ellis and Hardcastle (2002), the assimilated [ŋ]s are compared to the [ŋ]s that are underlyingly /ŋ/, to see if the two are different. If they are the same, we can conclude categorical assimilation by feature spreading, and if they are slightly different then we can conclude the assimilation is gradient and phonetic.

However, this kind of test is not especially relevant here, as /l/-darkening is not a form of assimilation, but lenition. It cannot be compared to any underlying canonical dark token, but must be compared to the canonical onset token, which would represent the lightest /l/. This makes distinguishing categoricity and gradience trickier, as to what phonetic extent does something have to be different in order for us to call it a separate category? Also, do we want to take articulatory magnitude as evidence of separate categories? Perhaps we could take this in one direction: a consistently large phonetic difference, in terms of articulatory magnitude or acoustic range, is likely evidence of two allophones, but not vice versa. That is, an absence of a large phonetic difference does not necessarily entail an absence of categoricity. We could compromise that categoricity could also be diagnosed on phonetic grounds in a more qualitative manner, by noting consistent patterns in articulation. For example, if English /l/ exhibited a categorical allophonic distribution, we would expect the light tokens to show a consistently similar articulation patterns of fronted tongue body and advanced tongue, whilst the dark tokens would show a retracted tongue root, possible velarisation, and a reduced tongue tip gesture. We may also expect articulatory discontinuity between the two sets of tokens.

If we find consistent differences, but for a context which is perhaps more susceptible to phonetic conditioning than phonological, then we may conclude that this effect is phonologised and consistently gradient. For example, if just one coda /l/, say phrase-finally, is showing darkened patterns but the other coda /l/s are not, this might be more consistent with a phonetically-defined utterance-final duration-driven definition of gradience. The process here is not conditioned by the phonologically defined syllable, but perhaps by a phrase-final lengthening effect, or something similar. We cannot tell for sure which patterns are cognitively controlled categorical phonological processes in the speaker's grammar, but we can tally up the evidence from phonetic data for and against such categorisations, and make intelligent judgement based on these.

More frequently it is agreed that the best evidence for determining whether a sound pattern is categorical comes for researchers who look at the phonetic facts, and conduct a thorough analysis of the statistical distribution from a quantitative perspective (Bermúdez-Otero and Trousdale 2012; Scobbie 2005). A phonetically gradient pattern may be indicated by a continuous quantitative distribution, whilst a phonological may show some discontinuity or cut-off between the two categories. Bermúdez-Otero and Trousdale (2012:7) argue that bimodality in the distribution provides compelling evidence of two categories. They cite work by Maye et al. (2002) which shows that babies exposed to speech sounds from a phonetic continuum could only differentiate between the two endpoints if the distribution was bimodal. Those exposed to a unimodal distribution could not differentiate between the two ends of the phonetic spectrum.

Some have pointed out the drawbacks of using bimodality as a diagnostic of categoricity. Bermúdez-Otero and Trousdale (2012:7) make the point mentioned above, that the bimodality criterion only works in one direction: a lack of bimodality does not confirm a lack of categoricity. A mixture of two equal normal densities results in bimodality if and only if the difference between the means is greater than twice the standard deviation (Schilling et al. 2002). Another aspect of speech cited as problematic for bimodality includes articulatory gradience being realised as abrupt in the acoustics (Browman 1995). Also this could in time result in categoricity through the listener as the source of sound change, it does not reflect categoricity in the speaker. This is not an issue for the present dissertation, however, which will apply bimodality tests to both articulatory and acoustic data (see the next chapter for details).

However, Yuan and Liberman (2009; 2011) provide an additional problem for bimodality if the bimodal distribution of positive and negative D-scores is overlaid with a gradient phonetic effect of duration, which may hinder the pattern from a statistical point of view. A way of testing the likelihood of categoricity accounting for both would be the kinds of linear models used above on the Sproat and Fujimura (1993) data. If category provides a good fit, either alone or alongside an additional gradient effect, it provides good evidence for categoricity.

Based on these observations, I tentatively suggest the following diagnostics for categoricity:

1. Articulatory discontinuity between two (or more) sets of splines.
2. Articulatory consistency within these sets.
3. Bimodality in the quantitative analysis of the articulatory data.
4. Bimodality in the acoustic data (keeping in mind this may not reflect the articulatory data).
5. A linear model with articulatory darkness as the dependent variable that provides a good fit for two categories.

It must be kept in mind that lack of evidence for categoricity does not necessary mean the pattern is gradient. There may be other elements to a categorical distribution which cannot be observed from the data. Although it may be difficult to know for sure, these empirical measures can be taken as cognitive categoricity on some level of the grammar.

4.5 Summary

This chapter has demonstrated that previous studies have reduced the variation in /l/-darkening to a false dichotomy between purely categorical and purely gradient effects, when all the evidence points to both operating within the grammar. This not only comes from descriptive suggestions in the phonetic literature, but also actual data from perhaps the most famous study of /l/-darkening. Fitting basic linear models to existing data shows that a model which accounts for both effects wins out.

In fact, studies of /l/-vocalisation widely discuss both effects of category, and the overlaid gradient phonetic effects, and this does not seem to be an issue for this more advanced stage of the lenition trajectory. In a study mentioned above, it could be the case that a dark [ɫ] is less 'salient' than a vocalised one. Salience i.e. something which speakers (or researchers) notice, could be the reason researchers seem less keen to ascribe categories for something they cannot hear reliably. Moreover, /l/-darkening is less clearly and dependently defined phonetically speaking. The allophonic darker variant is defined in terms of the initial variant: dark [ɫ] is more lenited, less consonantal than the light one, and has a more retracted tongue body. /l/-vocalisation, on the other hand, is clearly defined with no reference to other allophones in the speaker's grammar: it has no tongue tip gesture or lack of contact. This is an easy distinction for studies such as the EPG ones listed earlier, as if there is no alveolar contact, it can be noted as a vocalised /l/. For darkening, however, the delayed gesture is much more tricky to categorise.

The fact that /l/-vocalisation is a further stage of the lenition trajectory, means it is more advanced on the phonetic continuum. /l/-vocalisation is articulatorily distinct, in that it can be defined as total absence of palatal contact, and auditorily distinct, in that you can more or less hear it reliably in speech (see Hall-Lew and Fix 2012). Therefore, researchers are happy to consider it as a separate category, whilst non-controversially observing the extra phonetic effects overlaid. /l/-darkening's place, on the other hand, is less safe. Phonologists who see the allophonic distinction as necessary are reluctant to discuss the gradient effects as this does not fit into to their purely categorical model. Phoneticians see that dark [ɫ]s can have a range of darkness, and are reluctant or oblivious to note any evidence of bimodality in the patterns they find. Moreover, some may simply not have the data to be able to observe a categorical distinction in the first place, having sampled the distribution too coarsely.

This raises interesting questions for phonetics-phonology interactions. We are comfortable in assigning separate categories if the phonetics are suitably distinct. Do we want to distinguish between categorical and gradient processes by saying the former have a large phonetic distinction and the latter do not? In this chapter I have presented several possible diagnostics for categoricity, which will be detailed in a more methodological manner in the following chapter. I argue in this chapter, as many before have done, that although lack of categorical evidence from qualitative and quantitative approaches to the articulatory data may not be available, this does not necessarily entail phonetic gradience. Although we may never be sure of the exact nature of the speakers's grammar, the criteria listed in the previous section are a reasonable compromise.

The next chapter will outline the methodology used to investigate the proposals made in this chapter, alongside the predictions made in Chapter 3. This chapter will give an overview of the experimental procedure used in ultrasound tongue imaging, in particular to investigate /l/-darkening and vocalisation. Statistical methods for deciding between categorical and gradient processes are outlined, including linear models similar to ones used above, in addition to methods for quantifying articulatory data.

Chapter 5

Methodology

This chapter describes the methodology used in the experimental part of the thesis, the results of which are presented in Chapters 6 and 7. The present chapter is organised as follows. Section 5.1 recaps the ideas discussed in the previous two chapters which form the primary research goals of the thesis. Section 5.2 gives an overview of the experimental procedure, including the data collection process and equipment used (Sections 5.2.1 and 5.2.2); recording stimuli (Section 5.2.3), and information about the informants (Section 5.2.4). Thereafter, I discuss the coding and analysis procedures for the articulatory and acoustic analysis in Section 5.3 and Section 5.4 respectively, and the quantitative approaches used in Section 5.5, including Smoothing Spline ANOVA and Principal Components Analysis.

5.1 Research Questions

The goal of the experimental ultrasound tongue imaging work reported in this thesis is to address two questions in the phonological literature. The first is to look into reported phonological patterns of morphosyntactic sensitivity in /l/-darkening processes. To this end, ultrasound is used to compare /l/ realisations in different morphosyntactic positions to investigate the possible typology of /l/-darkening systems. The second question is to investigate whether allophonic realisations of /l/ are categorical, continuous, or whether we can find overlaid effects of both. This will be investigated by analysing /l/s across a broad range of phonological environments and correlating the findings with additional phonetic factors such as duration and acoustics.

Research questions

1. Do we find articulatory evidence for /l/-darkening operating at different levels of morphosyntactic sensitivity in different varieties of English?

   This question is investigated with ultrasound data from five speakers in Chapter 6

2. Is /l/-darkening categorical or gradient, or do we find evidence of both?

This question is investigated with ultrasound data from eight speakers in Chapter 7, using the following diagnostics for categoricity as discussed in Chapter 4:

i. Articulatory discontinuity between two (or more) sets of splines.
ii. Articulatory consistency within these sets.
iii. Bimodality in the quantitative analysis of the articulatory data.
iv. Bimodality in the acoustic data (keeping in mind this may not reflect the articulatory data).
v. A linear model with articulatory darkness as the dependent variable that provides a good fit for two categories.

The methods used to investigate these aspects are outlined below in Section 5.5.

Peripheral research questions In addition to the main research questions above, there are several secondary goals to the thesis, which can be investigated along the way.

1. Is the relative phasing of coronal and dorsal gestures the primary correlate of /l/ darkness, as suggested from Sproat and Fujimura's (1993) results? Can we gain insights to the realisation from one point in the articulation? Sproat and Fujimura (1993) also cite the displacement of the dorsal gesture as a primary correlate of darkness, so does this kind of analysis only suffice for our speakers?

2. What kind of secondary articulation strategies do speakers employ when producing a dark [ɫ]? Do we see any kind of conditioning for velarised and pharyngealised /l/, and does this vary within or across speakers?

3. Do the acoustic measures of F2-F1 suffice for this kind of data? How much of the articulatory information can be accounted for in the acoustics?

There are further peripheral aims of the thesis, which the experimental procedure will cover in due course. These largely concern an overview of dialectal variation in /l/-darkening, reporting on the systems of different varieties of English from an articulatory perspective. Not only will this give an idea of a wider range of possible /l/-darkening typologies, but also will provide a basis for future studies in sociophonetics and language variation and change.

5.2 Experimental Procedure

5.2.1 Recruiting participants

14 native speakers of English participated in the experiments, recorded in the University of Manchester's Phonetic Laboratory. Recordings took place over several months, although a recording for a particular speaker was always completed in one session. There

Figure 5.1: The Mindray DP-2200 ultrasound machine

are two speakers who took part in both experiments (see Section 5.2.4) the data for which were collected separately.

All speakers were unaware of the purpose of the study. None reported any speech or hearing pathologies. Participants were recruited by word of mouth, and by the University of Manchester's research volunteering website. Participants were rewarded £15 for carrying out the experiment. Informed consent was sought in all stages of the research reported on in this thesis (see Appendix A).

Three of the recordings for this project are not used in the thesis. In one case, this was through a microphone fault, in another case an informant reported feeling uncomfortable and wanted to stop the recording midway through. In the other case, the speaker did not image well enough to form a consistently good ultrasound picture across words and contexts.

5.2.2 Data collection

Data was collected by digital video recording of ultrasound tongue imaging, using the software program Articulate Assistant Advanced (henceforth AAA). The machine used was a Mindray DP2200, pictured in Figure 5.1. This machine was fitted with an electronic endocavity transducer (type 65EC10EB), set at a frequency of 5MHz. The transducer scanned 30 frames per second across a 120 degrees field of view, but was deinterlaced to 60fps. An Audiotechnica ATM10a microphone recorded the acoustic signal simultaneously. The audio and video were synchronised with the aid of a SyncBrightUp unit (Articulate Instruments Ltd 2010).

5.2.2.1 Probe stabilisation

The probe stabilisation headset (see Figure 5.2), designed by Articulate Instruments, was created to allow comparison of tongue imaging across the entire recording of a speaker without hindering speech production. It has been tested at Queen Margaret University for reliability (Scobbie et al. 2008). The headset holds the transducer in a fixed position under the chin, perpendicular to the jaw, and has the additional benefit of allowing the transducer to move with the jaw. This prevents the restriction of natural movement during speech, which some previous techniques for stabilisation were found to do (Davidson 2012). The headset is lightweight and portable, and has adjustable parts enabling it to be securely fit to the speaker's head, minimising lateral movement of the transducer and preventing rotation, while still allowing participants freedom of movement to speak. The positioning also prevents the beam reflected from the ultrasound from becoming refracted too much by the jaw or hyoid bone, both of which can create an acoustic shadow on the ultrasound image (Stone 2005).

The probe stabilisation headset also allows for more reliable tracing of the palate, which can be identified by having the speaker swallow some water before beginning the word-list. Once the palate is traced, the headset ensures that the positioning will not move, so the same trace can be kept for the whole analysis of that speaker.

Figure 5.2: Subject wearing probe stabilisation headset

5.2.3 Target Stimuli

As Figure 5.3 shows, target stimuli appeared on the screen in front of the informant, and they were instructed to read once the prompt background had turned green (meaning the audio and video were recording). Words and phrases were repeated a further

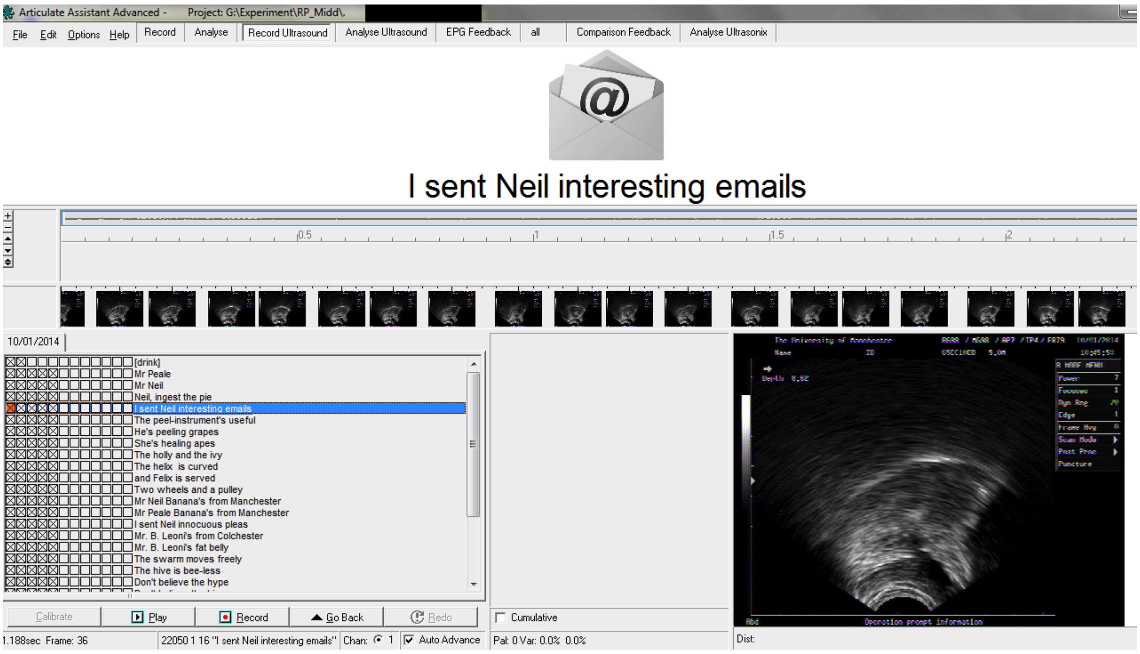

Figure 5.3: The AAA prompt interface. The prompts appear and participants read each phrase in turn. Audio and video are simultaneously recorded.

Experiment 1

Context	**Sentence**
1. word-initial	**_l_**_eap_
2. stem-medial posttonic	_he_**_l_**_ix_
3. stem-final presuffixal	_hea_**_l_**_-ing_
4. word-final pre-vocalic	_hea_**_l_** _it_
5. word-final pre-consonantal	_hea_**_l_**_C
6. utterance final	_hea_**_l_**

Table 5.1: The five/six phonological environments for Experiment 1 with example tokens

four times, resulting in five recordings per prompt. The target stimuli varied depending on whether the speaker was taking part in Experiment 1 or 2 (two speakers participated in both).

As discussed in Chapter 2, previous studies have shown that /l/ can be subject to a vast degree of coarticulation depending on its neighbouring vowels (Bladon and Al-Bamerni 1976; Lehiste 1964; Nolan 1983). For this reason it was important to keep the adjacent vowels the same, and high front vowels appear in all contexts, as found in Sproat and Fujimura (1993).

5.2.3.1 Experiment 1 stimuli

The example sentences for Experiment 1 are shown in Table 5.1 (full sentences are given in Appendix B). Each of the words or phrases in Table 5.1 was presented to the speaker, one at a time, interspersed with additional distracter words and phrases. For the first two speakers recruited in the study, the preconsonantal context was absent, as this was not identified as part of the study at the time. The sentences were presented in the same order each time, giving between 25-30 /l/s per speaker (as each was elicited five times) in addition to the distracters. Ethical considerations are important here, as the helmet can become extremely uncomfortable if worn for particularly long periods. However, it is important to get the recordings done in one session in order to have the tongue images be comparable to one another.

Experiment 1 was designed to elicit /l/s in different morphosyntatic domains, as well as analyse /l/ tokens of word-final prevocalic /l/ which were likely to be resyllabified (such as the ***heal*** *it*-type tokens). *A priori*, one might expect the relative linguistic strengths in the boundaries to be reflected in the ordering of tokens in Table 5.1.

5.2.3.2 Experiment 2 stimuli

As discussed previously, the primary aim of Experiment 2 is to address the issue of categoricity and gradience in /l/-darkening systems. Although Experiment 1 provides an overview of interactions between the morphosyntax and phonology, it may not give as clear a picture as we need for dialects which show less convincing evidence of categoricity through tongue splines. For this reason, we need to be able to quantify results from a large range of closely related phonological environments. In addition, the data need to hold up to the statistical tests described in 5.5, including interactions with duration, and linear regression.

Experiment 2 was designed for slightly different purposes than Experiment 1. Firstly, the 25-30 tokens of /l/ collected in Experiment 1 would be sufficient to analyse differences in different morphosyntactic domains, but not to conduct a thorough overview of the kinds of phonological environments which make up the full spectrum of /l/ realisations. Statistical analysis is preferable for such investigation, and more tokens are needed for this. Table 5.2 gives the environments collected for Experiment 2. Not only do we have 10 environments for this experiment, there were also two phrases included per environment (see Appendix B). At five repetitions per speaker, this resulted in 100 /l/s each, where possible.[1] Single word tokens were avoided for Experiment 2, as there were some issues with speakers drawing out the articulation of phrase-initial /l/s in Experiment 1.

An additional consideration of Experiment 2 was to attempt to avoid ambiguously light or dark tokens. That is, one of the prime targets of Experiment 1, collecting resyllabified /l/s to monitor their behaviour in this context, were specifically avoided in Experiment 2. This means that tokens which were word-final and prevocalic were placed

[1]For various reasons, 100 tokens could not be extracted for some speakers. This was usually down to the ultrasound image quality being particularly poor during the main frame of the /l/. In other cases, the speaker did not produce the sentence correctly. When this was noticed during the experiment, participants were asked to revisit the offending utterance at the end of the process, and repeat it once more. As can be seen from the figures in B, in some cases the original utterance was perfectly fine, and this results in an extra spline for this context. In other cases, the repetition per sentence may not reach 5, for the reasons just mentioned.

Experiment 2

Context	Example
1. word-initial	***l**eap*
2. stem-medial pretonic	*be**l**ieve*
3. suffi-initial	*free-**l**y*
4. intervocalic	*he**l**ix*
5. stem-final presuffixal	*pee**l**-ing*
6. compound boundary	*pee**l**-instrument*
7. word-final phrase-medial	*hea**l**#* V
8. phrase-final	*hea**l**#*, V
9. utterance final	*pee**l***
10. word-final pre-consonantal	*pee**l** bananas*

Table 5.2: The ten phonological environments studied for Experiment 2 with example tokens

before prosodic items such a verb phrases. In the sentence *I sent Neil interesting emails*, resyllabification is avoided by ensuring the /l/ is not part of the following prosodic unit. By doing this, the focus of the study remained firmly on finding a more consistent cross-speaker analysis of the darkest and lightest tokens, creating a full spectrum of comparable realisations. Again, *a priori*, one might expect the relative linguistic strengths in the boundaries to be reflected in the ordering of tokens in Table 5.2, although it is not clear where *pee**l** bananas*-type tokens might fall. They would not be expected to be as dark as phrase-final position, given durational effects in this position, but its darkness in relation to context 8 is unpredictable. The pause in the *hea**l**#]V*-type tokens might result in a darker /l/ through durational effects, or the following consonant might introduce more lenition through gestural overlap.

5.2.4 Participants

In this thesis, informants are referred to by their variety of English, usually ascribed by geographical location (with the exception of the location-less standard of RP). However, it is important to keep in mind that this is not a claim that patterns found are representative of the variety as a whole, as they are mostly based on just one speaker. It is entirely possible that speakers may show dialectal patterns subject to ongoing community change, variation based on sociolinguistic factors, or even idiolectal differences. Nevertheless, it is still insightful and interesting to note similarities and differences in general when comparing these speakers with descriptions from the existing literature.

All speakers were between 18 and 30 years old. The places where they are from in the UK are shown in colour on the map in Figure 5.4.[2] When more than one speaker from a particular variety is included, additional social information is used to differentiate

[2]The American speaker is not shown as this is to give an idea of whereabouts the regional varieties in the UK are with respect to one another. The RP speaker is not shown as the standard is meant to be 'regionless'.

them, such as gender or social class. Most speakers took part in either Experiment 1 or Experiment 2, but two took part in both (RP and Manchester WC). Around 4 months passed between the recording of Experiment 1 and Experiment 2. Therefore the two recordings cannot be cross-compared, as the probe would not have been in the exact same position.

5.2.4.1 Participants taking part in both experiments

5.2.4.1.1 RP

This informant is male and speaks a conservative form of RP, despite his age of 28 at the time of the experiment. He grew up in Yorkshire, but has attended private school since the age of six, which is where he acquired his more standard accent. Not only does he display more typical features of RP, such as a BATH/TRAP split, he also realises weak suffixal vowels as [ɪ] rather than [ə] during the experiment, which is a particularly conservative RP realisation (Fabricius 2002).

5.2.4.1.2 Manchester WC

This informant is female, and from a working-class area of East Manchester, in the North-West of England. She was 25 years old at the time of the experiment. She displays many features of WC Manchester English, such as *happ*Y-laxing phrase-finally (Turton and Ramsammy 2012), absence of GOAT-fronting (Baranowski 2013) and many more features discussed by Baranowski and Turton (Forthcoming).

5.2.4.2 Experiment 1

5.2.4.2.1 Middlesbrough

This informant is male and is from a working-class area of Middlesbrough, in the North-East of England. He was 24 years old at the time of the experiment.

5.2.4.2.2 Essex

This informant is female, and is from Colchester in Essex, in the South-East of England. She was 21 years old at the time of the experiment. She displays many features of Essex English, such as /l/-vocalisation (which is why she was approached to take part in the experiment, although she was not aware of this factor in her selection).

5.2.4.2.3 American English

This informant is a female speaker of General American English, and was 28 years old at the time of the experiment. She is from Texas, but does not show any typical features of a Texan accent, possibly because her parents are from California and she has spent much of her adult life living there.

Figure 5.4: A map of the UK highlighting where the participants are from.

5.2.4.3 Experiment 2

5.2.4.3.1 London Female

This informant is a female from East London, and was 19 years old at the time of the experiment. She was originally born in Jamaica and moved to London at 3 years old. She speaks English at home. She has some evidence of Multi-cultural London English features in her speech, such as *lett*ER backing and lowering, but this is not as auditorily salient as the typical speakers reported in Cheshire et al. (2011).

5.2.4.3.2 London Male

This informant is male, is from South-West London, and was 29 years old at the time of the experiment.

5.2.4.3.3 Manchester MC

This informant is female and is from Chorlton in South Manchester, North-West of England. She was 21 years old at the time of the experiment. She displays many features of MC Manchester English, such as lack of *happ*Y-laxing phrase-finally (Turton and Ramsammy 2012), GOAT-fronting (Baranowski 2013) and many more features as discussed by Baranowski and Turton (Forthcoming).

5.2.4.3.4 Belfast

This informant is a 19 year old male from Belfast in Northern Ireland.

5.2.4.3.5 Liverpool

This informant is male, is from Liverpool in the North-West of England, and was 30 years old at the time of the experiment.

5.3 Data analysis

After the data was collected, acoustic analysis was carried out in Praat, followed by ultrasound coding in AAA, and further analysis in R. Figure 5.5 shows the kind of data under analysis here. In all ultrasound frames, and all plots in the thesis, the tongue blade appears on the right edge of the screen, and the tongue root on the left.

5.3.1 Data labelling and coding

In order to keep the analysis process as efficient as possible, as well as allowing for the automatic collection of formant data, sound files were exported from AAA into Praat (Boersma and Weenink 2010) first and foremost. The sound files were annotated by

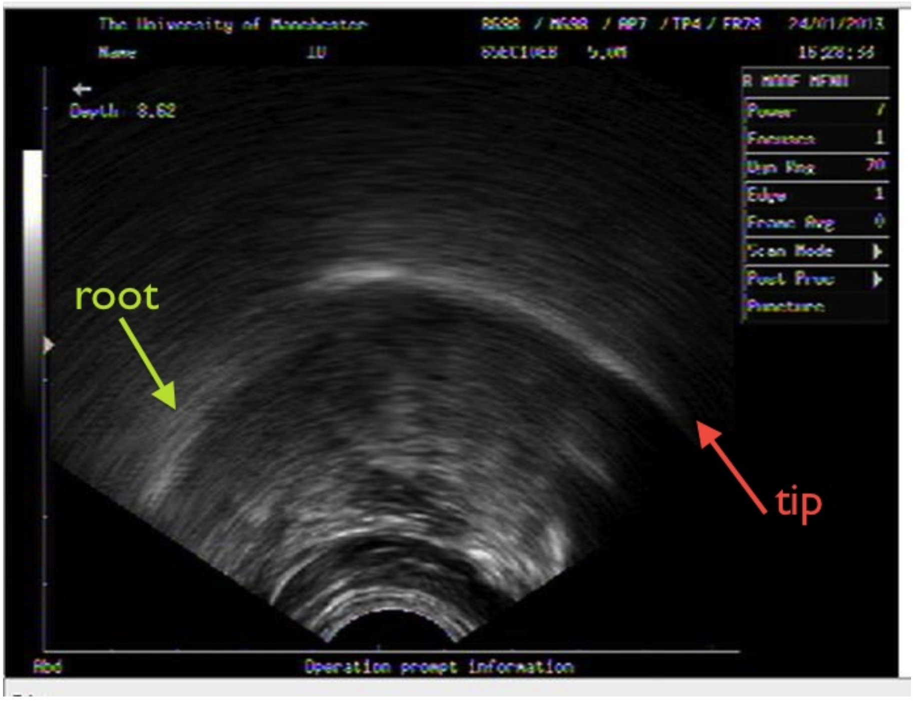

Figure 5.5: An ultrasound image of a tongue. The right side is the tongue tip and the left side if the tongue root (N.B. The tongue tip itself is not visible on the vast majority of images, but the terminology reflects a) the orientation of the tongue on the ultrasound image on and b) the estimated position of where the tip would be).

hand, identifying the /l/ boundaries and the boundaries of the preceding and following vowels. The acoustic identification of segments is discussed in Section 5.4. These sound files were then imported back into AAA so that all prompts in the ultrasound recordings were acoustically labelled.

Once the acoustic labelling was was fully imported and the /l/ intervals identified in AAA, the manual fitting of 'splines' to the tongue shapes in each ultrasound frame commenced, a feature which is built in to AAA. The definition of a spline in the Oxford English Dictionary is *a continuous curve constructed so as to pass through a given set of points*. In AAA there are 42 such points through which the splines passes, one for each radial axis of the grid. Each point is illustrated by a small orange cross mark, as shown in Figure 5.6. AAA has an in-built feature which draws a smooth line through the estimated area (aided by the user's inclusion of a maximum and minimum tongue position from the speaker template, discussed below). For very clear images, this works well, but less so for fuzzy ones.

It is useful to draw the palate trace prior to fitting tongue splines, which is done by viewing the video of the participant dry swallowing or swallowing water, as outlined in Section 5.2.2.1. Drawing the palate enables visualisation of the tongue position in reference to the highest possible point of the tongue. This is a useful point of reference, as there will be large differences between speakers due to both physiological reasons, as well as the positioning of the probe in a particular recording. AAA allows the creation of

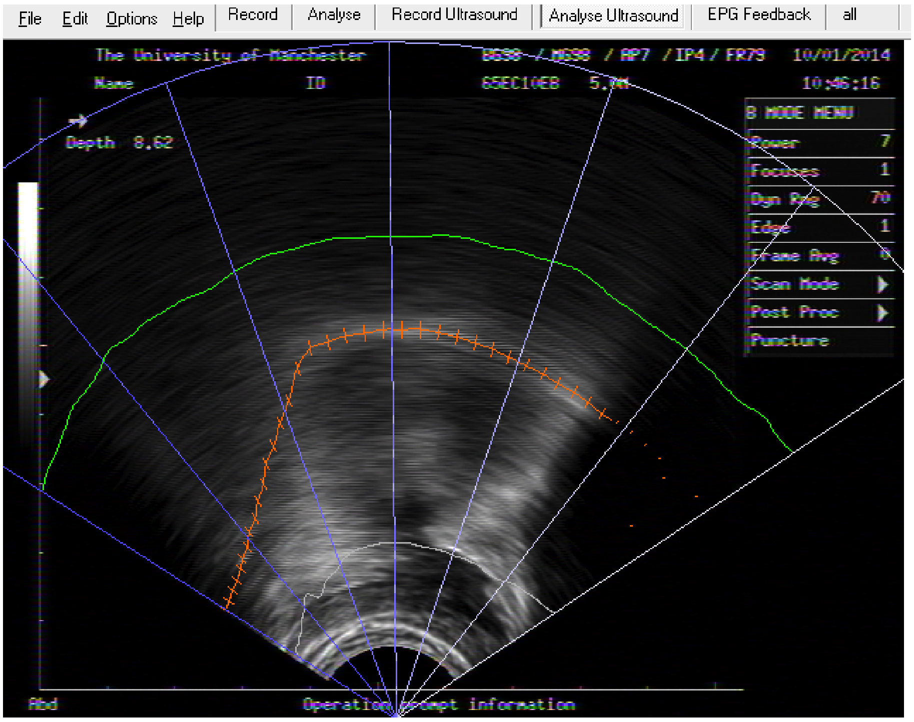

Figure 5.6: Spline fitting in the AAA prompt interface. The program suggests a spline of best fit which can be corrected by the user. As can be seen above, the image can often be somewhat fuzzy.

a template for each speaker, which saves the palate, a minimum tongue position, as well as an average tongue position which can be moved for each ultrasound frame. Once the speaker-specific template is added to the ultrasound image, the suggested spline should be more accurate, as refracted white pixels beyond the palate or below the minimum tongue will be ignored by AAA. The suggested spline can then be corrected or redrawn to fit the image. Usually images are only accurately fitted by AAA for part of the tongue shape, so corrections were needed throughout the analysis process. This can be done by hand with a mouse or trackpad, but a Wacom Bamboo pen graphics tablet was used to speed up the process.

As Figure 5.6 indicates, often the ultrasound image can be fuzzy in parts. In this particular image, the tongue root area is poorly represented. In such cases, the coder's best judgement is used to decide where that part of the tongue lies, usually in comparison with preceding and following frames, and with the recommendation of the automatic spline fitting tool. In cases where the position could not be confirmed with a high level of confidence, the image was removed from the analysis. In cases where just part of the tongue could not be confirmed with a high level of confidence, this area was not represented with a spline contour. Compare the medium-quality image in Figure 5.6 with the two contrasting images in Figure 5.7 and Figure 5.8. Figure 5.7 was taken from a speaker who produced a consistently clear image throughout the experiment period. Others, such as the speaker shown in Figure 5.8, do not image nearly as well. Images like that in Figure 5.8, for the most part, usually occur sporadically during the experimental procedure and are not necessarily reflective of every frame. However, it is certainly the case that some speakers image more clearly than others, for reasons which are not always apparent. The data from the speaker in Figure 5.8 was not used in this experiment, as the tongue trace was just too unreliable in the majority of frames.

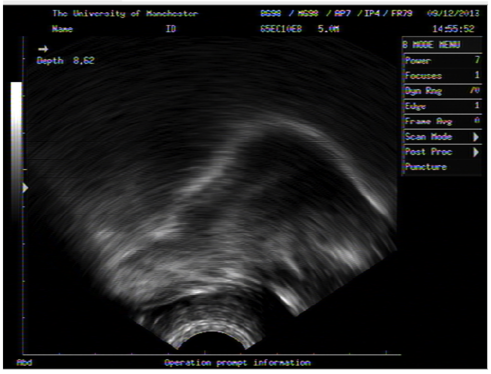

Figure 5.7: An example of a very clear image of the tongue.

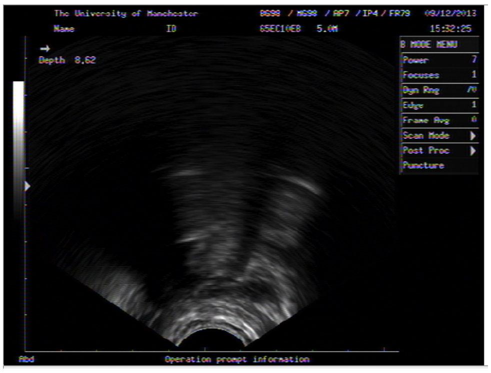

Figure 5.8: A particularly unclear ultrasound image of the tongue.

Although AAA comes with an in-built spline workspace to visually analyse comparisons between different tokens, due to the rigidity of some of the analysis techniques, the data points were exported out of the program for analysis in R (R Core Team 2014). Spline points were extracted in terms of cartesian (i.e. x and y) coordinates, across the 42 fan points. Although the exportation of polar coordinates is possible in AAA, this was only really exploited for the SS ANOVA part of the statistical analysis (see Section 5.5.1 below), and for this the polar coordinates were calculated from the extracted cartesian ones.

Once extracted, the database of 42 x and y coordinates for each spline was collated with all other speakers. Acoustic data were collected by running a Praat script, which measured the F1, F2, F3, duration and preceding segment duration for each /l/. For the main part of this thesis, only midpoints were analysed, although the script collected the data for 10% intervals within the /l/. The acoustics were then collated with the articulatory data using R's `merge` function for plotting and analysis. The result was one database per experiment of all speakers' acoustics and splines.

5.3.2 Spline selection

As mentioned above, this investigation will primarily report on results from the midpoint of the /l/. This is value is chosen somewhat arbitrarily as a consistent anchor point for all /l/s across all speakers. In their study of the morphological effects on intervocalic /l/, Hwang et al. (2010) and Lee-Kim et al. (2013) only compared the degree of retraction between phonological contexts and so took splines from the most retracted position of the tongue root during the production of the /l/. This is a fair measure, but it can be somewhat subjective, and with 30fps, it is possible that the most retracted point could be missed somewhat by the ultrasound. Machines with a much higher frame rate would do a better job, as would other articulatory methodologies.

5.3.3 Temporal analysis

For the main part, this study considers midpoint articulations of /l/. However, previous analyses have shown that temporal information can be important. Sproat and Fujimura's (1993)'s influential X-ray microbeam study has resulted in the claim that the primary correlate of a light or dark /l/ is the relative phasing of the coronal and dorsal gestures. They showed that, in addition to having greater retraction of the tongue dorsum in relation to light /l/s, dark [ɫ]s also have earlier retraction of the tongue dorsum relative to tongue tip gesture. With the kind of frame rate available on the ultrasound unit used in this study, this kind of temporal precision is just not available. However, there are between 3 and 5 splines per /l/ available for analysis, so a basic temporal account is given for two speakers in Chapter 6. In fact, this section will argue that, although a temporal analysis may be required to explain the difference between phonological environments of /l/ in American English, the British English speakers actually show magnitudinal differences which are much more significant and easily observable from the midpoint of the /l/. Speakers with a stark contrast between initial and final position show that /l/s stay within their allophonic target range, regardless of their position at different time points over the course of the /l/.

Lynch (2009) attempts to replicate the Sproat and Fujimura study using ultrasound, by identifying a relevant fan point relating to the coronal and dorsal gestures. The problem with this is, the fan point refers to a point on the ultrasound screen, not a point on the tongue. It is a fairly decent measure to use, but a very rough estimate which does not take into consideration what the rest of the tongue is doing. Also, with very few frame per /l/, we have a very rough calculation of the temporal displacement.

5.3.4 The problem of inter-speaker comparison

Aside from some descriptive observations, inter-speaker comparison on articulatory data is not attempted in this study. Scobbie et al. (2011) discuss the potential of using a bite plate in ultrasound to ensure speakers have the probe placed in the same position (similar to the reference coil method used in EMA). This bite plate is placed in the speaker's mouth prior to the recording, to see where the the surface is relative to the the ultrasound probe. Although the plate was not necessarily designed with normalisation in mind, it could help in developing a method of inter-speaker comparison in future, through calculating the portion of tongue flattened against the surface. The bite plates are not yet under regular manufacture and, despite obtaining traces for a subset of speakers, it is not used in this thesis as it was not available to all participants. Moreover, the main aims of the thesis: to investigate the interaction of morphosyntax and phonology, and to investigate the extent of categorical and gradient processes, are best conducted on a within speaker basis. The dialectological aims of the thesis can be achieved through descriptive analysis. To give an idea of how speakers compare, however, acoustic data is pooled in some aspects of the analysis, as discussed further below in Section 5.4.2 and Section 5.5.5.2.

5.3.5 Plotting splines

The majority of spline plots in this paper are created with the `ggplot2` package in R (Wickham 2009). The exception is the SS ANOVA plots which are created with R's base plotting features. Splines are plotted by phonological environment, however, each individual prompt is visually inspected before combining the two separate prompts per environment. This ensures the two are the same (in a minority of phrases in two or three speakers, this is not the case). These plots form the basis of the descriptive analysis. Further quantitative analysis is detailed in Section 5.5.

5.4 Acoustic measures

5.4.1 /l/-darkness

As discussed in detail in Chapter 2,the primary acoustic correlate of /l/-darkening is the difference between F2 and F1. Darker /l/s have smaller differences between the first and second formants (i.e. are more like an [ʊ] than an [i]). The acoustic results presented in this study will largely focus on the midpoint of the /l/.

Several acoustic measurements of /l/ have been considered by researchers over the years. Hawkins and Nguyen (2004) considered taking centre of gravity (COG) measurements, which has the advantage of full automation, without the need for hand-checking unreliable formants. It also gives one measure, rather than several formants, but this can be done anyway with the F2 F1 difference. Despite these advantages, Hawkins and Nguyen (2004) conclude that COG misses small coarticulatory effects which can be picked up by the formant measurements. Other phonetic studies have used F2 alone as the primary acoustic measure of darkness. F2 reflects an anterior or posterior place of articulation i.e. higher F2s indicate lighter /l/s. This measure has been used by Carter (2003); Carter and Local (2007); Recasens (2004); Stuart-Smith et al. (2011). However, back cavity length can also affect the F1 realisation, so F2-F1 can give us even more information about the /l/. Following Carter (2002), van Hofwegen (2010), Sproat and Fujimura (1993) and many more, this is the measure which is used in this study.

The place of measurement also varies from study to study. Espy-Wilson (1992) and Lee-Kim et al. (2013) take the lowest point of F2 (i.e. the highest point of constriction) as the place to collect formant information. Most studies, however, choose the midpoint as their measure to avoid possible coarticulation (van Hofwegen 2010; Huffman

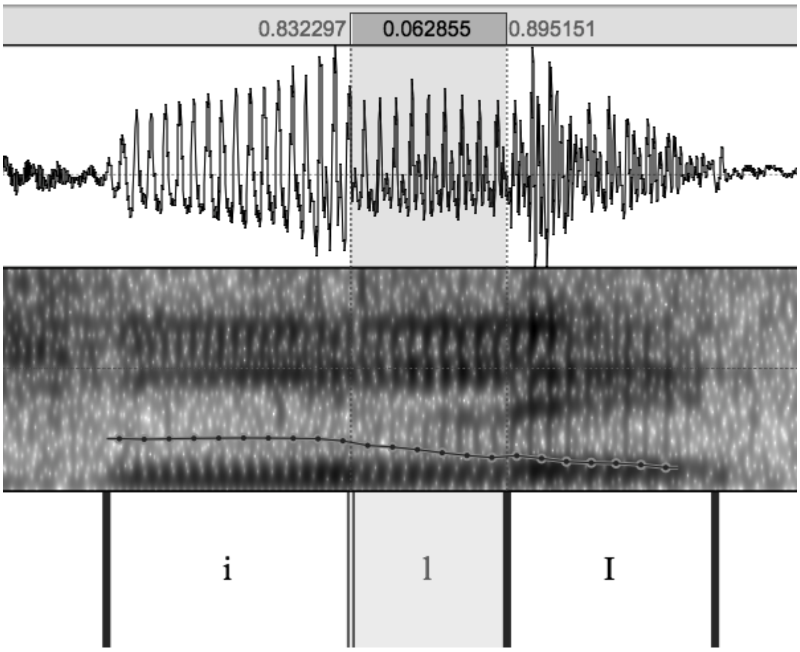

Figure 5.9: Praat segmentation of RP speaker's onset [l] intervocalically in *helix*. Note the spectral discontinuity between the [i] and the [l]

1997; Sproat and Fujimura 1993). The maximum point of constriction is an interesting choice, but because the ultrasound and the acoustics may not be exactly aligned to the millisecond, it is potentially not of use here. It makes more sense to have consistent measurements across tokens and speakers and opt for the midpoint of the steady state.

As mentioned above, the sound files were segmented in Praat. This allowed the running of a formant extracting script, to automatically collect formant measurements of /l/ and surrounding vowels for all speakers. Formants which looked unusual were checked by hand and corrected if necessary. /l/s were segmented where spectral discontinuity could be observed between the /l/ and the /i/ or /ɪ/. This is fairly simple for initial /l/, as in Figure 5.9, which shows segmentation of an intervocalic onset [l] token from the RP speaker's recordings. Here, the spectral discontinuity can be observed, as well as change in the waveform.

However, in final position, or for very dark [ɫ]s, segmentation can be trickier. Previous studies of /l/ have discussed the difficulty of segmenting between the actual /l/ and the preceding rime, as it can be completely unclear where one ends and the other begins. This can be seen to some extent when segmenting the RP speaker's final /l/, as Figure 5.10 shows for the phrase *peel bananas*. Here, the gradual transition between the preceding [i] and the [ɫ] makes the segmentation more complicated. For such tokens, segmentation between the vowel and the /l/ was placed immediately after the F2 transition down, where there was usually a clear break in formant structure. However, although this was fairly easy for varieties such as RP, it proved more difficult for speakers who have particularly dark [ɫ]s throughout the recording, which is what has been reported previously for American English (Umeda 1977:846; Sproat and Fujimura 1993:297). Transitions may

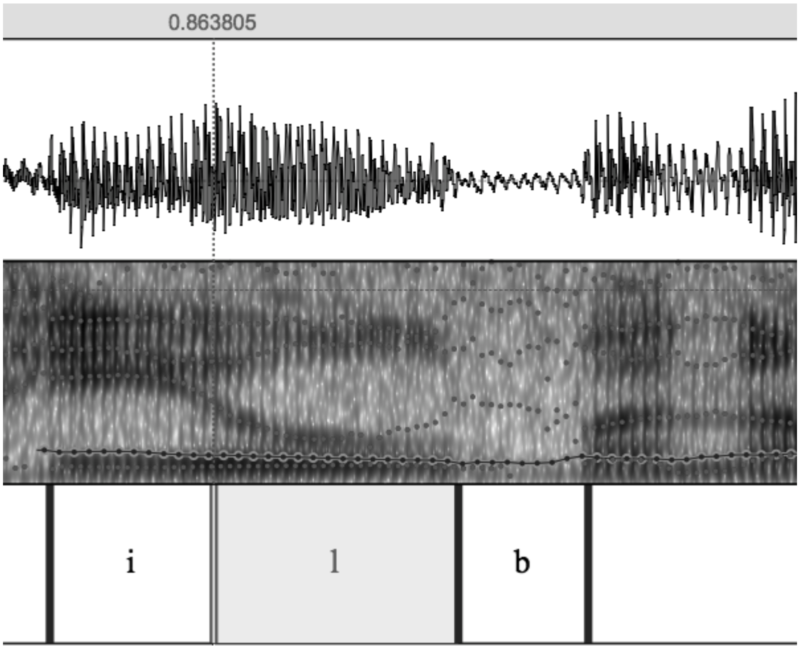

Figure 5.10: Praat segmentation of RP speaker's coda [l] preconsonantally in *peel bananas.*

be shorter or longer depending on the dialect, and the consistency of placement for the division between the vowel and lateral could be unreliable. However, the central point of the /l/ occurring in the F2 F1 steady state was confirmed for all tokens, ensuring that formant measurements would be consistent across speakers. One concern involves the reliability of the duration measurements, which are discussed in the following section.

5.4.2 Normalisation

Although the measure of F2-F1 has been used in raw form in previous studies, as it partially self-normalises, it does not take into consideration the differing vocal tracts of speakers. Previous studies of /l/ employ Bark normalisation to eradicate intra-speaker differences. Carter (2003) and van Hofwegen (2010) use the Bark Difference Metric (Syrdal and Gopal 1986), whereas Morris (2013) and Nance (2013) employ the basic Bark transformation (Traunmüller 1990). The latter was found to be a slightly better measure in a normalisation comparison by Flynn (2011) and is therefore the measure is used in this investigation, referred to as B1 and B2 in the plots. Note that the individual speaker analyses do not show normalised values. This is because Bark converts Hertz values into a scale that is more reflective of what the listeners hears. In order to have the best articulatory-acoustic comparison, the raw formant values are kept whenever there is no need for direct, quantitative cross-speaker comparison. Note that normalisation takes place before subtraction of F1 from F2.

5.4.3 Duration

The exact durational measures to use for analysis have been a point of conflict throughout this study. Originally, the duration of the /l/ itself was taken as the primary measurement, due to the ease of segmentation for many of the British speakers with a distinct dichotomy between initial and final positions. This is not the report given by many American phoneticians, particularly for word-final /l/, who tend to measure the 'rime' as a whole, also in consistency with Sproat and Fujimura (1993),[3] as it is difficult, or even impossible to define when the vowel ends and the /l/ begins as discussed above. This is also true of British speakers with very dark /l/s in this study. Because of the potential inconsistency in segmentation within and across speakers, the rime duration of /l/ and the preceding vowel is taken as the primary durational measurement in this thesis.

For statistical models, durational measures are log transformed from the original linear values. Because of the lower zero limit, linear values for duration result in a right skew, and it is common for phonetic studies to transform durational values to normalise the distribution (see Rosen 2005). As log scale treats smaller values in a comparable manner with larger ones (e.g. 10-20 milliseconds is weighted the same as 1000-2000 milliseconds) it allows us to analyse fine-grained critical differences that these smaller values hold. Furthermore, these can become more important for statistical analysis later on. However, this does mean that the measurement here is not consistent with Sproat and Fujimura (1993), who did not logarithmically transform their duration values.

5.5 Quantitative analysis

Qualitative approaches to articulatory data analysis are relatively straightforward for some aspects of ultrasound tongue imaging, and many other studies use these exclusively and successfully. Lawson et al. (2008; 2011) provide qualitative analysis only for their ultrasound data on Scottish /r/ realisations. They develop categorical possibilities for tongue shapes (e.g. bunched /r/, tip up etc.) and qualitatively decide which is which. For these distinct articulations, this kind of descriptive approach creates a succinct analysis of the articulatory possibilities. However, when tongue contours are very similar, quantitative approaches can help us decide whether small differences are important or statistically significant. This section gives an overview of the different quantitative procedures to be used in this project.

5.5.1 Spline comparison and Smoothing Spline ANOVA

AAA's spline workspace allows for quantitative analysis of tongue contours by means of a concentric grid which runs t-tests at each portion of the spline, calculating whether

[3]Note that in the present work, the 'rime' (i.e. the /i+l/ sequence) is put in quotations because under the theory of syllabification assumed, the /i/+/l/ sequence is not a rime in the first five phonological environments.

that particular section is significantly different from the same area in the other spline. This is the basic statistical method available in many different software packages and therefore used by most studies, including those at QMU (AAA; Wrench 2007), University of Maryland (EdgeTrak; Li et al. 2005), University of Toronto (Ultra-CATS; Bressmann et al. 2005) and UBC (Ultrax; Gick and Rahemtulla 2004).

However, Davidson (2006; 2012) argues that these methods are lacking in statistical sophistication as they do not analyse the entire tongue shape as one. Although the concentric grid method has been the most popular so far, she argues that there are limitations to running t-tests on a discrete number of points which can result in an incomplete method for characterising the tongue surface. If the most relevant constriction does not happen to fall on one of the fan's radii, the most important measurement is missed, she argues. Although this is unlikely given AAA's comprehensive number of fan points (42), Davidson details a more holistic method of comparing curves using the Smoothing Spline ANOVA (henceforth SS ANOVA; Wahba 1990). SS ANOVA is a statistical tool which can be employed for analysis of ultrasound tongue imaging data to ensure the entire tongue surface is taken into consideration. The SS ANOVA indicates whether the shapes of multiple curves are significantly different from each other by fitting 95% Bayesian confidence intervals to the curves. Overlapping confidence intervals mean the two are not significantly different. Rather than giving a p-value as an output, the statistical significance of two curves relies solely on the confidence intervals. In addition to providing a more sophisticated, holistic analysis overall, the output of the statistical significance is much easier to report and visualise with SS ANOVA, rather than reporting the results of 42 separate t-tests (however, if one wanted a p-value for a particular point in the curve, the t-test could do this). Full details of the SS ANOVA method can be found in Davidson (2006).

Figure 5.11 shows a basic SS ANOVA performed on two hypothetical phonological contexts. The dotted lines around the main splines form the 95% confidence intervals of the two phonological contexts; Context 1 and Context 2. They show no overlap, apart from the cross-over part in middle and the bottom of the tongue root area, showing that the two contexts are significantly different. Compare this to Figure 5.12, where the two splines are overlapping mostly.

SS ANOVA is used in this thesis primarily for Experiment 1, where differences between tongue shapes form the main part of the analysis. Elsewhere, R's `ggplot2` is used to plot splines and this automatically uses a slightly different kind of confidence interval calculation: the `loess` smoothing method. `Loess` also plots 95% confidence intervals and the difference between the two methods is negligible. The Bayesian confidence intervals plotted by SS ANOVA show an expansion of the original data, which is arguably better for predicting additional realisations. On the other hand, the `loess` smoothing method estimates from the actual values inputted. The two will give very similar results, with some slight differences at the edges of the plot. However, due to SS

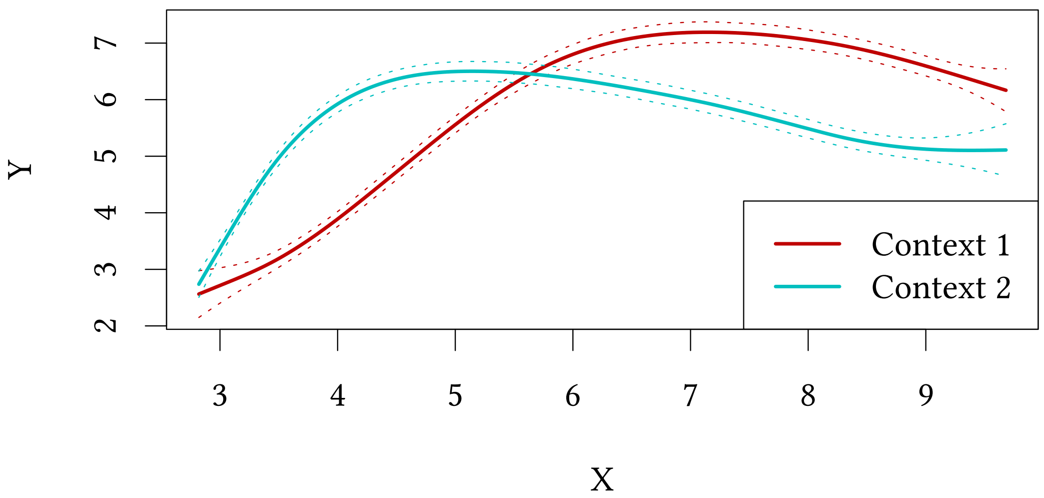

Figure 5.11: SS ANOVA plot of two hypothetical phonological contexts which are significantly different. N.B. the lines are not physiologically accurate.

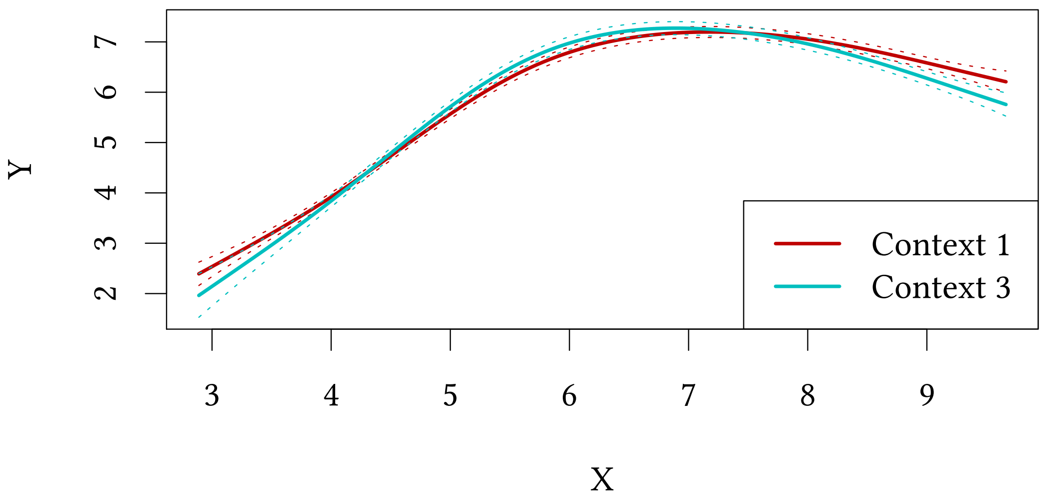

Figure 5.12: SS ANOVA plot of two hypothetical phonological contexts which are not significantly different. N.B. the lines are not physiologically accurate.

ANOVA's accepted use in the ultrasound literature, I will stick to this method when confirming statistical significances of tongue contours of different phonological contexts. This is largely restricted to Experiment 1. Elsewhere, the tongue shape is observed but additional factors form the main part of the argument (such as Principal Components Analysis and the other quantitative methods outlined in Section 5.5.2 below). Because of this, I will use the improved visualisation available in `ggplot2` in Experiment 2.

It should be noted that the SS ANOVA and `loess` ggplots do not have a fixed aspect ratio (i.e. they are not to scale) and do not represent physiological situation accurately. Although this does not provide a problem for the intra-speaker analysis, the reader should be aware these images may not be viewing an accurate 2D picture, as you would with an MRI scan.

5.5.2 Principal Components Analysis (PCA)

Another benefit of using a stabilising headset is the ability to carry out a Principal Components Analysis (henceforth PCA) on the raw data. As Slud et al. (2005:108) state, "The high dimensionality and lack of fixed landmarks in the human tongue make the parsimonious representation of its deformations a challenging problem." With a dataset which has 42 x and y coordinates per spline, how do we best analyse the data? It could be done by selecting a particular fan point for the tongue root and tongue tip, and running statistics on the variation there. However, this technique misses out the rest of the observations in that area of the tongue. Another technique is the visual selection of the maximum tongue tip advancement and maximum root retraction from the ultrasound images, but again this is focusing on one specific point on the tongue contour. An analysis which could reduce the high dimensionality of such datasets whilst considering all the data points together would be highly preferable. PCA is such a technique. PCA has additional benefits to SS ANOVA, as it produces a number for each spline, which can be compared in various ways, and can also be used for entry into linear regression.

PCA is a way of boiling down raw tongue contour data in order to extract areas of variance, and has been used previously with articulatory data (Slud et al. 2005; Stone 2005; Johnson 2011). It is a multivariate technique that allows a summary of systematic patterns of variation in data. PCA reduces the number of observed variables (in our case up to 84) to a smaller number of principal components which account for most of the variance of the observed variables using correlations of raw data, resulting in abstract components of correlated variation; the principal components. Overall, PCA identifies patterns in the dataset, highlighting similarities and differences. As Gould (1996:275) puts it, "Remember the high-school algebra exercise called 'factoring,' where you simplified horrendous expressions by removing common multipliers of all terms?" Although PCA has been employed by more recent studies of articulatory data, other explanatory factor analysis techniques, such as PARAFAC, stretch back to much earlier studies of tongue shapes (Harshman et al. 1977; Hoole 1999; Jackson 1988; Maeda 1990).

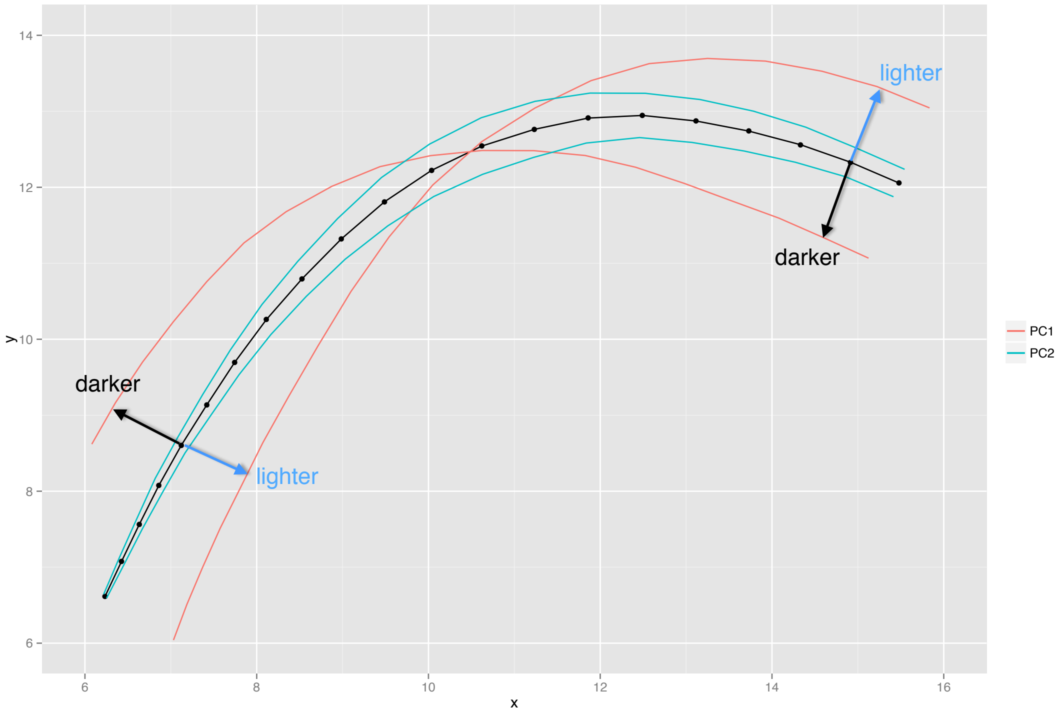

Figure 5.13: PCA on midpoint /l/s only for RP speaker

Previous studies have described the measurement techniques for conducting PCA, but fail to advise exactly what kind of raw data the input should be encompassed from. Originally, PCA for tongue shape data was used on the actual images of the midsagittal view of the tongue. This kind of technique used information from pixel data and is where the PCA is at its most powerful. The drawback for the present study is that the software used here (AAA) cannot perform such an analysis, and extra software such as MATLAB would have to be introduced to run the PCA on the ultrasound pixel images themselves. However, PCA has also been successfully used to analyse the coordinate data extracted, as opposed to the direct images (Johnson 2011; Pouplier and Hoole 2013).

5.5.2.1 Loadings plots and variable importance

Figure 5.13 shows the loadings plots of the PCA computed for /l/ midpoints only, for the RP speaker. The loadings matrix gives a score indicating how much of each principal component can be found in the /l/s tongue shape i.e. the weight by which each original value should be multiplied by to give the final component score. The loadings plots are these figures in graph form.

This is not an analysis whereby certain PCs correspond to height or backness, as the PCA does not have an input which is representative of the movement of the tongue, rather the midpoint. In Figure 5.13, PC1 is a proxy for both lightness (advanced tongue body, higher tongue blade) and darkness (retracted tongue root and lower tongue blade). This is demonstrative of all of the data we will see in the following chapters: variation

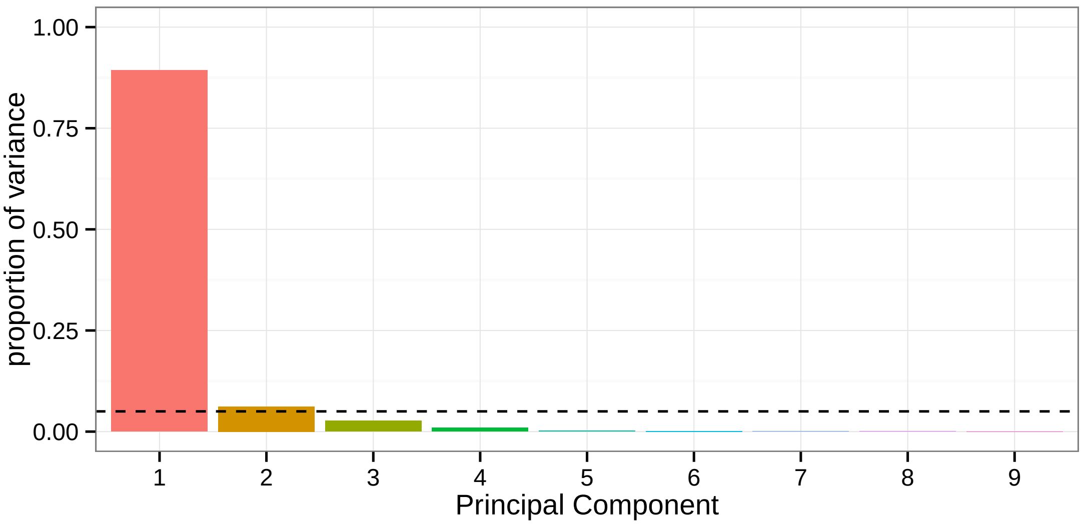

Figure 5.14: Proportion of variance. The first 9 PCs of a typical PCA on tongue spline data (taken from RP speaker)

in the dataset is represented by one principal component (PC1) and each tongue spline is assigned a score for PC1 which represents its shape. The PCA loadings translates this score into a coherent tongue shape.

For the dataset represented in Figure 5.13, the first principal component, PC1, accounts for 89% of the variation, and PC2 just 5%. Figure 5.14 demonstrates this result, and shows the proportion of variance for this tongue spline based analysis. Obviously, the first PC here accounts for most of the data. A 5% cut off for the remaining PCs is suggested (Baayen 2008:130). Any PCs which can account for more than 5% of the variance are significant. PC2 in this dataset accounts for 6.1% of points. However, Baayen also states that another rule of thumb with PCA is to cut off where there is a clear discontinuity. For the majority of speakers in this dataset, this occurs after PC1 and it seems that PC2 is simply soaking up the left over variance which PC1 cannot account for, usually at the spline crossover point. This is visualised by the dotted line in Figure 5.14.

5.5.2.2 Drawbacks of PCA on spline data

The quality of data needed to perform an interpretable PCA is not clear from previous studies. For example, if we are interested in the PCs of the midpoint of an /l/, is it sufficient to enter only /l/ midpoints as input to the PCA? Doing so results in summarising the variance in /l/ shapes, but not a true summary of tongue movement. Such an analysis, as represented above, allows no hybrid position i.e. no advanced tongue root and lowered tongue blade. Although this may be a fair summary of /l/s for this speaker, it is not a fair summary of the speaker's tongue shapes overall, which would be preferable.

An additional drawback of this method is when NAs ('not available' i.e. tokens which are missing in the dataset) occur in the data, which happens frequently with ultrasound

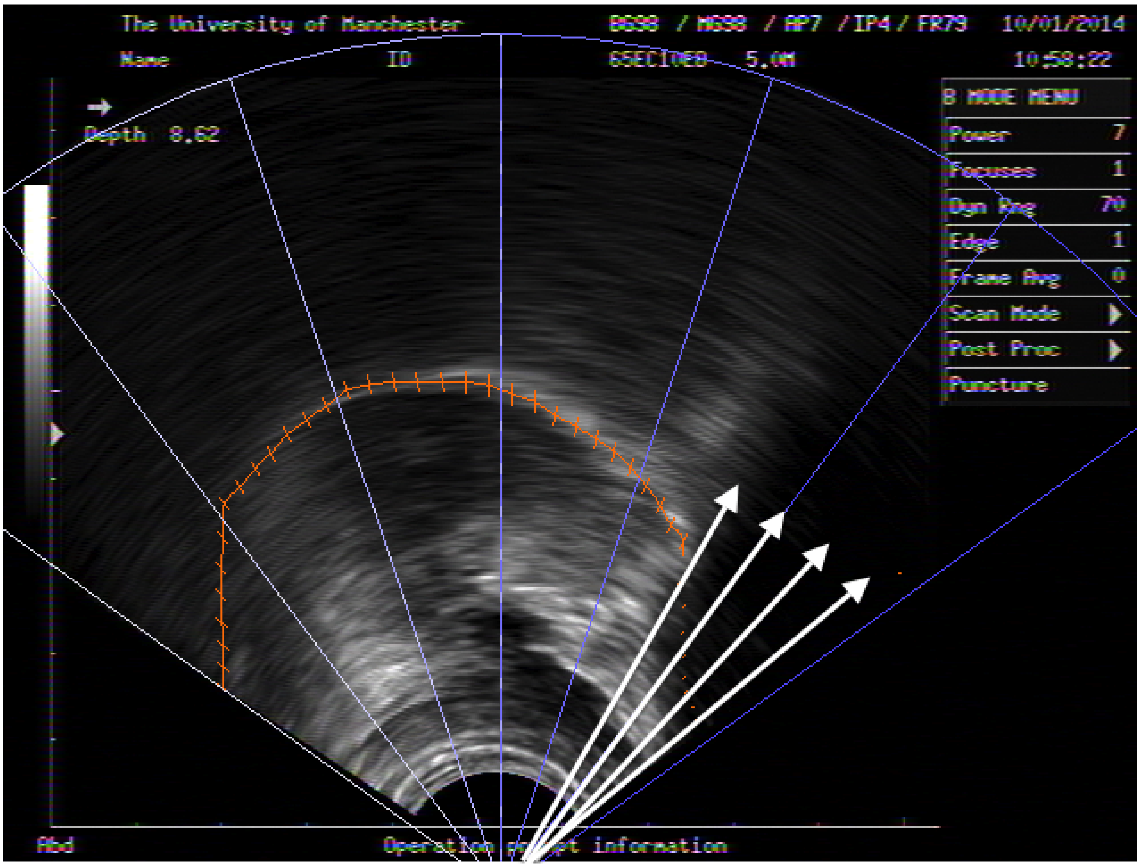

Figure 5.15: Ultrasound image demonstrating how NAs can occur in cells. White arrows demonstrate the fan points which will have NA values for this particular spline.

tongue splines. This is not just because the image is poor, but because the tongue may not be crossing a particular fan point during a particular articulation. Figure 5.15 is an example of such a spline, showing a backed tongue position with no activity in the frontmost fan points. The extreme fan points on the right side of the image do not have any clear information about the tongue. At the right edge, the tongue does not appear at all because it is sufficiently backed not to appear in this region. In a PCA of tongue contour data, these fan points receive an NA in the relevant cell. There is nothing wrong with this ultrasound image, but fan points from around Fan 33 onwards have no data available, as there is no spline information here. This then means that all data for this particular fan point has to be removed, even for splines which do have data available in Fan 33 onwards, as the PCA needs all rows of data to be NA-free. This results in the extreme edges of the points not being included in the PCA, instead giving a tunnel-vision outlook on the spline variation.

Because of the limited viewpoint, the results we get from the PCA of such tongue contour data is the first principal component (henceforth PC) being the only significant component. As adhered to above, for this study, PC1 accounts for all variation within a speaker. This is perhaps unexpected, given we may expect two significant PCs which correspond to tongue height and tongue backness respectively. This is what Johnson (2011) describes for his tongue pellet data, which has the advantage of having no missing data points, as all tongue pellets are constantly accounted for. In addition, analysis of pixel data comes across no such problem: the script simply reads pixels with no tongue image as black dots, resulting in no missing data. For a PCA of spline contours, however, this cannot be avoided. Despite these drawbacks, PCA still remains a very useful method for reducing the dimensionality of the dataset, and successfully picks up on patterns in the data, albeit not as succinctly as a pixel analysis might. As we shall see in the following chapters, its results are completely consistent with the original spline contours, and allow us to attach a value to each one, even if this technique is not as quite as detailed as we might like.

5.5.3 Statistical tests on PCA values

5.5.3.1 Tukey's Honest Significant Difference test

Tukey's Honest Significant Difference (HSD) test is used on the Experiment 2 data to calculate differences between continuous dependent variables and categorical predictors i.e. PC1 values by phonological context. Tukey HSD is a post-hoc test which uses ANOVA to compare means that are significantly different pairwise across all contexts. That is, for use in this thesis, the Tukey test compares the mean of each phonological context to every other individual phonological context pairwise, and any difference in the means that is found to be bigger than the expected error is significantly different. For the ten phonological contexts used in Experiment 2, the combination of pairwise comparisons results in 45 tests. Tukey HSD is a similar method to the perhaps more well-known Bonferroni correction. On the whole, the Bonferroni correction is usually reserved for controlling the family-wise Type I error rate (i.e. the incorrect rejection of the null hypothesis) with a small set or subset of comparisons, whereas Tukey has the edge when we are interested in all pairwise comparisons, which is the case here. Tukey HSD is also used on the acoustic data i.e. the difference between the first and second formants over ten phonological contexts. As the output to the Tukey HSD is so large, at 45 rows of data across eight speakers, as well as a table of the 45 p-values, it is reserved in full for the Appendices, and an overall summary of the significant differences are reported in the main body of the thesis.

5.5.3.2 Hartigan's dip test

Hartigan's dip test statistic (Hartigan and Hartigan 1985) is a measure of unimodality and bimodality in a given dataset. Using R's `diptest` package (Maechler 2013), this test is used in the results section of Experiment 2 to investigate bimodality in the PC1 results. Hartigan's dip test works by measuring bimodality and multimodality in a particular sample. It assesses potential 'dips' in the distribution and outputs a p-value reflecting whether or not the dip is due to chance. The test also outputs the dip statistic, or D statistic. The closer to 0 the statistic is, the more likely the distribution is to be

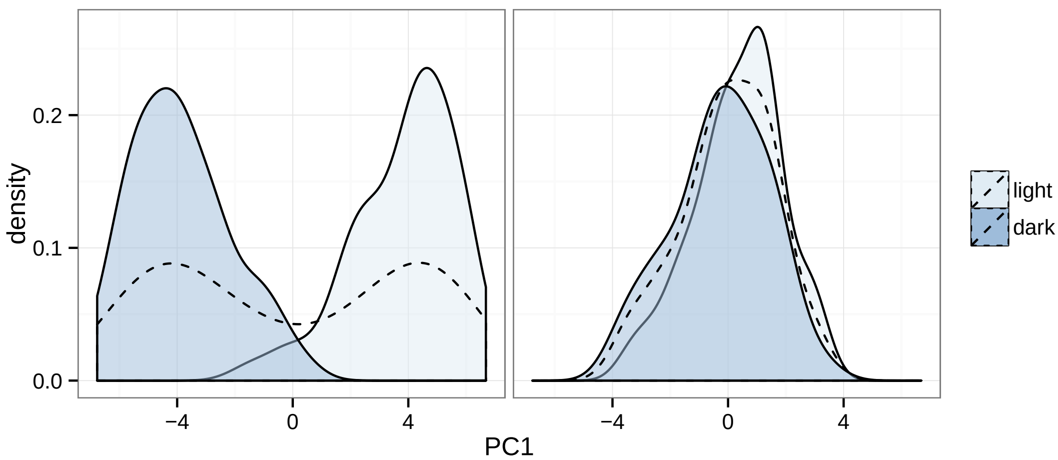

Figure 5.16: Density plots for a significant bimodal distribution by Hartigan's dip test (left panel) and a unimodal distribution (right panel). The dashed lines represent the overall distribution. Separate light and dark distributions are for visualisation purposes and the assignment of category is discussed below.

unimodal, or normally distributed. Some dips can be particularly shallow, but still significant. The dip test also outputs a p-value, indicating the statistical significance that the null hypothesis can be rejected, and that the dip is in fact significant (the cut-off for significance is placed at $p < 0.05$ here). The dip test is conducted both on the overall PC1 values as well as the F2-F1 acoustic measures. The test is accompanied by a density plot in the Experiment 2 results showing the overall distribution by a dashed line, and then separate category distributions by a solid line to accentuate any bimodality or lack of it.

5.5.3.3 Correlations

Pearson's product moment correlation coefficient r is used regularly to calculate and measure the linear correlation between darkness and duration. The closer to 1 or -1 the stronger the correlation is: a positive value indicates a positive correlation, i.e. that when one variable increases so does the other; a negative value indicates a negative correlation, i.e. when one value increases the other decreases. In this study, correlations are performed using R's basic `cor.test` function, which also produces significance values (R Core Team 2014).

5.5.4 Category formation

Category formation for the large part of the analysis in Experiment 2 is based on the results from the Tukey HSD tests. In order to test whether allophonic categories are a good way of accounting for the patterns found, categories are assigned to speaker tokens and the linear models indicate their fit. A good fit is an indication that allophonic categories provide a convincing explanation for the speaker's grammar, and a poor fit suggests the speaker does not have an allophonic distinction.

Categories are *a priori* generalised to consist of word-level onset and coda positions: i.e. a light/dark distinction based on syllable position at the word level. These are then confirmed or adjusted based on the results from the Tukey HSD post hoc test of significant differences between the ten contexts, as outlined above in Section 5.5.3.1. Tukey HSD tests are run on both the PC1 values, i.e. the articulatory correlate of darkening, and the F2-F1 values, i.e. the acoustic correlate of darkening. If the 'light' and 'dark' categories are significant across category but not within, then this provides evidence for treating them as two separate allophonic categories. The idea is, if these categories fit the data well, it provides evidence for treating them separately. If they do not fit the data well, these two categories may not be the best way of describing the speaker's system and they will have to be rethought, with the help of the Tukey HSD results.

As might be expected, with 45 comparisons, we may see some variation which is down to fluke (although Tukey performs a correction to account for the multiple comparisons, similar to Bonferroni). To be extra vigilant, the p-values from the articulatory and acoustic data are compared and consistency throughout both dependent variables is taken as enough evidence to treat this difference as a significantly different category. It may be pointed out that more recent studies (including this one) find patterns in the articulations which evade the acoustics. Any patterns in the PC1 data which are consistent and robust will be taken into consideration, even if the acoustics do not show anything. The only reason both are taken into consideration as a starter is so that fluke values do not slip through the net and form their own superfluous category.

5.5.5 Statistical models

Although the tests above provide ways of testing significant differences between environments, they do not provide a way of including several factors in the same model. For example, the previous literature suggests that duration is a factor in /l/-darkening, but also phonological category. By using the statistical methods described in this section, it is possible to include both in the same model.

5.5.5.1 Linear regression models

Linear regression models the relationship between the continuous dependent variable y, and one or more explanatory predictor variables x. Simple linear regression refers to models with just one predictor, and multiple linear regression refers to models with more than one predictor. Both will be used in this investigation. The linear regression used on a by-speaker basis in Experiment 2 attempts to predict the value of the response value, i.e. darkness, by its predictors, i.e. category or duration, or a combination of both. The models are presented with their adjusted r^2 values. r^2 is the coefficient of multiple determination, indicating how well the data fit the statistical model, and is simply the r value (i.e. the measure of correlation of the line) squared. The adjusted r^2 is better for comparing models of different sizes as r^2 alone automatically increases when extra variables are added to the model. The adjusted value takes account of this. The closer the adjusted r^2 is to 1, the better the fit. In these models, categorical predictors are handled by treatment coding, as it is more interpretable than sum coding in this dataset, and all factor levels have identical numbers of tokens. In the following chapters, the results of the linear regression are from the `lm` function in R, which uses Ordinary Least Squares (OLS) regression. The reader is directed to the following texts for further details on linear regression: Baayen (2008); Crawley (2012); Faraway (2004); Lewis-Beck (1980).

5.5.5.2 Mixed-effects linear regression models

An advancement on the models discussed above is employed in the final part of Chapter 7, when all speaker data is grouped together for the acoustics only. This number of data points allows for data modelling on a more sophisticated scale. With all speaker data in one model, it is paramount that speaker be included as a predictor. As the predictor of speaker is not repeatable, it is necessary to include it as a random effect. A predictor is repeatable if the set of possible levels is fixed (Baayen 2008:241). For example, the factor of gender is repeatable, as the set of possible options is fixed at two. An individual speaker or participant, however, is not. Mixed-effects models allow the inclusion of random effects into the model alongside the regular fixed effects. Linear mixed-effects regression is carried out with the `lmer` function in the `lme4` package in R (Bates et al. 2014).

Another benefit of running mixed-effects models is the ability to include random slopes. Including a random effect we are assuming that only intercepts vary with respect to each speaker. Random slopes, on the other hand, consider the differing behaviours of speakers with respect to another predictor. It works like an interaction in the model might, but with random effects and not fixed ones. Once the model with the random slope is configured (e.g. predicting darkness with a random slope of Speaker by categorical distinction) it is then compared with a nested model without the random slopes, in terms of deviance and AIC (Akaike Information Criterion; see below). If the inclusion of random slopes is significantly enhancing, there is evidence to keep it in the model. For further details on random slopes and intercepts, the reader is directed towards Gelman and Hill (2006).

A measure of fit displayed for mixed-effects models is the Akaike Information Criterion (Akaike 1974; henceforth AIC). AIC is a measures of goodness-of-fit, in order to assist choosing between models. Smaller values indicate better a model fit. AIC adjusts for the number of parameters in the model, typically yielding more parsimonious models where there are no superfluous predictors that do not explain much of the variance.

Tables show t values not p values. Values higher than +/- 1.96 are taken to be significant, equivalent to a p-value of less than 0.05. This is because the most recent version of the `lmer` function in R, using the `lme4` package (Bates et al. 2014), does not provide p-values.[4]

5.5.6 Model comparison

ANOVA comparison is employed in the results section (comparison tables are reserved for Appendix D) to compare whether a linear model is significantly better than a nested linear model, as the example in (17) demonstrates. A nested model is one which is included in a fuller model, with an extra predictor, so in (17) the nested model is (a) and the fuller model is (b).

(17) a. $y \sim x_1$
b. $y \sim x_1 + x_2$

In the ANOVA comparison tables, a p-value less than 0.05 indicates that the fuller model is significantly better, and significantly reduces the deviance, than the smaller model. In chapter tables, the adjusted r^2 value is used to compare model fit of linear models. The better the model fit, the closer to 1 the adjusted r^2 value will be. AIC is also used in some parts of the analysis. In the mixed-effects model, AIC is used solely.

5.5.7 Diagnosing categoricity with quantitative methods

Numerous conditions for exploring for categoricity in /l/-darkening data were outlined in Section 5.1, and discussed in detail in Chapter 4. The quantitative methods described above are employed to test these diagnostics, which are detailed below:

i. Articulatory discontinuity between two (or more) sets of splines will be tested for significance with SS ANOVA and `loess` confidence intervals: non-overlapping intervals show significant differences between contexts. Tukey HSD tests on the PC1 values will add further support to significances.

ii. Articulatory consistency within these sets will be explored with the same methods as above, with overlapping confidence intervals showing consistency between contexts. Similarly, the Tukey HSD might be expected to show high p-values when two contexts are the same.

iii. Bimodality in the quantitative analysis of the articulatory data will be tested using Hartigan's dip test on the PC1 values.

iv. Bimodality in the acoustic data will be tested using Hartigan's dip test on the formant values.

[4]p-values are not included because they can be unreliable for mixed effects models. Moreover, estimation of p-values by MCMC is not possible for models with random slopes. See Bates (2006) for an explanation of the problems associated with `lmer` and MCMC calculation of p-values (not yet implemented for models with random slopes).

v. A linear model with articulatory darkness as the dependent variable that provides a good fit for two categories will be tested on the PC1 values using linear regression.

in addition

vi. The acoustic values can also be subjected to the Tukey HSD test and will be used for a cross-speaker analysis in the mixed-effects linear regression section of Chapter 7.

5.6 Qualitative articulatory analysis

Although the quantitative methods discussed in the previous section encompass a large amount of the analysis, this supported with qualitative commentary throughout the next two chapters. The PCA value treats the tongue as a whole, but at most times it will be appropriate to discuss separate area independently. This may be in reference to the tongue blade's proximity to the palate trace, the retraction of the tongue root, or the visible velarisation of the tongue, and so on. Note that, in the majority of ultrasound images, the tongue tip is obscured due to the signal hitting air pockets in the region. Nevertheless, at some points in the qualitative analysis, the tip is referred to, either as a general marker of the front of the tongue, or in reference to the estimated position of the tip. A qualitative analysis is used frequently to make comparisons between spline plots, and across speakers in some situations.

5.7 Summary

This chapter has discussed the experimental methodology of ultrasound tongue imaging, employed as the prime empirical aspect of this dissertation. This chapter serves to define the research hypotheses. The two main goals of the thesis, morphosyntactic conditioning and categorical vs. gradient processes in /l/-darkening, are defined at the beginning of the chapter. These goals are subjected to empirical testing through the ultrasound analysis presented in the previous sections, and details of data collection and the experimental procedure have been provided. This is backed up by quantitative and statistical techniques which are outlined above, alongside the motivation for including such tests. The following two chapters are committed to discussing the experimental results for the two main research questions in turn. The results will be presented giving due consideration for implications of phonological theory of analyses of English /l/-darkening.

Chapter 6

Experiment 1: Morphosyntactic conditioning

This chapter seeks to investigate the potential /l/-darkening typologies that exist in English. As discussed in Chapter 2, although the allophonic conditioning of /l/-darkening is often reduced to a simple onset/coda distinction, previous studies indicate that different dialects vary in their morphosyntactic sensitivity, as Table 6.1 shows. The aim of this chapter is to provide experimental evidence for these different systems, using ultrasound tongue imaging. The articulatory evidence from ultrasound is supported with acoustic measures of F2-F1, the primary acoustic correlate of /l/-darkening (the smaller the distance between the two formants, the darker the /l/). The primary statistical method used in this chapter, as outlined in Chapter 4, is SS ANOVA. The 95% Bayesian confidence intervals fitted by the SS ANOVA to each phonological environment represent significant differences from other environments when there is no overlap. Overlapping confidence intervals indicate that there is no statistical evidence for the two environments being articulated differently.

This chapter will present results from speakers of several varieties, including RP, Manchester, Middlesbrough, Essex, plus a speaker of American English. Speakers read tokens corresponding to the headings in Table 6.1, with the extra phonological environment of word-final preconsonantal /l/.[1]

An additional methodologically inspired debate is also presented at the end of the chapter in Section 6.7, regarding the temporal properties of the /l/. In this section, I argue that midpoint splines give us more than enough information to conduct analyses of /l/-darkening, as tongue dorsum and overall tongue body retraction can be extracted from these splines for the majority of speakers.

Speakers were recorded producing /l/ in the phonological contexts in Table 6.2, hence-

[1]As mentioned in Chapter 5, the first two speakers recruited for this study were not given word-final preconsonantal tokens of /l/, as this was not considered part of the the scope of the study at the time. This was quickly corrected, but these tokens are missing for the speakers in Section 6.1 and Section 6.2. The *hea**l**_C* environment is preconsonantal /l/, with the following word beginning with a /b/ or an /h/. The Middlesbrough speaker /h/-dropped following /h/ initial words, so his tokens are comprised of pre-/b/ tokens.

	***l**ight*	*he**l**ium*	*hea**l**-ing*	*hea**l** it*	*hea**l***	
RP	[l]	[l]	[l]	[l]	[ɫ]	Cruttenden (2008); Jones (1966)
Am. Eng. 1	[l]	[l]	[l]	[ɫ]	[ɫ]	Sproat and Fujimura (1993); Gick (2003)
Am. Eng. 2	[l]	[l]	[ɫ]	[ɫ]	[ɫ]	Olive et al. (1993)
Am. Eng. 3	[l]	[ɫ]	[ɫ]	[ɫ]	[ɫ]	Hayes (2000); Yuan and Liberman (2011)

Table 6.1: /l/-darkening in different morphosyntactic environments. Adapted from Bermúdez-Otero (2007a).

Experiment 1

Context	**Sentence**
1. word-initial	***l**eap*
2. stem-medial posttonic	*he**l**ix*
3. stem-final presuffixal	*hea**l**-ing*
4. word-final pre-vocalic	*hea**l** it*
5. word-final pre-consonantal	*hea**l**_*C
6. utterance final	*hea**l***

Table 6.2: The five/six phonological environments for Experiment 1 with example tokens.

forth referred to by example token.

6.1 RP

Figure 6.1 corroborates the claims made by previous descriptions of RP (Cruttenden 2008; Jones 1966; Wells 1982): the Standard English variety has light [l]s whenever the sound is prevocalic.[2] The backed tongue body, reduced tongue-tip gesture, and retracted tongue root typical of [ɫ] is only found non-prevocalically, in the *hea**l***-type tokens. Whenever /l/ is followed by a vowel, regardless of its position the /l/ remains light (although see the next chapter for different effects of prosodic structure). This goes not only for word-level onset /l/s, such as the word-initial ***l**eap*-type tokens, intervocalic *he**l**ix*-type tokens and pre-suffixal stem-final *hea**l**-ing*-type tokens, but also when /l/ is resyllabified into the onset of the following word in *hea**l** it*-type tokens. The Bayesian confidence intervals fitted by the SS ANOVA technique confirm that the observed pattern is statistically significant, although in this case the distinct distributions are convincing enough without the statistical conformation. Nevertherless, the SS ANOVA does demonstrate that the first four contexts featuring onset /l/ are the same, as the confidence intervals are completely overlapping and therefore any small differences are not statistically significant.

The acoustics in Figure 6.2 confirm the articulatory range for this speaker, showing a clear cut-off range between the phrase-level onset /l/s and the the phrase-final *hea**l***-type

[2]It is worth reminding the reader at this point that the SS ANOVA plots are not necessarily to scale, but the scale can be seen on the x and y axes for this chapter.

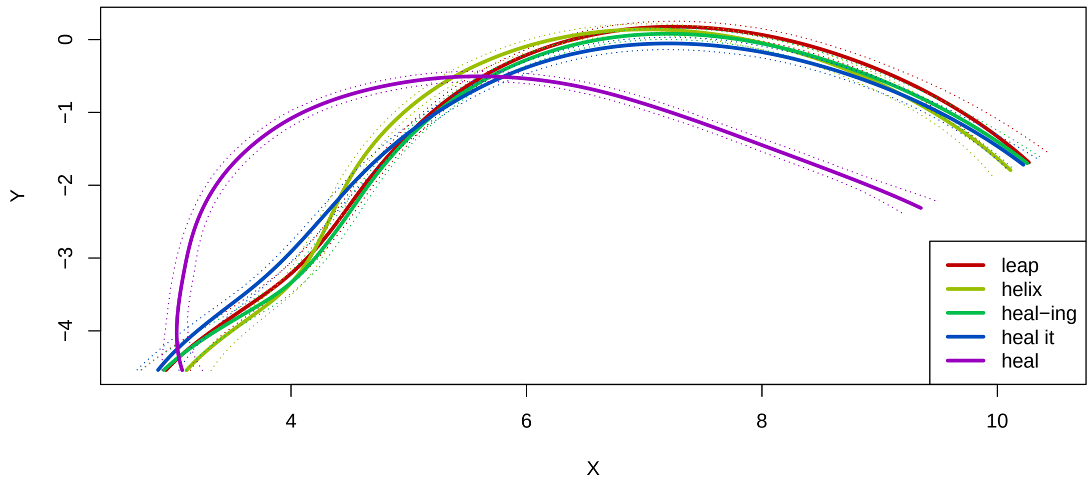

Figure 6.1: SS ANOVA of RP speaker's mean /l/ midpoints across five phonological environments (based on polar coordinates).

tokens. In addition, the acoustics show a small difference not observed in the articulatory splines: the *he**l**ix* and *hea**l**-ing*-type tokens are slightly lighter on the acoustic scale than the ***l**eap*-type tokens. This may be something to do with major morphosyntactic or prosodic boundaries from the word being utterance initial in this dataset, but given we only have five tokens per environment and the outliers for the two intervocalic contexts overlap with the ***l**eap* range, this may be down to chance. A Tukey HSD test confirms that the only significant differences in this plot are between the *hea**l***-type tokens and the other four contexts ($p < 0.01$ in all cases; see Appendix C). It is an interesting pattern nonetheless, and is most likely due to coarticulation, in that the intervocalic /l/s have *two* neighbouring high-front vowels, as opposed to the ***l**eap*-type tokens which only have the following vowel. Therefore, there is extra coarticulatory pressure on the intervocalic /l/s in comparison to the word-intial /l/s, which is the most likely explanation for this pattern. This is also observable for the *hea**l** it*-type tokens, but only with regard to the median bars; the ranges are very much overlapping and the potential acoustic leak effect is unsurprisingly not as strong given the following vowel is in another word. Overall, the pattern is consistent with the articulatory findings, with a clear distinction the phrase-level onset /l/s, which have a mean F2-F1 value of 2126Hz, and the phrase-final *hea**l***-type tokens, which have a mean value of 668Hz. Providing further support for the pattern, the minimum value for the light tokens in 1456Hz and the maximum dark value is 897Hz.

These results are consistent with the first stage of the life cycle, as summarised in Table 6.3. The life cycle can explain the difference between word-final /l/ in *hea**l*** being dark, but word-final *hea**l** it* being light. In this variety, darkening occurs at the phrase level, meaning that the syllable organisation of the phrase *hea**l** it* occurs before darkening applies. The [l] is already resyllabified into the onset of the following word when

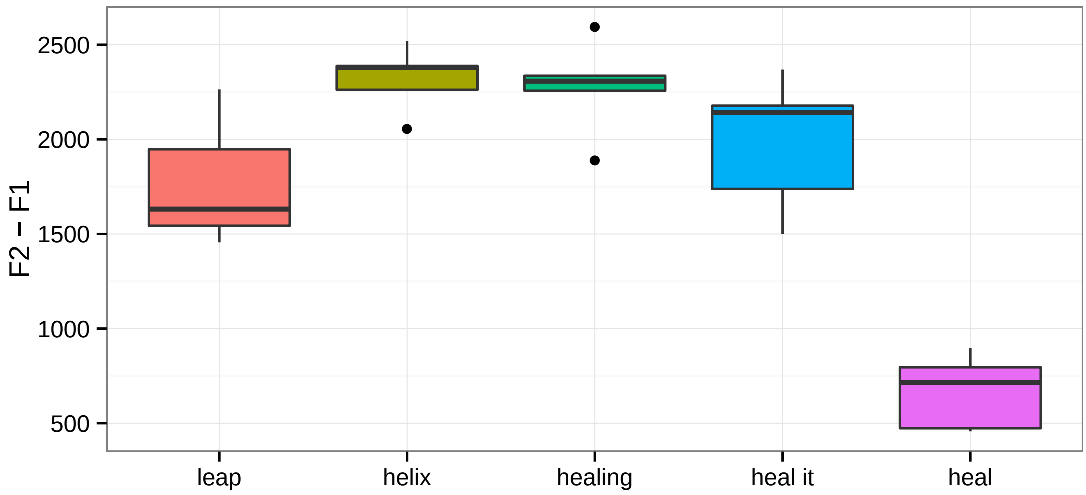

Figure 6.2: RP acoustics

	light	*helium*	*heal-ing*	*heal it*	*heal*
RP	[l]	[l]	[l]	[l]	[ɫ]
Am. Eng. 1	[l]	[l]	[l]	[ɫ]	[ɫ]
Am. Eng. 2	[l]	[l]	[ɫ]	[ɫ]	[ɫ]
Am. Eng. 3	[l]	[ɫ]	[ɫ]	[ɫ]	[ɫ]

Table 6.3: The RP speaker's system corroborates the first stage of the life-cycle: /l/ darkens in the coda at the phrase level.

darkening applies, and it is no longer susceptible because darkening applies to coda /l/s only.

6.2 Manchester WC

As discussed in Chapter 2, the variety spoken in Manchester (North-West England) is of interest, as the existing literature tends to conclude that Manchester /l/s are dark in all positions (Cruttenden 2008; Kelly and Local 1986). However, we have acoustic evidence that there is a small difference between initial and final position (Carter 2002). At first glance, the Manchester splines in Figure 6.3 seem to corroborate the claims of only one category in this dialect. All tokens have a lowered tongue tip and tongue body.

However, when examining the confidence intervals plotted by the SS ANOVA, there exists a small difference and lack of confidence interval overlap in tongue-root backing between phrase-final tokens and the other contexts. Although *phonetically*, the Manchester distribution is in contrast to the stark light/dark dichotomy found in RP, it is interesting to note that the *phonological* contexts pattern together, in that phrase-final environment is more retracted than the other four contexts. The F2-F1 differences in Figure 6.4 show that the ***l**eap*-type tokens are the lightest, and the *hea**l***-type tokens the darkest.

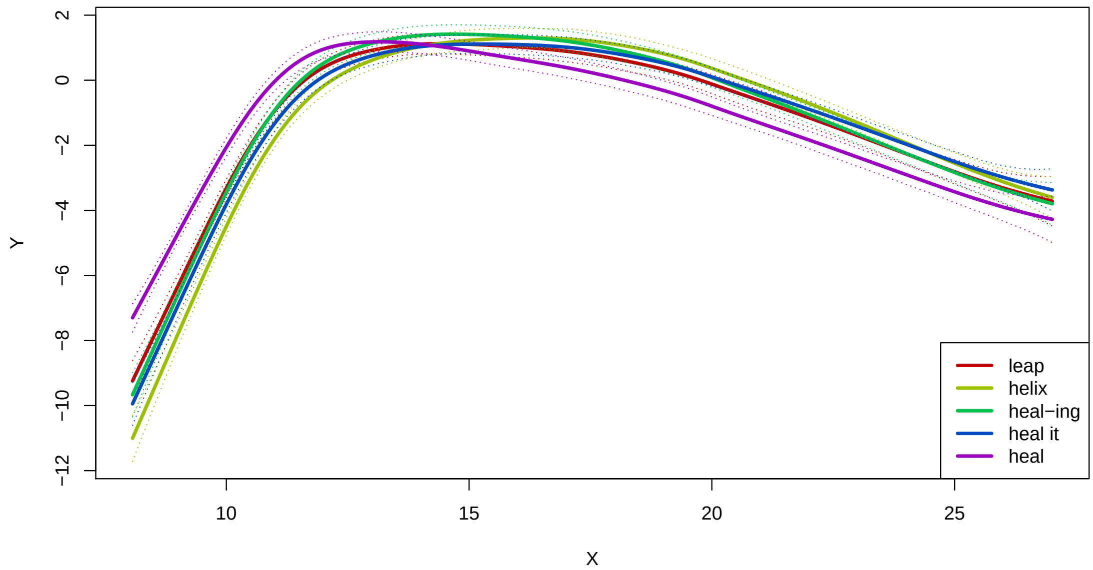

Figure 6.3: SS ANOVA of Manchester speaker

The acoustics show a difference between the four articulatorily lightest tokens, however, with ***l**eap* and *hea**l** it*-type tokens being lighter than *he**l**ix* and *hea**l**-ing*-type tokens. The Tukey HSD shows that the *hea**l**-ing* tokens are significantly darker than the ***l**eap* and *hea**l** it*-type tokens ($p < 0.05$ in both cases). What accounts for this is not clear, given this is in no way reflected in the average tongue splines in Figure 6.3. This would have to be investigated further to see if it is a true effect, or perhaps something prosodic (e.g. this speaker might show an effect triggered by enclisis of *it*.) This effect can be investigated further with more environments in the following chapter. These differences seem much larger in this plot, however, given that the frequency range over which different realisations of /l/ are distinguishable is much smaller for this speaker than for the RP speaker discussed above. Observe in Figure 6.2 that /l/ realisations for the RP subject span a 2000Hz frequency range, whereas /l/ realisations produced by the Mancunian subject span only a 600Hz range. This is not to claim that only large phonetic differences can indicate categorical differences, but is just to point out how tightly these acoustic ranges are clustered for this speaker. This point is discussed later in the chapter, in Section 6.6 (in particular see Figure 6.14).

Considering the evidence from articulatory and acoustic analysis, there is a difference between initial and final /l/s in Manchester, however, it is not clear from this picture whether Manchester has the categorical distinction so convincingly displayed by RP. Instead, we could be looking at a gradient effect of phrase finality, perhaps conditioned by duration, for this speaker. This question is addressed in the next chapter. For now, we can look to settle on one of two conclusions for the Manchester speaker:

i. Despite the phonetic disparity when compared with RP, Manchester patterns phono-

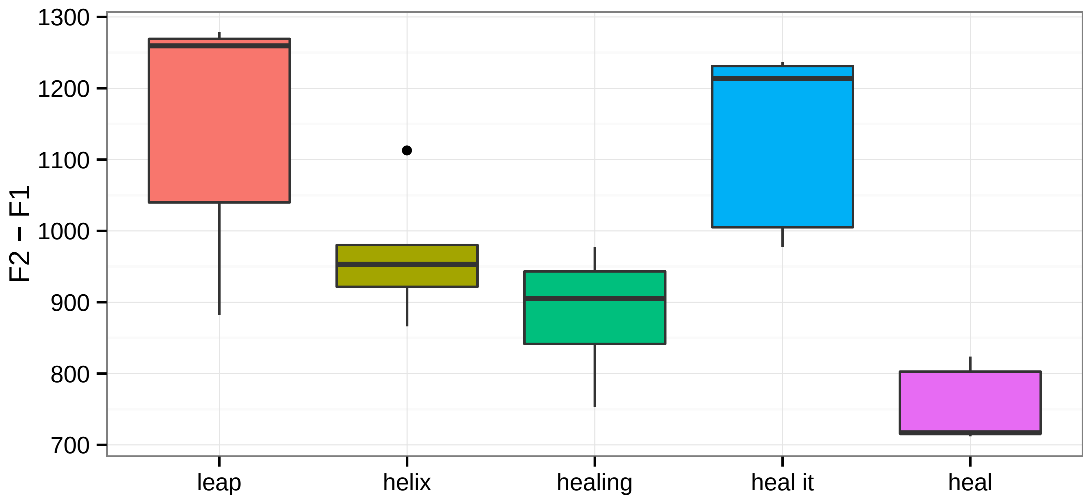

Figure 6.4: Manchester acoustics

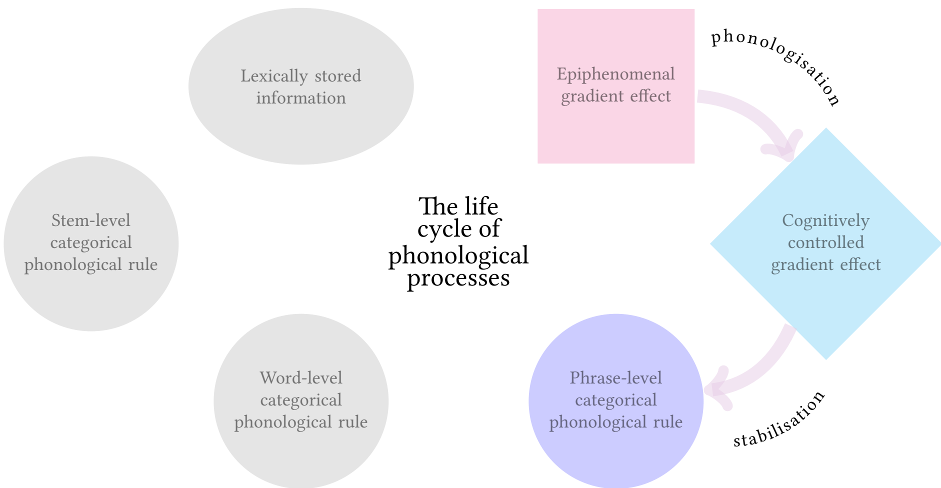

Figure 6.5: The life cycle of phonological processes: the Manchester speaker may be in any of the first three stages. Adapted from Ramsammy (forthcoming).

logically in the same way, and shows phrase-level darkening, as in Table 6.3. In terms of phonetics the speaker may have a system of dark [ɫ] vs. 'even darker' [ɫ]. This would be the third stage in Figure 6.5.

ii. This small amount of tongue root backing may be due to phonetic effects, such as the phrase-final extra durational effects resulting in additional time for the dorsal component to reach its full potential. This could be represented with one of the first two stages in Figure 6.5.

From this dataset, we cannot be sure. However, the following chapter deals with much larger datasets to specifically address such questions.

6.3 Middlesbrough

Middlesbrough is an interesting variety for a sociolinguistic perspective, given it was formally part of Yorkshire and is now considered part of the North-East. With respect to other features of the accent, Middlesbrough patterns more closely with those varieties found further North in Tyneside. However, the claim that /l/ is always light in the North-East has only been confirmed for towns and cities as far as Northumbria and Durham, with no description existing for Middlesbrough.

The splines in Figure 6.6 suggest that, unlike the reports for their neighbours on Tyneside, Middlesbrough do indeed have a distinction between initial and final /l/s. Although this distinction is not as stark as in RP, the magnitude between the two extremes does seem to be more convincing than in Manchester. In general, the phrase-level onset /l/s have a higher tongue body, and the phrase-level coda /l/s have a reduced tongue tip gesture and retracted tongue root. However, word-initial ***l**eap* is slightly retracted in the tongue root area for this speaker, showing an intermediate realisation between *hea**l***-type tokens and other onset realisations. Again, this is most likely due to the coarticulatory pressure on the intervocalic /l/s, which are neighboured by two high front vowels, as opposed to just one. However, there may be an extra effect here which is not present in all speakers. When analysing these tokens more closely, it is clear that some of these phrase-initial tokens of /l/ are articulated with a drawn-out run-up to the word, with a very elongated /l/ sound. These realisations have a backing effect on the overall mean spline. Such an effect could have been avoided by embedding the word-initial /l/s within a phrase so they were not phrase-initial. For Experiment 2, participants were given ***l**eap* tokens which were non-initial in the utterance, because of the risk of this happening in utterance initial position, but sadly this could not be avoided for Experiment 1.

Leaving aside the extra dark initial /l/s, we can see a clear three-way distribution in Figure 6.6. Syllable initial ***l**eap*-type tokens are the lightest, word-final *hea**l***-type tokens the darkest, and word-final preconsonantal /l/s are intermediate between the two. This could be consistent with a three-way distinction, but is most likely the result of a two-way allophonic distinction between phrase-level onset and coda /l/s, with an added gradient effect of phrase-final darkening.

Consideration of the acoustics in Figure 6.7 adds further support to the idea that the extra tongue root retraction in phrase-final *hea**l***-type tokens is simply a positional gradient side effect. The F2 and F1 difference between the word-final phrase-final and pre-consonantal contexts is almost non-existent. Once again, this demonstrates how the articulations may not always be clearly represented in the acoustics. Note that the average formant values for this speaker are much higher than Manchester, showing how much lighter the /l/s are overall. For this speaker, the data suggest that Middlesbrough is similar to RP in terms of its phrase-level coda /l/-darkening system, but phonetically the extremes of the distribution are on a much smaller scale (see Figure 6.14 for a demon-

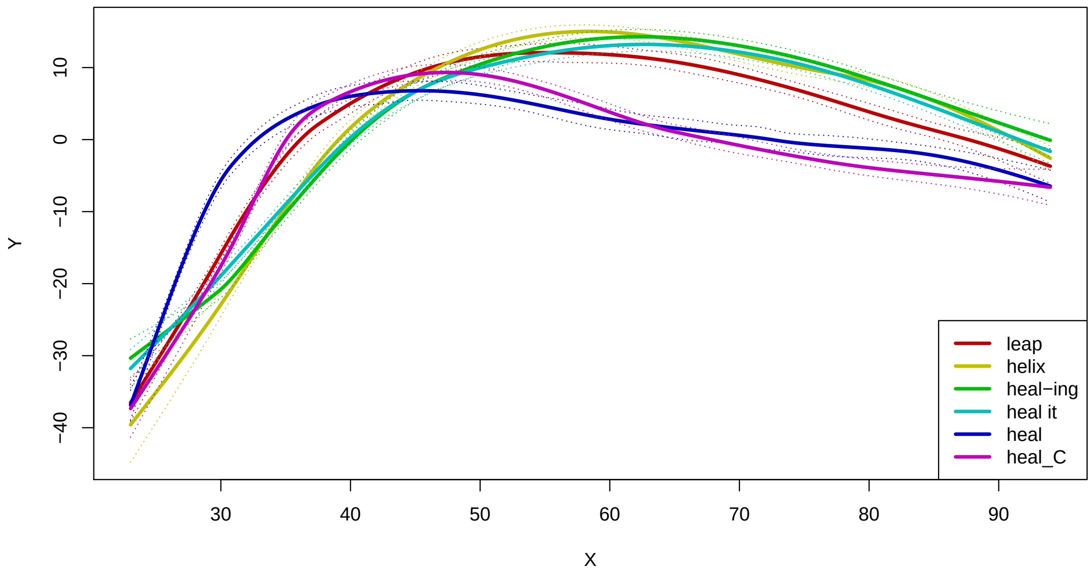

Figure 6.6: SS ANOVA of Middlesbrough speaker

stration of this point).

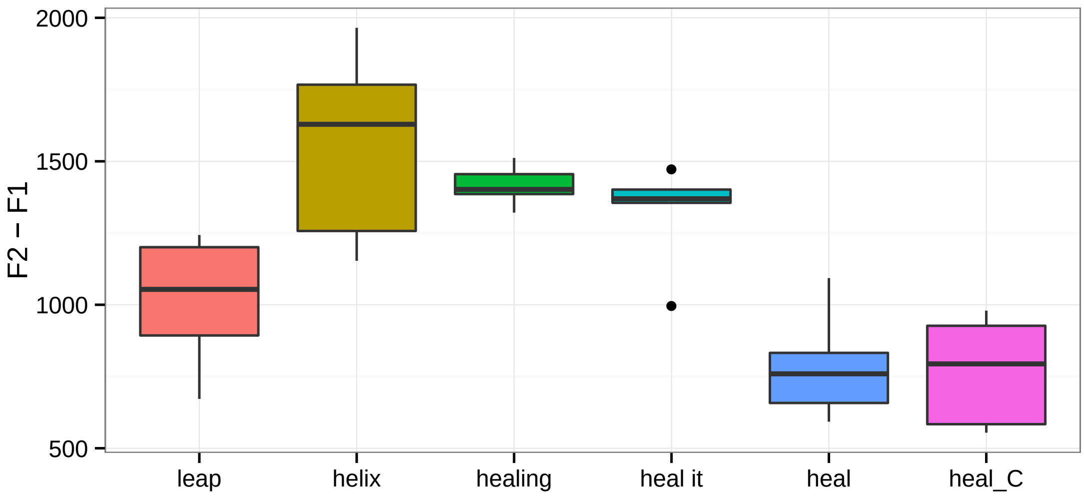

Figure 6.7: Middlesbrough F2-F1

6.4 American English

American English is often described in the literature as having dark [ɫ]s in all positions. Unlike the varieties of English spoken in England, numerous studies have covered the phonetics of American English /l/ both acoustically (van Hofwegen 2010; Olive et al. 1993) and articulatorily (Gick et al. 2006; Lee-Kim et al. 2013; Sproat and Fujimura 1993). These have shown that, despite American speakers showing very dark [ɫ]s in most positions, there is still a difference between different phonological contexts.

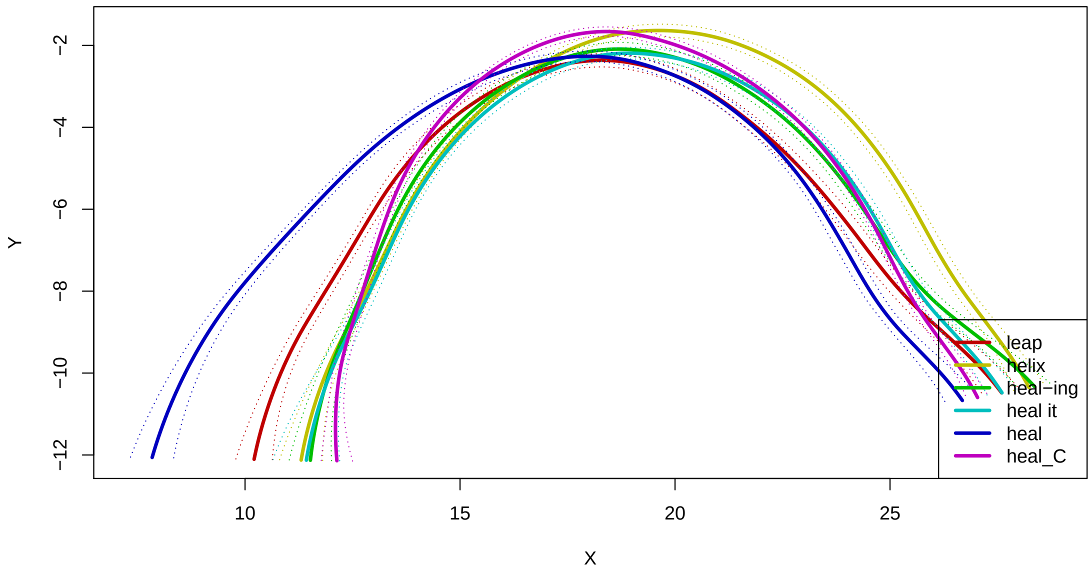

Figure 6.8: SS ANOVA of American English speaker

The first thing to notice about the American speaker in Figure 6.8 is that the splines are largely overlapping, but in a much more varied way than the Manchester speaker we saw in Figure 6.3. They also seem to be much more palatalised than the British speakers we have seen thus far. On closer inspection, there does seem to be a little more orderliness to the initially perceived chaos. Indeed, the phrase-final *hea**l***-type tokens are clearly more retracted for this speaker, in line with all previous speakers of British English. Surprisingly, word-initial ***l**eap* is the next most retracted spline for this speaker, showing an intermediate realisation between *hea**l***-type tokens and the rest of the distribution. As discussed in Section 6.3, this is likely a combination of the reduced coarticulatory pressure on this utterance initial /l/ with only one flanking high front vowel, as well as this speaker's tendency to articulate the /l/ with a drawn-out run-up to the word. These tokens drag the average ***l**eap*-type tokens' mean so that the spline looks more backed than perhaps expected.

Moving on from this token and considering the four remaining contexts, the SS ANOVA shows an additional distinction between intervocalic *he**l**ix*-type tokens on one hand, and *hea**l**-ing* and *hea**l** it* tokens on the other. The confidence intervals indicate a small but significant difference between tongue shapes, this time with a three-way distinction: the *he**l**ix*-type tokens have significantly advanced tongue-tips than the other environments. In addition, the phrase-final *hea**l***-type tokens have an extra amount of tongue-root retraction. In the pre-consonantal cases, we do not find the extra tongue-root retraction present in phrase-final *hea**l***-type tokens, and the tongue tip is not as low. Figure 6.9 shows the same mean splines plotted in `ggplot`, with the ***l**eap* tokens removed for clarity. Here we can see again that the *he**l**ix*-type tokens are indeed more advanced in the area towards the front of the tongue, and that this is significantly different from both the *heal-ing* and *heal it* tokens. There is some evidence for a three-way darkness distinction that is partially obscured by drawing-out of the odd *leap*-type token. From lightness to darkness this three-way distinction shows: *leap*, *helix*> *heal-ing*, *heal it*> *heal*.

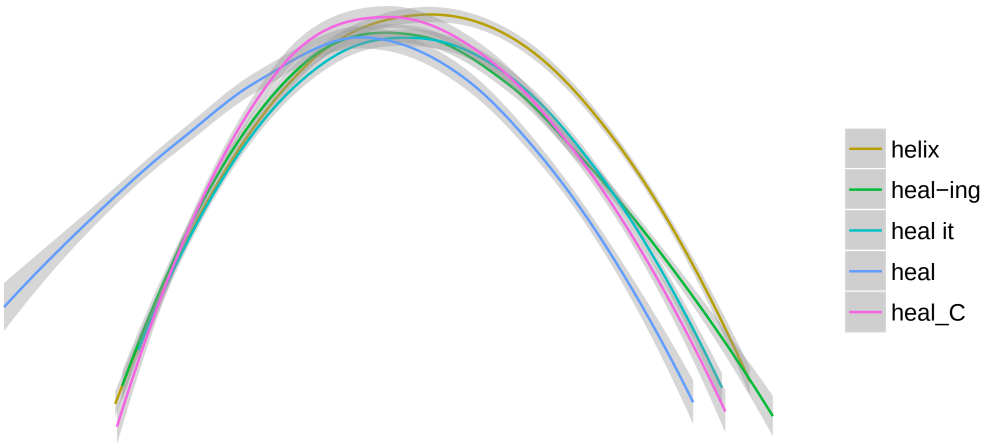

Figure 6.9: American English splines (with initial *leap* tokens removed for clarity)

The acoustics show a step-wise pattern from darkest to lightest, which could be interpreted as gradient, but the three potential categories do not overlap with each other in particular. Figure 6.10 shows this distribution in both box plot and jitter plot format (the latter shows the individual points). The Tukey HSD confirms the three-way distinction partially, but the *heal it*-type tokens are not significantly different to the phrase-final and pre consonantal tokens. That is, they are significantly darker than *leap* and *helix*, and show no difference with *heal-ing*, as expected, but they fail to show a significantly lighter realisation than the third possible category in the acoustics. This is interesting, but it is difficult to draw conclusion of a two-way or three-way distinction based on so few tokens. The effect that does remain, however, is the effect of stem-level darkening. Intervocalic /l/s show significantly different levels of darkness depending on whether there is a stem-suffix boundary present, giving darker /l/s in *heal-ing* than in *helix*.

Similarly to the Manchester data in Section 6.2, there are several possibilities for this speaker's system. Firstly, the speaker may show three categories of /l/, much like the lenition trajectories discussed in Chapter 3. Moreover, the differences in the root and tip area may be indicative of processes targeting different articulators. For example, the pattern displayed by the American English speaker may represent two synchronically overlaid processes in the grammar. The first process reduces the tongue-tip gesture and applies at the stem-level, darkening the pre-suffixal /l/ in *heal-ing*, but not in the intervocalic /l/ in *helix*, where the /l/ is in the onset at all levels. The second process affects tongue root retraction and applies at the phrase-level, further backing the tongue root in

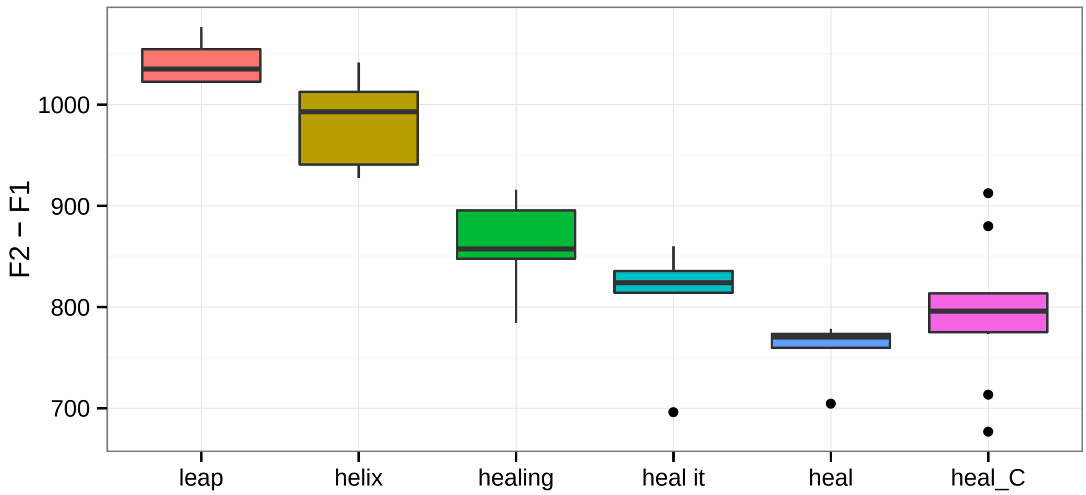

Figure 6.10: American English F2-F1

phrase-final *heal*. Lee-Kim et al. (2013) see three categories as problematic in their study of American English /l/. This is unlikely to be due to a dismissal of three categories altogether (although they may do), more a question of whether three cognitively defined categories can exist in such a small phonetic space. However, this three-way distinction is problematic for this data from an articulatory perspective, as the pre-consonantal tokens pattern with *heal-ing* and *heal it*-type tokens, not with *heal*-type ones. This does not show in the acoustics, however.

Therefore, a more convincing analysis would be that phrase final *heal*-type tokens exhibit an added durational effect created by rule scattering: the duration-driven gradient effect that was the diachronic precursor of categorical /l/-darkening remains synchronically active in the phonetic module. This applies across the board in the phonetics. Under this analysis, we still have two processes synchronically overlaid in the grammar, but rather than an older, milder form of categorical darkening applying in relatively narrow domains feeding a younger, harsher form of categorical darkening, we have a phonological process manipulating discrete categories overlaid with a phonetic process effecting gradient adjustments. This is represented in Table 6.4. The second row of the table, showing extra tongue root retraction, could be interpreted as a gradient phonetic process effecting phrase-final /l/s, or a categorical process reflecting the next stage of the lenition trajectory.

Reduced tongue tip	***l**eap*	*he**l**ix*	*hea**l**-ing*	*hea**l** it*	*hea**l***
Retracted tongue root	***l**eap*	*he**l**ix*	*hea**l**-ing*	*hea**l** it*	*hea**l***

Table 6.4: Interpretation of American speaker's pattern, dividing the distribution into three.

6.5 Essex

As a variety spoken in South-East England, Essex is well known for its /l/-vocalisation. Although vocalisation is traditionally associated with the working-class Cockney dialect of London, recent linguistic observations indicate that /l/-vocalisation is perhaps more widespread and advanced in Essex. Vocalisation is reported widely in phrase-final position and pre-consonantally in Essex, but also before vowels in advanced speakers.

As discussed in Chapter 3, /l/-vocalisation represents a third stage of /l/ variation on the lenition trajectory. Whereas in dark [ɫ] the consonantal tongue tip gesture is reduced, in vocalised /l/ the gesture is absent altogether. A drawback of ultrasound is that we cannot see what the actual tongue tip is doing and so cannot be sure of the presence, absence or overall magnitude of contact. However, although many articulatory studies seek to analyse gradience in /l/-vocalisation effects, they usually acknowledge that vocalisation has a categorical component, at least for some varieties (Hardcastle and Barry 1989; Scobbie and Wrench 2003). This is certainly the case for Essex, where the vocalised element is audibly vocalic, and very different to the kind of dark [ɫ] /found in Northern varieties or in Standard English. Although it is not possible to represent visually here, audible identification is helpful to confirm that /l/ is indeed vocalised, as acoustically it is difficult to distinguish between a dark and vocalised /l/ (Hall-Lew and Fix 2012) and articulatorily we would need a method such as EPG to be 100% sure the tongue tip contact was truly absent.

This speaker was approached to take part in the experiment due to her strongly vocalised /l/s in natural speech. Of course, the speaker was not made aware of this fact and did not know what the experiment was testing. Figure 6.11 shows the her mean splines with `loess` smoothers. This Figure is shown to demonstrate the speaker's pre-consonantal /l/s, the *heal_C* environment. During the recording, it was audibly clear that the speaker was 'hypercorrecting' in this environment. That is, her usual naturally vocalised /l/s were replaced with an unnatural sounding lighter variant. As unexpected as this was, it should perhaps not be surprising, given previous studies of British English have reported the exact same effect in their wordlist data. Hughes et al. (2012:81) describe this effect in one of their London speakers as sounding 'oddly exaggerated', which is an apt description of the tokens produced by this Essex speaker. This unusual realisation is reflected in the articulation, shown in Figure 6.11.

This is interesting from a sociolinguistic and a sociophonological perspective. Firstly, speakers do not tend to style-shift with /l/-darkening; at least, no other speakers in this experiment have shown such a pattern. There are no other reports of style-shifting with /l/-darkening in the existing literature. This could be because it is less audibly salient, but it is more likely that this third stage of the lenition trajectory is subject to style-shifting for a number of reasons. Firstly, it is a newer linguistic change, reflected in the fact that it tends only to occur finally at the phrase-level and that it functions as a more

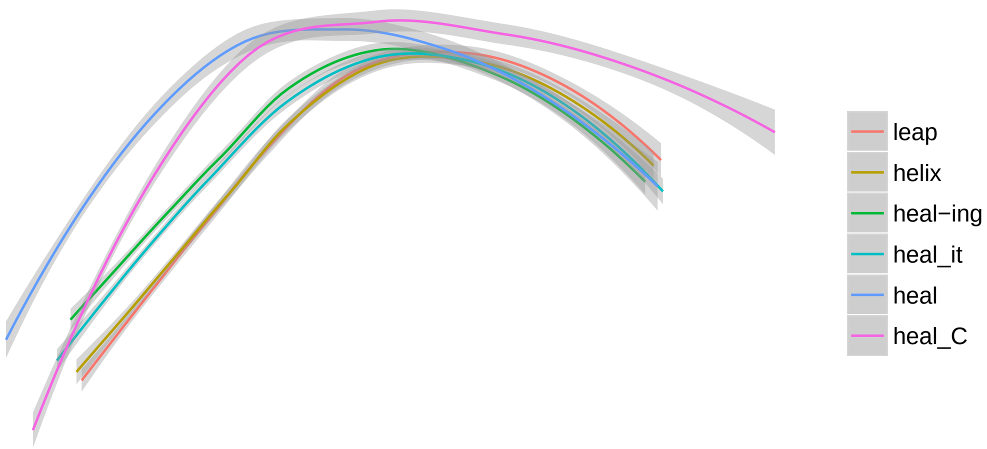

Figure 6.11: Essex splines

advanced stage of /l/-darkening. In addition, reports from the sociolinguistic literature indicate that this is a change is progress, in that younger speakers are vocalising much more often than older speakers, and is found increasingly more in different areas of the UK (Britain 2009; Docherty and Foulkes 1999). Luckily, the speaker retains her natural vocalisation in phrase-final ***heal***-type tokens. This is interesting, as one may expect both vocalising environments to behave the same, although this has not necessarily be found by all e.g. Scobbie and Pouplier (2010). However, as we have seen to some extent in RP and Manchester, and more clearly in Middlesbrough and the American English speaker, it is the phrase-final tokens which tend to have the darkest realisation overall. It looks as though this environment is where categorical vocalisation originates and the Essex speaker is potentially reflecting this fact in her style-shifting methods. This interaction between sociolinguistics and the life cycle of phonological processes is a topic for future research. However, for this study, due to the unnaturalness of the word-final prevocalic tokens produced by the Essex speaker, this environment was removed from the SS ANOVA test in Figure 6.12.

The SS ANOVA in Figure 6.12 shows the distribution of the remaining tokens. The first thing to note is just how backed the vocalised /l/s in comparison to the remaining four contexts.

It is therefore important to highlight that these realisations may well involve quite robust pharyngeal constriction. The degree of pharyngealisation involved in phonetic production of phonologically vocalised /l/ is not something that has ever been subjected to rigorous instrumental study. On the one hand, if /l/ vocalisation is defined just as the loss of the tongue-tip gesture, then one might not expect to observe greater or lesser amounts of tongue backing in vocalised tokens as compared to consonantal dark realisations. On the other hand, if vocalisation is defined in phonological terms as the loss or change of feature values, then one might expect to observe more robust phonetic differences between vocalised /l/ and consonantal dark /l/. For this speaker, it is clear, after all, that the magnitude of tongue-root retraction is considerable, whereas the tongue tip area does not seem to differ in any significant way from /l/ tokens in other phonological environments. Further research on the articulatory consequences of vocalisation must therefore be undertaken before this effect can be fully understood.

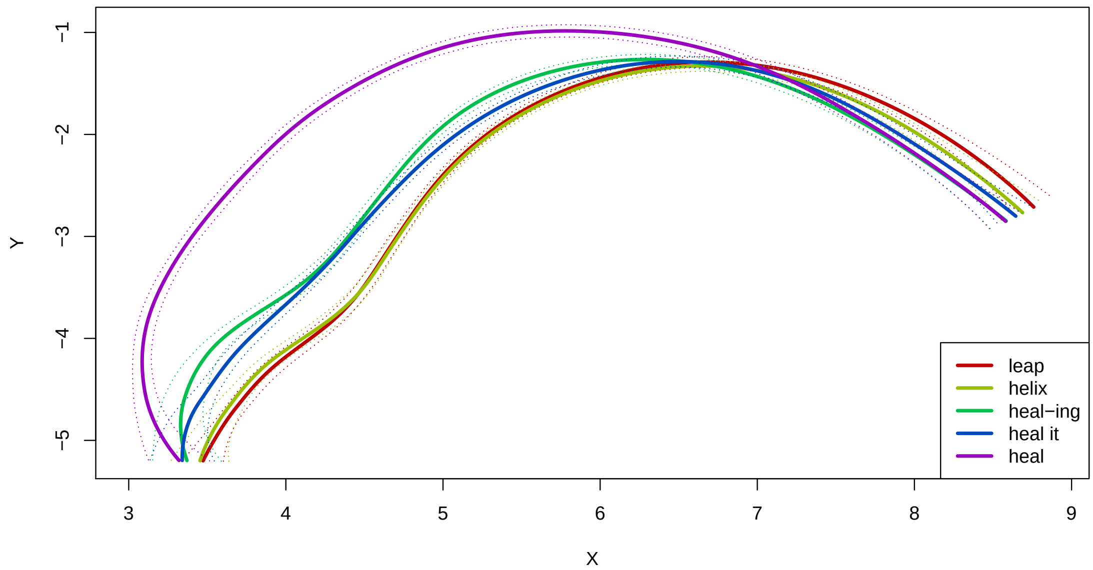

Figure 6.12: SS ANOVA of Essex speaker

In spite of this, we do observe very similar articulatory configurations with a more fronted tongue root in the remaining four environments where /l/ is in the onset at the phrase level. However, the SS ANOVA shows a distinction in this area between the ***l**eap* and *he**l**ix*-type tokens on one hand, and *hea**l**-ing* and *hea**l** it* type tokens on the other. This suggests a three-way distinction. This can also be observed, albeit not as clearly, in the acoustics plot in Figure 6.13. The Tukey HSD only finds a two-way difference here, between the *hea**l***-type tokens and the other four contexts. The extra tongue root retraction in the *hea**l**-ing* and *hea**l** it* environments is not picked up in the acoustics.

Combining auditory, acoustic and articulatory evidence from the remaining tokens, it seems convincing that the Essex speaker shows a three-way lenition trajectory in /l/ realisations. There are similarities with the American speaker: both have a three way distinction between ***l**eap*-type and *he**l**ix*-type tokens, *hea**l**-ing* and *hea**l** it* type tokens, and *hea**l***-type tokens. The question is whether or not the data here can be claimed to show three categories, as for the American speaker the evidence for three categories was not particularly strong. It is worth noting that Lee-Kim et al. (2013) see three categories as problematic, but our Essex speaker provides support for this possibility. However, the American pattern is much less distinct phonetically and conclusions of categoricity and gradience can not be made from these data alone, although tentative conclusions could be argued for. The Essex pattern could reflect the operation of two overlaid cate-

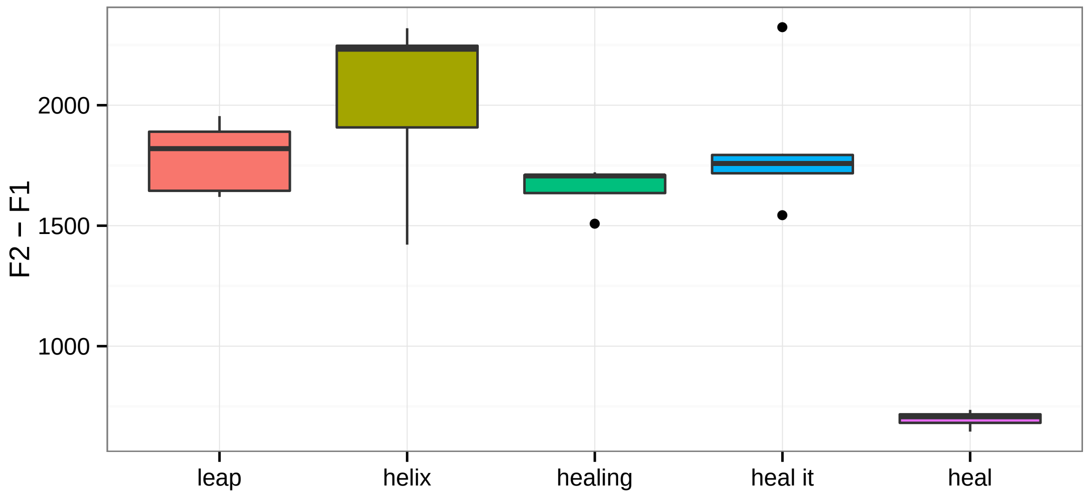

Figure 6.13: Essex F2-F1

gorical phonological processes: one controlling tongue-root retraction, the other linguo-alveolar contact. The American pattern may be down to categorical darkening, superimposed on a gradient duration-driven adjustment in final position. In both scenarios, we have two processes synchronically overlaid in the grammar. This pilot study has the same problem as Lee-Kim et al.'s: such a coarse sample of the possible phonological environments makes it difficult to distinguish between categorical and gradient effects reliably, at least in dialects whereby the articulations are so similar. However, the vocalisation in the Essex speaker's recording is clearly audible. In the case of this speaker, it is not the distinction between the intermediate and darkest tokens which seem the weakest, but the distinction between the initial tokens and the intermediate ones.

	__l__eap	*he__l__ix*	*hea__l__-ing*	*hea__l__ it*	*hea__l__*
American English	[l]	[l]	[ɫ]	[ɫ]	[ɫ]
Essex	[l]	[l]	[ɫ]	[ɫ]	[ɤ̈]

Table 6.5: Interpretation of Essex and American English patterns.

6.6 Acoustic comparison

Although the articulatory data give us a unique insight into variation, alongside some patterns which obscure the acoustics, one drawback is the difficulty with inter-speaker comparison. The acoustics, on the other hand, can give us a very clear idea of each individual speaker's range of darkness in comparison to other varieties. Figure 6.14 shows the speaker's distributions on the same scale, using the Bark normalised formant values (B1 and B2 for F1 and F2). Table 6.6 shows the original formant values used to make the individual plots in the sections above, and Table 6.7 shows the values after they have undergone the Bark normalisation transformation. As discussed in Chapter

5, Bark normalisation is recommended to eradicate potential inter-speaker differences, and so is a better measure for the cross-speaker comparison in Figure 6.14. However, it is also more reflective of the auditory scale i.e. what the listener hears, when ideally we want to compare the acoustics with the articulatory data, so we are more interested in the original values for the individual speaker plots in the previous sections.

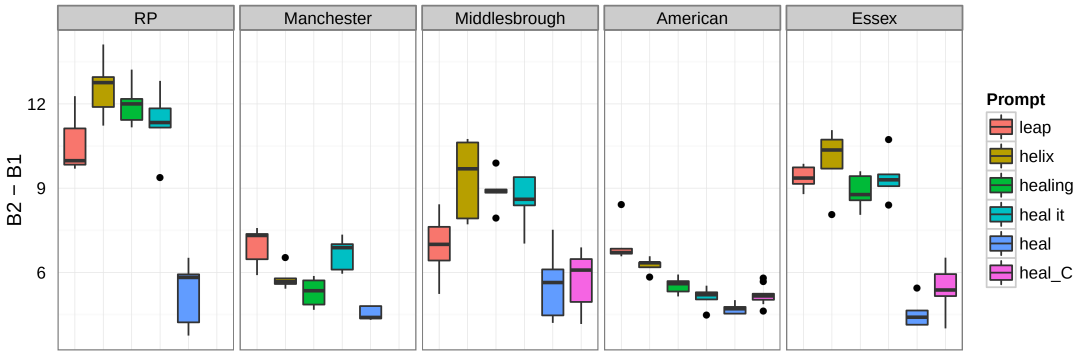

Figure 6.14: Bark normalised F2-F1 across all speakers

	leap	*helix*	*heal-ing*	*heal it*	*heal*	*heal_C*
RP	1784	2321	2277	1986	668	
Manchester	1146	967	884	1133	754	
Middlesbrough	1013	1554	1415	1319	787	768
American Eng.	1042	983	860	806	757	795
Essex	1786	2026	1656	1827	697	

Table 6.6: Mean F2-F1 measurements for each context for all speakers

	leap	*helix*	*heal-ing*	*heal it*	*heal*	*heal_C*
RP	10.65	12.59	12	11.31	5.25	
Manchester	6.93	5.8	5.29	6.66	4.54	
Middlesbrough	6.94	9.34	8.9	8.56	5.59	5.71
American Eng.	6.7	6.26	5.54	5.11	4.71	5.19
Essex	9.38	9.98	8.88	9.4	4.55	

Table 6.7: Mean B2-B1 measurements for each context for all speakers. Bark normalised values of formants.

The first thing to note Figure 6.14 is the RP speaker's considerably wide range between initial and final /l/, which only Essex comes close to matching. Given that the word-final /l/s in Essex are vocalised, we might expect such a stark difference between this context and the onset /l/s. The difference between the RP's speaker's range and the American is truly considerable. This kind of perspective gives an insight into why so many phonetic studies of American English /l/, such as Sproat and Fujimura (1993) and Lee-Kim et al. (2013), conclude a lack of allophony for this variety. Firstly, it is problematic to dismiss categoricity in /l/-darkening as a possibility by looking at one variety, and if one is to do this, many phonological environments are needed.

6.7 Preliminaries to a temporal analysis

As discussed in the introduction and at several other points in the thesis thus far, due to the frame rate of the ultrasound machine used for the present study, a temporal investigation proves problematic. Rapid gestural movements may be missed by the coarse monitoring, resulting in a limited analysis. This kind of limited overview is possible, however, and these results are presented in this section. A temporal analysis may not be as all-important as may have been suggested by previous studies, and I will demonstrate this claim using speakers from this chapter.

Sproat and Fujimura (1993) show that for their American speakers, dark [ɫ] in comparison to light [l] has a) a larger dorsal displacement and b) a delayed coronal gesture relative to the dorsal gesture. We can infer from this that the articulatory correlate of darkening may rely on a temporal analysis, and many since have worked off this assumption. However, in this section I will demonstrate that the overall dorsal retraction magnitude suffices as the primary articulatory measure, at least from some speakers.

Figure 6.15 shows initial and final tokens over the full course of the /l/, selected from the RP speaker's data. As can be seen, we only have three splines per /l/. The splines are coloured on a gradient scale reflecting the time point during the /l/: blue splines reflect the earliest tongue shapes, black reflect the later tongue shapes, with gradience in between. As Figure 6.15 shows, regardless of gestural ordering, the splines during the course of the /l/ tend not to move much. The main facts reflect that the initial and final /l/s are an entirely different shape throughout the course, and the tongue dorsum retraction remains stable. It is not the case that initial and final tokens consist of the same gestures which are mirrored temporally over the course of the /l/, at least not for this speaker. Of course, it would be preferable to have the entire temporal phasing on top of this information, but Figure 6.15 demonstrates that the midpoint is perfectly suitable to indicate the articulation of the /l/ in RP. There are no covert gestures that the midpoint measurement is missing. However, it does highlight the fact that small differences in a certain area, such as the tongue root, may be due to catching the tongue at a slightly different point in the /l/ from frame to frame. This is worth pointing out, but is a minor concern, as the final splines are averaged out over all tokens and the statistical tests can account for small deviances such as these.

Although the RP speaker shows a clear and consistent distinction throughout the course of the /l/, it is not possible to draw conclusions about all speakers from Figure 6.15. Sproat and Fujimura's speakers were primarily American, or American influenced, which are often described as having very dark [ɫ]s throughout the spectrum of possibilities. In fact, as discussed in Chapter 4, many phoneticians have claimed there is no allophonic distinction when studying American English. Therefore, it might be useful to compare the temporal properties of speakers who do not have this articulatory categorical distinction between initial and final position. In this case, the temporal

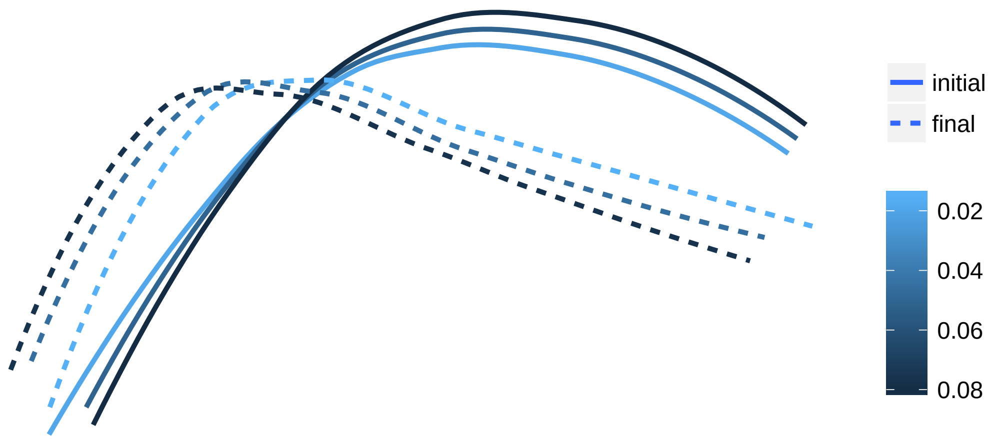

Figure 6.15: Tokens of initial and final RP splines over the course of the /l/.

analysis may seem more of a necessity than for RP.

It would have been preferable to use the American speaker for this comparison, but her data image was particularly poor and only two splines per /l/ could be extracted on for most tokens, due to fuzzy imaging, which is not really enough to show presence or absence of phasing. This is disappointing, and is certainly something that will be attempted again, either with new American speakers, or on an ultrasound machine with a higher frame rate. For now, we shall use another speaker with a small articulatory difference between initial and final /l/.

Figure 6.16 shows the possible gestural phasing in the Manchester speaker as an alternative. This is still a worthy comparison, as we are comparing the speakers with the most extreme and least extreme phonetic continuum. As discussed, we may expect varieties such as Manchester, with little difference between initial and final /l/s, to rely more on gestural phasing as a way of signifying the allophonic category of the /l/. Figure 6.16 provides some evidence for this. Although the phrase final /l/s, with the dashed lines, seem to be remain consistent temporally, the initial /l/s do show some movement. Again, the closer to blue the colour is, the earlier in the /l/ it was articulated. Therefore, the initial shape of the initial token is further back and almost looks velarised in the tongue dorsum area. As the sound continues, the splines become gradually black, and the position becomes more fronted in the mouth. This is perhaps unexpected given the Sproat and Fujimura (1993) evidence: the coronal gesture precedes the dorsal gesture in initial /l/. However, it could be argued that the Manchester /l/s are all dark, even if some are not as dark as others, so we expect this raising towards the velum to occur first, and the most front shape of the tongue being achieved last of all. Although, looking at the RP data, the same pattern (albeit to a less convincing extent) as the Manchester speaker can be observed. In initial /l/, the most front articulation is achieved last, and in final position, the most retracted position is achieved last. It would not be wise to argue for an alternative description of the temporal properties of British English /l/ here, as the picture is far too coarse, and there could be tiny gestures which have been completely missed by the ultrasound. However, these patterns would benefit from a full articulatory analysis which has its focus on temporal properties. See the discussion for future research in Chapter 8 for more information on this. For now, the temporal investigation is concluded by arguing that the midpoint of the spline gives us a highly interpretable insight into the articulatory properties of /l/. For many speakers, this may be the standalone primary articulatory correlate of darkening.

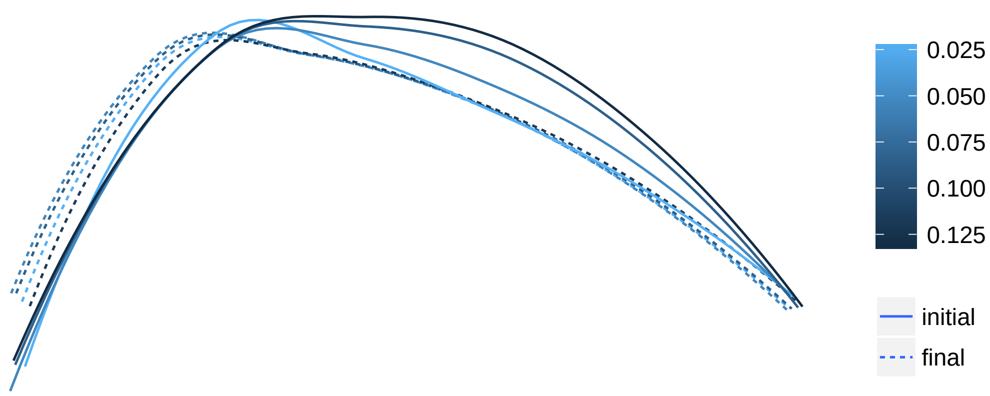

Figure 6.16: Tokens of initial and final Manchester WC splines over the course of the /l/.

6.8 Summary

This chapter has shown that dialectal diversity has been vastly underestimated in the existing literature on /l/-darkening. We need a theory that can account for the evidence that categorical darkening domains may differ in size between dialects. The life cycle of phonological processes can make sense of such facts, with domain narrowing accounting for the coexistence of categorical and gradient effects, as well as accounting for lenition trajectories such as /l/-vocalisation. The evidence for rule scattering is seen through extra effects of darkening in phrase-final position, but it is difficult to be sure of these claims with so few tokens. This issue will be considered properly in the following chapter. The suggestions of such processes here, however, can be accounted for under the life cycle. Overall, this chapter has shown that the wide range of dialectal diversity in /l/ lenition processes, for which this chapter provides only a small subset, shows a great deal of orderliness if considered from the viewpoint of the life cycle.

In addition, the temporal analysis presented in this chapter provides evidence that measuring the articulatory midpoint of /l/ performs very well at providing the primary articulatory correlate of darkness. For speakers with a clear articulatory distinction between initial and final contexts, a temporal analysis may even be superfluous. However, for those speakers with a much more fine grained phonetic difference, a temporal perspective may be of use. This may have been even more useful in the following chapter when trying to distinguish between categorical and gradient processes, but sadly it is not possible in this study.

In the following chapter, we shall follow up on some of the points raised by patterns found in this chapter, namely, how can we conclude whether a distinction is categorical and phonological, or gradient and phonetic?

Chapter 7

Experiment 2: Categorical vs. Gradient Processes

7.1 Introduction

The primary aim of this chapter, and one of the main goals of the thesis overall, is to address the ongoing debate in the literature on whether variation in /l/-darkening is categorical or gradient. As we have seen, traditional phonological approaches to /l/-darkening posit two categories, light and dark /l/, but this categorical distinction has been challenged by phonetic studies, whose findings suggest the allophones are simply two ends of the same continuum. Such claims have implications for the architecture of grammar: for example, if /l/-darkening is purely gradient and at the same time sensitive to morphosyntactic structure, then it follows that phonetics must have a direct interface with morphosyntax, bypassing phonology (Kawahara 2011:2290-1; cf. Bermúdez-Otero 2013:2.4).

In Chapter 6, we saw that quantitative tests such as SS ANOVA can help us to decide whether two or more splines are significantly different from one another. However, the small number of phonological environments studied in Experiment 1 made it difficult to determine whether two significantly different splines represented two allophonically distributed phonological categories, or merely significant effects of phonetic gradience. This chapter presents data from Experiment 2, which attempts to address the question of categoricity and gradience by presenting data from 10 phonological contexts: the idea is that, by sampling the range of morphosyntactic and prosodic environments for /l/-darkening much more finely, we will be able to decide whether two significantly different realizations belong to discrete categories, or merely represent different points in a single continuum. This section will address each speaker in turn, weighing up the evidence for each side of the debate by considering the evidence from ultrasound tongue imaging, supported by acoustics, interactions between darkness and duration, and quantitative methods including Principal Components Analysis (PCA) (Section 5.5.2) and linear regression (Section 5.5.5).

Experiment 2

Context	Example
1. word-initial	***l**eap*
2. stem-medial pretonic	*be**l**ieve*
3. suffix-initial	*free-**l**y*
4. intervocalic	*he**l**ix*
5. stem-final presuffixal	*pee**l**-ing*
6. compound boundary	*pee**l**-instrument*
7. word-final phrase-medial	*hea**l**#* V
8. phrase-final	*hea**l**#*, V
9. utterance final	*pee**l***
10. word-final pre-consonantal	*pee**l** bananas*

Table 7.1: The ten phonological environments studied for Experiment 2 with example tokens

7.1.1 The experiment

In this chapter, the results of Experiment 2 are presented, drawn from eight speakers from different places in the UK. Speakers read sentences with tokens of /l/ framed in ten phonological contexts, as shown in Table 7.1. Previous studies have shown that a wide range of phonological contexts is needed to provide insight into the entire spectrum of /l/-darkening possibilities. The full details of the experiment are outlined in Chapter 5, but to summarise, speakers were presented with two different sentences per context, and these were shown five times each, resulting in 100 /l/s per speaker. In this chapter, the phonological contexts are referred to by the example token in the right column of Table 7.1.

7.1.2 Data analysis

This subsection summarises the methodology which is fully explained in Chapter 5. Each section of this chapter will firstly present the midpoint /l/ splines across 10 phonological contexts for the speaker in question. Quantitative analysis is performed on the articulatory data using the PCA values as a proxy for darkness. A high PC1 indicates a light /l/, and a low PC1 a dark one. Density plots are used to visualise the distribution of PC1 in order to decide whether it shows bimodality, a possible indicator of two categories. This conclusion is supported by Hartigan's dip test, which is applied to quantify any potential modality, and provides a p-value of statistical significance. Correlations of darkness and duration are represented by means of the scatterplot, and correlations of Pearson's r value are added to smoothing lines. Acoustic findings show the F2-F1 difference, which is the primary acoustic correlate of /l/ darkness (high values indicate lighter /l/s, low values darker /l/s).

Category assignment in this chapter is done by two means. In the general instance, category assignment is done *a priori*. Thus, two categories were initially set up: one consisting of tokens of /l/ which are in the onset at the word level, the other consisting of tokens of /l/ which are in the coda at the word level. The word-level onset class corresponds to environments 1 to 5 in Table 7.1, on the assumption that syllabification is onset-maximal at the stem level, and that at the word level there is resyllabification before vowels in the same prosodic word (for arguments against ambisyllabicity, see Section 3.2.1). In turn, the word-level coda class comprises environments 6 to 10 in Table 7.1. The expectation is that word-level onsets are likely to be light, and word-level codas will be dark. The extent to which expectation is borne out can then be exactly gauged using linear modelling: data which shows a categorical distinction along the expected lines will provide a close fit, whereas other distributions will display a loose fit.

In addition to this, however, evidence for categories is investigated through a Tukey HSD post-hoc test of significant differences between the ten contexts, articulatory and acoustically (outlined the Chapter 5). If the *light* and *dark* categories are significant across category but not within, then this provides evidence for treating them as two separate allophones. For the most part, the Tukey HSD confirms the *a priori* decision to split the ten environments in half, giving a predicted *light* label to the first five contexts and a predicted *dark* label to the second five. As we shall see, this distinction is not the best for all speakers, and is shifted for those showing obvious alternative patterning. For example, if by Tukey HSD a speaker shows that the *pee**l**-index*-type tokens (context 6) are significantly different to all of the *dark* contexts, but not significantly different to the *light* ones, this provides a strong case for placing this context into the *light* category instead. In some cases, the results may indicate the possibility of three categories. As the Tukey HSD test involves 45 comparisons per speaker, as well as a p-value table comparing the 10 contexts, the full results are displayed in Appendix C, and a summary is provided in this chapter.

To avoid the assignment of individual categories based on one potentially sporadic p-value resulting from 45 possible comparisons of ten contexts, a Tukey HSD is run on both the articulatory and acoustic correlates of darkness to ensure consistency in potential additional categories, or category switching. However, if there were to be robust and consistent patterns in the articulations which do not appear on the acoustics, this will also be taken into account. Obviously, one of the benefits of ultrasound tongue imaging is the ability to observe patterns which evade the acoustics. The aim here is to not only conduct category formation on the basis of both production and acoustics, but to weed out any potential flagging of significance which is really just a fluke of the data, resulting in robust and reliable category assignment.

The conclusions based on a categorical distinction do not end here. As mentioned above, the closest possible categories for each speaker are passed to a linear model, and the fit of this model is assessed: i.e. a strong fit indicates that category may play a role. The combination of all of these analyses is used to diagnose categoricity in the summary.

7.1.2.1 Excluded tokens

Some of the carrier sentences chosen resulted in the production of unexpected vowels which had a coarticulatory effect on the /l/. For example, the word *linoleum* is listed in the OED as /lɪnəʊlɪəm/, but some speakers produced the first vowel as a schwa, rather than a high front /ɪ/. This has an coarticulatory effect on the /l/, so we observe a small but significant difference in the realisation of the /l/ in this word, compared with the other carrier word *Leoni* for most speakers. When this effect was spotted, the remaining speakers were recorded with a new sentence which included the carrier word *Leoni.* For the earlier speakers who pronounced *linoleum* with a schwa in the first syllable, these five tokens were removed from the analysis, as they were not deemed to be consistent with the other environments. The full figures can be viewed in Appendix B. This highlights the importance of keeping vowels consistent across tokens, as neighbouring vowels have a significant effect on the quality of the /l/.

7.2 RP

In Chapter 6, we saw that the RP speaker recruited for the experiments displayed a conspicuously bimodal distribution between phrase-level onset and coda /l/s, in articulatory terms. His onset /l/s displayed a typically light realisation, with an advanced tongue tip and fronted tongue root. In contrast, the coda tokens, which were represented by just one phonological environment in Experiment 1 (phrase-final ***peel***-type tokens), showed a retracted tongue root, lowered tongue body and reduced tongue tip gesture. One of the drawbacks of Experiment 1 was that, with so few tokens (just 25 for the RP speaker) it was difficult to conduct any quantitative analysis other than the SS ANOVA, which was accompanied by a descriptive overview. Additional quantitative commentary, such as correlations and linear models, is simply unreliable for such small datasets. For the RP data in Chapter 6, this is not too much of an issue due to the clear and convincing distribution in the spline contours, but we saw much more ambiguity for other speakers. The SS ANOVA for RP showed the statistical significance between the phrase-level onset and codas, as well as confirming no significance with the remaining phrase-level onset tokens. However, with only five phrase-level coda /l/s for this speaker in Experiment 1, it was impossible to investigate factors such as duration interacting with darkness reliably, and one coda environment does not provide the comprehensive insight into the full /l/-darkening system. The 100 tokens of /l/ collected from the same in Experiment 2 will enable us to have a closer look at the speaker's system overall.

7.2.1 Ultrasound splines

Figure 7.1 shows the midpoint splines across ten phonological environments in the RP speaker's recording for Experiment 2, once again suggesting a bimodal distribution between the contexts where /l/ is in the onset at the word level (contexts 1-5) and where it is in the coda at the word level (contexts 6-10).. The retracted tongue root, lowered tongue body and reduced tongue tip gesture typical of dark [ɫ] is found in the latter contexts only. The palate trace taken before the recordings accentuates how consonantal these light variants are for this speaker, showing a very close proximity to the hard palate.

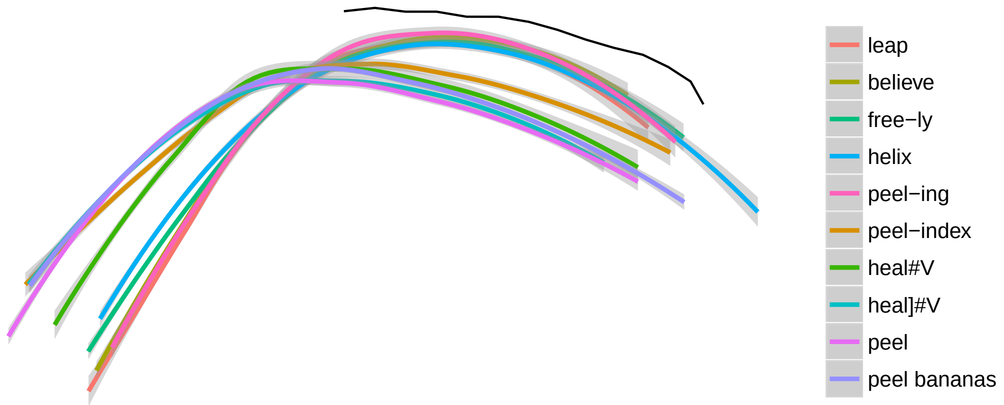

Figure 7.1: RP splines at /l/ midpoint across phonological context.

The splines for the word-final prevocalic tokens *peel-index* and *heal*#V look a little intermediate, in that the tongue tip is not as low as the other dark [ɫ]s for the former, and that the tongue-root is not as retracted for the latter.[1] If we consider the individual splines in Figure 7.2, we can see that this intermediacy is not down to gradience, but due to the splines in Figure 7.1 representing a mean of the variance i.e. these contexts are not articulatorily intermediate, but are intermediate in terms of the mean. In other words, for the RP speaker, word-final prevocalic tokens where the /l/ precedes an internal prosodic boundary or a compound boundary can variably result in an /l/ the same as a canonical onset, or the same as a canonical coda. The variability in the dataset for Experiment 2 is discussed in more detail in Section 7.10.2. The mean spline value gives the misleading picture of an intermediate phonetic realisation. Figure 7.2 shows that *pee**l**-index*-type tokens can range from almost as light as the mean ***l**eap*-type splines in the tongue tip area, to even darker than the mean *pee**l***-type spline in the tongue root area. The *hea**l**#V*-type splines do not vary as wildly, but still show one spline which is particularly light in comparison to the others. When looking at the individual splines in 7.2, we can see that this intermediate mean is a result of variance.

The fact that this speaker has variable pronunciation in word-final prevocalic position in Experiment 2 does not conflict with the results from the same speaker in Experiment 1. In Experiment 1, it was concluded that RP has phrase-level darkening, so

[1]Note that we largely cannot see the tongue tip in these images, but its position can be inferred. Tongue blade may be a more appropriate term generally.

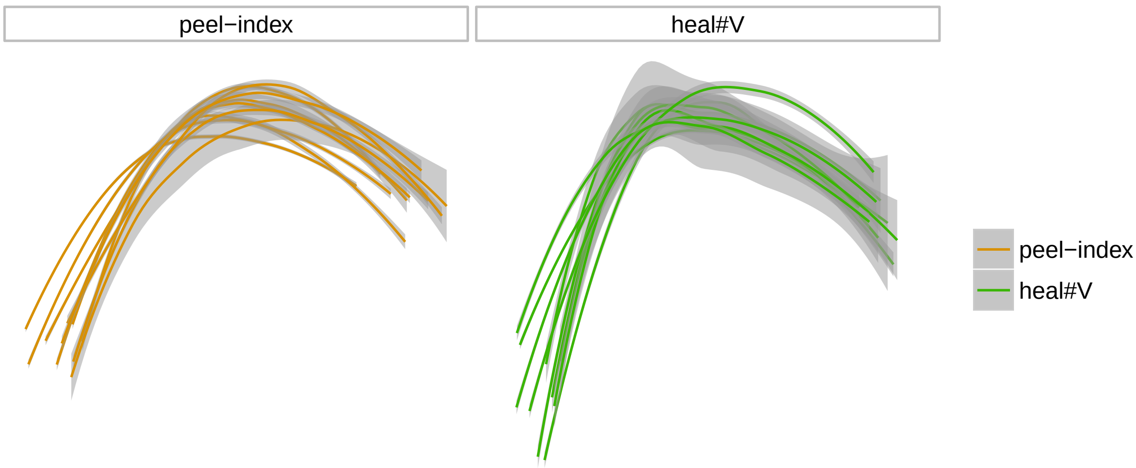

Figure 7.2: RP splines at /l/ midpoint showing variation in *peel-index* and *heal*V tokens.

that /l/ darkens in the coda at the phrase-level only and remains light prevocalically in phrases such as *heal it*. The word-final prevocalic tokens in Experiment 2 have different prosodic boundaries, which were designed to coerce speakers into produce an /l/ which tended towards one extreme of the continuum. In Experiment 1, /l/s which obligatorily resyllabify, e.g. in a phrase such as *heal it* where the following vowel is part of the cliticised morpheme, are used in order to examine the morphosyntactic sensitivity of /l/s which are in the coda at the word-level, but the onset at the phrase-level. For the large part in Experiment 2, speakers treat the *peel-index* and *heal*#V tokens as canonical coda-like /l/s. Nevertheless, it brings up the issue of what really constitutes a word-final prevocalic consonant, and this is discussed in more detail in Section 7.10.3.

7.2.2 PCA

7.2.2.1 Variable importance and loadings

The PCA for the RP speaker results in one principal component (PC1) accounting for 89% of the variation, as can be seen in left panel of Figure 7.3. PC2 accounts for 6% and is unlikely to be significant. As mentioned in Chapter 5, and following Baayen (2008), it is advised that the cut-off point for significant PCs is either:

i. those which account for less than 5% variation

ii. where there is a large discrepancy between two PCs.

The right panel of Figure 7.3 shows the range of PC1 and PC2, demonstrating just how little PC2 contributes for this subset of the data. Therefore PC2 will not be used for this speaker. As we shall see in subsequent sections, PC2 makes a similarly small contribution for all speakers. The reasons for this were discussed previously in Chapter

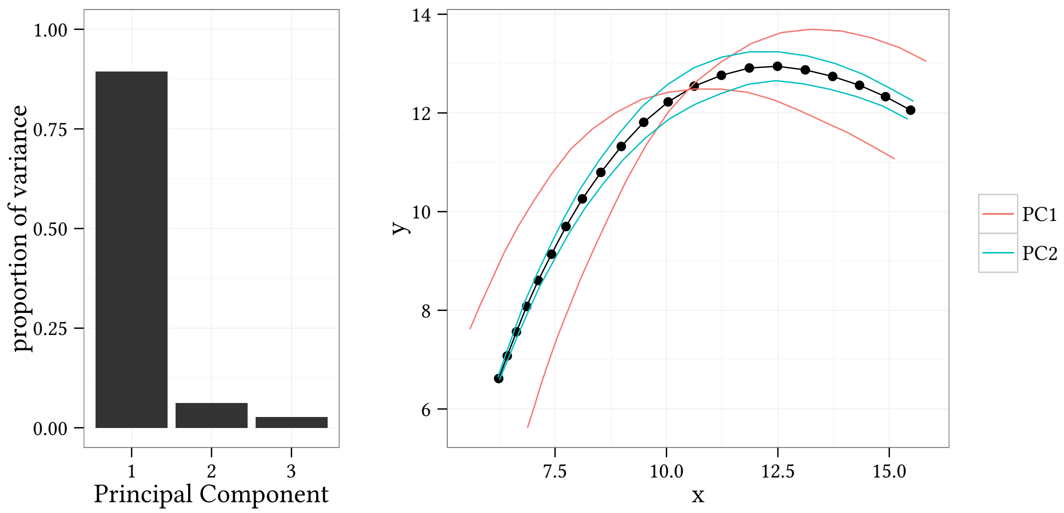

Figure 7.3: PCA summary. Variable importance and loadings plots from the RP PCA output

5. The PCA in this section (and throughout the Chapter) will focus on PC1 only, which ranges from the lightest light [l] (high PC1), to the darkest dark [ɫ] (low PC1).

7.2.2.2 PCA results

The box plot in Figure 7.4 shows the PC1 values for the 10 tokens in each phonological context. The first five contexts show high PC1 values, indicating lighter /l/s, and the latter five contexts show low PC1 values, corresponding to their darker realisation. As observed in the splines in Figure 7.1, there is a clear distinction between these two groups, adding support for a categorical interpretation of the RP speaker's data.

Further support for the bimodality of this distribution is provided by Hartigan's dip test. When the PC1 values in Figure 7.4 are subjected to the test, the dip value is 0.078, which shows a significant chance of the distribution being bimodal ($p < 0.001$; the closer to 0 the dip value is, the more likely the distribution is to be unimodal). Figure 7.5 is a density plot of the PC1 values. The overall distribution, represented by the dashed line, demonstrates the bimodality dip for this speaker. Individual distributions for the light and dark categories are also plotted on top of the general distribution, accentuating the bimodal peaks.

A post-hoc Tukey HSD test between the contexts, based on the PC1 values visualised in Figure 7.4, shows that there are no significant differences between any of the first five contexts, and each of the first five contexts is significantly different to each of the second five contexts (p < 0.001 in all cases). The same is found for the acoustics. This is further evidence that the two distributions should be treated as two separate categories. In the final five phonological contexts, the only intra-category significant difference in articulation is between the *pee**l***-type and *pee**l**-index*-type tokens (p = 0.04). This is because of the relative lightness of the mean *pee**l**-index* spline, and the slightly retracted phrase-

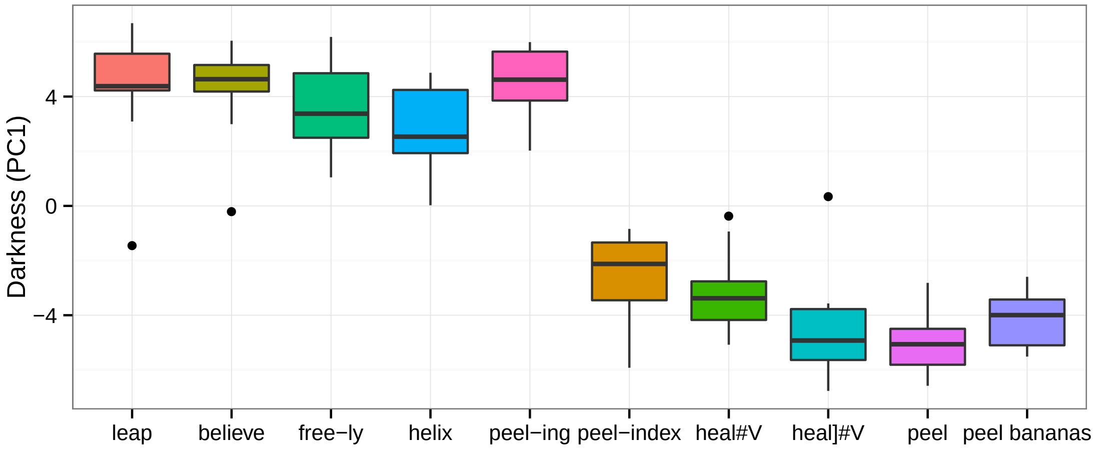

Figure 7.4: PC1 values for RP speaker across phonological context (higher values represent lighter /l/s)

final realisation. This distinction is not found in the acoustics and is not found between any other tokens, so is not classified as a separate category, but the intermediate realisation of tokens is discussed later in Section 7.10.2. The full results of the 45 contextual comparisons conducted under Tukey HSD test are presented in Appendix C.

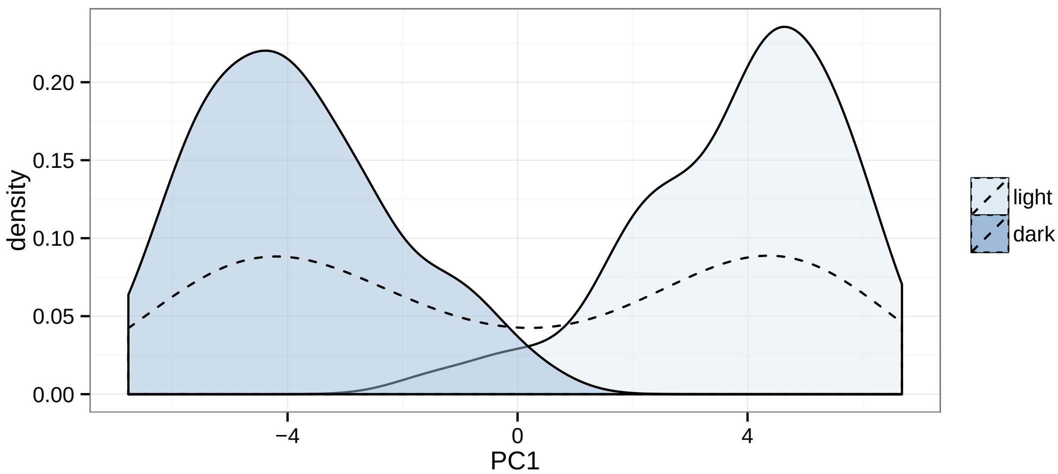

Figure 7.5: Density plot for RP showing separate distributions for light and dark categories (categories determined by phonological environment). Dashed line represents overall distribution.

7.2.3 Acoustics

The box plots in Figure 7.6 show the F2-F1 value across phonological context for the RP speaker, and show a very familiar pattern, similar to the PC1 box plots in Figure 7.4. Again, this information provides support for a categorical allophonic distinction. The first five phonological contexts have a high F2 and low F1, resulting in a large difference between the two formants, characteristic of light /l/. In contrast, the final five contexts have a very small difference between F1 and F2, characteristic of back vowels and a darker /l/ sound. The dip statistic for the difference between the first and second formants is significant at $D = 0.073$ ($p < 0.001$).

Although the formants do a good job of representing the articulatory data for RP, there are a few minor areas of variation that the acoustics do not capture. Nevertheless, for this speaker, the acoustic data give a reliable picture of the articulatory pattern. We shall see that the acoustics can be much muddier for other speakers later in this chapter.

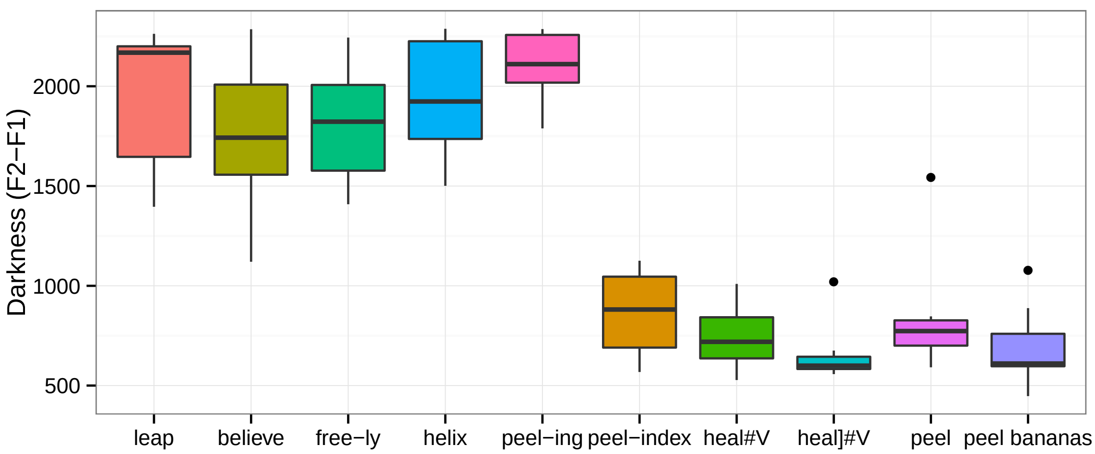

Figure 7.6: RP speaker's F2-F1 values across phonological context

7.2.4 Darkness vs. duration

Thus far, we have seen there is little evidence for the absence of categorical allophony in /l/ realisation in RP, with two separate categories emerging from the statistical analysis of both the articulatory and the acoustic data. However, this does not necessarily mean that duration plays no part in this speaker's system. In order to compare darkness with duration, it may be useful to runs correlations and comparisons with darkness, using the PC1 values. This will help address whether:

i. duration can fully account for the darkness patterns (as per Sproat and Fujimura 1993).

ii. duration has a minor low-level effect or plays no role whatsoever and the two categories account for all variation (as per Hayes 2000).

iii. duration correlates with darkness for the dark [ɫ]s only (as per Yuan and Liberman 2009; 2011).

Figure 7.7 shows darkness, as represented by PC1, plotted against the actual duration of the 'rime' (i.e. the /i+l/ sequence). As discussed in the methodology, the 'rime' is used instead of the /l/ itself, as for some speakers the segmentation between the two is impossible. This is not the case for RP, but the 'rime' is used to remain consistent with other speakers throughout the chapter. Correlation lines between darkness and duration are fit separately to each category. We can see from the plot that both categories have a weak but significant correlation: for the light /l/s Pearson's r = -0.296 (p =0.04), and for the dark /l/s r = -0.32 (p =0.025). These values show a significant correlation between the two measures for both categories, in that the longer the 'rime' the darker the /l/. Unlike the findings for American English from Yuan and Liberman (2009; 2011), there does not seem to be a strong correlation between darkness and duration for the RP speaker, or any distinction between how each category treats the interaction. Most notably phrase-final /l/, which was predicted to be the most strongly correlated of all tokens, shows no correlation between darkness and duration. These tokens cluster into two separate shorter and longer regions, but this does not change their darkness, which remains relatively stable.

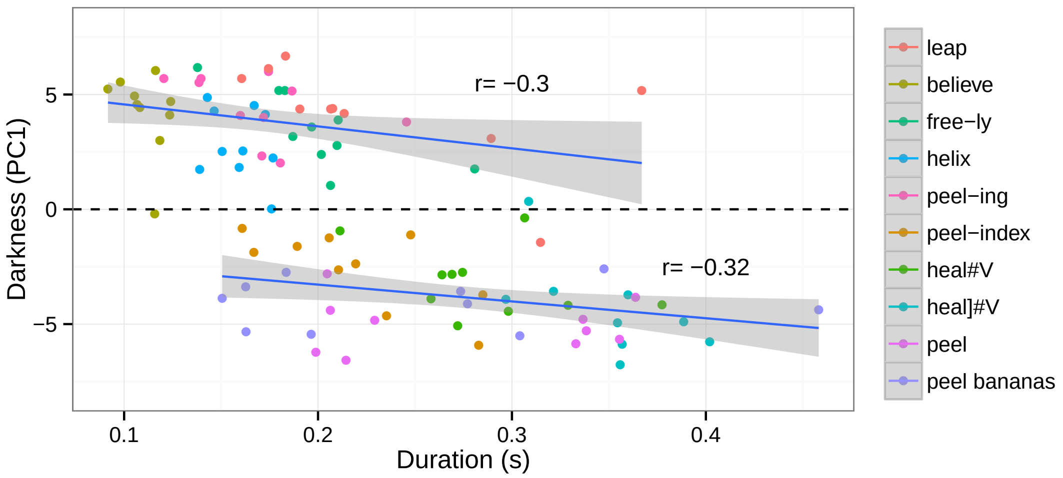

Figure 7.7: RP PC1 values against rime duration

Although significant, the correlations are very moderate and a very crude way of summarising the data. Pearson's r provides a basic line of best fit, but is a problematic measure for on which to base final conclusions. Its high sensitivity to extreme values, and its lack of indication of the direction of the relationship between two variables highlight the need for more sophisticated methods of modelling data. Linear models provide an improved way of looking at all of the predictors in the dataset. It may well be the duration is important for accounting for American English /l/, but there are other more important factors to consider before we conclude this for RP.

7.2.5 Models

Table 7.2 shows the adjusted r^2 values for four different linear models fit to the dependent variable of darkness (i.e. PC1). A perfectly fit model has an adjusted r^2 value of 1. As outlined in the introduction, category was split into the factor levels 'light' or

model	adj. r^2
1. Darkness ∼ log(Duration)	0.449
2. Darkness ∼ Category	0.832
3. Darkness ∼ Category + log(Rime duration)	0.845
4. Darkness ∼ Category * log(Rime duration)	0.843

Table 7.2: Comparing duration correlations for RP speaker alongside adjusted r-squared of models with a category duration interaction

'dark' and the Tukey HSD test indicated that the *a priori* splitting of the contexts between 1-5 and 6-10 was the best way of classifying the data. We shall now assess the goodness of this fit by running linear models including other factors such as duration.

We can immediately see that duration alone provides a weak fit for RP (model 1). However, any model which takes category into account performs very well (2). A series of ANOVA comparisons (see Appendix D) shows that a model which takes duration as well as category into account (3) performs significantly better than category alone ($p < 0.005$), but adding an interaction between the two (4) creates next to no difference ($p=0.96$). This suggests that, although duration plays a role to a small extent, it is not conditioned differently for light and dark variants, but applies across the board. Models 3 is the optimal model (in that it has the highest adjusted r^2 value without adding superfluous insignificant interactions), showing that including both categorical and gradient effects provides the best fit for RP.

Table 7.3 shows the coefficients of linear model 3. The negative estimate for the dark category shows that a dark category gives negative PC1 in comparison to the intercept of a light category. The negative estimate for the log of the 'rime' duration shows that a longer token reduced the PC1 value, i.e. gives a darker /l/.

	Estimate	Std. Error	t value	Pr(>|t|)
(intercept)	0.562	1.078	0.522	0.603
category: dark	-6.836	0.433	-15.776	0.000
log(Rime duration)	-1.839	0.586	-3.138	0.002

Table 7.3: RP model 3, Darkness by Category + log(Rime duration)

In summary, the results of the quantitative analysis for RP suggest two things. Firstly, category can easily account for the majority of the variation in this model, producing an excellent fit of the data with no added factors. Secondly, there is some effect of duration, but this is weak and applies across the board, rather than just applying to the dark [ɫ]s. However, although duration only improves the models a tiny amount, it does so significantly, and its effect cannot be ignored.

7.2.6 Summary

Overall, the results for RP suggest that there is a categorical allophonic distinction between word-level onset and coda /l/s. This conclusion is supported through various analyses of the dataset. Firstly the articulatory evidence from ultrasound tongue splines shows a clear distinction in the expected direction: onset /l/s are articulated towards the front of mouth with an advanced tongue body, and coda /l/s have a diminished tongue tip gesture and pharyngealised tongue root. Moreover, there is a large articulatory distinction between the two extremes. The non-overlapping confidence intervals, fit by the `loess` smoothers, do not overlap at all, indicating the articulations are significantly different, whilst backing up the evidence from Experiment 1. A clear categorical distinction in the tongue splines is supported by the PC1 box plots and the significant effect of bimodality from the dip statistic. Moreover, the Tukey HSD test shows that none of the five phonological environments in the light category are significantly different to one another, whilst all being significantly different to the environments within the dark category. With one exception of the two extreme values being significantly different to one another, the same is true of the dark category. The acoustic data from the difference of the first and second formants show the same pattern as the PC1 values, for the dip test and the Tukey HSD.

Nevertheless, there still seems to be evidence of gradient effects of duration active in this speaker's system. However, these apply across the board and suggest an epiphenomenal gradient effect, rather than a gradient phrase-level phonologised effect under cognitive control. Perhaps this would be more strongly correlated with temporal data on gestural phasing, which we do not have reliable access to with the ultrasound unit used in this project. However, the lack of significant interaction of category and duration in the linear models suggests that RP does not show strongly phonologised effects of duration, which perhaps gives us some insight why this dialect remains linguistically conservative, in that it has stayed immune from the apparently natural tendency towards darkening of lengthened coda /l/. This is seen both phonologically, with respect to the stability of the lenition processes applying in the coda at the phrase level,[2] and also phonetically, in that it shows no further lenition processed applying to the /l/, such as vocalisation, whilst maintaining a clear magnitudinal distance between initial and final position.

7.3 London Female

As outlined in Chapter 2, the London vernacular is well-known for its widespread /l/-vocalisation in final position. The two Londoners in this study, as expected, have audibly vocalised /l/s accompanied by visible lip rounding.[3] Henceforth, this auditorily-convincing vocalisation in coda position will be discussed in terms of categoricity and gradience depending purely on the ultrasound tongue body image. It must be noted

[2] Recall from the previous chapter that this speaker has no darkening in ***heal** it*-type tokens, so exhibits phrase-level darkening.

[3] Note, this study did not have access to a lip camera and so we do not have data on this, but the vocalised /l/s were visibly and audibly accompanied by labialisation. The magnitude of this labialisation would be an excellent source for further study.

that confirmation of a total lack of palato-alveolar contact would be preferable, but the sound files themselves are more than enough to convince the author that there could not possibly be tongue-tip contact during these articulations. Further study of the multi-dimensional phonetic space would be needed to address questions of labilalisation and tongue-tip contact thoroughly.

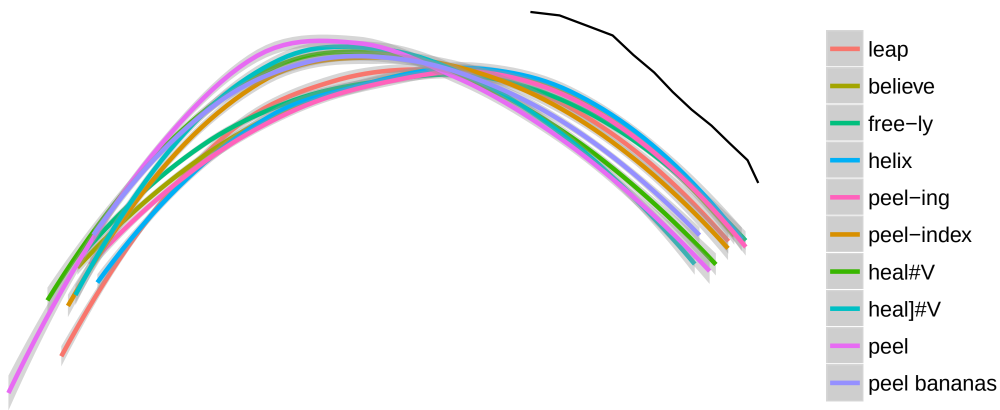

Figure 7.8: London Female's splines at /l/ midpoint

In Chapter 6, we saw a speaker from nearby Essex, another vocalising dialect of the South-East, with a potential three-way distinction between light, dark and vocalised /l/. In addition, the Essex speaker showed a large magnitudinal difference between her initial and final /l/s, and also exhibited style-shifting pre-consonantally. It will be interesting to observe whether the London speaker shows similarities phonologically, in terms of a three-way distribution; phonetically, in terms of large tongue root retraction in vocalised /l/s; and sociolinguistically, in terms of style-shifting.

7.3.1 Ultrasound splines

Figure 7.8 shows the midpoint splines for the London Female across ten phonological environments. Unlike the Essex informant, this speaker did not style-shift or produce any unnatural sounding tokens at any point during the recording. Her midpoint splines show two separate distributions, with the same phonological pattern as the RP speaker, in that she splits the ten environments into the same two groups of five. Articulatorily, however, the distinction between the two extremes is not as striking as in RP. This is interesting given the large difference noted between the Essex speaker's initial and final /l/s in Chapter 6. It seems the London Female is not as advanced as the Essex speaker in the phonetic magnitude of /l/-vocalisation. The small magnitudinal difference between the two distributions in Figure 7.8 may also seem surprising given the RP differences. There is some backing/raising of the tongue dorsum here, but only a small amount. The distinction between the light and dark /l/s in the RP speaker in Figure 7.1 seems clearer than the distinction between light and vocalised for the London Female. One might expect a vocalised /l/ to be more back than a regular dark realisation, resulting in a bigger difference between initial and final /l/ for a vocaliser than a mere darkener. This is certainly what was found for the Essex speaker in Chapter 6. This has also been found in other studies, such as Wrench and Scobbie (2003), who found that their 'impressionistically obvious' vocaliser had the most extreme magnitude difference than the other speakers. However, they were considering tongue tip height only in this EMA study, the part of the tongue which gives the worst and most unreliable image in ultrasound data. This poses a question for future work on /l/-vocalisation, then: to discover whether the backing of the tongue is consistently an articulatory correlate of vocalisation at all, or whether it is merely a movement of the tongue tip which creates the difference. The difficulty in distinguishing between dark and vocalised /l/ on the spectrogram (Hall-Lew 2011) might suggest the latter, but the existence of several strategies of production in articulating vocalised /l/s is certainly possible.

However, it is difficult to compare absolute magnitude in this way between speakers. It raises the issue of inter-speaker comparison and normalisation in ultrasound studies and other articulatory work. The PCA works well for comparing intra-speaker phonological contexts, but not for comparing the extent of extreme magnitude between, for example, the London Female with the RP speaker. Differences in vocal tract size, as well as position of the probe mean that the two datasets are not comparable. This speaker may have a decent magnitudinal distance comparative to the size of her mouth, but we cannot be sure from the data we have. For now, we can see the potential variance in the front of the tongue by looking at Figure 7.9, which shows the ***l**eap* and *pee**l***-type tokens compared with the tongue position in inter-dental fricative [ð] (a position in which the tongue tip is expected to be very front) and back rounded vowel [ɔ] (with a very typical low and back tip). We can see here that the vocalised /l/s are not as low in the blade area as the back rounded vowel [ɔ]. For future study, it would be worthwhile to see how close these vocalised /l/s are to [w], which is often how London final /l/ is transcribed.

Another way in which the London Female does not seem as advanced in comparison to the Essex speaker is in terms of the three-way distinction between light, dark and vocalised /l/s. There is next to no difference between the *he**l**ix*-type and *pee**l**-ing*-type tokens for the London speaker, as shown in Figure 7.10. The Essex speaker showed variation across these two contexts: she had a stem-level darkening rule, such that /l/ darkens in the coda at the stem before the affixation of the /ɪŋ/ suffix. This resulted in a significant difference in tongue root backing between *he**l**ix* and *pee**l**-ing*-type tokens by SS ANOVA. In this sense, the Essex speaker's darkening is more advanced, as it applies at a higher level of the grammar than the London speaker's.[4]

[4]Note, a drawback to Experiment 2 is that we cannot see how the resyllabified /l/s behave for this speaker. They could indeed show an intermediate realisation for *hea**l** it*-type tokens, but this environment was not collected.

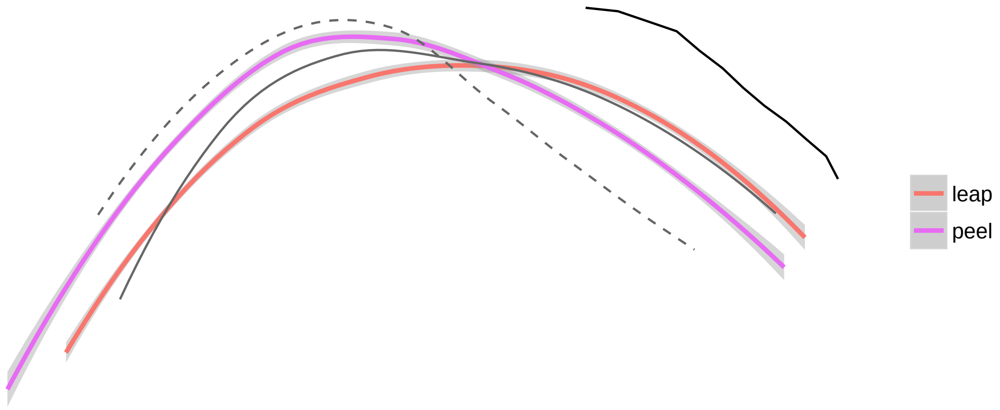

Figure 7.9: London Female's /l/s overlaid with front token of [ð] (solid line) and back token of [ɔ] (dashed line) to demonstrate potential tongue tip variance

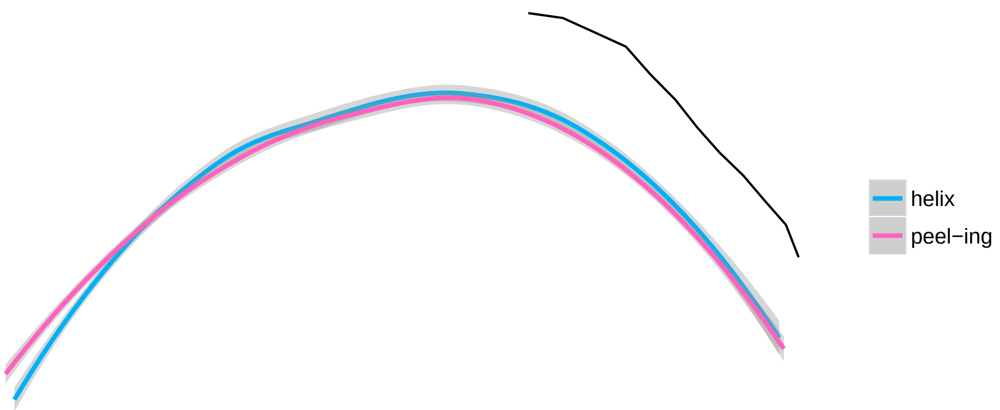

Figure 7.10: London Female's splines at midpoint showing no difference between intervocalic /l/s in monomorphemes and pre-suffixal contexts.

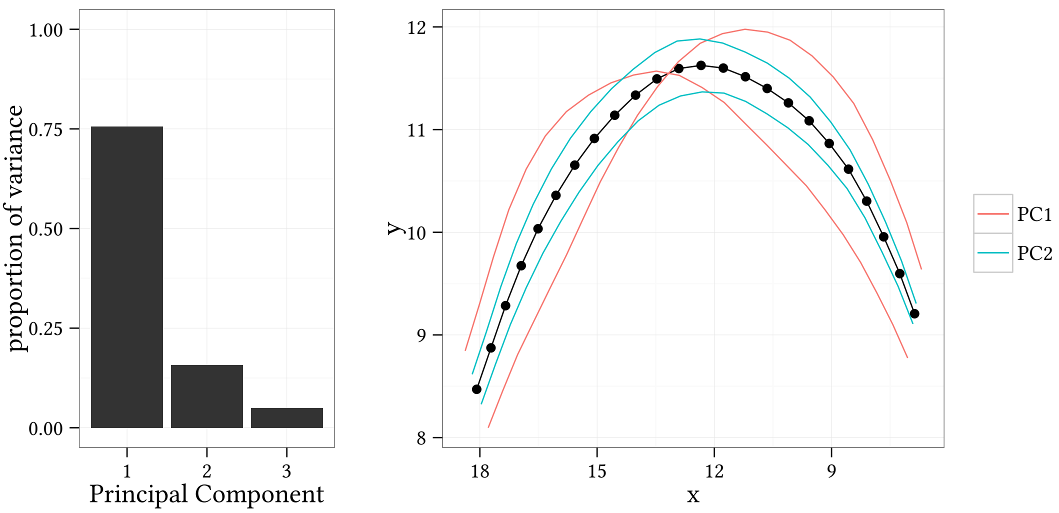

Figure 7.11: PCA summary. Variable importance and loadings plots from the London Female PCA output

7.3.2 PCA

7.3.2.1 Variable importance and loadings

The details of the PCA performed on the London Female tongue splines are plotted in Figure 7.11. PC1 accounts for 76% of the variation, and PC2 16%. Although PC2 accounts for a somewhat larger amount of the variation than we saw with the RP speaker, the steep cut-off between PC1 and PC2 suggests that this PC is capturing only a very small part of the variation, and this is confirmed by the loadings plot in the right panel.

7.3.2.2 PCA results

Despite the smaller magnitudinal distinction observed for the London Female in Figure 7.8, Figure 7.12 shows that the London speaker has a small but definite cut off between the first and last five phonological environments. There is next to no overlap, with all interquartlie ranges remaining distinct, as predicted from the spline plot's confidence intervals. Support for the bimodality of this distribution is provided by Hartigan's dip test. When the PC1 values in Figure 7.12 are subjected to the test, the dip value is 0.058, which shows a significant chance of bimodality (p =0.016). Figure 7.13 demonstrates the bimodality dip for this speaker. Phrase-final tokens are the darkest of all, showing a small but considerable difference between this position and pre-consonantal position. Two categories are supported by a Tukey HSD test (full significance tables are available in Appendix D). Phonological environments 1-5, i.e. those in the light category, are not significantly different from one another, but are all significantly different from each of the environments in 6-10, i.e. the dark category. Those in the dark category are all significantly different to those in the light, but there is a small amount of intra-category variation with the ***peel-index*** tokens being significantly lighter than all but the ***peel ba-***

nanas-type tokens. Phrase-final darkening can be seen in that the *pee**l***-type tokens are significantly darker than the *pee**l** bananas*-type tokens, and is discussed further in Section 7.3.6.

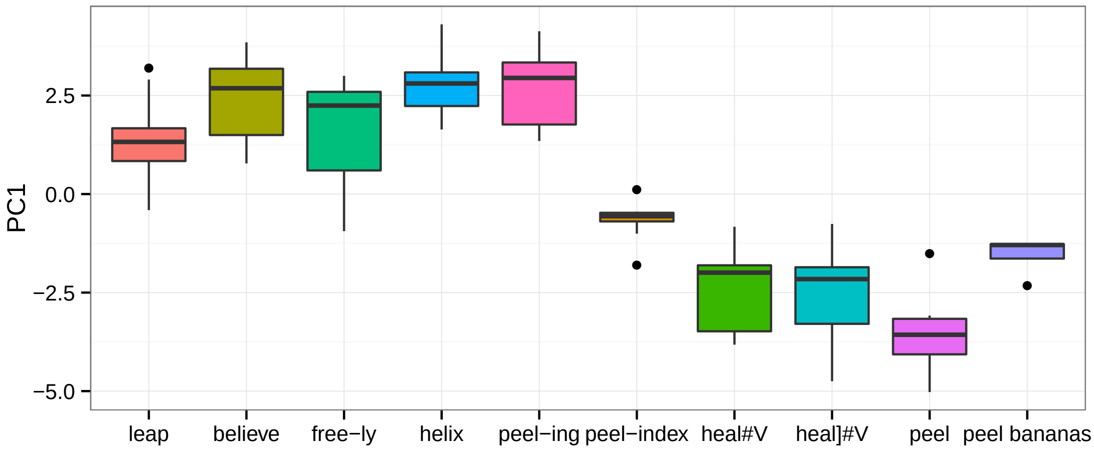

Figure 7.12: London Female PC1 values across phonological context

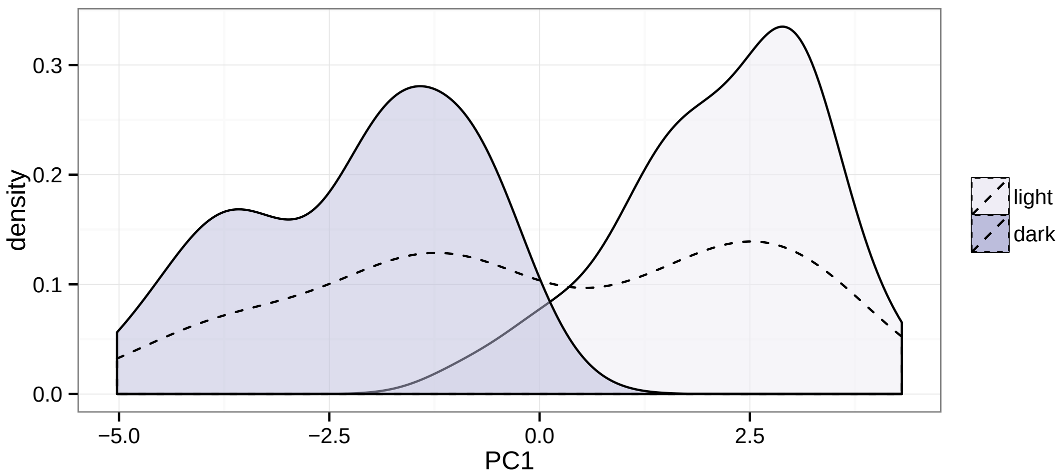

Figure 7.13: Density plot for London Female showing separate distributions for light and dark categories (categories determined by phonological environment). Dashed line represents overall distribution.

7.3.3 Acoustics

The acoustics for the London Female show a striking resemblance to the PC1 values in Figure 7.12. The clear split between light and dark tokens, again with no overlap, is present, as well as the darkest tokens found in the phrase-final *pee**l***-type tokens, which display the lowest F2-F1 difference. The Tukey HSD test shows mostly the same distribution as the articulatory data, including the distinction between the *pee**l**-index*-type tokens and the significantly darker *pee**l***-types, and *pee**l***-type tokens being significantly darker than the *peel bananas*-types. The other distinctions do not come out in the acoustics, however. The F2-F1 values result in a significantly bimodal dip statistic $D = 0.074$ ($p < 0.001$), confirming this observation.

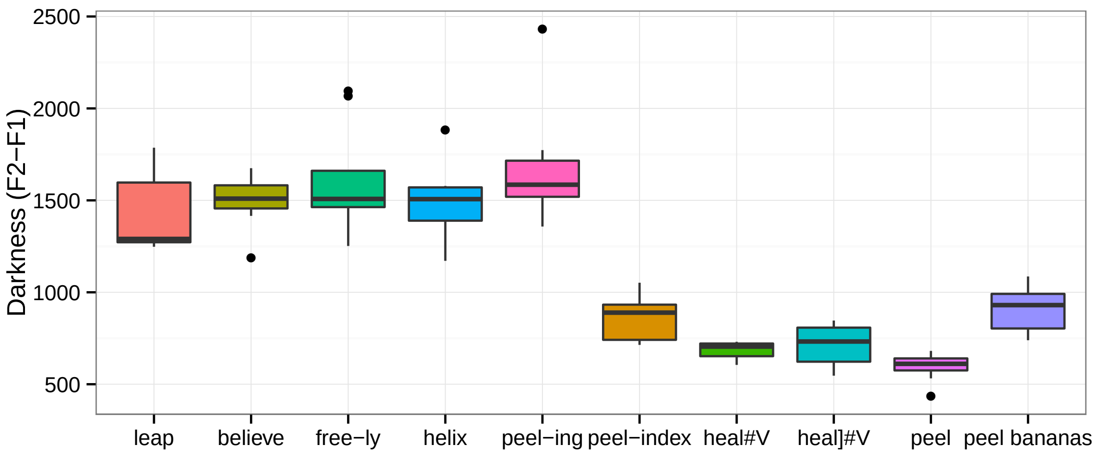

Figure 7.14: London Female F2-F1 across phonological context

7.3.4 Darkness vs. duration

Thus far, the evidence for the London Female showing a categorical allophony between light and dark /l/ comes from several sources: the spline midpoints, the PCA results, the dip statistic and the acoustics. This is also in line with auditory impressions. Nevertheless, as we have seen in earlier chapters, evidence for categoricity does not necessarily mean we should abandon all investigations of gradient effects. The following plots show the interaction of darkness and duration for this speaker. Unlike the RP speaker, the London Female does indeed accord with the claims made by Yuan and Liberman (2009; 2011) in that only the dark [ɫ]s show a strong correlation with duration, with a Pearson's value of r -0.755 ($p < 0.001$)as opposed to the /l/s which have a much weaker and non-significant value of $r = -0.285$ ($p > 0.05$).

7.3.5 Models

Table 7.4 shows the r^2 values for four different linear models fit to darkness (i.e. PC1). Again, duration alone provides a weak fit, although not as weak as we saw for RP. Category performs well on its own (2) but adding duration significantly improves this (3; see ANOVA comparisons in D). Unlike RP, the interaction of darkness and duration (4) significantly improves on a model without the interaction, suggesting that duration behaves differently for the two categories for the London Female.

Although the Tukey HSD test indicated that *peel-index* tokens behaved differently from some of the other categories, this was only significant across acoustics and articulation for the *peel*-type tokens, but not for anything else. There was also a difference

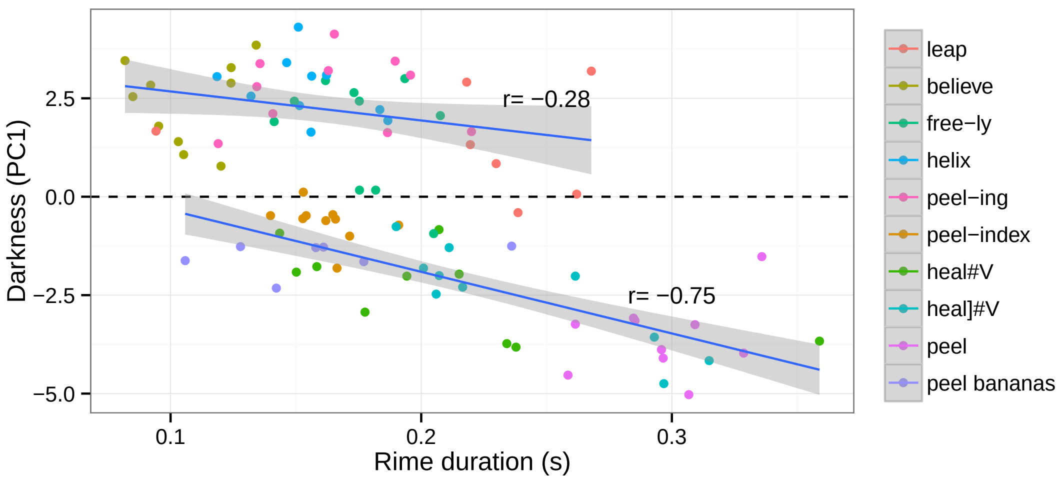

Figure 7.15: London Female PC1 values against rime duration. Correlation lines calculated separately and show Pearsons r value.

model	adj. r^2
1.Darkness ∼ log(Duration)	0.371
2. Darkness ∼ Category	0.746
3. Darkness ∼ Category + log(Rime duration)	0.813
4. Darkness ∼ Category * log(Rime duration)	0.83

Table 7.4: Adjusted r-squared of darkness modelling for London Female

between ***peel***-type and ***peel** bananas*-type tokens. Given the preditable effects of boundary strength, coarticulation and phrase-final lengthening, it does not seem prudent to treat it as a separate intermediate category here, but this point will be returned to later in Section 7.10.2.1.

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	0.39	0.99	0.39	0.70
Catdark	-7.80	1.29	-6.06	0.00
log(Rimeduration)	-0.99	0.52	-1.88	0.06
Catdark:log(Rimeduration)	-2.35	0.73	-3.21	0.00

Table 7.5: London Female model 4, Darkness by Category * log(Rime duration)

Table 7.5 shows the coefficients for linear model 4. This demonstrates that, once the interaction is added, the sole effect of duration generates a non-significant p-value, highlighting how duration plays a bigger role in the conditioning of the dark category than the light. Again, the estimates show the dark category has a negative effect on PC1, as does a longer ‘rime’.

7.3.6 Summary

The results from the London Female data provide an interesting comparison with those from the standard RP speaker, as well at the results from Essex in the previous

Chapter. Firstly, we know that London is susceptible to /l/ lenition processes of some kind, because of the widespread reports of /l/-vocalisation. This speaker is a vocaliser, which can be heard clearly in her recordings. Vocalisation does not only occur phrase-finally and before consonants, but also before vowels in following verb phrases and when the /l/ is part of a compound.

Overall, the results from this speaker's data point to an obvious categorical distinction between the light and vocalised distributions. This is evidenced by numerous tests and analyses. The ultrasound tongue splines presented at the beginning of the section show two separate, albeit tightly clustered, distributions, with non-overlapping `loess` confidence intervals, for the most part. This pattern is clarified in the PC1 box plots, as swell as the acoustics. The Tukey HSD tests show that the observed patterns are significantly different in the expected regions, and not significantly different within category (with one or two exceptions).

However, this speaker shows compelling evidence for a gradient interpretation of durational effects overlaid on top of the categorical differences. The correlation of darkness and duration is only significant for the dark tokens, and the increased model fit of an interaction between category and darkness indicates it applies differently to each category. The data are consistent with a situation where /l/-darkening processes have increased the magnitude of their lenition, becoming completely vocalised, applying to coda /l/s at the phrase level. This process has supplanted a potential previous system where dark [ɫ]s occurred in intermediate contexts, as we found in Essex. For this speaker, a two-way distinction exists. The gradient precursor of darkening, longer duration, still exists as a phonologised gradient process, applying to all dark [ɫ]s. This speaker also has a small distinction between phrase-final and word-final preconsonantal /l/s, showing the extra effect of phrase-final position. There is no overlap between the two box plots of these contexts, articulatorily (Figure 7.12) or acoustically (Figure 7.14), and this is confirmed by the Tukey HSD test. We did not find this in RP, which showed consistently dark values for both of these contexts. For the London Female, this is evidence of a gradient phonetic effect at the phrase-level. It is probable that this extra effect of duration is not the kind of stable cross-contextual effect we found for the correlation between darkness and duration for the dark [ɫ]s, as it only applies in this context. Rather, this phrase-final lenition is likely to be an epiphenomenal gradient effect, not under speaker control. This speaker show several diachronic levels of the life cycle, synchronically active in the same system: a phrase or word level categorical rule; a phonologised cognitively controlled effect of duration in coda position; and a epiphenomenal effect of duration phrase-finally.

7.4 London Male

7.4.1 Ultrasound splines

Like his female counterpart, the London Male shows a two-way distinction between the initial light [l] found in onsets, and the audibly vocalised /l/ found in codas.[5] Again, there is no evidence for the three-way distinction found in Essex for this speaker's distribution. Auditorily, this speaker's vocalised /l/s do not sound as strong as the London Female's, but this is observation does not come out when comparing the articulations, due to the difficulty involved in conducting reliable inter-speaker analysis of ultrasound tongue images in this way. This extra impressionistic effect of vocalisation may come from other phonetic details, such as lip rounding, which was not considered in this experiment. The comparison of ***l**eap* and *pee**l***-type tokens with the inter-dental fricative [ð] and the back rounded vowel [ɔ] can be seen in Figure 7.17. The tongue blade position of the initial and final /l/s are very similar to these front and back tokens, although the overall magnitude is greater for these extreme tokens.

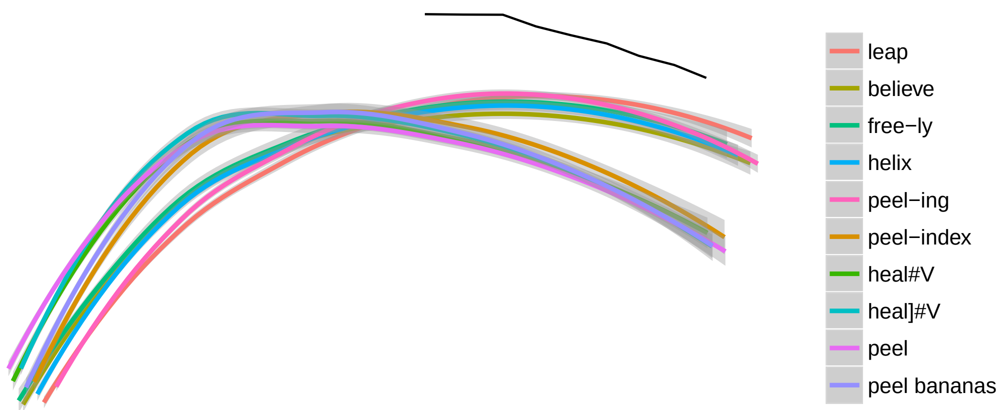

Figure 7.16: London Male speaker's splines at /l/ midpoint across phonological context

7.4.2 PCA

7.4.2.1 Variable importance and loadings

The details of the London Male's PCA are plotted in Figure 7.18. PC1 accounts for 86% of the variation, and PC2 6%. PC2 accounts for next to nothing here, as confirmed by the loadings plot in the right panel.

7.4.2.2 PCA results

[5] Note, this speaker has five less splines for the initial ***l**eap*-type tokens, for the reasons covered in the introduction.

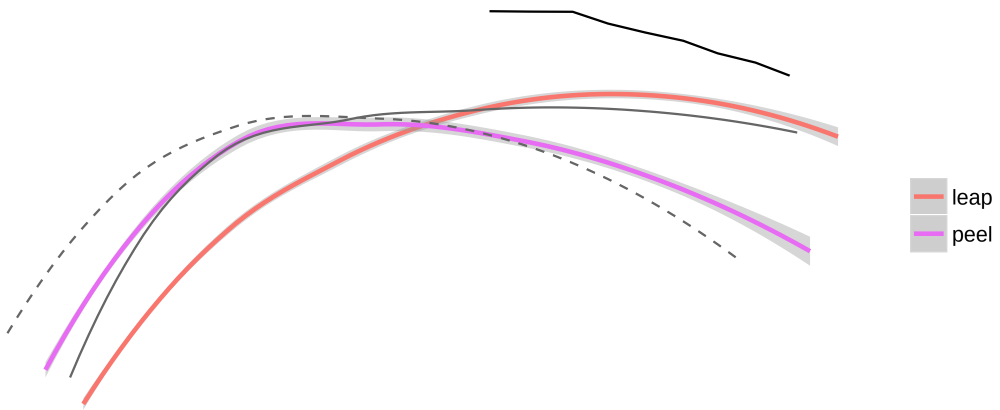

Figure 7.17: London Male's /l/s overlaid with front token of [ð] (solid line) and back token of [ɔ] (dashed line) to demonstrate potential tongue tip variance

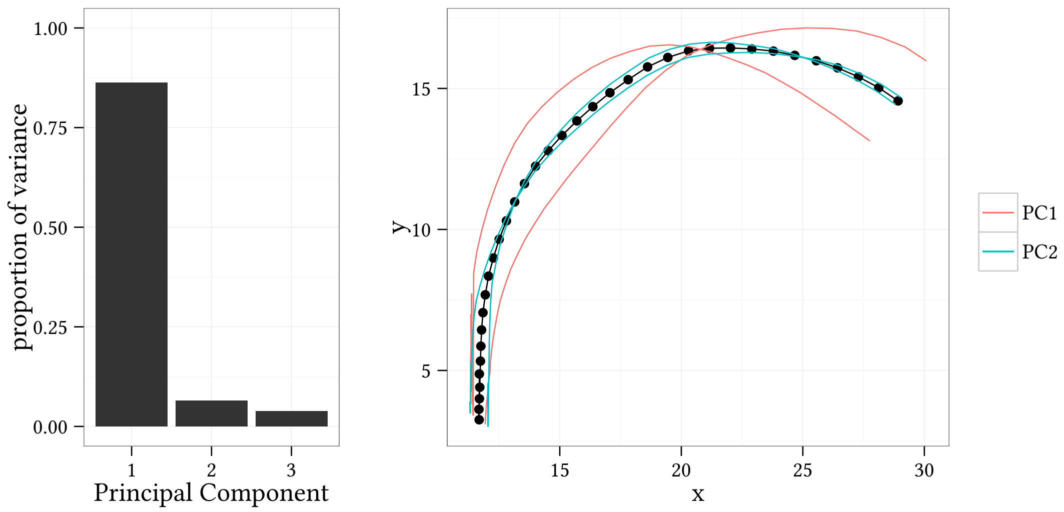

Figure 7.18: PCA summary. Variable importance and loadings plots from the London Male PCA output

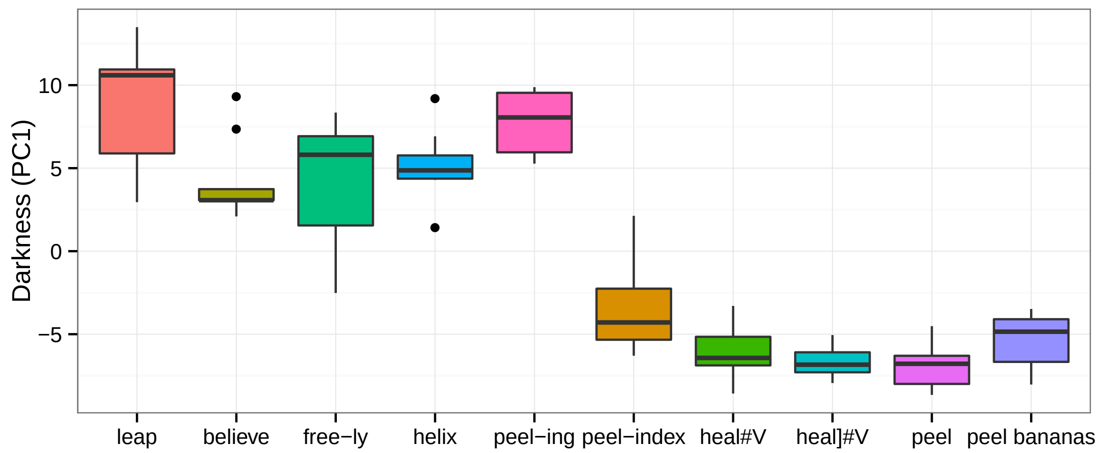

Figure 7.19: London Male PC1 values across phonological context.

Figure 7.19 shows that the London speaker has a small but convincing cut off between the first five phonological environments, and the latter five. Aside from one or two extreme values, there is no overlap with the PC distribution between the word-level onset and coda /l/s, suggesting a clearly bimodal distribution for this speaker. Bimodality is supported by Hartigan's dip test, with the PC1 figures giving a significant dip value of 0.066 ($p < 0.001$). Figure 7.20 shows the significant bimodality dip for this speaker. Note that this speaker shows a greater range in his light realisations than the dark ones. This may be suggestive of some kind of burgeoning process affecting stem-level /l/s, causing lenition in higher morphosytactic domains. However, this remains speculation at this stage, as within-context the splines are very tightly clustered.

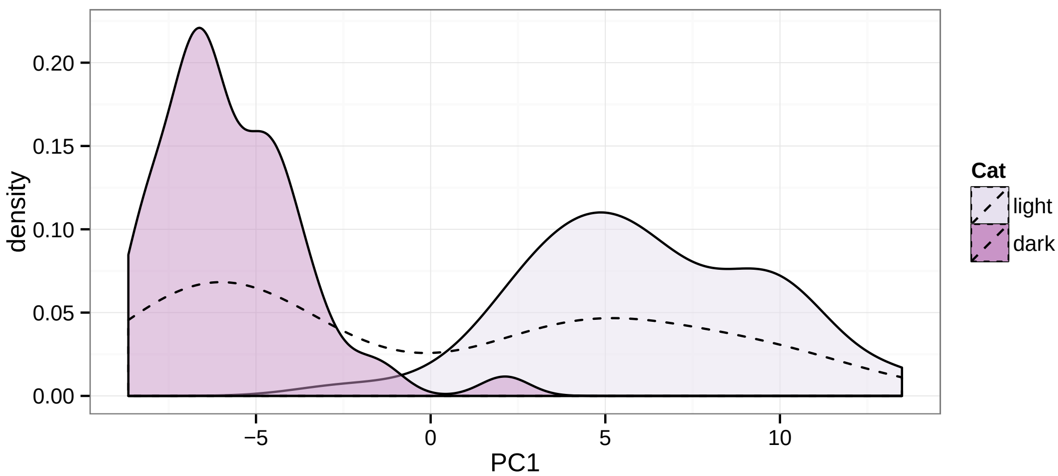

Figure 7.20: Density plot for London Male showing separate distributions for light and dark categories (categories determined by phonological environment). Dashed line represents overall distribution.

The Tukey HSD test, like the London Female, is not as easily interpretable as the RP speaker. However, where the London Female showed more complicated patterns within the dark category, the London Male shows more variation for the light tokens. In general, the pattern is the same for both speakers, with the first five phonological environments being significantly different to the latter five. The dark category shows no intra-group differences, and none of the five environment comparisons generate a significant p-value. The light category shows some small in-group differences for the articulatory data (likely due to several missing tokens for this speaker, from the *linoleum* tokens, for example), but these are not borne out in the acoustics below, and so two categories are used to fit the data. The full statistical tables can be found in Appendix C).

7.4.3 Acoustics

Once again, the acoustics corroborate the PC values for this speaker but to a less detailed extent, as shown in Figure 7.21. There is no difference in the acoustics of ***peel*** and *peel bananas*-type tokens for the F2-F1 values. Although the difference in the darkness values in Figure 7.19 showed a much smaller distinction than for the London Female, it was still present, but the acoustics do not capture this. This accentuates how slight the effect must be for this speaker, as well as showing how the articulatory data can pick up patterns which evade the acoustics. Supporting the observations of bimodality, the dip statistic is significant at D = 0.096 (p < 0.001).

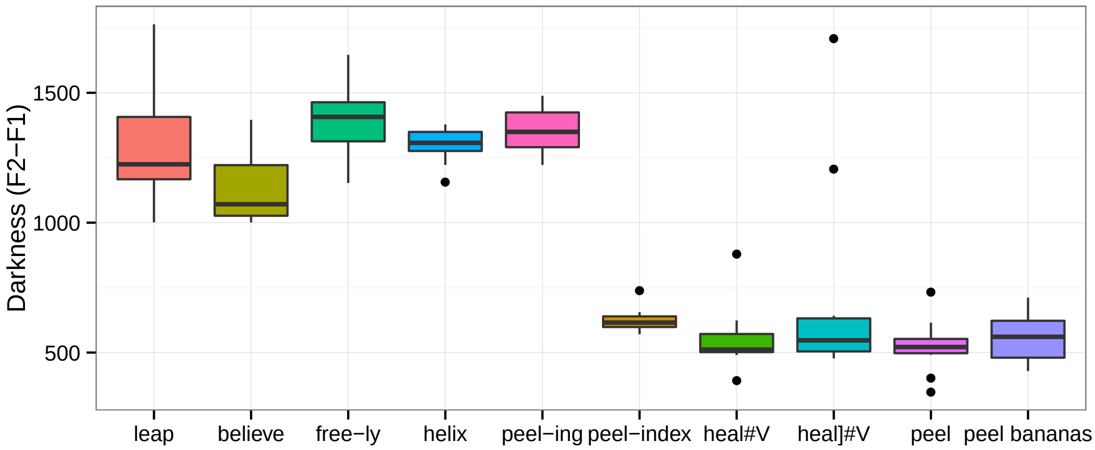

Figure 7.21: London Male F2-F1 values across phonological context

7.4.4 Darkness vs. duration

In the previous section we saw that the London Female's /l/ darkness correlated with duration for the dark [ɫ]s only, as had also been inferred from the studies of /l/ in American English. Figure 7.22 shows that this trend is also observable for the London Male. Although the trend for the dark [ɫ]s is not as strong (r =-0.383, p = 0.007), the lack of any correlation whatsoever for the light [l]s makes up for this. The large durational difference between the *peel* and *peel bananas*-type tokens is not present for the London Male, however. This relative lack of distinction can also been seen in the PC1 plots in Figure 7.19 and is discussed further in Section 7.4.6.

7.4.5 Models

As discussed above in Section 7.4.2, the Tukey HSD indicates that the light/dark distinction is the most parsimonious categorisation for this speaker (although light/vocalised may be more apt labels here). Table 7.6 shows the r^2 values for four linear models predicting darkness for the London Male. Aside for the model with only duration in it, all models provide a strong fit to the data, including category alone. In fact, category performs so well that, when using the original light/dark categories, adding duration makes no significant difference, although it does come close (ANOVA comparison in Appendix D; p=0.08). Even an interaction with duration does not make a significant improvement

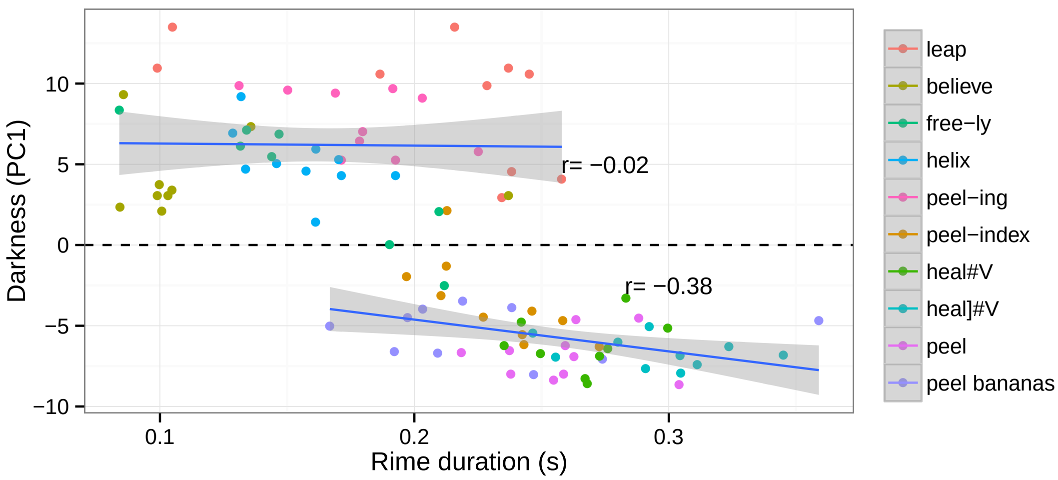

Figure 7.22: London Male PC1 values against rime duration, with Pearsons r correlation values.

model	$r - squared$
1. Darkness ∼ log(Duration)	0.437
2. Darkness ∼ Category	0.817
3. Darkness ∼ Category + log(Rime duration)	0.817
4. Darkness ∼ Category * log(Rime duration)	0.821

Table 7.6: Adjusted r-squared of darkness modelling for London Male

on category alone. Although duration seems to form a moderate significant correlation for this speaker, it cannot explain anything that category cannot account for.

Table 7.7 shows the best model for this speaker, a simple model with darkness modelled by category only.

	Estimate	Std. Error	t value	Pr(>|t|)
(intercept)	6.2037	0.4132	15.0148	0.0000
category: dark	-11.9328	0.5783	-20.6333	0.0000

Table 7.7: London Male model 7, Darkness by Category

7.4.6 Summary

Overall, the London Male shows a clear distinction in articulation and acoustics between the first five and last five phonological environments studied in Experiment 2, suggesting a categorical approach is the best way of summarising this data. The PCA also corroborates this pattern, and is confirmed as significantly bimodal by Hartigan's the dip test. The Tukey HSD comparisons suggest, for the most part, that the first five and final five phonological environments are behaving differently across category, but the same within. The linear models suggest that category can account for the patterns found, and that adding duration does not significantly affect the model. Although its effects can be seen through the significant correlation in Figure 7.22, category alone can account for all the variance found.

7.4.7 Comparing London speakers

In general, the two London speakers are very much in-step with one another. Both exhibit convincing evidence for a categorical distinction of allophonic variation in their /l/ systems, with light [l] found in phrase-level onset position,[6] and vocalised variants found in phrase-level coda position. However, they do vary from each other in a couple of ways, both of which are associated with duration. The London Female is more sensitive to durational differences, as we have already seen. The London Male shows significant effects of duration, but these drop out when entered into more sophisticated statistical analysis. This could be evidence of the London Female being more advanced. Overall, then, it seems plausible to suggest that the London Male and the London Female may be at different points in the same diachronic trajectory, with the incipient durational effects shown by the Male having become clearly phonologised as a gradient rule targeting the dark category in the Female.

Another way in which the two Londoners vary is their treatment of word-final preconsonantal /l/. As Table 7.8 shows, the RP and London speakers' *pee**l***-type tokens are almost the same, but the pre-consonantal tokens are significantly different for the London Female. The London Male does not show this large durational discrepancy between the two contexts. However, it does not seem to be that the London Female's phrase-final /l/s are much longer (although they are), but that the preconsonantal /l/s are much shorter. This could be something to do with /l/-vocalisation. The tongue tip gesture is absent usually from such contexts and when a following consonant comes next then the /l/ may be deleted altogether. This flags an inconsistency in the terminology of vocalisation of liquids. /r/-vocalisation in non-rhotic varieties of English today refers to a complete deletion of any /r/-like sound, although related processes of breaking and lengthening of the preceding vowel are present. With /l/-vocalisation, there is still some gesture associated with the /l/ present. It could be the case that these gestures are absent altogether in preconsonantal contexts for the London Female, but this would need to be investigated with a thorough temporal analysis. Whether this kind of advanced /l/-vocalisation will be found for future generations of vocalisers will be interesting to monitor. Indeed, the word *peel* in the preconsonantal target does sound as though it could be identical to *pill* for the London Female, but not for the London Male. Unfortunately as this potential vowel merger was not predicted before the experiment, it was not tested, but could certainly form grounds for future research. It is well reported that /l/-vocalisation results in mergers of preceding back vowels, but less has been done regarding front vowels. Nevertheless, this merger of *peel* and *pill* is mentioned for vocalising dialects by Harris

[6]This could be potentially be word-level but we would need the ***heal** it*-type tokens to be sure, which is one drawback of Experiment 2.

	*pee**l***	*pee**l** bananas*
RP	0.278	0.252
London F	0.296	0.158
London M	0.259	0.231

Table 7.8: Comparing mean durations for RP speaker alongside London Female for *pee**l*** and *pee**l** bananas*-type tokens

(1994:267).

This is another potential indicator that the London Female is more advanced than the male in her vocalisation, as she has extra durational differences phrase-finally, and when a consonant follows a word-final /l/, half of its duration is cut off. The London Female is a decade younger than the London Male, so perhaps this is a gradual change in progress, which would be expected to be led by women anyway. On the other hand, it could be an idiolectal difference which would fall out when a larger number of speakers were analysed.

Whether the London speakers show a less advanced pattern of darkening and vocalisation than Essex is because the phenomenon is more recent in London, or whether /l/ lenition has propagated more rapidly through the Essex speech community is not clear here. Nevertheless, we can note that the London speakers seem more consistent and stable than the Essex speaker from the previous chapter. The younger more aggressive lenition process of vocalisation seems to have supplanted darkening at the word/phrase level, resulting in a two-way distinction between light and vocalised /l/ for this variety. The behaviour of the Essex speaker, in contrast, suggests an intermediate stage in an ongoing process of change: it shows sociolinguistic sensitivity, as well as an /l/-darkening process which has advanced to the stem-level, so that the next stage on the lenition trajectory also needs to advance through more levels before supplanting it. Moreover, some sociolinguistic research suggests that Essex speakers may be vocalising prevocalically and even stem-finally before back vowels in words such as *falling* (Gibb 2014). If so, this would provide even more evidence for the predictions of the life cycle.

7.5 Manchester WC

The Manchester speaker analysed in Chapter 6 showed that the descriptions in the the existing literature, claiming Manchester /l/ is dark in all phonological environments (Beal 2008; Cruttenden 2008; Kelly and Local 1986), have articulatory corroboration to an extent. This Manchester speaker in Chapter 6 realised all /l/s in the five contexts studied in Experiment 1 with a backed tongue root, lowered tongue body and reduced tongue tip gesture. However, on closer inspection, the SS ANOVA statistical tests showed a small but significant difference between phrase-final *hea**l***-type tokens and the other phonological environments studied. The problem with the experiment was that there were no other contexts in which a darker variant could occur for this speaker. This made it impossible to tell whether the extra tongue-root retraction phrase-finally was a separate, although articulatory similar, category, or whether this speaker was displaying a gradient, possibly duration driven, phonetic effect which would only ever occur in phrase-final position, and not in other environments. Although the SS ANOVA would not have fit such tight confidence intervals if this was a fluke, the fact that this potential analysis rested on just five tokens was a concern. For this reason, the speaker was invited back to participate in Experiment 2.

7.5.1 Ultrasound splines

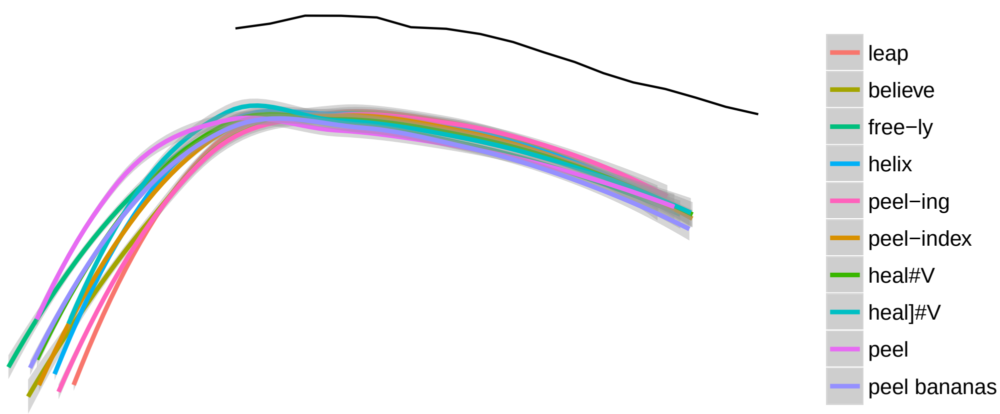

Figure 7.23: Manchester WC splines at /l/ midpoint across phonological context

A quick glance at Figure 7.23 shows a very similar picture to that in Chapter 6, but with more variation in the tongue root area due to the five extra phonological environments collected in Experiment 2. The palate trace helps to visualise just how reduced the tongue tip gesture is at the midpoint for this speaker. This is accentuated in Figure 7.24 where the initial ***leap***-type tokens are compared to the high front vowel [i]. Although it is rather difficult to pick out in Figure 7.23, there do seem to be differences in tongue root retraction in different environments. This is exemplified in Figure 7.25, which focusses on ***leap*** and ***peel*** tokens only.

Figure 7.25 shows the individual splines for each of the 10 ***leap***-type and 10 ***peel***-type tokens. The extra tongue root retraction in the ***peel***-type tokens is convincing and consistent, and shows no overlapping confidence interval with the initial tokens, showing that the pattern observed for the speaker in Experiment 1 was not down to chance. However, this extra tongue retraction cannot be seen to the same extent for the ***peel*** *bananas*-type tokens, which seem to fall somewhere in between initial and final environments.

Turning back to Figure 7.23, this speaker shows a different overall pattern from that observed in RP (Figure 7.1) and London (Figures 7.8 and 7.16) in terms of phonological distribution. Although these speakers also showed a difference between their initial

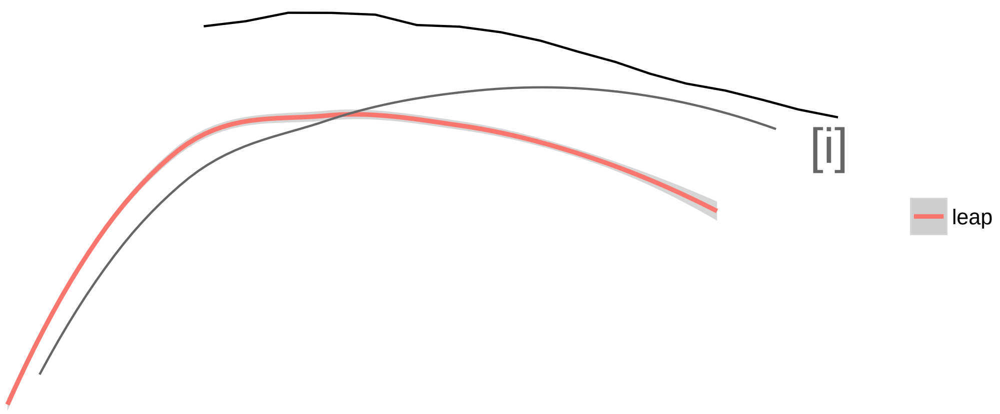

Figure 7.24: Manchester WC comparison of a high front vowel /i/ with /l/ midpoints of word-initial tokens.

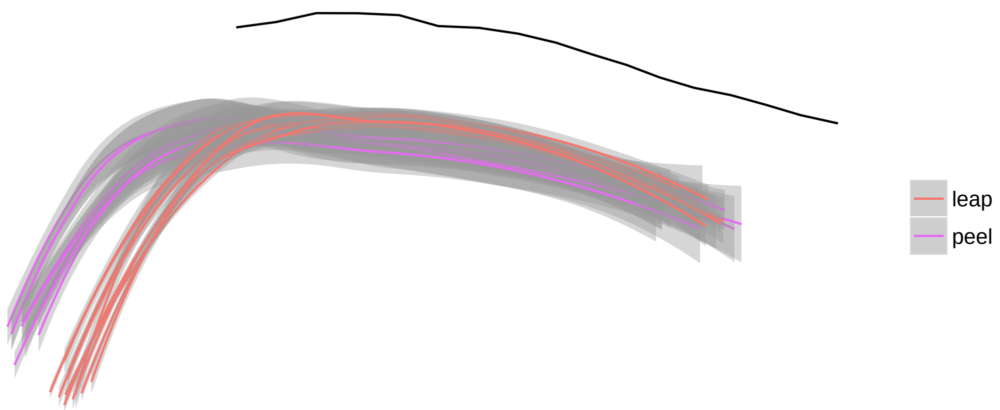

Figure 7.25: Manchester WC speaker's individual splines at /l/ midpoint in initial and final positions

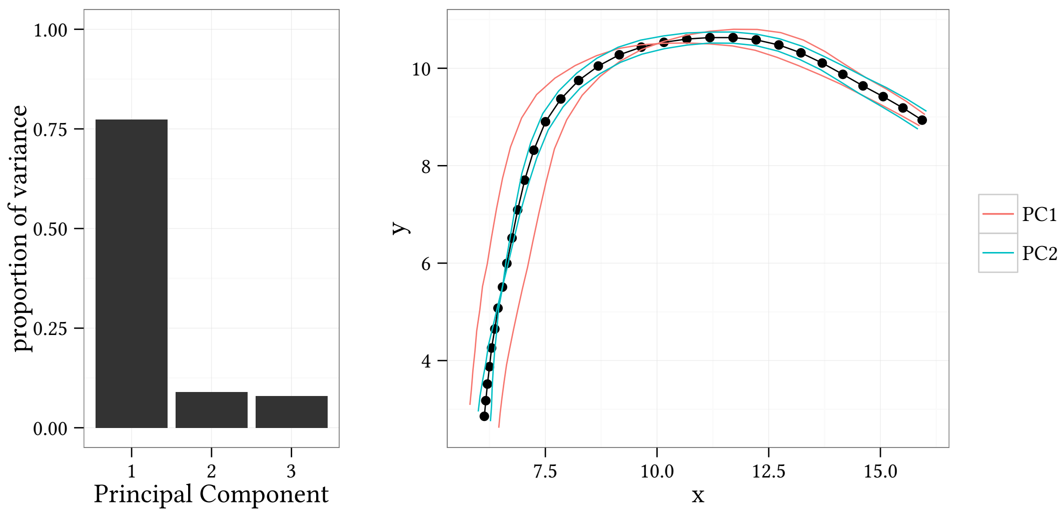

Figure 7.26: PCA summary. Variable importance and loadings plots from the Manchester WC PCA output

and final /l/s, the other phonological environments patterned with the expected extreme depending on whether the /l/ was in a typical onset or coda position. This meant that these speakers' categories consisted of a five to five split down the middle for the ten phonological environments in Table 7.1. With the Manchester WC speaker, the distribution does not match this: initial and final /l/s are significantly different, but the other phonological environments pattern somewhere in between these two extremes. This pattern was not clear from Experiment 1, where only one phonological context with coda [l]s was elicited, highlighting the importance of collecting a wide range of possible environments when conducting such studies, particularly if the question of categoricity and gradience is important to the investigation. These observations raise doubt as to whether this speaker does display two categories. On the other hand, they make the case for the RP and London speakers' categorical distributions even stronger, as they show that this is not simply an effect found in all speakers. The question whether this speaker shows an effect of category will now be addressed with additional quantitative information.

7.5.2 PCA

7.5.2.1 Variable importance and loadings

The Manchester WC speaker's PCA findings are plotted in Figure 7.26. As we have already noted, the main area of variation for this speaker is in the tongue root area, and the PCA accounts for this. A high PC1 corresponds to a fronter tongue root and a low PC1 to a backer tongue root. PC1 accounts for 77% of the variation, and PC2 9%. Again, the loadings plot in the right panel confirms that PC2 does not play much of a role here.

7.5.2.2 Results

The PCA in Figure 7.27 allows us to observe this distribution from a quantitative perspective. One immediate observation is that this speaker does not show the nice linearly trend seen previously. The *free**l**y*-type tokens are usually amongst the lightest in other speakers, but this environment falls out of line here. From inspection of the spectrogram and relistening to the recordings, it is clear that this is coarticulation due to this speaker's very lax pronunciation of the *happ*Y vowel. As shown by Turton and Ramsammy (2012), Mancunians exhibit an opposite pattern to the widely reported phenomenon of *happ*Y-tensing found in the South of the UK, whereby the *happ*Y vowel lowers and backs in phrase-final position to become more like [ɛ̈]. That means, for this speaker, /l/ is not flanked by two high front vowels in this position. From this point on, the *free**l**y* tokens are removed from the analysis for this speaker. Setting the *free**l**y*-type tokens aside reveals a stepwise gradient pattern, but still one that holds only approximately. Reflecting the results from the spline plots in Figure 7.23, the phrase final *pee**l***-like tokens have a lower PC1 than the *pee**l** bananas*-type tokens.

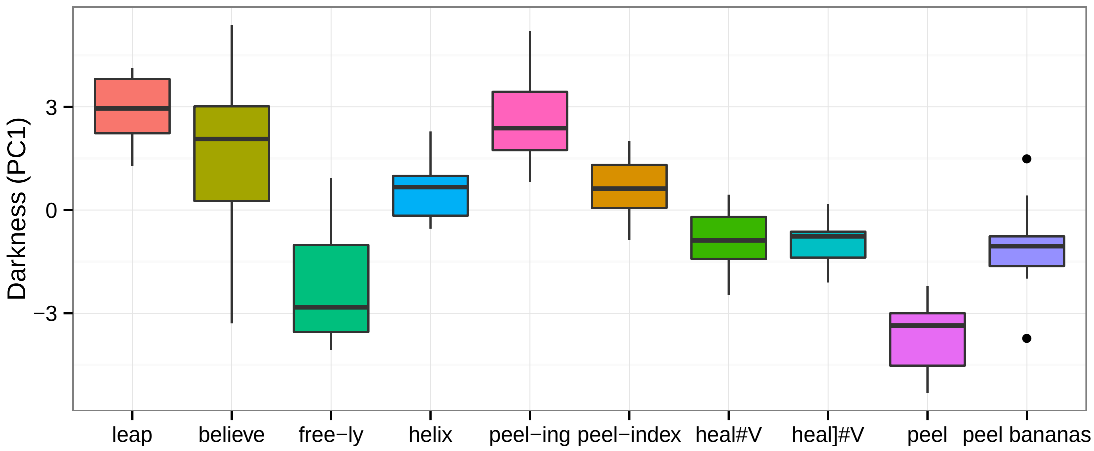

Figure 7.27: Manchester WC PC1 values across phonological context

Nevertheless, this speaker does not seem to be showing any evidence of a clear bimodal distribution, as the density plot of the PC1 values in Figure 7.28 demonstrates. Following the convention I have used for similar plots in this chapter, the darker range represents tokens of /l/ in the coda at the word level (which for previous speakers were categorically dark), whereas the lighter range represents tokens of /l/ in the onset at the word level (which for previous speakers were categorically light). In this case, however, the overall distribution, indicated by the dashed line, shows absolutely no dip. Indeed, there is no statistical support for a bimodal treatment, with Hartigan's dip test giving a low dip value of 0.025 (p= 0.97). The two superimposed categories do show a tendency of being respectively lighter and darker, but this is not bimdoal, with too much overlap of the ranges.

The Tukey HSD results for this speaker are difficult to summarise and interpret.

There are inconsistencies on the whole but overall the picture points towards a distinction between the canonical categories seen for the previous speakers, but with extra significant differences between phrase-final *pee**l***-type tokens and everything else. Phrase-final *pee**l***-type tokens are the only environment to show significant differences to all other categories throughout the articulatory data (and in the acoustics for the most part; see Section 7.5.3). The main inconsistency is through the *pee**l**-index*-type tokens, which are difficult to categorise. They are significantly different to ***l**eap*, *pee**l**-ing* and *pee**l***-type tokens, but not anything else. This shows an intermediate distribution which is significantly different to the extreme categories, but which overlaps with many other intermediate contexts in the range. The majority vote puts it in with the dark tokens, which is how it will be categorised for the linear models in Section 7.5.5.

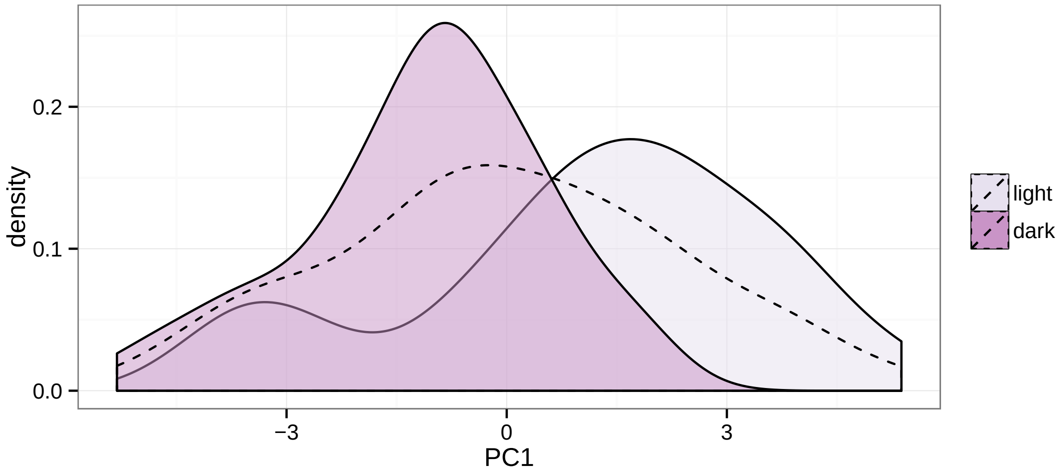

Figure 7.28: Density plot for Manchester WC showing separate distributions for light and dark categories (categories determined by phonological environment).

7.5.3 Acoustics

The acoustics paint a similar picture to the spines and PC1 plots, resulting in a non-significant dip statistic of $D = 0.031$ ($p = 0.753$). However, there is one clear discrepancy between the acoustics and articulatory data, in the *be**l**ieve*-type tokens. In the splines, these are one of the lightest phonological contexts, but in the acoustics one of the darkest. For example, the *be**l**ieve*-type tokens are one of the lightest in the articulations, but patterns with the dark [ɫ]s in the acoustics. This may be an effect of the preceding weak vowel's deletion resulting in a branching onset which has an effect on the acoustics (this effect is discussed in more detail in Section 7.8). However, one would expect to see this effect in the splines also. One interesting aspect of the formant plot for Manchester WC is the lack of difference between the *pee**l***-type tokens and the other dark ones. In the articulations, this difference was clear and distinct. In the acoustics, it is muddied, showing once again that ultrasound tongue imaging and articulatory data can pick up on patterns not found in the acoustics.

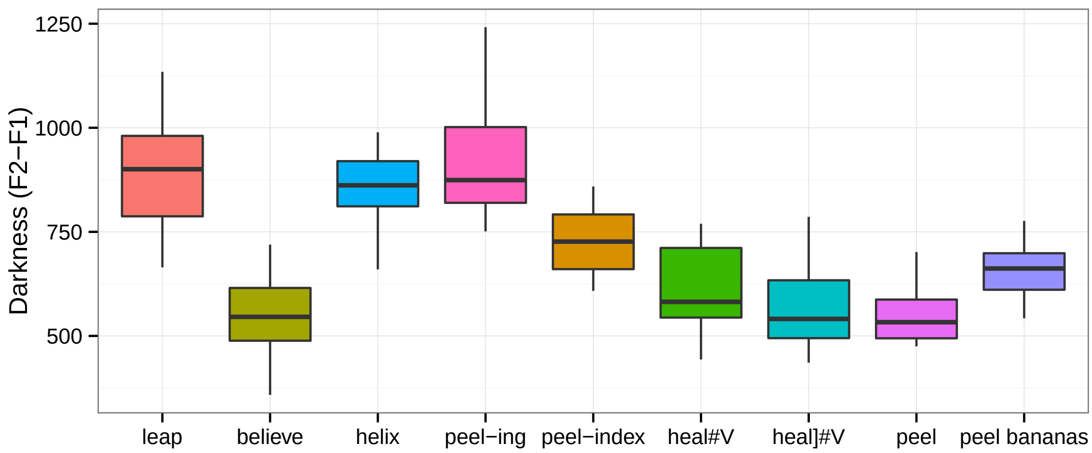

Figure 7.29: ManchesterWC F2-F1 values across phonological context

The acoustic data from the Manchester WC suggest that, although there is clearly some variation dependent on phonological contexts in this dialect, it is very gradual across contexts. The speaker has a clear phrase-final effect of /l/-darkening, possibly dependent on duration. Whether this is a cognitively controlled process affecting phrase-final /l/s only, or an epiphenomenal effect of duration is not clear, but perhaps the linear models later on in the chapter can shed some light on this further.

7.5.4 Darkness vs. duration

The PC1 box plot above in Figure 7.27 shows a somewhat gradient pattern from lightest to darkest. As we have seen from previous studies of American English, this kind of pattern may be expected to show an effect of duration. However, that is not what Figure 7.30 indicates, where PC1 does not seem to correlate with duration at all. Separate smoothers are fit to presupposed light and dark categories, based on whether the /l/ is in the onset or coda at the word level. The assigned light category has near to no correlation, with a Pearson's r of 0.115 (p= 0.486). Note that the *free-ly* tokens are removed from the analysis. The assigned dark category shows a slightly higher correlation, but not significantly so with a Pearson's r of -0.204 (p= 0.16). This is perhaps unsurprising given Carter (2002:158)'s acoustic results from Manchester /l/, which found no correlation between darker variants and longer duration. In fact, Carter 2002 noticed a slight trend in the opposite direction: the greater the duration of the initial /l/, the lighter it was. However, this trend does not reach significance in his acoustic study due to a small amount of tokens. The 90 tokens plotted here show that this was probably down to chance.

As mentioned above, the phrase final *pee**l***-like tokens have a lower PC1 than the ***peel** bananas*-type tokens, which is reflected in the spline plot in Figure 7.23 and the PC1 box plot in Figure 7.27. This is the one environment whose tokens look as though darkness may correlate with duration. This environment shows a very high and significant

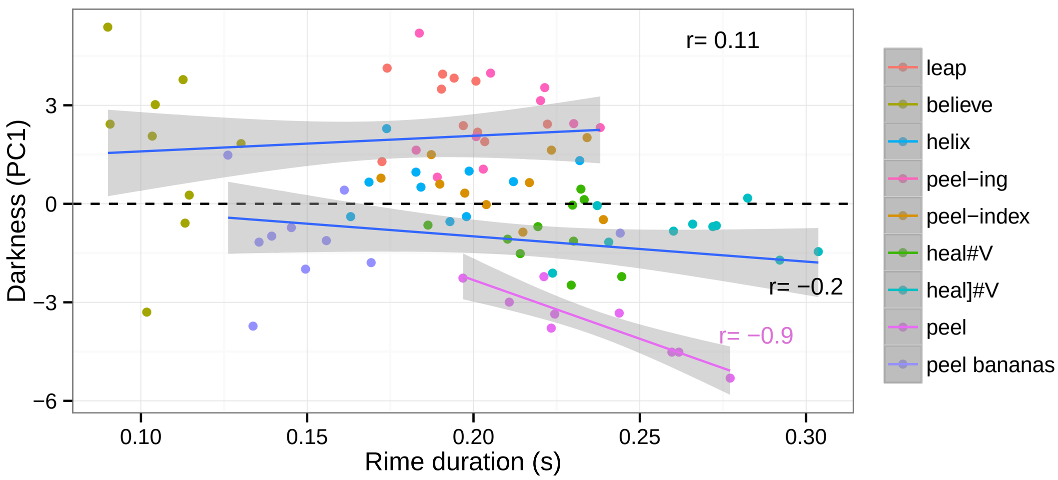

Figure 7.30: Manchester WC PC1 values against rime duration. Separate smoothers fit for 'light' and 'dark' categories, and phrase-final tokens.

correlation, with a Pearson's r of -0.904 (p< 0.001).

Thus, there is only limited support for the hypothesis that, in the absence of a sharp categorical distinction between light and dark allophones (as in RP and London English), Manchester would show a strong correlation between darkness and duration (as found in previous studies of American English). The correlation with duration holds only in a very narrow set of environments: utterance-final ones.

This brings us to Yuan and Liberman (2009; 2011), who found a correlation between darkness and duration for dark [ɫ]s only in their acoustic study of American English. As we have seen, our London informants exhibit a similar pattern, which suggests that Yuan and Liberman's description of American English may well be correct. On the other hand, Yuan and Liberman did not distinguish between subcontexts within their dark category. As a result, there is a risk that the correlation they report may not in fact hold for all dark /l/s, but only for phrase-final tokens: these could be doing most of the work, as in our Manchester WC informant. Once more we see that firm conclusions are possible only in experiments using a fine-grained sample of the whole range of possible morphosyntactic and prosodic environments.

7.5.5 Models

Table 7.9 shows the adjusted r^2 values for different linear models fit to darkness (i.e. PC1). The number of models fit for this speaker reflects the difficulty in analysing this pattern. A model based solely on duration (1) performs very poorly indeed. This might be particularly surprising given that Manchester is a very dark dialect, like American English, but as we have seen in the previous section, there really seems to be little correlation between darkness and duration at all in this variety. Model 1 confirms this further. Category, based on a binary distinction between word-level onsets and word-level codas, does poorly too (2), and adding duration to it either as standard predictor (3) or as

model	adj. r^2
1. Darkness ∼ log(Duration)	0.079
2. Darkness ∼ Category	0.221
3. Darkness ∼ Category + log(Rime duration)	0.44
4. Darkness ∼ Category * log(Rimeduration)	0.445
5. Darkness ∼ Category2	0.439
6. Darkness ∼ Category2 + log(Rimeduration)	0.433
7. Darkness ∼ Category2 * log(Rimeduration)	0.451
8. Darkness ∼ Context	0.665
9. Darkness ∼ Category3	0.03

Table 7.9: Comparing duration correlations for Manchester WC speaker alongside adjusted r-squared of models with a category duration interaction

an interaction (4) does nothing to improve it, statistically speaking. A different way of classifying the data is shown in models 5-7, where the phrase-final ***peel***-type tokens are separate from the other word-level codas as a third category, but this fails to add anything. Because these models perform so poorly, separate phonological context is tried as a possible predictor in model (8), and although it performs much better it is still a weak fit overall. One would expect a model allowed the freedom of fitting ten phonological environments to perform very highly indeed. Adding duration in any capacity (5, 7) to a pure contextual model does nothing in terms of significantly improving the model. Reclassifying all of the categories into light with phrase-final as dark does very badly as a model (9), showing that this is not just a case of dark [ɫ] everywhere and darker [ɫ] in phrase-final position, but more of a gradual relationship through the system, and this categorisation will be left here. All of the comparisons can be observed in Appendix D.

7.5.6 Summary

Quantitatively, it seems that none of these predictors do very well at summarising the darkness patterns in Manchester. This is most likely because there is one clear pattern here which differs from everything that we have seen so far. This speaker has no distinction between any of the phonological environments apart from phrase-finally. This is the only context which shows a significant correlation with duration. However, treating this as a separate category does not provide a well-fitting model and so there is no reason to think of the /l/-darkening system here as one with an added effect of /l/-darkening targeting the tongue root in phrase final position only. For now, it is safe to say that there is no convincing evidence of two allophonic categories for this dialect, although temporal data could help us investigate the phasing of relative gestures here, and a potential investigation of tip delay. This may very well show correlations with duration, but for this dataset, duration plays an even smaller role than in a variety with a clear allophonic distinction, such as RP or London.

Overall, this dataset has given us an insight to some of the phonetic spectra which

American linguists are dealing with when analysing variation in English /l/. A very small distance between the first and second formants (see Section 7.10.4 for a clearer comparison), alongside a gradient scale of articulatory darkness could understandably lead the research to conclude there are no categories for this speaker. There are two things to keep in mind before drawing this conclusion:

i. Categorical effects may be present in a temporal analysis which takes into account the relative phasing of gestures.

ii. There may well be no categorical distinction in the particular variety studied. Yet, as we have seen elsewhere, it would be entirely unwarranted to generalise this conclusion to other varieties.

7.6 Manchester MC

The data from Manchester that we have considered so far matches the impressionistic reports found in the descriptive and dialectological literature, which hold that Lancashire and Manchester /l/s are dark in all environments. The previous section has shown that this is generally the case, although there are some fine-grained differences in realisations that articulatory analysis can tease out. However, we know from sociolinguistic work that /l/ variation may differ across social groups. It is therefore a matter of interest to establish whether the pattern of dark /l/ everywhere is found throughout the social scale. The speaker studied in this section is female, and of a similar age to Manchester WC, but from a more middle-class socioeconomic background. As mentioned in the speaker summary in Chapter 5, this speaker shows many features of a more middle class Mancunian accent.

7.6.1 Ultrasound splines

The Manchester MC /l/s show a different distribution to the widely reported 'all dark' pattern typical of the Manchester accent, and the distribution in Figure 7.31 shows a rather different situation to the WC Manchester speaker[7]. Rather than all contexts overlapping one another, the splines for this speaker tend to be staggered, although perhaps not as convincingly as we saw for the the RP speaker. Some of the realisations look a little more intermediate, such as the tongue root retraction in the *free**l**y*-type tokens. These will be analysed further below.

[7]This speaker produces a schwa in the first syllable of *linoleum* which has a slight by significant coarticulatory backing effect on the /l/, and so these tokens have been omitted from the analysis for this speaker.

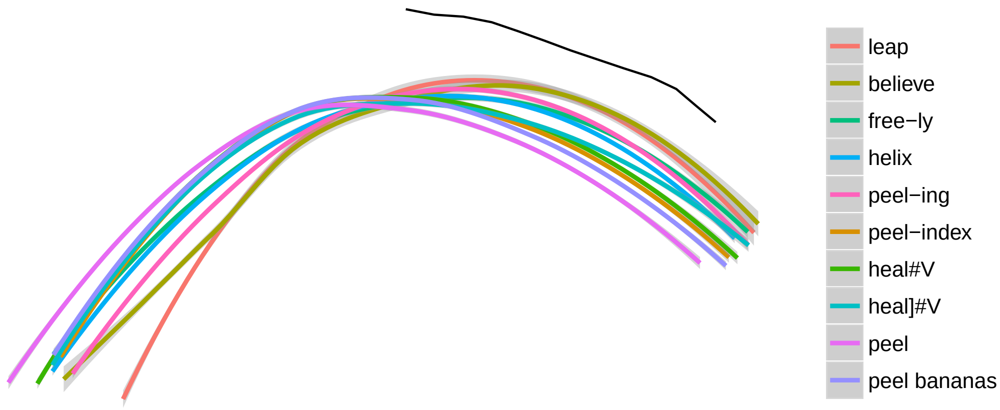

Figure 7.31: Manchester MC speaker's splines at /l/ midpoint across phonological context

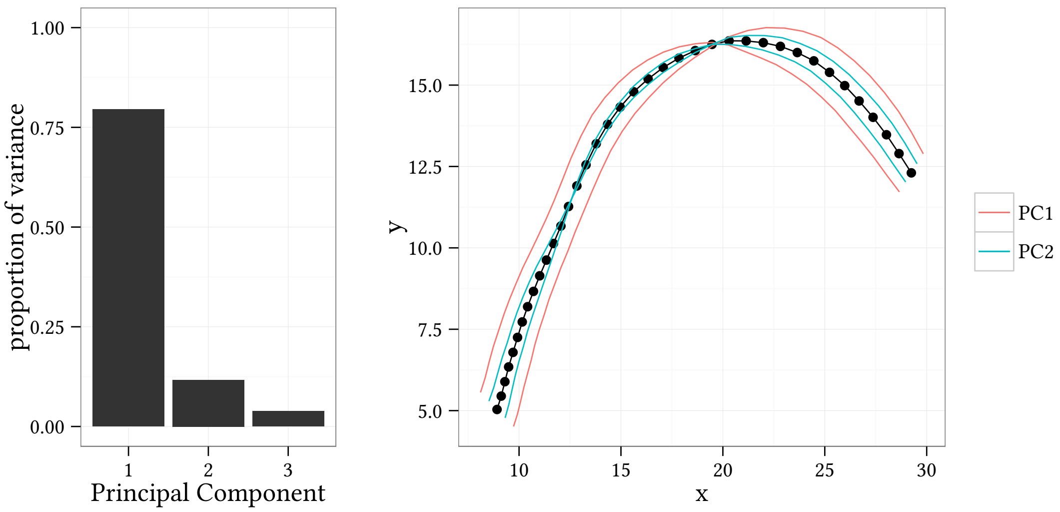

Figure 7.32: PCA summary. Variable importance and loadings plots from the Manchester MC PCA output

7.6.2 PCA

7.6.2.1 Variable importance and loadings

The Manchester MC speaker's PCA findings are plotted in Figure 7.32, with PC1 accounting for 79% of the variation. PC2 does not play much of a role, as shown in the loadings plot on the right, accounting for just 11% of the variation.

7.6.2.2 Results

The staggered pattern shown by the splines in Figure 7.31 becomes clearer in the PCA box plot (Figure 7.33). On visual inspection, the distribution appears compatible with a gradient scale of darkness. If we attempt to sort the environments into groups, the result is a rather fine subdivision: the word-level onsets seem to split into two groups

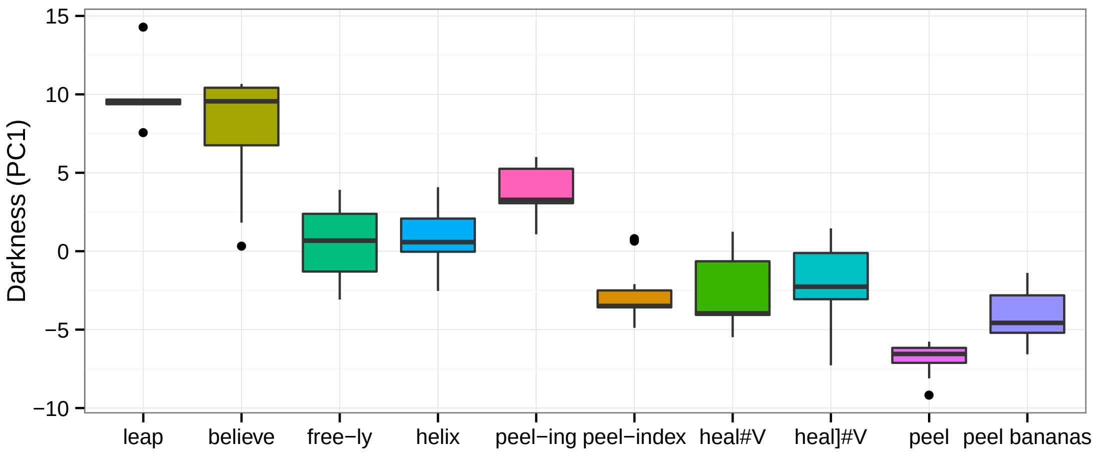

Figure 7.33: Manchester MC PC1 values across phonological context

(***leap*** and ***believe*** vs. all the others), whereas ***peel***-type tokens differ from all the other word-level codas. All in all, then, the case for category differentiation looks weak on a preliminary inspection, just as it was for Manchester WC. The main difference seems to lie in the fact that Manchester MC exhibits a broader articulatory range, allowing for less overlap between environments. This may conceivably reflect a tendency on the part of this speaker to use hyperspeech in the formal setting of the lab.

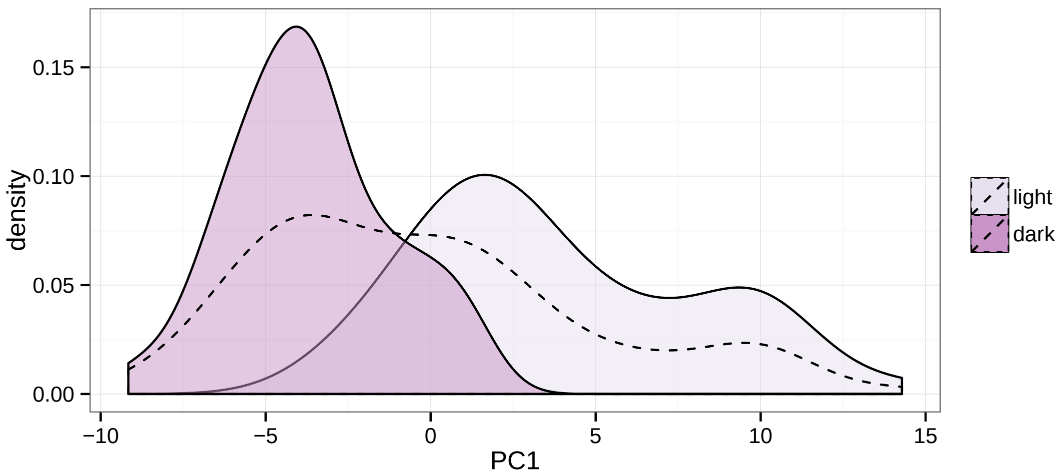

Figure 7.34: Density plot for Manchester MC showing separate distributions for light and dark categories (categories determined by phonological environment) and overall distribution indicated by dashed line.

The dip statistic shows that evidence for bimodality is weak, with $D = 0.031$, and a non-significant p-value (p =0.708). This isn't necessarily surprising, given that the distribution is very gradual rather than bimodal. However, the dashed line in the density plot in Figure 7.34, which treats the entire distribution as one, does seem to show three bumps, suggesting some kind of multi-modality which does not reach significance.

The Tukey HSD tests show a complex but robust pattern which is consistent through both the PC1 and acoustic tests. It is the same pattern suggested by the box plots in Figure 7.33. The distinction in the significance values from lightest to darkest is as follows: ***l**eap*, *be**l**ieve* > *free**l**y*, *he**l**ix*, *pee**l**-ing* > *pee**l**-index*, *hea**l**#V*, *hea**l**#]V*, *pee**l** bananas*> *pee**l***. This suggests Manchester MC has the same phonological pattern as Manchester WC, but implemented over a somewhat broader articulatory range. This may possibly reflect a greater tendency to use hyperspeech, in turn conceivably driven by sociolinguistic factors (see my suggestion in 7.6.6). The significant differences found by Tukey's HSD test are probably artifacts of this stretched articulatory range.

7.6.3 Acoustics

The acoustics in Figure 7.35 show evidence of looking more like two distinct categories, with the usual distinction of word-level onset vs. coda /l/s. Conducting a dip test on the acoustics give $D = 0.066$, and a significant p-value (p < 0.001). This suggests that the acoustics show a bimodal distribution similar to that found in London, due to the secondary distinction in the lighter tokens being smoothed out in the acoustics, as well as the extra dark *pee**l***-type environment. This also might suggest that the acoustics cannot see some of the fine-grained differences which the articulatory analysis brings out. However, the Tukey HSD tests do show a significant four way distinction in the acoustics also, again suggesting that a categorical distinction may be unlikely given the gradient pattern, and perhaps that the Manchester MC has the same phonological pattern as Manchester WC, but implemented over a wider continuum.

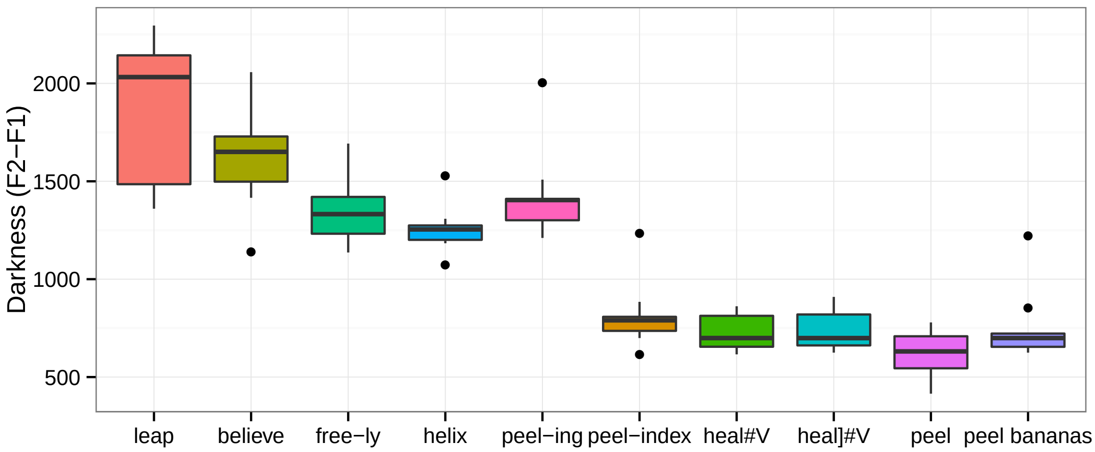

Figure 7.35: ManchesterMC F2-F1 values across phonological context

Perhaps interactions with duration can shed more light on these complex patterns. We saw in Section 7.5 that duration showed no correlation with darkness apart from the phrase-final tokens. Figure 7.36 shows the correlation of PC1 values with the'rime' duration. The input to category for the correlations is the same light/dark split we have seen in all previous speakers.

7.6.4 Darkness vs. duration

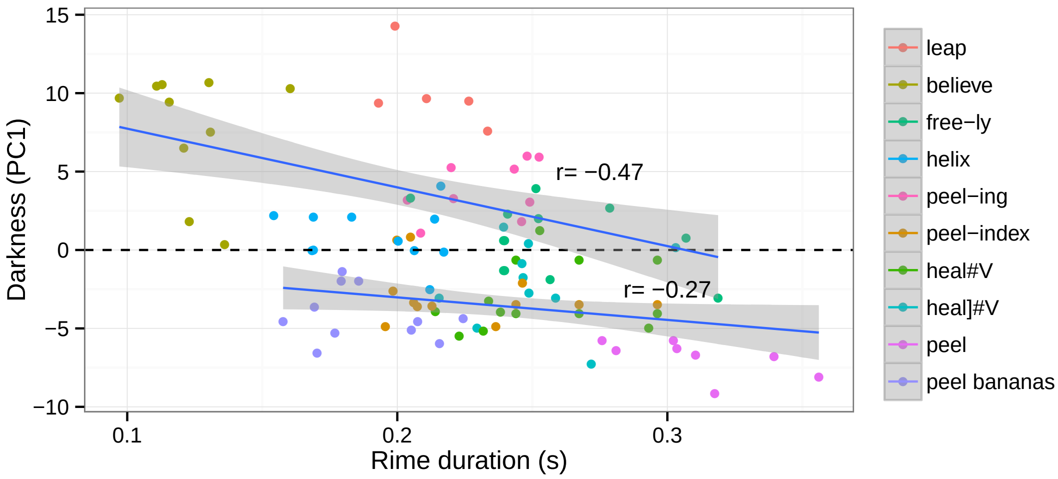

Figure 7.36: Manchester MC PC1 values against rime duration

There seems to be significant correlations for both sides of the plot when considering the ten phonological contexts as two categories. The lighter category, primarily in the upper panel shows a correlation of r = -0.47 which is significant (p < 0.001), as is the correlation of r = -0.27 for the darker variants on the lower panel (p = 0.048). However, as just discussed, it seems the distribution here does not neatly fit into two categories. This can be seen in the plot, creating a false sense of correlation for the so-called 'light' categories in environments 1-5. The *be**l**ieve*-type tokens are all very light and very short, and the *free**l**y*-type tokens have an intermediate duration and darkness with respect to the entire distribution, forming a neat correlation line through blank space.

7.6.5 Models

As the articulatory, acoustic and statistical evidence above shows two categories struggling to capture all of the fine-grained information, linear models give us a better idea of what is going on in this dataset, and any possible interactions. Table 7.10 shows the adjusted r^2 values for seven different linear models fit to darkness (i.e. PC1). Again, duration performs very poorly in a model alone (1). Category performs moderately well alone (2) and this is significantly improved when adding duration (3), although an interaction makes no improvement (4).

'Category 2' indicates the a different categorisation method which relies on a four-way split of the data. Although we cannot test the significances of non-nested models by ANOVA comparison, we can see that the adjusted r-squared indicates a much better fit when splitting the data this way. However, this is not surprising, as allowing the model more freedom to fit different estimates is sure to result in a better fit. We cannot conclude that this is the best way of fitting the data just by this r^2 value. Adding duration in

model	adj. r^2
1. Darkness $\sim$ log(Duration)	0.272
2. Darkness $\sim$ Category	0.549
3. Darkness $\sim$ Category + log(Rimeduration)	0.618
4. Darkness $\sim$ Category * log(Rimeduration)	0.623
5. Darkness $\sim$ Category2	0.774
6. Darkness $\sim$ Category2 + log(rimeduration)	0.774
7. Darkness $\sim$ Category2* log(rimeduration)	0.771

Table 7.10: Comparing duration correlations for Manchester MC speaker alongside adjusted r-squared of models with a category duration interaction

any capacity does not provide a significant improvement to this model (see the ANOVA comparisons in Appendix D).

	Estimate	Std. Error	t value	Pr(>|t|)
(intercept)	-5.3356	2.1647	-2.4649	0.0154
category: dark	-6.3887	0.6706	-9.5267	0.0000
log(Rime duration)	-5.6415	1.2990	-4.3428	0.0000

Table 7.11: Manchester MC model 3, Darkness by Category and log(Rime duration)

	Estimate	Std. Error	t value	Pr(>|t|)
(intercept)	8.5094	0.6191	13.7440	0.0000
category2: onset	-6.8351	0.7503	-9.1093	0.0000
category2: coda	-11.5733	0.7130	-16.2325	0.0000
category2: final	-15.3834	1.0498	-14.6537	0.0000

Table 7.12: Manchester MC model 5, Darkness by Category2

In conclusion, the linear models in Table 7.10 show that, as was the case for Manchester WC, duration is a poor predictor of darkness for Manchester MC. A two-category analysis fits the articulatory data rather poorly. As one would expect, allowing four ad hoc categories improves the fit, but, as shown by the adjusted r^2 values, still captures less of the variation that a simple two-category analysis did in the case of RP, London Male, and London Female. Overall, a monocategorial gradient analysis seems justified, although, again, the gradience is driven by factors to be discovered–not by duration.

7.6.6 Summary

The balance of evidence suggests that Manchester MC is a monocategorial speaker, like Manchester WC, with purely gradient darkening. Figure 7.37 compares the two Manchester speakers' acoustics side by side. This accentuates the pattern found where Manchester MC uses a considerably broader articulatory range, which produces artifactual bumps in the distribution as the ten environments under study are pulled apart (Figure 7.34). One could speculate that this use of a broader articulatory range is driven by sociolinguistic evaluation: perhaps the speaker tacitly knows that the 'all dark' pattern exhibited by Manchester WC is nonstandard. If so, the interesting result is that, acoustically, she does succeed in approximating an RP-like or London-like bicategorial-sounding distribution (Figures 7.32 7.35). Yet the articulatory data suggest that this acoustic approximation is achieved by a rather different grammar, one which retains the monocategorial structure of the Manchester vernacular. In other words, Manchester MC has not restructured her inventory of surface categories. Of course, in the absence of more detailed sociophonetic investigation, with better control of register and style, this interpretation remains speculative. Nonetheless, it is again noticeable that the complex nuances of the situation come into view only when one looks at a fine-grained sample of articulatory data.

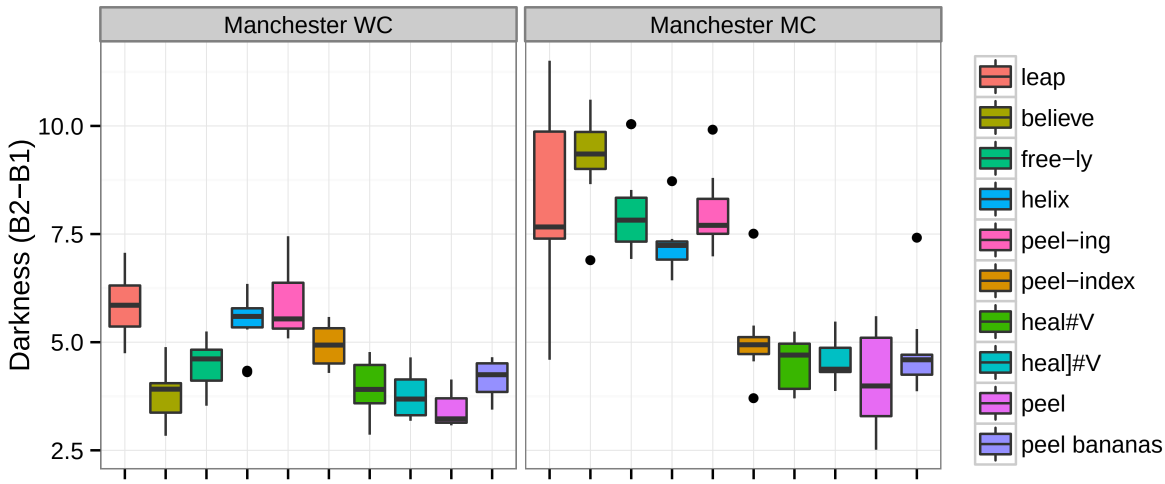

Figure 7.37: Manchester B2-B1 for both WC and MC speakers.

7.7 Newcastle

In Chapter 6 we saw partial corroboration of the dialectological claims about North-East /l/ through the Middlesbrough speaker, who had much lighter /l/s in all positions than other dialects studied in the chapter. However, the speaker still had a difference between initial and final contexts. In the existing literature, the 'all light pattern' is largely reported for areas further North, such as Northumbria and Newcastle. In this section, we turn to the biggest city in the North-East to see if the claims of 'all /l/s are light' is true here, and what comparisons can be drawn with our previous speakers.

7.7.1 Ultrasound splines

Figure 7.38 shows the midpoint splines for the Newcastle speaker. Although they are very tightly clustered, it is possible to observe a small but distinct split between some of the environments. Close inspection of the splines shows that the first five phonological environments are not significantly different from one another in terms of confidence interval overlap, and they pattern together as the lightest variants. The final three phonological environments are also not significantly different from one another and pattern together as the darkest variants. The mean splines of the *pee**l**-index* and *hea**l**#V* environments, however, show some intermediacy between the two categories, and from each other, with *pee**l**-index* patterning separately from everything else in the tongue root area (but patterning with the onset /l/s in the tip area) and *hea**l**#V* tokens patterning separately for a very small portion of both the tip and root area. As Figure 7.39 shows, this is down to variation in these two areas, rather than an intermediate realisation. Variability and intermediate realisations are discussed further in Sections 7.10.2 to 7.10.2.1.

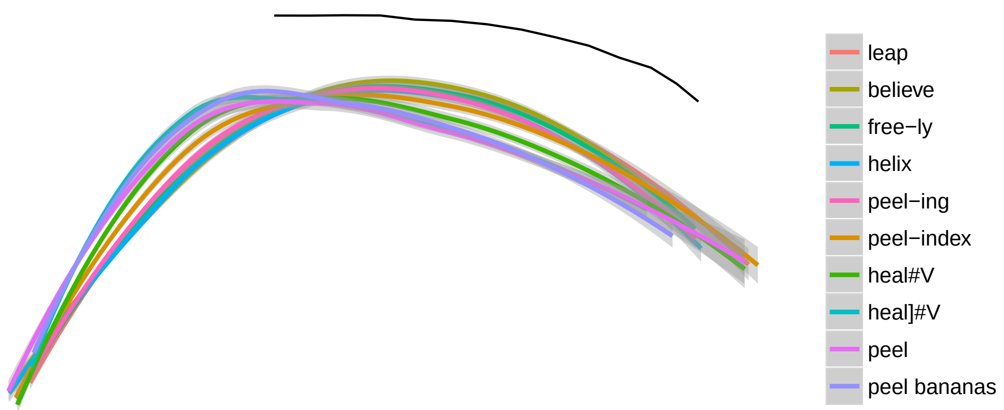

Figure 7.38: Newcastle splines at /l/ midpoint across phonological context

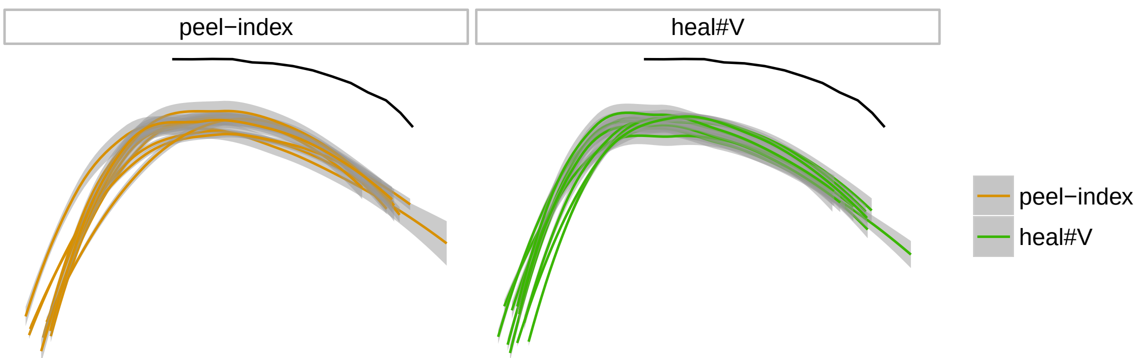

Figure 7.39: Newcastle splines at /l/ midpoint in *peel-index* and *healV* tokens, with initial and final means added for reference.

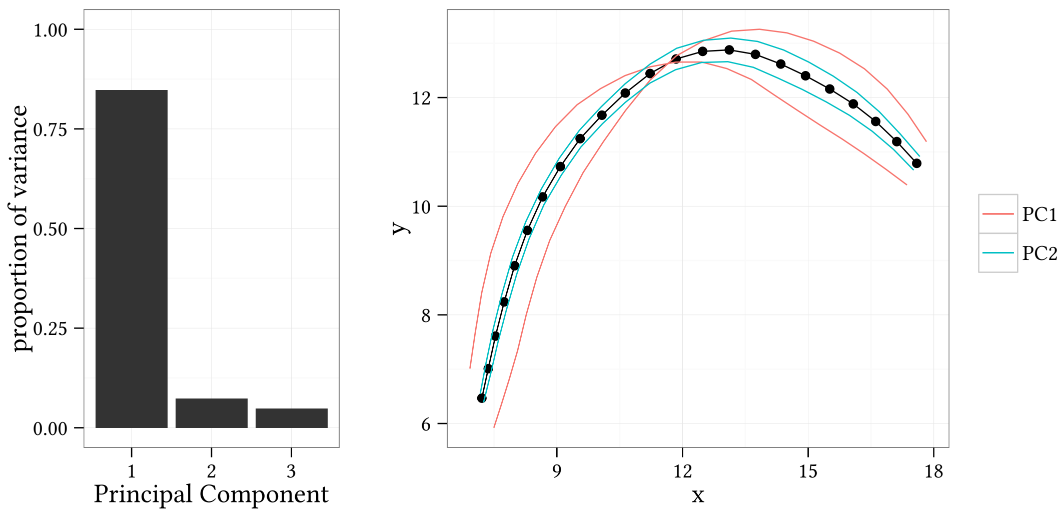

Figure 7.40: PCA summary. Variable importance and loadings plots from the Newcastle PCA output.

7.7.2 PCA

7.7.2.1 Variable importance and loadings

PC1 accounts for 84% of the variation (PC2 just 7%) and therefore the analysis of this speaker will focus on the first component, with higher values reflecting lighter /l/s.

7.7.2.2 Results

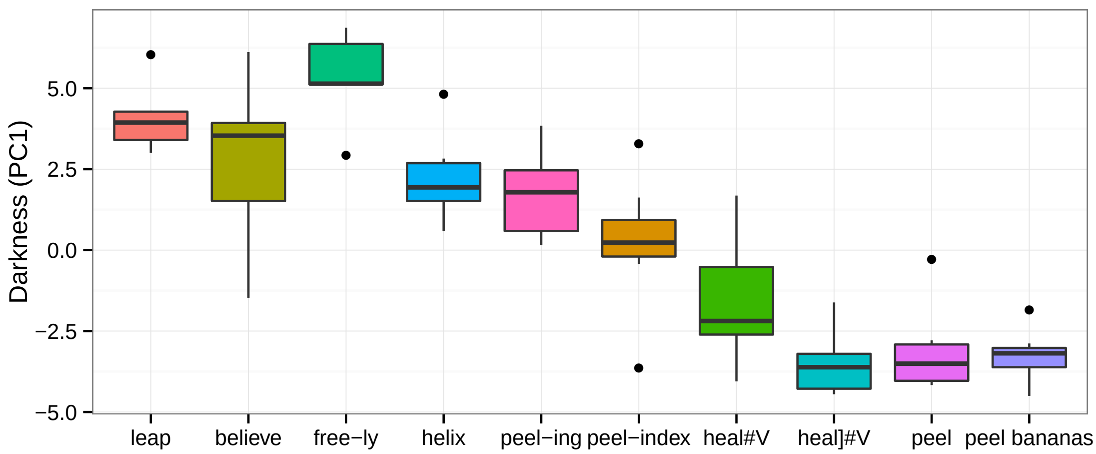

Figure 7.41: Newcastle PC1 values across phonological context

Figure 7.41 shows the PC1 values plotted for each context. The box plots indicate a relatively stable lightness in the first five contexts and last three, with gradience in the *peel-index* and *heal#V* tokens. This plot adds further weight to the evidence from Figure 7.38 that not all Newcastle /l/s are the same, even though the range is articulatorily smaller and less distinct than varieties such as RP. The gradience in the final five contexts is a little misleading, however, given the range of the ***peel**-index* and ***heal**#V*-type tokens. This can be seen from the outliers in the ***peel**-index* box and the wide stretched range of the ***heal**#V* box. A jitter plot is perhaps a better way of visualising the data for this speaker, as in Figure 7.42, which shows the distribution of individual tokens rather than the average. This plot allows us to note patterns which were impossible or difficult to detect in Figure 7.41 and Figure 7.39. The dashed line indicates a possible PC1 cut-off point between two potential categories, rather than the stepwise gradient pattern suggested by Figure 7.41. There is just one ***peel**-index*-type token which takes a darker realisation, the rest tend to pattern with the onset tokens. This is the first speaker to show consistent resyllabification and lightening of compound /l/s. Everyone else thus far has treated these /l/s as canonical coda realisations. The RP speaker, who we know from Chapter 6 shows light [l] in ***heal** it*-type tokens, has a dark [ɫ] in this environment. This raises the question of the interaction between compound prosodification and phonological stratification. Do all speakers resyllabify the /l/ in ***heal** it*-type tokens because the following word forms a clitic? These points are discussed later in the chapter, in Section 7.10.3.

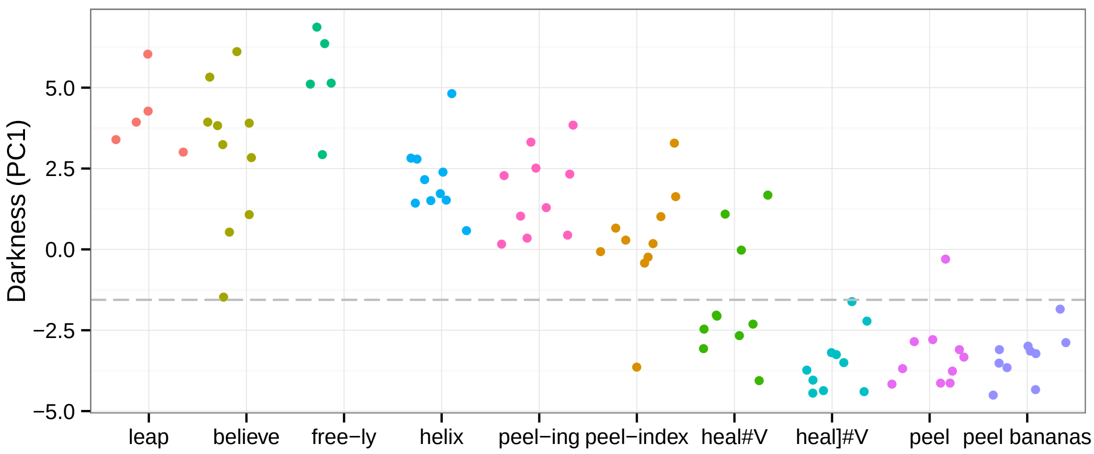

Figure 7.42: Jitter plot of Newcastle PC1 values across phonological context.

The Tukey HSD test indicates a split between the first six environments and the final four. This is fairly clear from the analysis (full results are in Appendix C) and, while not as clear as the RP and London speakers, it is a much simpler situation than with both of the Manchester speakers. As mentioned above, the difference between previous analyses and this one is that the ***peel**-index*-type tokens pattern with the lightest tokens, not the darkest ones. The Tukey HSD confirms this, with one or two significant differences which fall out of line with this analysis: ***peel**-index*-type tokens are significantly darker than ***be**lieve*-type tokens, and not significantly lighter than ***heal**#V* tokens This shows that the intermediacy of this context should probably be considered further (see Section 7.10.3).

The dip statistic shows that there is evidence of bimodality for this speaker, with a significant figure of $D = 0.052$ (p =0.067). This very shallow dip is demonstrated by the dashed line in the density plot in Figure 7.43, and accentuated by the separate ranges for imposed ‘light’ and ‘dark’ categories. Note that the *pee**l**-index*-type tokens are assigned to the light category from now on, rather than the dark one. A couple of bumps in the light range demonstrate that this speaker shows some variability (see Section 7.10.2). So, despite some unclear evidence from the spline and box plots, the dip statistic has confirmed that this speaker has a bimodal distribution in /l/ realisation, albeit one with a much smaller range (note that the p-value just makes the cut-off). The /l/s are certainly very light (as the distance from the splines to the palate show), but there is a distinction here.

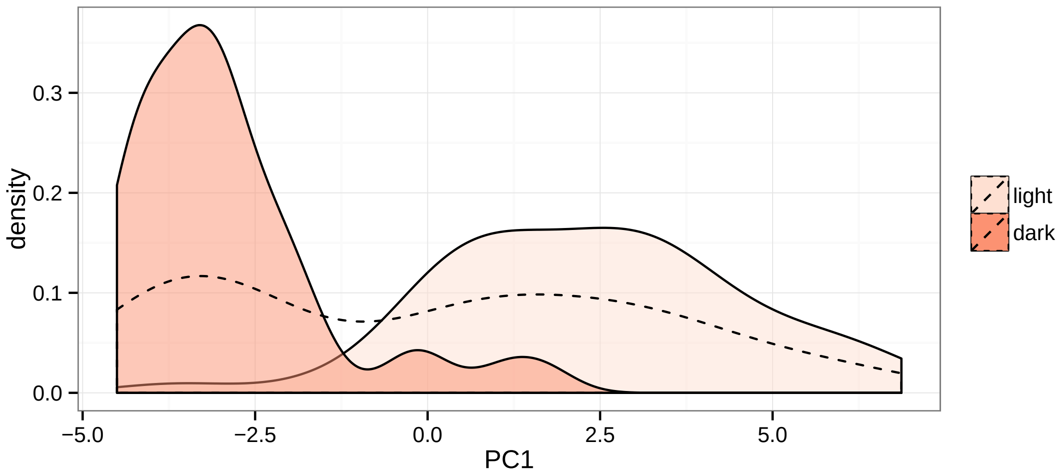

Figure 7.43: Density plot for Newcastle showing separate distributions for light and dark categories (categories determined by phonological environment) and overall distribution indicated by dashed line.

7.7.3 Acoustics

At first glance, the acoustics appear to support a picture of gradient orderliness, represented by the box plots of the F2-F1 values in Figure 7.44. However, there is a small cut-off point with no inter-quartile range overlap between the *pee**l**-index* and *hea**l**#V*-type tokens, i.e. the cut-off between the potential light and dark categories. Moreover, the first four contexts, where /l/ is in the onset at the stem-level, have almost identical median values. The stepwise effect begins with the *pee**l**-ing*-type tokens. Note the relative lightness of all tokens, with most ranging from around 700Hz to 1500Hz, whereas the Manchester speaker ranges from 500Hz to 1000Hz.

The Tukey HSD test on the acoustics shows the same distinction between contexts as the test for the PC1 values did, but with an additional distinction. The *pee**l**-index* tokens are significantly different to all other contexts (borderline with the *pee**l**-ing*-type tokens; p = 0.054), creating an intermediate category. Perhaps this intermediacy is picked up by the acoustics and not the articulatory data due to the Hertz scale being more fine-grained, although the acoustics have generally been more blunt that the articulatory data thus far. No differences within the final four contexts reach anywhere near significance, with most values approaching or reaching 1.

Although the Tukey of the formant values suggests a three-way distinction, with the *pee**l**-index* tokens acting as an intermediate category, we can see from Figure 7.42 and in Figure 7.45 that this probably does not warrant a separate category for the linear models, as there is a lot of overlap with other contexts. Therefore two categories are taken as input to the models below, with *pee**l**-index*-type tokens falling into the 'light' category as opposed to the dark, leaving the goodness of fit decisions with the linear models.

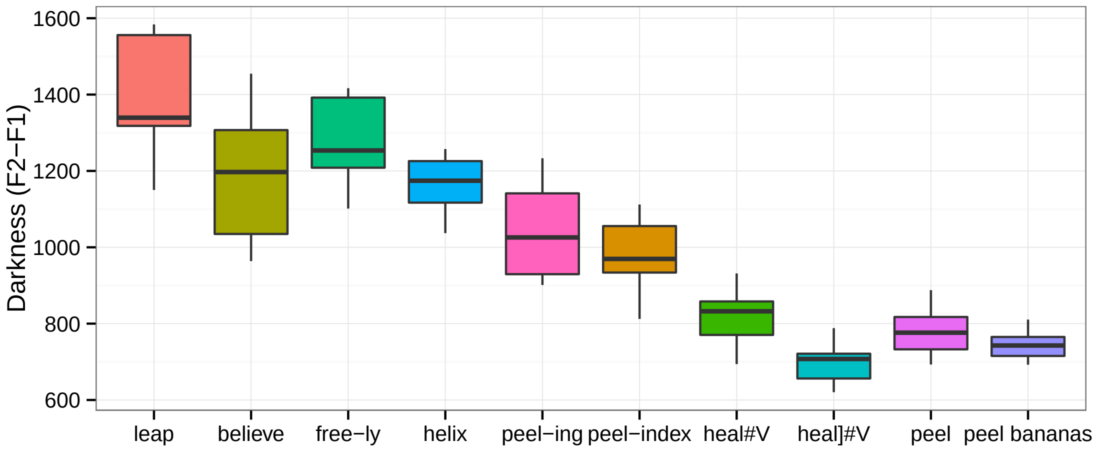

Figure 7.44: Newcastle F2-F1 values across phonological context

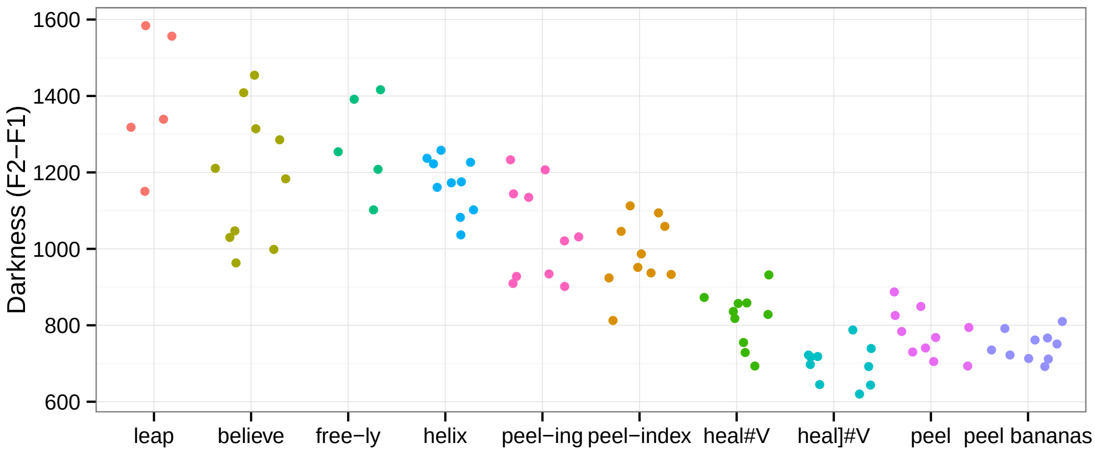

Figure 7.45: Newcastle F2-F1 values across phonological context

7.7.4 Darkness vs. duration

We have seen that Newcastle /l/s are all very light, but still have a bimodal range. This section will investigate whether duration can account for any of the patterns found in the data by investigating its correlation with PC1.

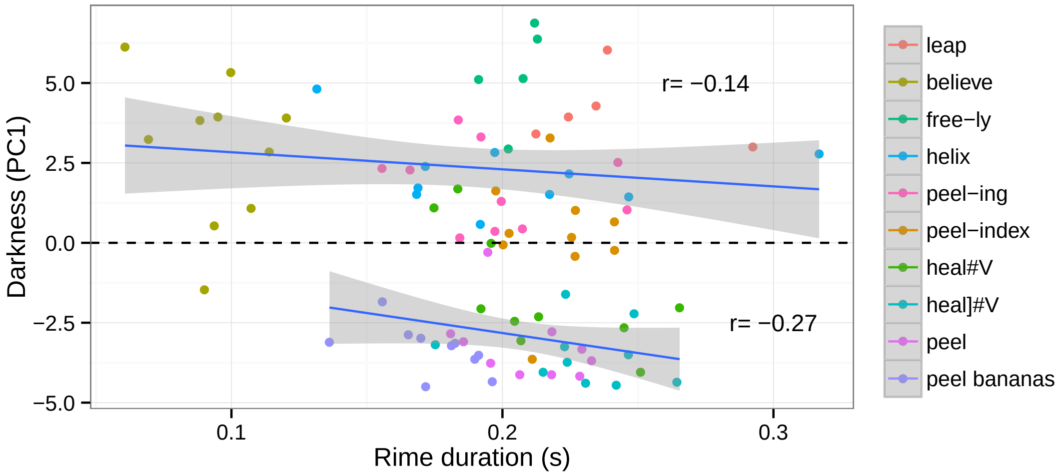

Figure 7.46: Newcastle PC1 values against rime duration (new categories shown in correlation).

Both correlations for light and dark tokens are extremely weak and do not prove significant (light r = -0.141, p= 0.328; dark r =-0.266, p=0.097). The durational effect is not too convincing when we eyeball the scatterplot in Figure 7.46. The *believe* tokens seem to be carrying the correlations, as they are very short and averagely light. Indeed, removing these from the correlation weakens it to the level of non-signficance, giving a Pearson's r = -0.03 (p=0.18). Nevertheless, the weakness of the Pearson's r in general has already been pointed out, and it is merely used as a rough measure in this section. The true relationship between duration and darkness is better left to the linear models in Section 7.7.5.

7.7.5 Models

Table 7.13 shows the adjusted r^2 values for four different linear models fit to darkness (i.e. PC1). The model categories reflect the patterns observed in articulation. Again, duration gives the worst fit on its own (1), with an extremely poor fit close to zero for Newcastle. Category (2) alone performs moderately well, and adding duration improves the model (3). An interaction provides no further improvement (4). This suggests, like we found for the RP speaker, duration is at play here, but does not apply differently to different categories. This is hardly surprising given the equal correlations. However, the overall fit here is a little more modest than we found for the other speakers with a categorical distinction.

The best model is shown in Table 7.14. This shows that the dark category results in a significantly lowered PC1 value in relation to the light intercept, and that a higher 'rime' duration also results in a lowered PC1 value. However, the r^2 value suggests that, although there is some bimodality in two categories for the Newcastle speaker, this

model	adj. r^2
1. Darkness ~ log(Rimeduration)	0.078
2. Darkness ~ Category	0.664
3. Darkness ~ Category + log(Rimeduration)	0.672
4. Darkness ~ Category * log(Rimeduration)	0.67

Table 7.13: Comparing duration correlations for Newcastle speaker alongside adjusted r-squared of models with a category duration interaction

	Estimate	Std. Error	t value	Pr(>|t|)
(intercept)	0.3778	1.1869	0.3183	0.7510
category: dark	-5.1041	0.4032	-12.6598	0.0000
Log(Rime duration)	-1.1481	0.6681	-1.7183	0.0893

Table 7.14: Newcastle model 3, Darkness by Category

model does not fit as well as some of the others. The effect size of category is much smaller than we have seen for other speakers.

7.7.6 Summary

Overall, this speaker presents a situation where bimodality is clear, but in a less distinct way than we have seen for RP and London. This speaker has a distinction in the non-overlapping `loess` confidence intervals fit to the tongue splines in Figure 7.38, which is supported by the PC1 box plots, Hartigan's dip test, and the F2-F1 acoustic results, all of which are confirmed further by the Tukey HSD. We find an effect of duration, but this is weak, and certainly nowhere near enough to argue for a purely gradient approach. In contrast, we find a lot of overlap on both the articulatory and acoustic spectra, calling the categorical interpretation of this speaker's data into question.

However, this could just be because this speaker is operating on a much more reduced phonetic spectrum. This speaker's range is much smaller than the RP or London speakers', and this means that the category effects just are not as big. This raises the question whether we should expect categoricity to be accompanied by a large phonetic difference. Conceptually, there is no obvious reason why the two should go hand in hand. In fact, the results from Manchester suggest that categorial structure and phonetic range can vary independently of each other. (The same point is demonstrated more clearly, of course, by situations of incomplete neutralisation and of Labovian near merger, where two lexical categories subsist but with a great deal of phonetic overlap). Take Figure 7.47 as an example. This shows the acoustics for the London Male and the Newcastle male speaker side by side (using Bark normalised values). This highlights a few things. Firstly, it is hard to argue against a categorical distinction for the London speaker given the range of /l/s studied, and the output of the F2-F1 values. Secondly, it may look plausible enough to suggest a gradient interpretation of the Newcastle speaker's system when viewing the results in comparison to the London Male's, but the acoustic range is a little more restricted for this speaker. It does not follow that this speaker does not have a categorical distinction just because the values are not as distinct.

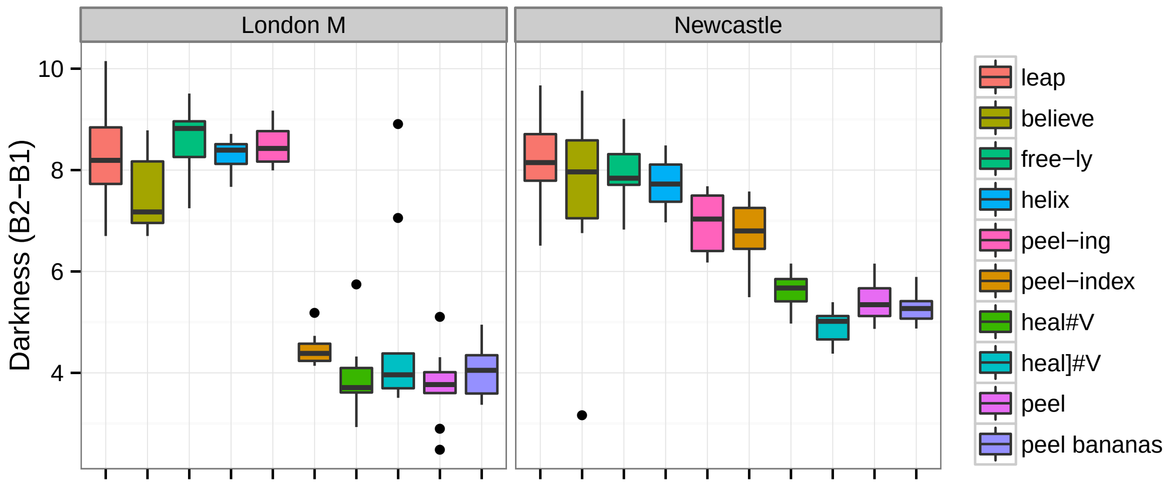

Figure 7.47: Newcastle B2-B1 values compared with London Male.

It is important to point out that this speaker may not be representative of the Newcastle population in general. He is an upwardly mobile young student, living away from home. Who is to say that a speaker from the heart of the Geordie vernacular *would* have a distinction? This speaker could show similarities with the Manchester MC speaker, and have a small distinction in initial and final tokens, whereas a working-class speaker may truly have non-distinguishable light [l]s in all environments as the dialectological literature suggests. Again, this calls for a thorough sociophonetic sample in future work. More generally, it may be that sociolinguistic studies of England have hitherto overlooked a large-scale phenomenon involving the realisation of /l/: the growing influence of the bicategorial South on different monocategorial northern vernaculars.

7.8 Belfast

Irish varieties of English have been described as 'strikingly light in all positions' (Wells 1982:431), although some of the sociolinguistic literature has suggested this might be changing for speakers in Northern Ireland (McCafferty 1999). As we have seen for Manchester and Newcastle, sweeping generalisations such as these are often not borne out fully in the phonetics. This section investigates the accent of Belfast, in an attempt to see if all /l/s are indeed light or, like Newcastle there is evidence for some kind of bimodality.

7.8.1 Ultrasound splines

The plot in Figure 7.48 shows the most consistently light set of splines thus far. There is next to no variation in these tongue contours, making the Manchester WC plot look variable by comparison. All tokens have an advanced tongue tip and high tongue dorsum, with little retraction in the tongue root area.

It can be noted that one of the mean splines has a slightly lowered tongue body and is slightly backer than the others. This is the token *believe*, where the /l/ occurs in foot-initial position. For all of the other speakers in the experiment, this token stands as the lightest, or one of the lightest, of all contexts, so it is unusual that the speaker with the consistent light [l] pattern would have a slightly different realisation in this prosodically strong position. However, on listening to the tokens and and examining the spectrogram and acoustics, it is clear that this token is subject to pre-stress contraction (Zwicky 1972:283). Often, particularly at fast speech rates, the /l/ becomes part of a branching onset so rather than a realisation such as [bəˈliːv], we find something more like [ˈbliːv], with no audible vowel and a darker realisation of the /l/ (Huffman 1997:118). These /l/s are the shortest of all, in general, with a mean'rime' duration of 0.1s. This is not necessarily surprising, as we may expect consonants in clusters to shorten to achieve a 'common average' alongside other sole consonants (Lehiste 1980:18). This is problematic for comparison, as this token is no longer flanked by two vowels. Thus, this context is removed from the rest of the analysis of the Belfast speaker.

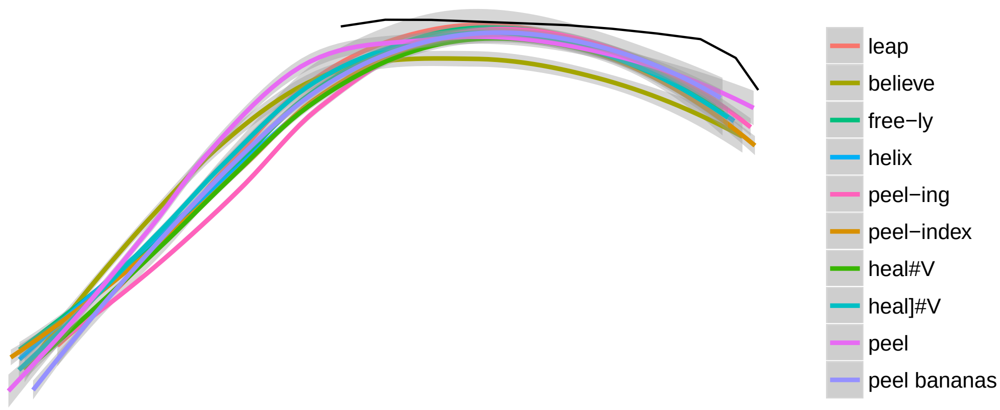

Figure 7.48: Belfast splines at /l/ midpoint across phonological context

Removing the ***bel**ieve*-type tokens from the distribution, it is easier to visualise the slight but significant tongue root retraction of the *pee**l***-type tokens in Figure 7.49. This is unsurprising, considering we have seen this extra retraction in phrase-final tokens for almost all speakers. There is no evidence of a more distinct light/dark pattern as reported for Derry (McCafferty 1999). However, phrase-final *peel*-like tokens show significantly more retraction (note that word-final pre-consonantal *peel bananas* tokens do not). This is a very similar distribution to that found for the Manchester speaker, although at the light end of the spectrum, rather than the dark one.

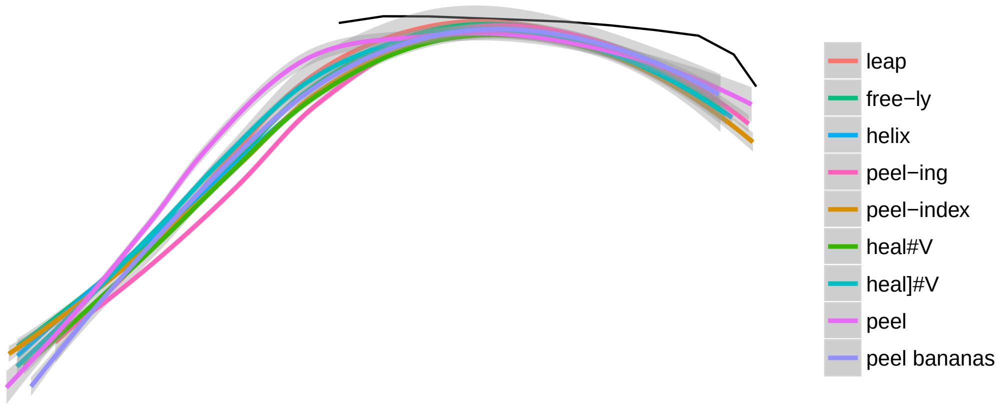

Figure 7.49: Belfast splines at /l/ midpoint across phonological context removing believe tokens

7.8.2 PCA

7.8.2.1 Variable importance and loadings

The PCA seems almost purposeless for this speaker, due to the tiny amount of overall variation in the input. However, it will allow us to visualise tiny tongue movements on a more interpretable scale. The lack of variation for this speaker is demonstrated in Figure 7.50, which is the first variable importance plot to show as low a PC1 (at 53%) and others PCs which stand a chance of inclusion (PC2 is at 23%, but PC3 only 10%). However, the loadings plot in on the right panel shows that these are doing very little (PC3 is not plotted as it would only clutter the crowded range). Although PC1 is also doing very little here in the way of magnitude, we can still see a limited but interpretable continuum of tongue retraction, and this component does allow us to quantify this in the same manner as the other speakers in this chapter.

7.8.2.2 Results

The Tukey HSD for the PC1 values show very little significant differences between environments. The exception is the *pee**l***-type tokens, which are significantly darker than all other environments, apart from the ***l**eap*-type tokens, interestingly enough. Other significant differences are found in the *pee**l**-ing*-type tokens when compared with *free**l**y*, *hea**l**#]V* and *pee**l** bananas*-type tokens. It seems like this speaker shows his darkest variants when there is a following pause, even if this pause is only slight. The lightest tokens appear intervocalically. Figure 7.52 demonstrates the lack of bimodality in this dataset, confirmed by the dip test; $D = 0.035$ ($p = 0.66$)

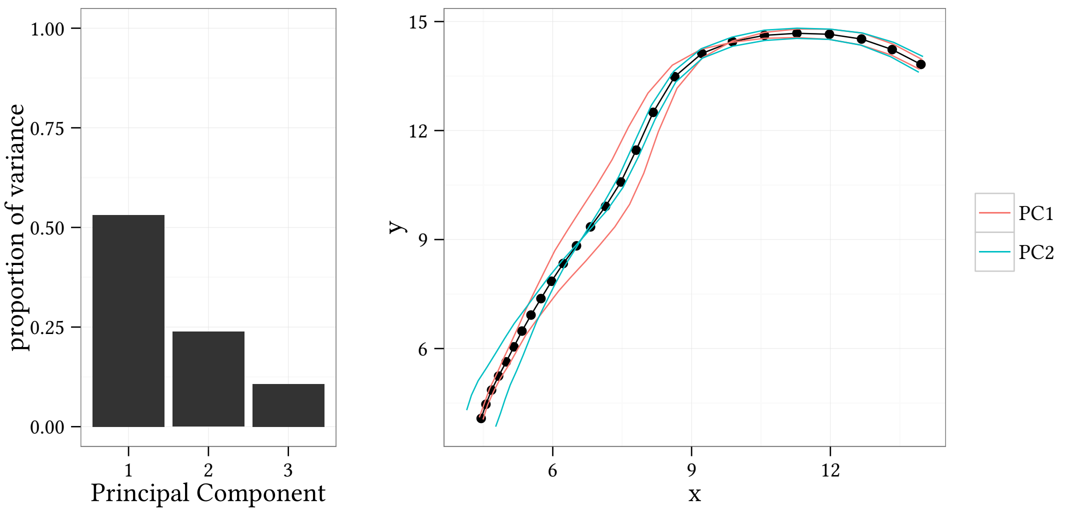

Figure 7.50: PCA summary. Variable importance and loadings plots from the Belfast PCA output

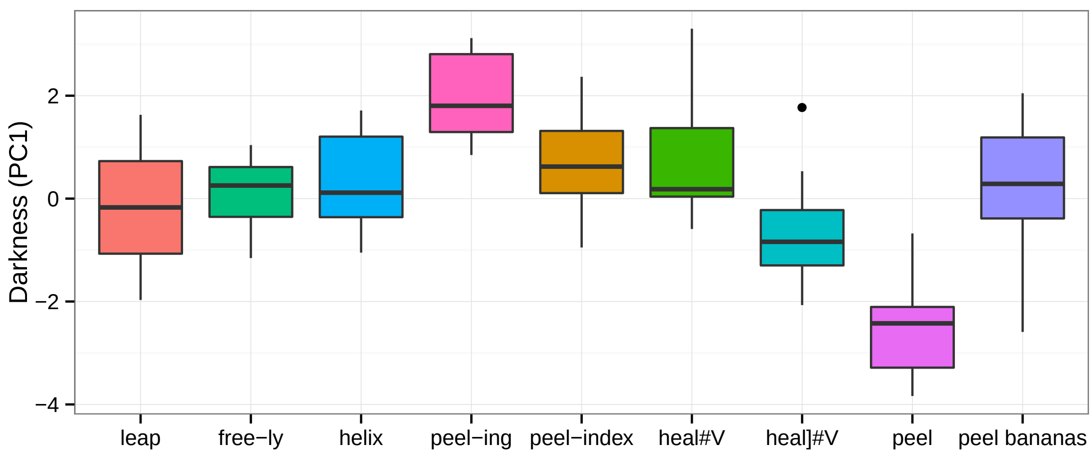

Figure 7.51: Belfast PC1 values across phonological context

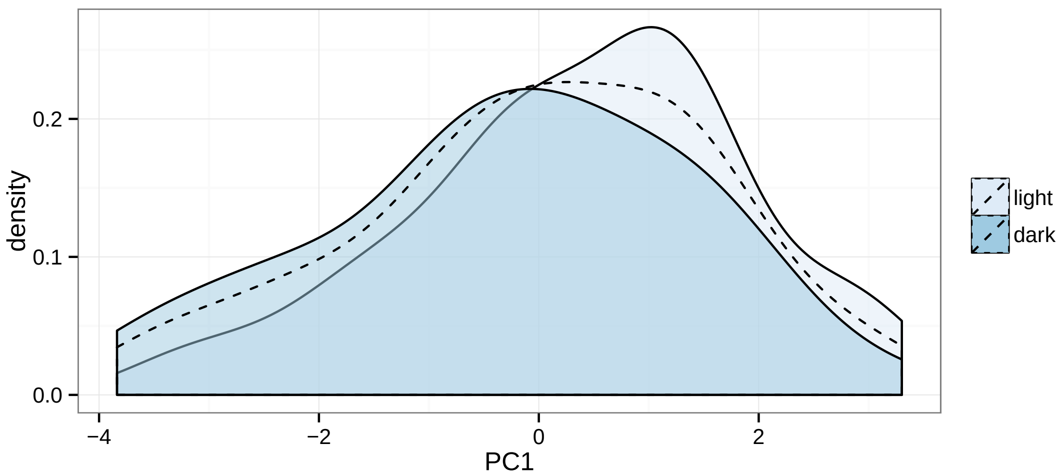

Figure 7.52: Density plot for Belfast showing separate distributions for light and dark categories (categories determined by phonological environment) and overall distribution indicated by dashed line.

7.8.3 Acoustics

The Belfast acoustics are impressively tight, and cluster together. The box plots in Figure 7.53 do give the impression of having some kind of range, but when this is compared against the RP speaker's acoustic range in Figure 7.54 (using Bark normalised values), it shows just how light the Belfast /l/s really are. The dip statistic is unsurprisingly non-significant (D = 0.027, p= 0.965). The Tukey HSD results for the acoustics show significant difference between the ***peel**-ing*-type tokens and all of the typically dark environments, as well as the ***l**eap* environment. These are the only significant differences, however and there is no indication of two categories. The standard five vs. five will be used in the model later in the section.

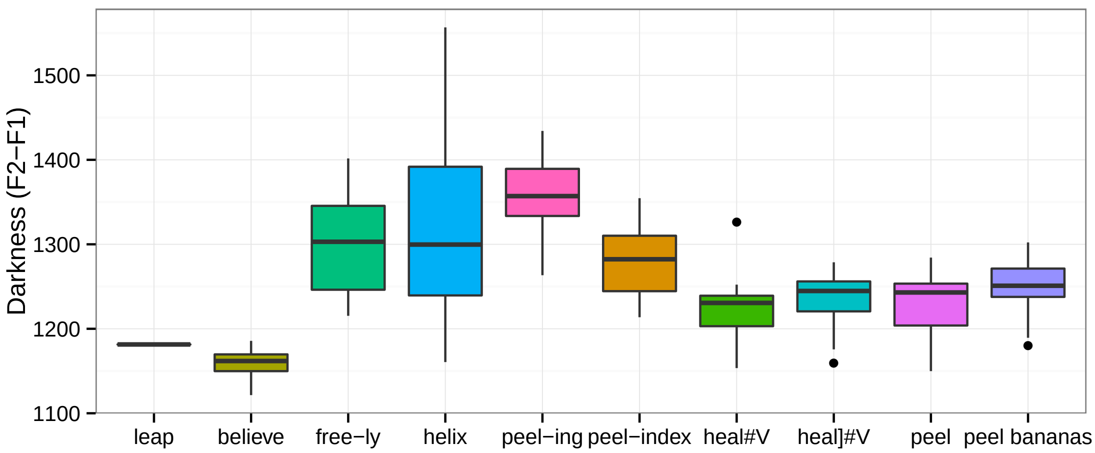

Figure 7.53: Belfast F2-F1 values across phonological context

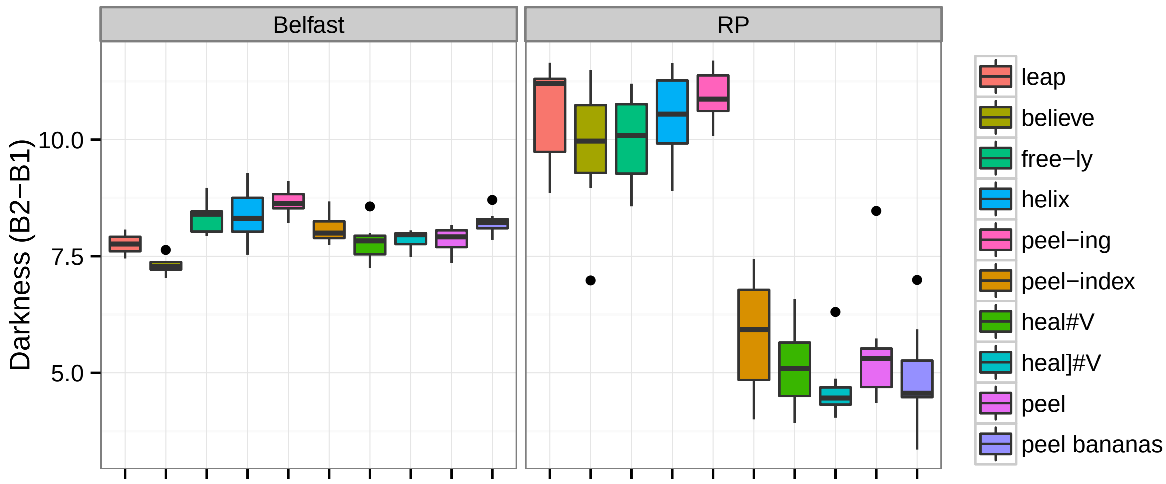

Figure 7.54: Belfast B2-B1 values across phonological context compared with RP

7.8.4 Darkness vs. duration

The duration plot for Belfast (Figure 7.55) shows that, although the phrase-final *peel*-type tokens are indeed darker, they are actually just as short as the initial tokens. There is no indication of a better category distinction, so the *a priori* distinction between word-level onset and coda /l/ will be used for the correlations.

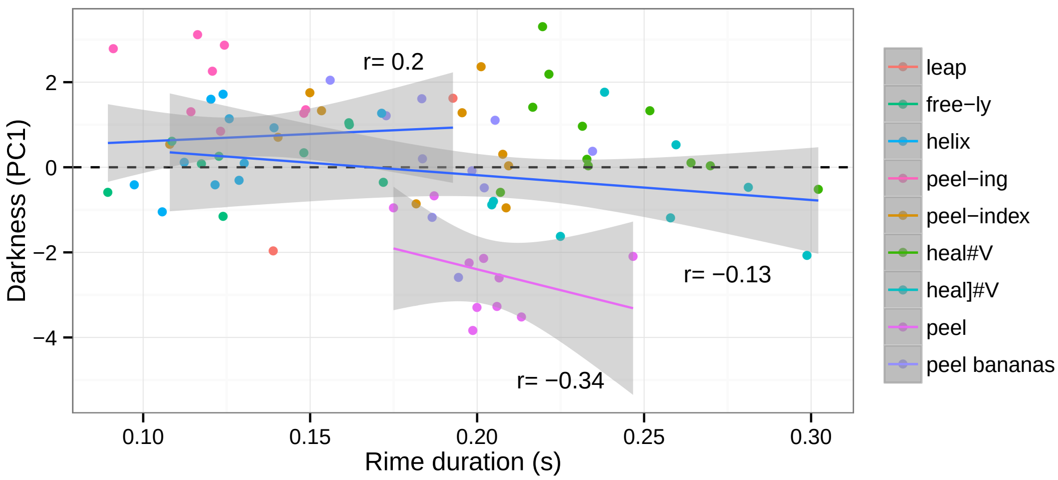

Figure 7.55: Belfast PC1 values against rime duration

For light [l] we find a position correlation, which is the inverse of the others, $r =$ 0.201, in that the lighter the /l/ the longer it is. This is not statistically significant, however, p= 0.254. The correlation for the dark [l]s is almost non-existent, $r =$-0.134, p=0.358). Even though the ***peel***-type tokens show a slightly stronger correlation ($r =$-0.345), this is not significant (p=0.329). The Manchester WC speaker showed a similar pattern, with all tokens very similar and extra tongue root retraction for the ***peel***-type tokens, but the correlation between duration for these was very strong. That is not the

model	adj. r^2
1. Darkness ∼ log(Rimeduration)	0.011
2. Darkness ∼ Category	0.022
3. Darkness ∼ Category + log(Rimeduration)	0.01
4. Darkness ∼ Category * log(Rimeduration)	0.027

Table 7.15: Adjusted r-squared of models on PC1 for Belfast speaker.

model	adj. r^2
1. Darkness ∼ log(Rimeduration)	0.011
2. Darkness ∼ Category	0.112
3. Darkness ∼ Category + log(Rimeduration)	0.157
4. Darkness ∼ Category * log(Rimeduration)	0.368

Table 7.16: Adjusted r-squared of models on formant values for Belfast speaker.

case here. That could be an inherent prediction of the phonetic realisation of the /l/: darker /l/s have a stronger dorsal gesture, which is more cumbersome and takes longer. The light /l/s of Belfast are unlikely to show a significant phrase-final effect of duration, even if they are the darkest, as they are still very light within a crossdialectal typology..

7.8.5 Models

As indicated above, as there is no indication of a better category distinction that the *a priori* distinction between word-level onset and coda /l/, this will be used for the models, if only to highlight how bad a fit they provide. [8] Table 7.15 shows the adjusted r^2 values for four models fit to the Belfast speaker's PC1 values. All models provide a very poor fit.

Although, it is fair to point out that PC1 is a particularly rough measure of darkness for this speaker, as the PCA had very little variation to work with in the first place. This highlights the need for a much more multi-dimensional dataset when conducting PCA. The measure is crude, but fine for datasets like RP, but is not particularly insightful when there is next to no variation. Given the potentially bad measurement provided by PC1, it may be prudent to run a model on just the formant values for the Belfast speaker. These results are shown in Table 7.16. An interaction between duration and category does nowhere near as well as other speakers, but we can see that it is better than any other models. The quantitative results for Belfast suggest that category plays no role in this speaker's /l/ grammar.

7.8.6 Summary

Overall, there is no evidence for a categorical distinction in the Belfast speaker's /l/ system. Duration cannot account for darkness, and in fact, there is next to no variation

[8] Models were also fit using a separate category of all light and dark for phrase-final *pee**l***-type tokens, but the model fit was so poor (r^2 < 0.01) that there is little point reporting them here.

at all. This speaker's results show just how convincing the bimodal distribution of the Newcastle splines were after all. Moreover, one might be tempted to revisit the analysis of Manchester WC, and specially of Manchester MC, in the light of Belfast's impressively monocategorial pattern. It is not clear that this would be warranted, however, for, as we have seen, there a good a priori reasons for expecting greater variance in an 'all dark' than in an 'all light' system. One suggestion for the Belfast system might arise from the clues coming from the fact that this speaker's lightest /l/s are intervocalic, and significantly lighter than the ***leap***-type tokens in F2-F1 values. This would suggest that all /l/s in Belfast are the same regardless of phonological environment, but there is some coarticulation due to adjacent vowels. Perhaps an /l/ preceding a back vowel would be much darker. Unfortunately, different vowel environments were not included for most speakers, as it was not identified as being in the scope of the dialect investigations.

In summary then, in seems that this speaker shows us what a system with truly no categorical distinction between initial and final /l/ looks like. There is no gradient effect of duration, cognitively controlled or epiphenomenal, at play here. The only variation found in this speaker's /l/ system is from coarticulation. Belfast /l/s will happily move in the wind of the coarticulation from the neighbouring vowels. If their neighbours have a larger F2-F1 value and are fronted, then so shall they be. If their neighbours are backed, then the /l/ will look darker, but this is not under cognitive control in the way that a positional conditioning rule would be. If a token is phrase-final, it gains some additional darkness. Overall, this speaker shows Stage 0 of the life cycle: no /l/-darkening at all.

7.9 Liverpool

The articulatory data for Liverpool are of particular interest, given the conflicting predictions that can, and have, been made about its realisation. Surrounded by the dark laterals of Manchester and Lancashire in geographical terms, yet heavily influenced by Irish English linguistically (with their 'strikingly light /l/s'), the realisation is difficult to predict. This goes for the phonological patterning, i.e. will it show a lack of an obvious phonological distinction as Lancashire and Irish English do, albeit to opposite extremes, as well as the phonetic realisation, i.e. whether it be towards the lighter or darker end of the scale. Moreover, the dialectological reports are conflicting, with some saying it is very light due to the influence of Irish English (Jones 1966:92), and others saying it is velarised in all positions (Knowles 1973:256).

7.9.1 Ultrasound splines

The spline plot in Figure 7.56 shows the realisation across ten environments for the Liverpool speaker in this experiment. The splines indicate a bimodal distribution. This is interesting, given the two accents that could influence Liverpool do not have anything close to this kind of distinction between the ten phonological contexts. However, the

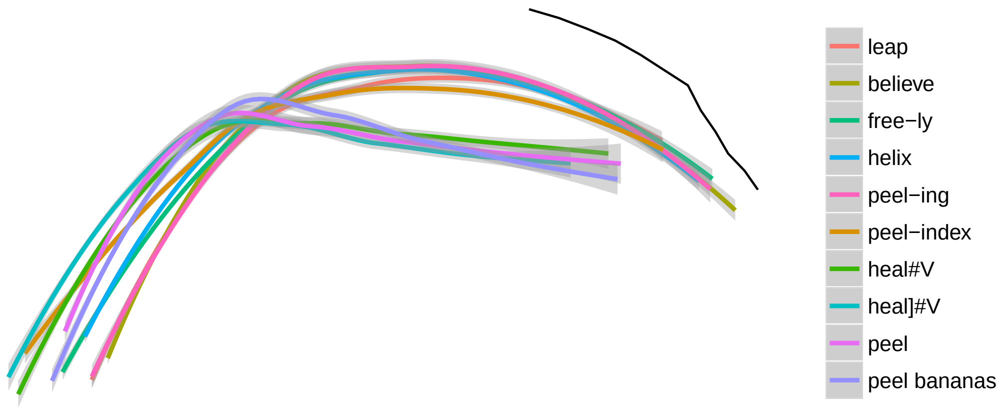

Figure 7.56: Liverpool splines at /l/ midpoint across phonological context

most exciting observations to be made from this plot is the obvious velarisation of the darker tokens. This is in contrast to all of the other speakers, who show pharyngealisation of their dark [ɫ]s, with backing in the tongue root area, but no raising toward the velum. An exception to this may be a minority of the tokens uttered by the Mancunian speaker. The Liverpool accent is described by dialectologists as velarising throughout speech (Hughes et al. 2012:114; Knowles 1973), so it may be interesting for future articulatory work to see if this is true of other sounds in the accent, as it is for the dark [ɫ]s. In Figure 7.57, the ***peel***-type tokens are compared to a [w] token uttered by the speaker, in order to demonstrate what dorsal raising and retraction towards a known velar target looks like. The tongue blade area is higher and fronter in this token, taken from the word *swarm*, most likely due to the coarticulation from the preceding /s/.

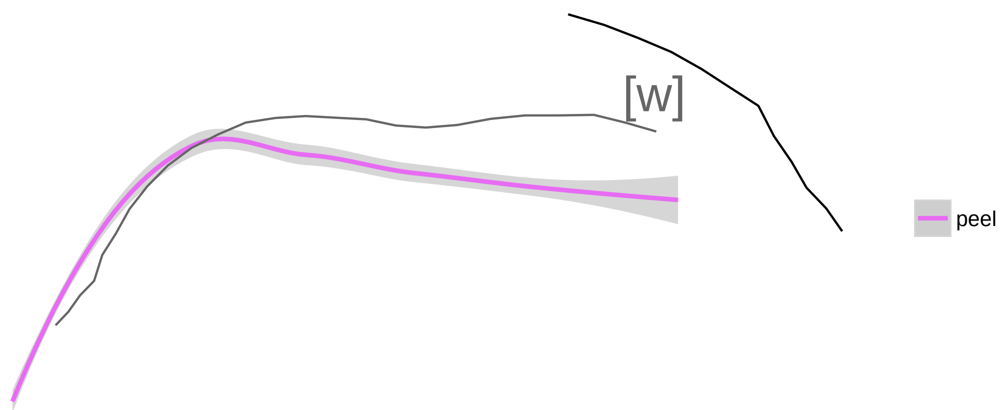

Figure 7.57: Liverpool speaker producing a /w/ alongside word-final splines at /l/ midpoint.

For the most part, this speaker patterns with the majority of other speakers in this section, in that environments with typical onset /l/ show a slightly more advanced tongue root, higher tongue body and tongue tip, whereas the typical word-final coda /l/s are more retracted, and in this particular case, velarised. The exception to this is the *pee**l**-index* tokens, which pattern with the lighter ones instead of the darker ones, showing the same pattern displayed by the Newcastle speaker. These tokens are a little more intermediately positioned in the tongue root area, however, as Figure 7.58 shows. There is also one token which looks velarised, showing that this position may be subject to variation. Prosodic factors varying from repetition to repetition may account for why the speaker resyllabifies the /l/ in some repetitions but not in others.

Figure 7.58: Liverpool splines at /l/ midpoint in *peel-index* and *heal*V tokens, with initial and final added for reference.

7.9.2 PCA

7.9.2.1 Variable importance and loadings

The Liverpool speaker displays the familiar pattern displayed by most speakers: PC1 accounts for 89% of the variation and PC2 just 6%, as shown in Figure 7.59 and the analysis will focus on the first component, with higher values reflecting lighter /l/s. The velarised dark [ɫ]s are reflected in the loadings plot in the right panel of Figure 7.59, which demonstrates the large effect that the velum raising has on the front of the tongue.

7.9.2.2 Results

Figure 7.60 shows the box plots of the PC1 values, highlighting the finding in the splines that the compound internal *peel-instrument*-type tokens pattern with syllable initial tokens. Unlike Newcastle, there is no overlap here with the darker tokens. However, the value is more intermediate between the two, despite having a positive PC1 value. This is discussed further in Section 7.10.2.1. Overall the PC1 plot provides a convincing picture of a categorical distinction. Subjecting the PC1 values in Figure 7.60 to the dip test gives a value of $D = 0.08$. This shows a significant chance of the distribution

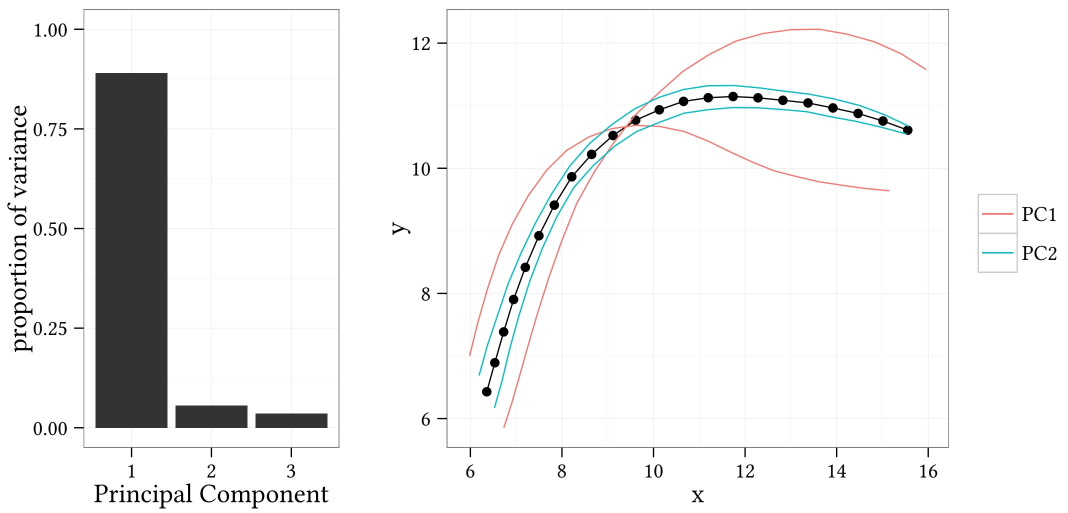

Figure 7.59: PCA summary. Variable importance and loadings plots from the Liverpool PCA output

being bimodal (p < 0.001). Figure 7.61 demonstrates the bimodality dip for this speaker, with the dashed line representing the overall distribution.

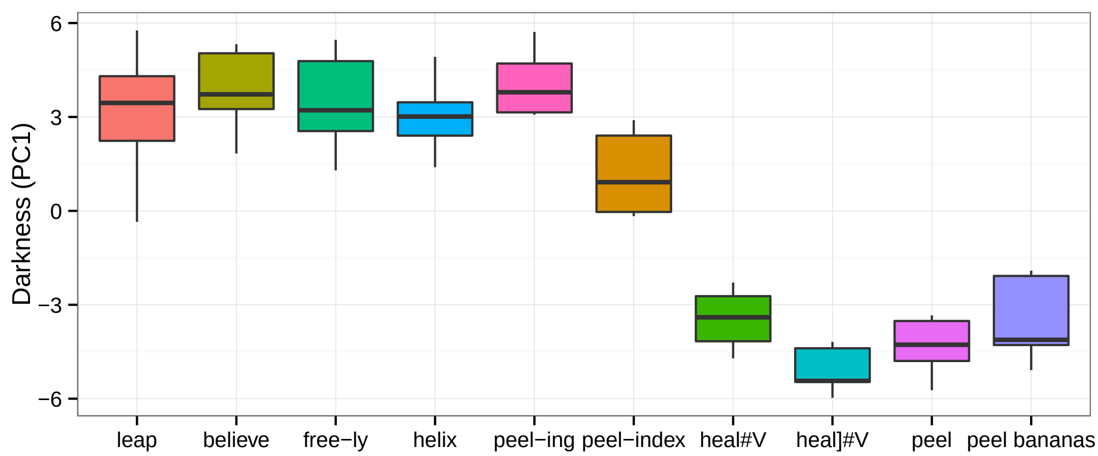

Figure 7.60: Liverpool PC1 values across phonological context

Also showing no support for a unimodal distribution, the Tukey HSD test suggests a three-way division for the Liverpool speaker's /l/. Not only are the ***peel**-index*-type tokens significantly different to the light tokens, they are also different to the dark ones. This would suggest a third intermediate category of darkness, but we have argued that this is only through variable prosodification in compound environments.

7.9.3 Acoustics

Figure 7.62 shows the F2-F1 figures for the Liverpool speaker according to context. The clear articulatory split can be seen in the acoustics, although it does accentuate the

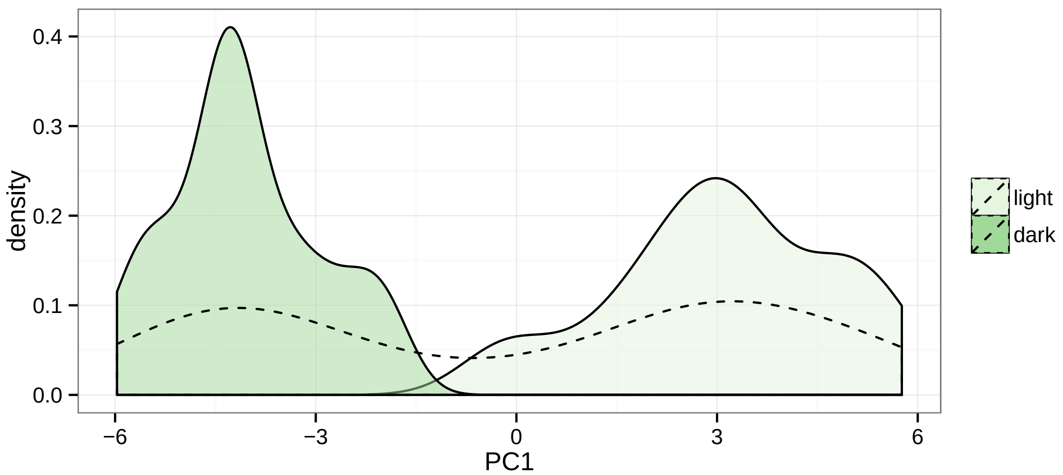

Figure 7.61: Density plot for Liverpool showing separate distributions for light and dark categories but with new category type 6-4 (categories determined by phonological environment)

intermediate position of the *pee**l**-index* tokens. However, the Tukey HSD does not pick up on the intermediate realisation of this context as clearly as the articulatory data does (apart from the *pee**l**-index* and *hea**l**#V* difference). In fact, the acoustics for the most part group these tokens with the darker ones, which can be seen in the box plot in Figure 7.62. The splines above in Figure 7.58 seem to fit with the initial tokens in the tongue tip area but the root area is as retracted as final tokens. This makes the situation even more difficult to diagnose and this is reflected in Hartigan's dip test on the acoustics, which fails to generate a significant chance of bimodality (D = 0.046, p= 0.126). This must be due to the intermediate values bridging the gap between the two distributions.

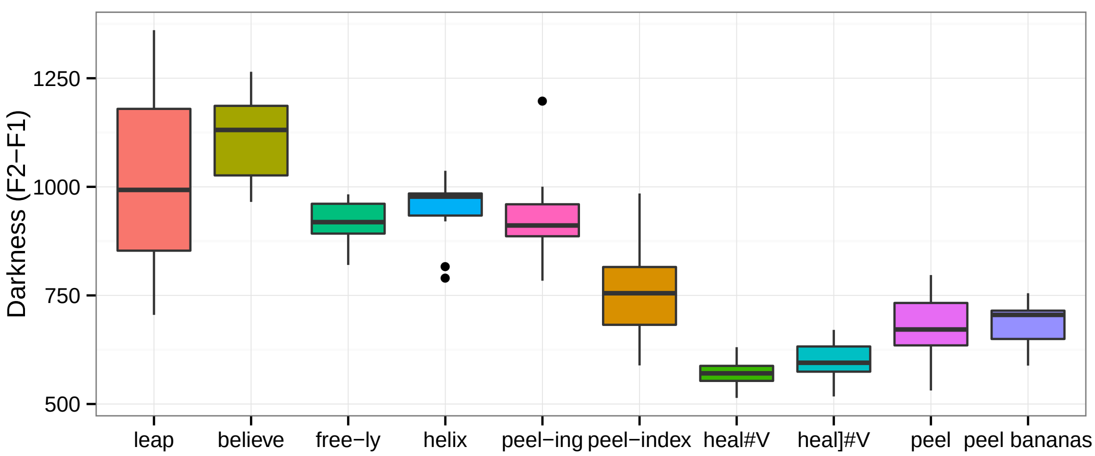

Figure 7.62: Liverpool F2-F1 values across phonological context

7.9.4 Darkness vs. duration

Thus far, we have seen a convincing split between phonological contexts for the Liverpool speaker for the most part, with the exception of the ***peel**-index* tokens, which possibly should be considered separately. This will be addressed in the next section, but for this section these tokens will be included with the other light ones, as their articulation looks to resemble these tokens more, and there is no velarisation (with the exception of one token). Figure 7.63 shows the PC1 values plotted against 'rime' duration for the Liverpool speaker. The correlations and lines are fit to the first six and last four categories for this speaker.

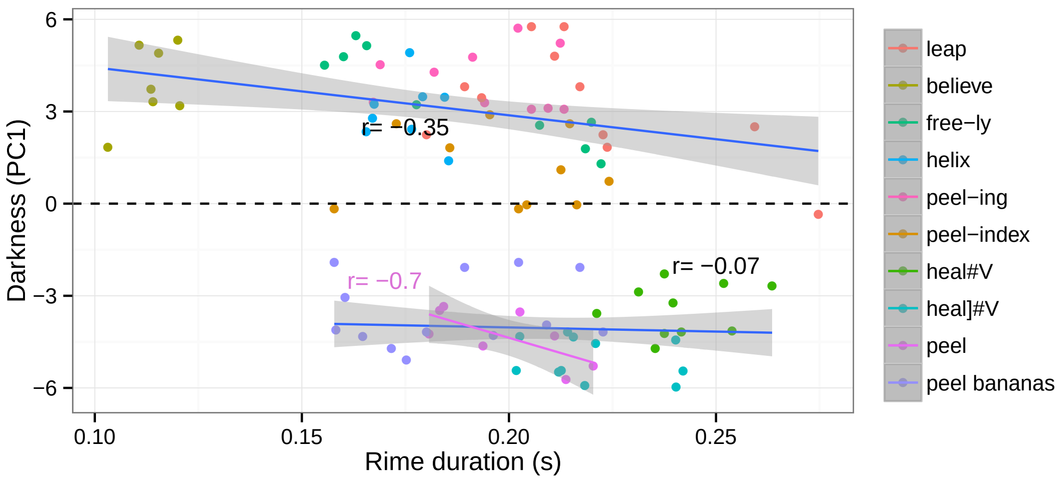

Figure 7.63: Liverpool PC1 values against rime duration

For light [l] we find r = -0.346, which is statistically significant at p= 0.01. However, the spread of the points looks a little problematic here. There is one unusually long and slightly dark ***peel***-type token dragging the correlation, as well as the cluster of very short and very light ***believe***-type tokens. The correlation for the dark [l]s is almost non-existent, r =-0.067, p=0.673). It seems that duration does not play much of a part in darkness for this speaker, but possibly for the light [l]s, however, we do see a significant correlation once again for the ***peel***-type tokens (r =-0.7, p=0.05) in phrase final position.

7.9.5 Models

Table 7.17 shows the results of the linear models fit to the Liverpool darkening data in terms of adjusted r^2. Models 2-4 use the two-way category distinction, models 5-7 use the three-way distinction (Category 2). Duration cannot account for the variation in the model, giving a very low value. In the two-way distinction, category does well alone (2) but adding duration (3) creates a significant improvement (p = 0.04; full ANOVA comparisons in Appendix D). An interaction does not, which suggests that if duration is effecting anything in this speaker's system, then it applies across the board, rather than differently to light and dark variants. We find the same thing for the three-way

model	adj. r^2
1. Darkness ~ log(Duration)	0.161
2. Darkness ~ Category	0.738
3. Darkness ~ Category + log(Rimeduration)	0.746
4. Darkness ~ Category * log(Rimeduration)	0.744
5. Darkness ~ Category 2	0.892
6. Darkness ~ Category 2 + log(Rimeduration)	0.896
7. Darkness ~Category2 * log(Rimeduration)	0.894

Table 7.17: Adjusted r-squared of darkness modelling for Liverpool

distinction, that duration improves, but interactions do not. The three-way distinction provides a better fit in terms of adjusted r^2, but this to be expected with finer grained factor levels, as it allows the model more freedom in fitting different contexts. Whether we can take this as an indication of three categories for this speaker is not clear from the statistics alone.

	Estimate	Std. Error	t value	Pr(>|t|)
(intercept)	-0.3440	1.3258	-0.2594	0.7959
category: dark	-6.8804	0.3052	-22.5439	0.0000
log(Rime duration)	-2.0189	0.7696	-2.6233	0.0102

Table 7.18: Liverpool model 3, Darkness by Category 1 + log(Rime duration)

	Estimate	Std. Error	t value	Pr(>|t|)
(intercept)	1.0393	1.2075	0.8607	0.3916
category2: intermediate	-2.2532	0.4396	-5.1259	0.0000
category2: dark	-7.3682	0.2872	-25.6582	0.0000
log(Rime duration)	-1.4480	0.6922	-2.0918	0.0392

Table 7.19: Liverpool model 6, Darkness by Category 2 + log(Rime duration)

7.9.6 Summary

Overall, this speaker provides some interesting results from several perspectives. Firstly, we gain a phonetic insight into the /l/ system of Liverpool, showing that speakers do seem to have a categorical light/dark distinction in this dialect. Evidence for this comes from the spline plots, the Tukey HSD tests for both the PC1 values and the acoustic values, as well as a significant Hartigan's dip test value indicating a bimodal distribution. All of these findings are supported by linear models, which show that category provides a good fit to the data, with duration adding a small but significant effect in the expected direction.

An additional exciting find from this speaker is the evidence of velarised /l/s. Not only does this corroborate Knowles's (1973)'s claims about Liverpool having velarised /l/s, it also shows how comparatively rare this darkening strategy is in this dataset, as most speakers just show tongue root retraction and not any kind of velarisation. In addition, it drives home the point that velarisation is a form of darkening, and not a synonym for it. Whether there are acoustic correlates indicating differences between a velarised and pharyngealised /l/ is not clear. It has been said that such differences do not show up on the spectrogram (Ladefoged and Johnson 2014:245) and there are no clear differences in the acoustics here, but perhaps a better picture may come from a speaker who is more variable than this one. It raises questions for future articulatory work on Liverpool English, to see if speakers really do velarise everywhere, as Knowles suggests.

There is also the issue of the intermediate realisation of the *pee**l**-index*-type tokens for this speaker. These tokens pattern with the light splines in the tongue tip area, but closer to the dark splines in the tongue root area. The clue to dark splines in this dialect is velarisation, and two splines do show this tendency, indicating that the intermediate mean spline is partially down to variance. However, they are mostly the same shape as the light /l/s, just a little backer. It could be the case that there are several processes in this speaker's system, targeting different gestures in different levels of the phonology. The first process affects the backing of the tongue root and applies to coda /l/s at the word-level, affecting *pee**l**-index* /l/s onwards. The second process effects the raising of the velum and targets /l/ in the coda at the phrase level. Because of the prosody of the sentences chosen in Experiment 2, the final four categories cannot resyllabify, ensuring that they remain dark.

7.10 General observations

In this section, the dataset will be analysed as whole, considering all speakers alongside one another. This includes an overview of all the evidence for categorical processes for all speakers (Section 7.10.1) and an overview of gradience and variance across and within speakers (Section 7.10.2). At the end of this section (Section 7.10.4), a cross-speaker acoustic comparison is conducted, which includes measuring the effect of duration, and perfoming mixed-effects linear regression on all of the data points collected for Experiment 2.

7.10.1 Evidence for categoricity in Experiment 2

Sections 7.2 to 7.9 above provide a summary of the findings for each speaker on an individual basis. This section will attempt to summarise the general patterns and differences overall, with respect to dialect variation, phonetic differences in articulation and acoustics, and most importantly the evidence for categoricity and gradience in the speakers' systems. Table 7.20 is a summary of all the quantitative and statistical approaches conducted in the previous sections. These are:

- **Splines:** whether the splines show two separate distributions of light and dark variants, and whether the `loess` smoothers confirm they are significantly different by having non-overlapping confidence intervals.

- **PC1:** the appearance of the PC1 box plots, whether they form two separate distributions and if this plays out in the Tukey HSD results.
- **Dip test:** does the dip statistic get a statistically significant p-value (indicating that the distribution is not unimodal)?
- **Acoustics:** the appearance of the F2-F1 box plots, whether they form two separate distributions and if this plays out in the Tukey HSD results of the acoustics.
- **Duration correlations:** what was found in the duration correlations and does this support a fully gradient interpretation? Does duration correlate with just light/dark tokens, or all tokens?
- **Linear models:** which predictors result in the best overall fit of the data?

Generalising across the dataset with the aid of the summary in Table 7.20 results, we can observe several different possible systems for our speakers:

a. Speakers with a clear categorical allophonic distinction between light and dark variants. These include:
   i. Those which show no effect of duration (London M)
   ii. Those which show durational effects for the darker category only (London F)
   iii. Those which show general durational effects across the board (RP, Liverpool)

b. Speakers which show some tendencies of categoricity, but where the picture is not as distinct as the varieties mentioned above and there is a lot of phonetic overlap (Manchester MC, Newcastle)

c. Speakers which show little evidence for a typical categorical distinction, including:
   i. Those who show significant phrase-final darker variants, but not a typical light/dark allophone between onset and coda /l/ (Manchester WC).
   ii. Those who show next to no variation, and no categorical distribution (Belfast).

In this section, I will outline how all of these systems can be accounted for under the life cycle of phonological processes.

For the varieties in a), speakers with a clear categorical allophonic distinction, we have several possible situations. The RP speaker has a conservative system in which /l/ darkens in the coda at the phrase-level,[9] with a moderate durational effect across the

[9] We know this speaker does not have typical word-level darkening from Experiment 1, where he showed light [l]s in phrases such as ***heal** it*. Aside from Manchester WC, we do not have this information for any others speakers in this section so it is difficult to conclude whether the process has moved up to the word-level, based on the *peel-index* tokens alone. Because of this, we shall just refer to speakers with the first five phonological contexts behaved separately to the second five as having phrase-level darkening.

board. This shows that duration is at play, and that darker /l/s do have longer 'rimes', but there is no clear evidence that this durational effect has undergone significant incrementation under grammatical control.

The London Female, on the other hand, does show a durational effect conditional on category, in that only the darker (or vocalised) /l/s are affected by duration in her data. This suggest that the durational effects are under grammatical control and show a strong effect with the /l/s that take the longest to articulate. Regardless of correlation, the plots in the previous eight sections show that the dark tokens are the longer ones. Thus, though the pattern shown by the London Female has become phonologised, it remains a natural one. The London Male is not as advanced as the London Female, as he has the same distribution, but without the extra effect of duration in the dark tokens. Now, of course, the Londoners are both more advanced than the RP speaker anyway in terms of lenition trajectories, as they have vocalised /l/s. The vocalised /l/s could have come about through one of two possible scenarios:

i. /l/ was categorically dark at the phrase-level in London, as it is in RP and many other dialects. This had overlying gradient effects leniting the /l/, resulting in gradient vocalisation. Over time, the gradient effects became stable in that /l/ lost its tongue-tip gesture and became vocalised, the next stage in the lenition trajectory. As darkening had not climbed any further levels, only phrase-level /l/s were targetted by the next stage.

ii. /l/ darkening existed in a gradient, non-stable form as some kind of delayed tongue tip gesture in the London grammar. The next generation, or under the next stage of the life cycle, this was treated as being vocalised and /l/s went from light to vocalised in one step.

Although it is possible for the diachronic pre-cursor of darkening to co-exist on top of the stabilised phonological process, as the London Female shows, it could also just supplant it, as the London Male shows. Both are possibilities under the life cycle.

We might think that ii) is more likely for London as there does not seem to be much magnitudinal difference between light and vocalised /l/. It would also make sense for darkening to escalate to vocalisation very quickly, as the process does not have chance to rise through the hierarchy. In Essex, tokens of /l/ in the coda in deeply embedded cyclic domains, e.g. *hea**l**-ing*, are dark and consonantal, whereas utterance final /l/ is vocalised. It seems likely, therefore, that vocalisation arrived in this variety considerably later than darkening: as a result, darkening had time to climb to the stem level, whilst vocalisation remains at the phrase level, where it targets only the tokens most exposed to lenition.

In b) we find the in-between speakers such as the Newcastle speaker and the Manchester MC speaker. These speakers behave very differently to one another, but what they have in common is that their evidence for categoricity looks weak in comparison to RP, but compelling in comparison to Belfast. We already know that duration is not the answer to this, and cannot account for the variation found. What these speakers also seem

Variety	Splines	PC1	Dip test	Acoustics	Duration correlations	Linear models
RP	Categorical	Categorical	Bimodal (categorical)	Categorical	Moderate durational effects across the board	Category + duration
London F	Categorical	Categorical	Bimodal (categorical)	Categorical	Strong durational effects for dark [ɫ]s	Categorical, plus durational effects for dark category only.
London M	Categorical	Categorical	Bimodal (categorical)	Categorical	Moderate durational effects for dark [ɫ]s	Only category matters
Manchester WC	Root backing in *pee**l*** tokens only	Gradient	Unimodal	Gradient	Correlation for *pee**l*** tokens only	Neither category or duration provides a good fit. Context alone only provides a moderate fit.
Manchester MC	Categorical tendencies with some overlap	Two or four way distinction	Not significantly bimodal	Categorical	Moderate durational effects across the board	Category provides a moderate fit with duration, four-way distinction improves fit without need for duration.
Newcastle	Categorical?	Categorical/ gradient	Categorical	Categorical/ gradient	No correlation	Category plus very small durational effect
Belfast	Very small root backing in *pee**l*** tokens	Slightly darker *pee**l*** tokens	Unimodal	All equal	No correlations	No combination provides a good fit
Liverpool	Categorical (velarisation in dark tokens)	Categorical, with *pee**l**-index* tokens intermediate	Bimodal (categorical)	Categorical, with *pee**l**-index* tokens intermediate	Correlations with light tokens and *pee**l*** tokens	Category with durational effect across the board provides best fit. *pee**l**-index* as separate category improves fit.

Table 7.20: Summary of quantitative tests and the evidence they provide for categoricity. Cells state whether the evidence from a particular tests points towards categoricity, gradience, or something in between, as well as additional effects, such as duration.

to have in common is that gradient overlap in the phonology and phonetics could be taken as no significant categorical difference. However, this is because the scale is much smaller for these speakers. Small phonetic differences do not mean a lack of a categorical distinction. One possibility is that these are systems in diachronic transition, with an incrementation of the phonetic range heralding the rise of a bicategorial system, possibly under the influence of southern varieties: speculative suggestions along these lines were made in Section 7.6.6 in connection with the particularly unclear pattern exhibited by Manchester MC. All in all, it may be best to suspend judgment about these dialects, pending more detailed investigations with better control of the sociophonetic factors.

Finally, we turn to the speakers in c). Again, these two show two very different patterns from one another but are united in that neither variety gets an acceptable goodness-of-fit in the linear models. Belfast, on the one hand, does not show any variation, so it is clear this variety does not have a distinction determined by phonological environment. Manchester WC, however, is problematic, as there is some difference here. The phrase-final tokens are significantly darker and do correlate with duration. We have seen that this is not just something that happens in all speech, as some speakers (e.g. RP) show no correlation of phrase-final tokens with duration. We have also seen that a model considering phrase-final *pee**l***-type tokens as a separate category does not perform well. However, it could be the case (as discussed in the Manchester WC summary) that the midpoint may not be a sufficient analysis for this speaker, and that a temporal analysis could be needed. That is, we cannot dismiss categoricity altogether for the Manchester WC speaker . There could be several things going on here:

i. phrase-final /l/ is showing the beginnings of a new darkening process, targeting /l/s in the coda at the phrase-level. This is a phonologised duration-driven effect active to /l/s in the phrase-final coda. It only affects phrase-final ones because these are so dark and delayed (because they've been darkened twice)

   or

ii. The Manchester WC speaker does have a categorical distinction, but this is only visible through gestural phasing, as the articulatory space utilised by this speaker is so small that the midpoint results in the same tongue shape.

What is interesting for these speakers is that we don't see the durational effects increase as the effects of categoricity decreases. We actually see the inverse here: the dialects which seem to show no categorical effects (Manchester WC), or slightly weaker ones (Newcastle) actually show no durational effects. The ones which show convincing effects of categoricity are the ones where duration seems to also be at play.

It is possible that some of these effects may increase or fade once subjected to more token collection, but under the results we have here this is the interpretation.

7.10.2 Variance and gradience

The issue of variability often gets tied up with gradience, when really they are two very different things. However, it is understandable why the two can become confused when considering some results of this experiment. We have seen in the previous sections that what looks like intermediate gradience can often be the mean output of variable results from opposite ends of the spectrum. It does bring up the question, however, how variable speakers are at producing light and dark /l/s in different environments.

7.10.2.1 Intermediate realisations and gradience

Many speakers in this section have intermediate averages for the realisation of the compound internal /l/s i.e. the *pee**l**-index*-type tokens. However, as has been discussed in many of the previous subsections, this usually does not warrant a third intermediate category. In the previous chapter, we saw that the Essex speaker has a consistently intermediate realisation of her ***heal** it* and *he**l**ix*-type tokens in comparison to initial ***l**eap* and final *hea**l***-type tokens. This affected the tongue root area and was confirmed by SS ANOVA. No speakers in this data set show this pattern as clearly as the Essex speaker. Most intermediate splines are a result of variation: the phonological context is made up of some light splines and some dark ones, and the plotted mean makes it look like it is in the middle overall.

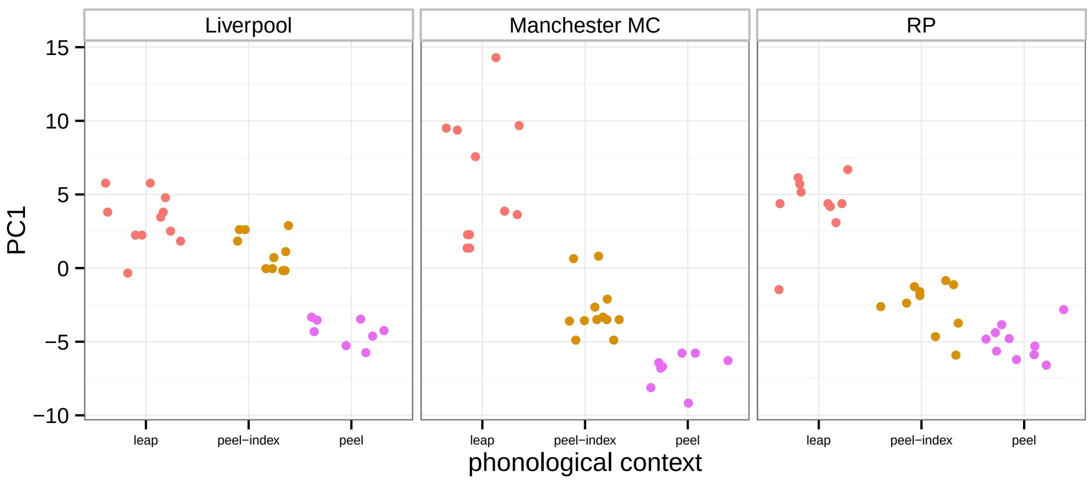

Figure 7.64: Potential intermediacy: comparing peel-index tokens with initial and final for three speakers (categories based on PC1 results).

Figure 7.64 shows three speakers and their realisations of *pee**l**-index*-type tokens in relation to initial and final tokens. A jitter plot is used to show individual tokens, as a box plot obscures any variance, making it look like gradience. We can see here that, although *pee**l**-index* tokens do look intermediate at first glance, they usually pattern with one of the extremes. For RP, they pattern with the dark tokens, but the fact that they are not as light gives a picture that they are intermediate. However, it is not surprising that this would happen given that, articulatorily, this compound /l/ is flanked by two high front vowels. There is bound to be some coarticulation which makes this sound lighter. The Liverpool speaker shows the opposite pattern, with the compound internal /l/s patterning with the light ones. We can see that the lightest *peel-index* token does not reach as high as the lightest ***l**eap*-type one, however, its range is completely encompassed by the ***l**eap* range, showing that it is not intermediate. One speaker who does show an intermediate range from these tokens is the Manchester MC speaker. Recall, however, that all her other intervocalic /l/s show the same behaviour, so that, in fact, it is conceivable that this is a monocategorial speaker with an expanded continuum. She and the Essex speaker show that intermediate realisations are possible, and this is discussed further in Section 7.10.3.

7.10.2.2 Variance

As we have seen from the previous section, a lot of the perceived gradience in the *pee**l**-index* context (i.e. the compound internal environment) is actually down to category mixture. This is the route more generally adopted by Hayes (2000) in his analysis of variation in American English /l/-darkening. This paper was followed up with a Stochastic OT approach to the problem (Boersma and Hayes 2001), where frequencies of discrete light and dark /l/ categories were modelled in different environments. The paper is discussed in more detail in Chapter 2, but the general idea is that canonical onset /l/s are light 100% of the time, canonical codas are light 0% of the time, and intermediate contexts result in intermediate frequencies. There is some evidence of this from Experiment 2, but more from sentences where the prosodic context could be variable, resulting in different strategies of syllabification. That is, most environments show next to no variation.

Figure 7.65 shows the distribution of tokens that are light or dark for each phonological context. Categories are taken by the results of the PCA, *not* on the *a priori* onset/coda distinction used for categorisation in the previous sections. This definition of a category instead emerges from the results, with positive PC1 values indicating a light /l/ and negative ones indicating a dark /l/.[10]

Figure 7.65 shows that most speakers and most contexts show little to no variation.[11] The Manchester MC speaker is quite variable due to the categorical postitive/negative PC1 split not being the best measure for this speaker. For example, her *he**l**ix*-type tokens are considerably darker than her ***l**eap* ones, and this results in the PC classifying these acoustically and articulatory light tokens as dark by comparison. Aside from this, there seems to be little variation in the more typical phonological environment. Two contexts which are very variable, however, are the ones which may be subject to varying prosodic

[10] Of course, it must be stressed that this way of defining category is not used in any of the correlations, linear regressions or PCA analyses, as they would be circular (e.g. working out the fit of light/dark category as a predictor of PC1 when the category was calculated by PC1 in the first place would be ridiculous).

[11] Note than the Manchester WC and the Belfast speaker have been removed from this plot, as their PC1 values may be unreliable due to the lack of much variation between splines.

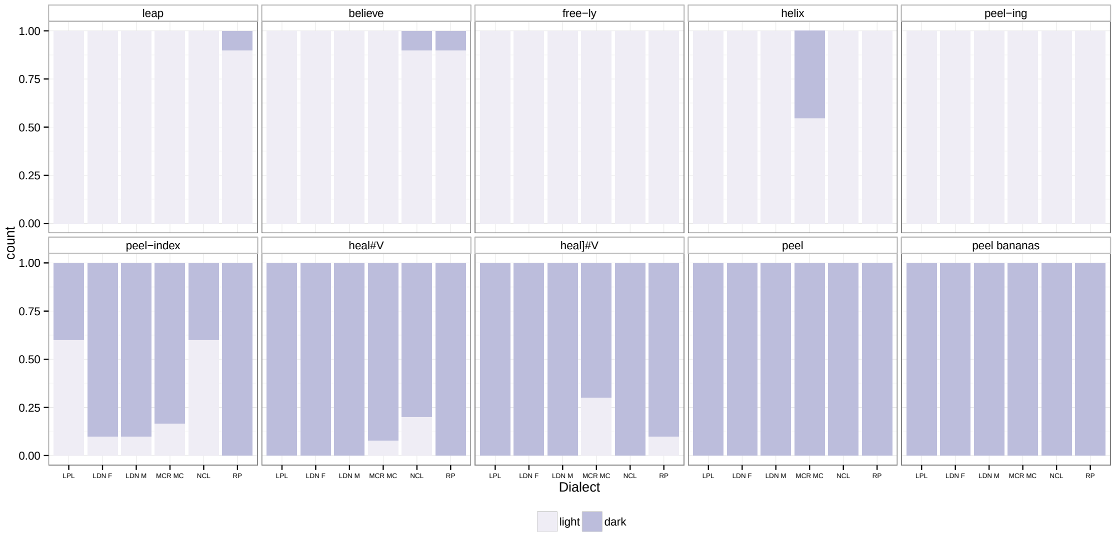

Figure 7.65: Distribution of light and dark /l/ in each phonological context (categories based on PC1 results). The categories are represented by their respective shades.

pressures. These are the ***peel**-index* and ***heal**#V*-type tokens, which have been discussed previously with respect to different speakers due to their seemingly intermediate realisation. These are the focus of the plot in Figure 7.66.

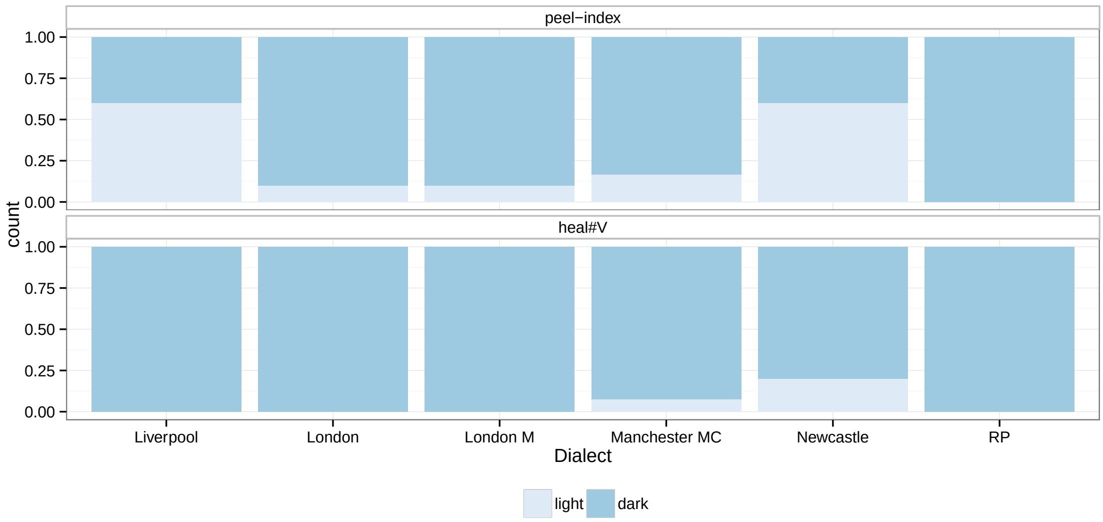

Figure 7.66: Distribution of light and dark /l/ in variable phonological contexts (categories based on PC1 results). The categories are represented by their respective shades.

On the whole, speakers treat ***heal**#V*-type tokens as being coda-like, that is, they are not resyllabified into the following word. Figure 7.66 shows the frequency of dark and light tokens for each speaker, based on the PCA results. With this context, it seems to be the result of how the speakers treat the sentences, and break them down prosodically. The two sentences in this environment, *I sent Neil interesting emails* and *I sent Neil innocuous pleas* vary from and within speaker. For example, the RP speaker utters all of his sentences very carefully and has a short but clear break between *Neil* and the following vowel. He produces dark [ɫ]s 100% of the time in this context, as do the two London speakers and the Liverpool speaker. Although the Newcastle speaker does this 80% of the time, on a couple of occasions he utters this sentence very fluidly, allowing for resyllabification and resulting in the /l/ being more ***l**eap*-like than *pee**l***-like. Such patterns show the importance of prosody on resyllabification strategies. This is discussed further in the following section.

The *pee**l**-index*-type tokens display much more variation in this phonological context. No speakers show a consistent pattern 100% of the time for these tokens. Only the Liverpool and Newcastle speakers produce lighter tokens for the majority part. The more standard sounding speakers of RP and London treat these as dark (or vocalised) /l/s. This is in contrast to the claims made by Halle and Mohanan (1985:65), who say that /l/ is resyllabified in compounds, e.g. *the seal-office*, but not in verb phrases e.g. *the seal offered a doughnut*. They claim the former displays a light [l] and the latter a dark [ɫ]. Indeed, Sproat and Fujimura (1993:303) speakers display dark [ɫ] in such verb phrases, as they do in the similar *hea**l**#V* phrases in this study. However, *pee**l**-index*-type compound tokens are dark (i.e. significantly retracted in comparison to initial tokens) for most speakers in this study, refuting the claims made by Halle and Mohanan. That is, it is not that they are wrong, necessarily, but are not representing all possible speaker systems. They are right for the Liverpool and Newcastle speakers, but not the others. An approach which pays due consideration to the predictions of the life cycle of phonological processes can account for both. However, it does bring us on to the discussion of what constitutes a word-final prevocalic consonant, and whether resyllabification automatically results in an onset-like realisation. This ties in with the intermediacy findings in Section 7.10.2.1 and is discussed further in the next section.

7.10.3 Word final prevocalic /l/s

The aim of Chapter 6 was to investigate what happened to word-final prevocalic consonants in different dialects, using prompts such as *hea**l** it*, where the /l/ is followed by a cliticised pronoun. In this chapter, such /l/s were avoided and instead two other word-final prevocalic contexts were included. These were the *pee**l**-index* and the *hea**l**#V*-type tokens. This has shown that word-final prevocalic consonants have to be treated carefully, as different syntactic structures result in different resyllabification strategies. Phrases such as *hea**l** it* result in a cliticised pronoun, which obligatorily results in resyllabification.

On the other hand, we find that non-cliticised following words beginning with a vowel vary within and across speakers. Sproat and Fujimura (1993)'s sentences *Beel equates the actors* elicit dark [ɫ]s for the majority of their speakers (i.e. the dorsal gesture precedes the coronal). In this study, sentences where /l/ is followed by a verb phrase also produce dark [ɫ]s for the most part, as we saw in Figure 7.66. In Chapters 2 and 3, this was analysed as a word-level darkening process, whereby /l/ darkens in the coda at the phrase level, regardless of what follows and regardless of what resyllabification may occur at the phrase level. However, closer attention paid to the prosodic treatment of sentences by speakers may be needed in order to diagnose the true syllabification of such sentences. As shown by Cho et al. (2014), a prosodic boundary prevents the resyllabification of a final consonant across word boundaries. Listening to the sentences from this experiment, speakers often display a short prosodic break before the phrase *interesting emails*, which arguably assists with them in treating this as coda /l/. The upshot is that distinguishing between word-level and phrase-level processes can be unexpectedly difficult owing to the variability of phrasal prosodification, and to the fact that different processes seem sensitive to different prosodic boundaries. Further work in this area will be needed in the future.

7.10.4 Cross-speaker analysis

Many studies of liquid consonants have shown that the most reliable data is obtained through articulatory measures, and that although the acoustic can give us a good insight, there may be patterns which evade the acoustics. In Experiment 2, the match between acoustic and articulatory data has been impressive for the most part. Although we see time and time again that some articulatory differences leave no immediately apparent acoustic trace, the F2-F1 measures give us a nice picture of the distribution across different phonological categories. It must be stressed that this is due to the neighbouring environments being kept constant across all speakers. If different vowels were used across different contexts, the acoustic results would have been completely unreliable.

One thing which is striking is the similarity between the PC1 plots and the acoustic values, showing that the midpoint of the splines and the midpoint of the F2 values minus the F1 values tally up nicely (although cf. the Manchester WC speaker's *be**l**ieve*-type tokens). This is reassuring, as the synchronisation between the ultrasound images and the acoustics is always a concern in such studies. This does not seem to be a problem here, given the acoustics were segmented in Praat and then imported into AAA for spline drawing. Indeed, the synchronicity is hardly surprising as the unit has been subjected to a thorough taptest, as outlined in Chapter 5.

A benefit of the acoustics over the articulatory data is the ability to conduct cross-speaker comparisons in the same analysis. In the next section, we take another look at the durational effect to see if the combination of tokens shows some kind of overall effect which was not convincingly displayed on a speaker-by-speaker basis.

7.10.4.1 Duration

Perhaps surprisingly, given the strong statement made in Sproat and Fujimura (1993), duration does not seem to have much of an influence over darkening in this dataset. The common theme in most speaker data is that darkness does correlate with duration, but this correlation is only moderate and often drops out in the linear models. Generally, however, the effect can be observed.

As discussed extensively in Chapter 4 and at points in this chapter, Yuan and Liberman (2009; 2011) find a correlation for dark [ɫ]s and duration, but not for light [l]s. Bermúdez-Otero and Trousdale (2012) use this to argue that categorical and gradient effects can be simultaneously active in the same grammar.

Some speakers in this dataset, such as Manchester WC, have results which suggest that actually only phrase-final /l/s show this correlation. Perhaps this was the case for Yuan and Liberman's results, as they did not look at a range of darkness, and so could only conclude observations from phrase-final tokens anyway. It does seem to be the case the typical Manchester /l/ system has similarities with the American system, and this could be investigated more closely in future with more American English data. Sproat and Fujimura (1993) on the other hand, as discussed in Chapter 4, claim to find a correlation for all of their environments, although this does not look particularly convincing for the light tokens, as shown in Chapter 4. They do not have any absolute phrase-final tokens in their study. The closest to this environment, with /l/ in the coda at the phrase-level is their pre-/h/ tokens. However, as we have seen here, these tokens are often not correlated with darkness at all for most speakers, so it is impossible to test this for this dataset.

Furthermore, it could be argued that Sproat and Fujimura's measure of darkness is inherently correlated with duration. They use tip delay as their measure of darkness: the longer the tongue tip gesture is delayed, the darker the /l/. The dorsal and coronal gesture make up the /l/ for these speakers, and occur almost simultaneously for a light /l/. For a dark /l/, however the tongue tip gesture is delayed. Therefore it will take more time to articulate the segment overall, but notably the measure of darkness is inherently linked to duration here. Perhaps a durational effect would be stronger if this study had used temporal data, but it raises the question as to whether this measure is really the best gauge of /l/-darkening. As mentioned above, it is the very speakers for which we would expect duration to play a bigger role (i.e. the ones who show no or less evidence for categoricity) where the durational effects are the least significant.

However, lighter /l/s do tend to be shorter, as can be seen in the combined jitter-box plot in Figure 7.67. Perhaps combining the data together and looking at the data set as a whole will give us more information. Figure 7.68 shows darkness against duration for all speakers. This time darkness is based on the F2-F1 measures. Duration is the 'rime' duration again. The correlation lines are fit slightly differently here, based on the results of the PCA. Like the variability results in Section 7.10.2 above, a positive PC1 is assigned a 'light' label, and a negative PC1 is assigned a 'dark' label. These could not be used in previous sections, as the darkness value was based on PC1 itself, so the correlation lines had to be fit to the pre-decided categories. Although these were informed by the data, it did not allow for variable contexts, where the /l/ may be light in some tokens but dark in

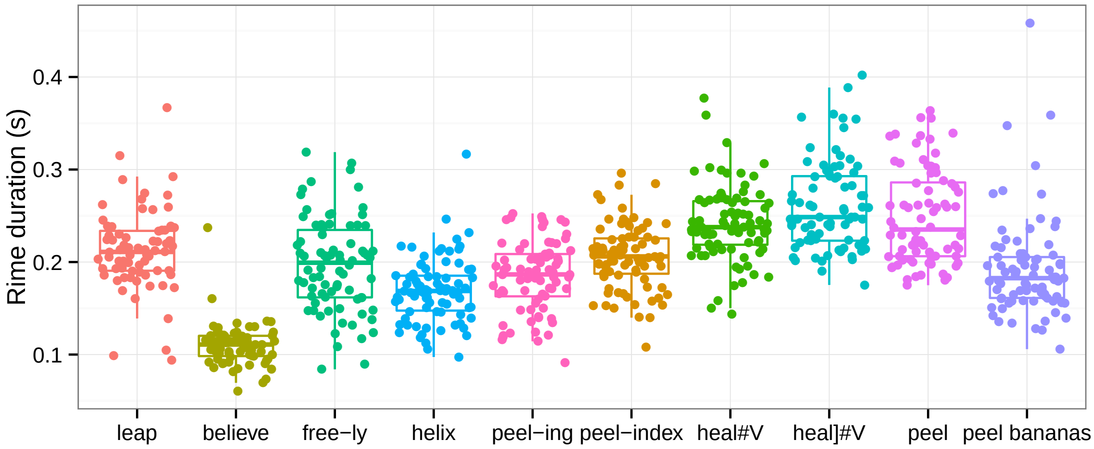

Figure 7.67: Average rime duration across phonological context for all speakers.

others. The correlation lines in Figure 7.68 are purely based on the articulatory outcome.

We can see a weak correlation for the light tokens (r = -0.164; p= 0.001), and a much stronger one for the dark tokens (r = -0.292; p= 5.402×10^{-9}). Both are statistically significant and show that the longer the 'rime', the darker the /l/. However, with a sample size of almost 1000 tokens, Pearson's r will tend to result in significant values. A more robust model including all factors is presented in the next section.

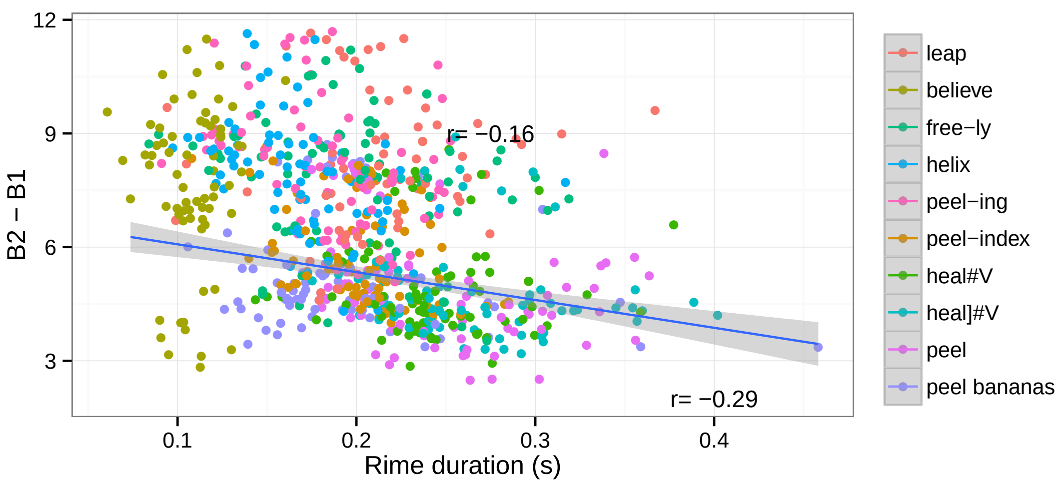

Figure 7.68: Bark normalised F2-F1 (i.e. B2-B1) against log of Rime duration for all speakers. Correlation lines are fit to the PCA categories (a positive PCA is light, and a negative PCA is dark).

7.10.4.2 Mixed-effects linear regression of acoustics

One benefit of using the acoustic data for analysis is the ability to conduct cross-speaker comparisons. Although it is not possible with the data collected in Experiment 2 to enter all articulatory information into the same statistical model, due to physiological differences and differences in probe position, it is possible to run larger models on the

model	AIC
1. B2-B1 ∼ log(Rimeduration) + (1|Dialect)	2511
2. B2-B1 ∼ Category + (1|Dialect)	3019
3. B2-B1 ∼ Category + log(Rimeduration) + (1|Dialect)	2493
4. B2-B1 ∼ Category * log(Rimeduration) + (1|Dialect)	2437
5. B2-B1 ∼ log(Rimeduration) + (Category | Dialect)	2106
6. B2-B1 ∼ Category + log(Rimeduration) + (Category | Dialect)	2097
7. B2-B1 ∼ Category *log(Rimeduration) + (Category | Dialect)	2078
8. B2-B1 ∼ Context + (1|Dialect)	2500

Table 7.21: Comparing mixed effects models fit to B2-B1 difference for all speakers. AIC shows goodness-of-fit.

acoustics. Mixed effects models also allow us to include Speaker or Dialect as a random effect, accounting for different behaviour by individual speakers.

Once again, in the cross-speaker part of the analysis, the Bark normalised values were used as the dependent variable, with F1 and F2 converted to B1 and B2 in order to eradicate inter-speaker differences. Mixed-effects models using Dialect as a random effect were fitted to category alone, category plus duration, and an interaction of the two. The models fitted, alongside their AIC values are shown in Table 7.21 (the lower the AIC, the better the fit). Taking advantage of the possibilities of mixed effects models, a random slope of Category and Dialect was also fitted, which allows the model to specify separate intercepts for Category for each Speaker. When including this alongside an interaction of Category and Duration, this model created the best fit in terms of AIC, as shown in Table 7.21. The ANOVA comparison also shows a significantly better result for adding Category by Dialect as a random slope ($p < 0.001$; see Appendix D).

Table 7.22 shows the coefficients of the fixed-effects for the best model, model 7. As explained in Chapter 5 this coefficient table shows t-values instead of p-values. Values higher than +/- 1.96 are taken to be significant, equivalent to a p-value of less than 0.05.

	Estimate	Std. Error	t value
(Intercept)	7.9934	0.6182	12.9297
Catdark	-4.5721	0.6988	-6.5429
log(Rimeduration)	0.0524	0.1526	0.3433
Catdark:log(Rimeduration)	-1.2005	0.2588	-4.6395

Table 7.22: Summary of best mixed-effects linear regression on acoustic data (model 7)

Table 7.23 shows how the random slope fits separate intercepts for Category for each speaker. The figures highlight why this provides a much better fit, given some speakers have a much greater distance in formants between the two categories. Running Context as a predictor does not provide as good as fit as Category, suggesting that a Categorical analysis fits the data better than one which just considers phonological environment.

	(Intercept)	Catdark
Belfast	2.0000	-0.1297
Liverpool	0.3952	-1.8389
London	2.6357	-3.7551
London M	2.0670	-3.8757
Manchester WC	-1.0294	-0.9712
Manchester MC	2.0901	-3.4799
Newcastle	1.5346	-2.0040
RP	4.2056	-5.0244

Table 7.23: Random slope coefficients for Category and Speaker,

7.11 Summary

This chapter has investigated the realisation of English /l/ in ten phonological environments using ultrasound tongue imaging. The investigation of the data was conducted in order to address the research question set out earlier in the thesis regarding categoricity and gradience. On the basis of these results, we have seen that previous approaches to /l/-darkening have been overly dismissive in their interpretation of categorical allophonic distinctions between light and dark /l/.

This chapter has made several contributions to our understanding of /l/-darkening in English, including categorical vs. gradient effects, allophonic categorical distinctions and lack of them, durational effects and their correlation with darkening, articulatory correlates of /l/-vocalisation; and dialectal variation. Moreover, this chapter has sought to account for the variation found in this subset of dialects using the predictions of the life cycle of phonological processes.

The main point this chapter seeks to make is that it is not helpful to dismiss categoricity outright, particularly when the opportunities to observe such patterns are overlooked. A full analysis of a wide range of phonological contexts is needed, but even then, fine-grained phonetic differences may make it difficult to diagnose the true pattern. One glance at the stark difference between initial and final /l/ realisations in dialects such as RP may be a convincing enough an argument for categoricity, but the issue becomes trickier when observing small phonetic differences such as the Newcastle and Manchester MC patterns.

The results from the Manchester WC speaker demonstrate the kind of phonetic range that American phoneticians may be dealing with in their investigation into English /l/-darkening. One may understandably become more sympathetic to their dismissal of categoricity when observing the tiny phonetic differences and overall homogeneity displayed by the Manchester WC midpoint splines. However, the results for RP and London demonstrate that, although studies of American English may not show such clear patterning, claiming that /l/-darkening displays no allophony is a short-sighted approach to the variability in the English language. Some dialects may not have a categorical allophonic distinction between light and dark /l/. This does not mean we should dismiss the idea of categoricity altogether.

This chapter has also shown that a wide range of phonological contexts are needed in order to appreciate the full spectrum of possible /l/ systems. This is also required if one wishes to make claims about categoricity and gradience in general, as zooming in on just a subset of possible realisations does not give the full picture, and is not a suitable dataset on which to base such claims of phonetics-phonology interactions.

We have seen that the effect of duration has been overestimated in previous studies. The contribution of duration in realisation of English /l/ originated with a study which looked at a durational-intrinsic measure of tip delay as its primary correlate of darkness. Duration does not seem to play as an important role here. However, the general trend is confirmed: longer /l/s are darker, or darker /l/s are longer.

This chapter has shown that different dialects of English display different phonetic implementations and phonological realisations in their /l/-darkening systems. We can still have categorical distinctions over a small phonetic space. A speaker whose system operates on a relatively light scale can still show consistent differences between initial and final tokens which may not be as articulatorily or acoustically as stark as a standard English speaker, but still shows a categorical distinction.

We have also seen the first articulatory evidence for /l/-darkening systems in different dialects of English, many of which are novel findings which raise questions for future research in sociophonetics. The comparatively small distance between light and vocalised /l/s in London poses the question as to whether vocalisation is simply the loss of the tongue-tip gesture in this variety, which can be analysed further using EPG data. The velarisation of Liverpool dark [ɫ]s has shown that there is clear difference between pharyngealised and velarised /l/s which evades the acoustics. As the other six speakers showing dark [ɫ]s did not velarise, this might suggest that this strategy for darkening is in the minority in English varieties of English, but this will have to be investigated further. We have also seen social class differences in Manchester, in that the working class speech corroborates (to a certain extent) the dialectological claims, whereas the middle class speaker approximates a more standard system.

Overall, this chapter has shown that dialectal diversity has been vastly underestimated in the existing literature on /l/-darkening. The phonetics-phonology interface effects have thus far been reduced to a false dichotomy between either a purely categorical, or purely gradient approach, when in fact, both exist within the same grammar. The effects of differing sensitivity to morphosyntax as outlined in Chapter 6 have not been discussed widely in this chapter; however we need a theory that can account for the evidence that categorical darkening domains may differ in size between dialects. The life cycle of phonological processes can make sense of such facts, with rule scattering and domain narrowing accounting for the coexistence of categorical and gradient effects, and lenition trajectories such as /l/-vocalisation. The wide range of dialectal diversity, for which this chapter provides only a small subset, shows a great deal of orderliness if considered from the viewpoint of the life cycle.

Chapter 8

Conclusions and future directions

The process of /l/-darkening in English, and its subsequently related processes of lenition, shows effects of morphosyntactic sensitivity, categorical allophony, phonetic gradience, and dialectal diversity. The current investigation has shown evidence of these effects across numerous dialects of English. Accordingly, the analyses conducted using ultrasound tongue imaging contribute to several ongoing debates in phonology, phonetics and language variation and change. It has been shown that previous approaches to /l/ allophony have underestimated the diversity of the phenomenon, not only idiolectally but also dialectally. We can also see that, on the whole, the speakers in this study are extremely consistent, demonstrating that the amount of categorical intra-speaker variation in /l/-darkening may have been vastly overestimated thus far (Hayes 2000).

The first research goal of the thesis was to investigate evidence of the life cycle of phonological processes through articulatory means. The first stage of the life cycle concerns the predictions made with respect to morphosyntactic conditioning and variation in domain size. In the existing literature, it is possible to find evidence of different levels of morphosyntactic sensitivity by collating findings of separate studies. However, these studies are often based on descriptive and auditory materials, or acoustic data, all of which can be problematic in studying /l/ realisation reliably. This investigation has shown within consistent articulatory parameters that speakers exhibit differing levels of morphosyntactic conditioning depending on their variety of English. In line with the life cycle's predictions, we would predict that /l/-darkening was able to apply at the stem, word and phrase levels, and this study confirms such domain specific application. In Chapter 6, it was shown that the Essex speaker seems to have a stem-level darkening process, with more tongue root retraction in *hea**l**-ing* and *hea**l** it* than in ***l**eap* and *he**l**ix* i.e. more retraction when the /l/ is in the coda at the stem level. Chapter 7 shows various speakers who have word-final dark [ɫ] prevocalically demonstrating the possibility of a word-level darkening process (although it is noted that prosodic structure and speech rate may be interfering with potential resyllabification here). In turn, we have speakers who show light realisations in word-final prevocalic /l/s, indicating a phrase-level process. The stratal architecture predicts that /l/-darkening can apply at all three levels, and this is what we find. Moreover, we have seen that all patterns fall in line with the predictions of the Russian Doll Theorem (Section 3.4), in that overapplication in one domain implies overapplication in the next wider domain i.e. lenition in *hea**l**-ing* implies lenition in *hea**l** it*.

However, as just mentioned, one aspect of this experiment which should be taken into consideration in future studies is the potential for inter and intra-speaker variation of prosodic phrasing, treatment of boundaries and syllabification. As discussed in Chapter 2, it has been shown by articulatory means that the placement of prosodic boundaries prevents expected resyllabification (Cho et al. 2014). This means that, for some of the speakers in Chapter 7, we do not know for sure whether they have a word-level darkening rule or a phrase-level darkening rule. The word final prevocalic /l/s were treated differently across speakers and, in some cases, even within speakers. Factors such as stylistic variation and speech rate are difficult to predict, and alternative stimuli or possibly alternative speech extraction methods should be considered if such phonological contexts are of interest in future work. However, the fact that speakers may treat word-final prevocalic tokens differently depending on whether the second word is cliticised or not would also be an interesting avenue for future research to investigate in line with the domain narrowing stage of the life cycle.

The life cycle also makes reference to the interaction between gradient phonetics and categorical phonology. Given the predictions of rule scattering, we would expect to see synchronically overlaid gradient and categorical effects of the same process operating within the grammar, where the categorical effects have arisen diachronically from the gradient effects by stabilisation. We find many speakers with a categorical distinction, who also display overlaid gradient effects, such as extra tongue root retraction and longer duration. Such patterns fit in with the second research goal of the thesis: to investigate the debate in the existing literature surrounding categoricity and gradience. As we have seen, several studies have interpreted the allophonic distinction between light and dark variants as a gradient effect where /l/ variation lies on a continuum. Given the clear morphological effects, such an analysis is problematic for a modular approach, as it relies on phonetics and morphology sharing an interface. The question of categoricity and gradience in /l/-darkening was investigated in Chapter 7, using both qualitative descriptions of spline plots and several quantitative diagnostics, as outlined in Chapter 5. The results from the investigation show that the phonetics-phonology interface effects have thus far been reduced to a false dichotomy between either a purely categorical, or purely gradient approach when, in fact, both exist within the same grammar. Many speakers show a clear categorical distribution between certain contexts, whilst also displaying gradient effects of duration. Such effects are predictable under the life cycle by rule scattering, whereby categorical phonological processes which stabilise do not replace the phonetic effects from which they emerge, but coexist with them. In this sense, the present analysis is more in line with the findings of Yuan and Liberman (2009: 2011), who find evidence of both types of effects operating in the same grammar in the manner predicted by rule scattering. The categorical allophony displayed by many speakers also disproves Sproat and Fujimura's (1993) conclusions that /l/-darkening is always gradient. Overall, I have argued that in order to observe effects of categoricity and gradience properly, one needs an insight into the entire spectrum of possibilities, and we cannot draw conclusions from just two or three phonological contexts. The patterns displayed in the ten phonological contexts used here reinforce this point clearly.

The life cycle makes a further prediction concerning rule scattering which is also corroborated in the thesis: the interaction of darkening and vocalisation. As vocalisation is a more drastic form of lenition, it is a separate stage of the lenition trajectory and may apply at a different level of the grammar. Crucially, when such interactions are found, the life cycle predicts the older milder process applies at a higher level (within a narrower domain) than the younger harsher process. This is what we find in the Essex speaker. The younger harsher process of darkening is phrase level, but the older milder process of darkening has existed for longer and thus has had time to climb up to the stem level.

One aspect of the life cycle which has not been accounted for in this study is rule generalisation. Rule generalisation involves the progression of a phonological process through the prosodic domains, for example, advancing from lenition in the coda, to lenition anywhere outside of foot-initial position. This means that, theoretically, there is the possibility that some dialects may have dark variants in all foot-medial position e.g. *helix, yellow*. Previous studies such as Hayes (2000) and Yuan and Liberman (2009: 2011) would lead us to believe that this could certainly be the case. However, no speaker in the present study showed any form of /l/ lenition in this position, with *he**l**ix*-type tokens patterning with ***l**eap*-type tokens for all participants. It is likely not a coincidence that both of the studies mentioned above concern American English. The foot-based darkening pattern has not been vindicated by articulatory evidence thus far, and is a compelling issue for further research.

This thesis has also shown that dialectal diversity has been vastly underestimated in the existing literature on /l/-darkening. Although the mass of data may be difficult to interpret when considering the morphophonological effects alongside evidence for categoricity and gradience, the patterns show a great deal of orderliness if considered from the viewpoint of the life cycle. We need a theory that can account for the evidence that categorical darkening domains may differ in size between dialects. The effects of /l/-vocalisation, in addition, may coexist as a lenition process with darkening or replace it altogether. The life cycle of phonological processes can make sense of such facts, with rule scattering and domain narrowing accounting for the coexistence of categorical and gradient effects, as well as lenition trajectories accounting for the presence of /l/-vocalisation alongside darkening.

We have also seen that there are several aspects in the ultrasound data that evade the acoustics. Firstly, more fine-grained articulatory properties may not be picked up in the acoustics, causing a more abrupt effect of bimodality as warned by Browman (1995). This can be seen in the Manchester MC speaker who seems to have a four-way distinction in the articulatory data, but a two-way distinction in the acoustics.

On top of the fine-grained differences, there are two other findings in the articulatory data collected in this study which the acoustics miss altogether. The first of these is the presence of velarisation in the Liverpool speaker's dark [ɫ]s, a difference which is not reflected in any of the acoustic measurements. Although some speakers show occasional evidence of velarisation, this secondary articulation does not occur consistently for any other participant in either experiment. This shows that *velarisation* is not an accurate synonym of darkening, more a hyponym of the possible forms. It also raises questions as to the distribution of velarised vs. pharyngealised forms in British English as, in this dataset, just one of ten British speakers showed this velarising distribution. The velarisation of Liverpool /l/s poses further questions which can only be investigated by articulatory measures for this accent. Is it true that Liverpool velarises throughout, not just in dark [ɫ]s, as Knowles (1973) suggests? As this does not seem to apply to the light /l/s, would we predict that velarisation would just apply to sounds produced with the back of the tongue? These questions provide many prospects for future study, not only for Liverpool English, but also in investigating additional dialects which display velarisation instead of pharyngealisation. Further study of transatlantic comparisons may shed more light on this, as many American English dialects are often described as having velarised /l/s. This study has provided an overview of several dialects of British English, but only one speaker of American English. Given some of the interesting patterns found in previous studies, it would be worth re-running the experiments with American speakers, to see if we can find evidence of velarisation, as well as accounting for more levels of morphosyntactic and prosodic sensitivity, or contrasting temporal patterns with British English.

In addition to secondary articulation effects of velarisation, we also see variation in terms of the degree of pharyngealisation involved in /l/-vocalisation. As discussed in Chapter 6, the degree of pharyngeal constriction involved in vocalised /l/s is not something that has ever been investigated instrumentally. We saw that the Essex speaker had a considerably retracted tongue root in the vocalised tokens, but the London speakers did not, and this difference was not reflected in the acoustics. The difference between the vocalisation patterns in London and Essex highlights the need for a full-scale quantitative analysis of /l/ variation in these communities. This would not only provide us with a more in-depth overview of the link between darkening and vocalisation, but more importantly it would show us on a wider scale how lenition trajectories interact with domain narrowing, which could be applied to all kinds of phonological phenomena. Similarly, the differences in tongue root retraction in vocalised /l/ could inform our knowledge of articulatory and acoustic comparisons, and also how articulatory factors slot into the stages of lenition trajectories. If /l/ vocalisation is defined as the loss of tongue-tip contact only, then we may not be surprised that the London speakers have such a small magnitude between light and vocalised /l/, and the realisation may be comparable to that of a regular dark variant. However, if vocalisation is defined in phonological terms as the loss or change of feature values, then we might expect to observe more robust phonetic differences between vocalised variant and consonantal dark variant. Whether a speaker has two categories (as the London speakers seem to show) or three (as the Essex speaker provides evidence for) may also affect this aspect of the articulation.

Not only has this thesis made a theoretical contribution, there are also several methodological contributions worth pointing out. Firstly, I have shown that Principal Components Analysis is a viable method of quantifying tongue splines, albeit in a comparatively crude manner to pixel data. Nevertheless, the convenience of quantifying a tongue shape with one number is an advantage of this more general approach. Secondly, I have demonstrated that a temporal analysis may not be necessary when accounting for differences in /l/-darkening data. Although it is not claimed here that temporal data is superfluous, I do argue that midpoint data is largely sufficient for describing and quantifying patterns produced by the majority of British speakers. Since Sproat and Fujimura's (1993) seminal study, many have cited the relative phasing of gestures as the primary articulatory correlate of darkness, but here I argue that their other diagnostic, the magnitude of the dorsal component, suffices for many dialects. As outlined at several points in the thesis, this study did not aim to conduct an in-depth analysis into the temporal properties of /l/-darkening processes. This was a concern from the beginning of the study, as the ultrasound unit used in this investigation gives a basic frame rate of just 30 fps, which is not high enough for a thorough analysis. It does give us some indication of temporal effects, however, and this in shown in Chapter 6, where we see that the midpoint of the /l/ functions sufficiently as an overall measure of darkness. The RP speaker is used as the primary example, demonstrating that the initial context varies a small amount over the course of the /l/, but remains completely separate from the trajectory of the final /l/. What we are not seeing for these speakers is two gestures which mirror one another in different positions: for RP the two allophonic contexts involve very different articulations throughout. However, it must be pointed out that a thorough temporal analysis could shed light on either a) small gestures we cannot see in 30 fps ultrasound, and b) more information for speakers who don't seem to show a clear categorical distribution. It could just be that, for speakers such as Manchester WC, we need access to a temporal data in order to tease apart the true categoricity. For these reasons, it is important to carry out a follow-up study which can address these questions, possibly using automated temporal analysis tools such as optical flow (Moisik 2013: 2014). Finally, I have used a combination of statistical techniques, such as the dip test for bimodality, as well as simple and mixed-effects linear regression to highlight patterns in the data.

The results of this study provide numerous opportunities for further research. This does not only include investigations such as foot-medial darkening, patterns of velarisation, and pharyngeal constrictions in vocalisation, as discussed above, but the potential to research new aspects of /l/ variation, as well the analysis of other phonological variables. The data from the Belfast speaker suggests that coarticulation with neighbouring vowels may be important for this dialect, which shows the need to repeat the present investigation with a wider variety of vowel sounds. The excrescent schwa found between /l/ and the preceding vowel in words such as ***heal*** (e.g. [hiːəɫ]) would also be an avenue for future research, as the acoustics clearly show that speakers have varying degrees of insertion, bringing up further questions of categoricity and gradience. The effect of /l/ on preceding vowels has been noted in this experiment, with many speakers showing possible mergers before /l/. Distracter sentences have been collected for future analysis on some of these speakers, including words such as *Paul, pool, pull* and *pole*.

The most obvious direction for future work on a new phonological variable would be the realisation of /r/ in different phonological environments. We have plenty of evidence of the articulatory realisation of American English /r/ (Gick 1999; Campbell et al. 2010), but little evidence of whether non-rhotic English dialects exhibit the same amount of lenition in contexts such as word-final prevocalic position. A temporal analysis may be preferable for such study, and ultrasound could help us discover whether tongue shapes vary in different positions in /r/ in non-rhotic varieties. In addition, by considering several aspects of the multidimensional phonetic space, we could investigate whether the lip gesture is smaller in resyllabified /r/s, and whether the tongue-root gesture is timed earlier, as found by Campbell et al. (2010:289) for North American English. Moreover, we could compare findings with intrusive /r/ realisations (e.g. *saw*[ɹ] *it*). Such data would allow comparisons with /l/, and prompt discussions of whether the life cycle of phonological processes continues to provide a successful account of the patterns found. Thus far, the ability to connect the seemingly varied processes of phonologisation, stabilisation, domain narrowing, rule scattering and rule generalisation under one model by providing a diachronic account of synchronic patterns in language makes the life cycle of phonological processes the choice theory to account for the data.

Bibliography

Akaike, Hirotugu. 1974. A new look at the statistical model identification. *Automatic Control, IEEE Transactions on* 19:716–723.

Articulate Instruments Ltd. 2010. Syncbrightup users manual: Revision 1.10. Edinburgh, UK.

Ash, Sharon. 1982. The vocalization of /l/ in Philadelphia. Doctoral Dissertation, University of Pennsylvania.

Baayen, Harald. 2008. *Analyzing linguistic data: A practical introduction to statistics using R*. Cambridge: CUP.

Baranowski, Maciej. 2013. The sociolinguistics of Goose and Goat in Manchester. Paper presented at NWAV42 conference. 17-20 Oct 2013, Pittsburgh.

Baranowski, Maciej, and Danielle Turton. Forthcoming. Manchester English. In *Researching Northern Englishes*, ed. Raymond Hickey. Amsterdam and Philadelphia: John Benjamins.

Barry, Martin C. 2000. A phonetic and phonological investigation of English clear and dark syllabic /l/. *Les Cahiers de l'ICP. Bulletin de la communication parlée* 77–87.

Bates, Douglas. 2006. [R] lmer, p-values and all that. url https://stat.ethz.ch/pipermail/r-help/2006-May/094765.html.

Bates, Douglas, Martin Maechler, Ben Bolker, and Steven Walker. 2014. *lme4: Linear mixed-effects models using eigen and s4.* URL http://CRAN.R-project.org/package=lme4, r package version 1.1-6.

Beal, Joan. 2008. English dialects in the north of England: phonology. In *Varieties of English. vol. 1: The British Isles*, eds. Bernd Kortmann and Clive Upton, 122–144. Berlin: Mouten de Gruyter.

Benua, Laura. 1995. Identity effects in morphological truncation. In *Papers in Optimality Theory*, eds. Jill Beckman, Suzanne Urbanczyk, and Laura Walsh Dickey, University of Massachusetts Occasional Papers in Linguistics 18, 77–136. Amherst, MA: GLSA, University of Massachusetts.

Benua, Laura. 1997. Affixes classes are defined by faithfulness. *University of Maryland Working Papers in Linguistics 5. Selected Phonology Papers from Hopkins Optimality Theory Workshop 1997 / University of Maryland Mayfest 1997*, 1–27.

Bermúdez-Otero, Ricardo. 1999. Constraint interaction in language change: quantity in English and Germanic [opacity and globality in phonological change]. Doctoral dissertation, University of Manchester and Universidad de Santiago de Compostela.

Bermúdez-Otero, Ricardo. 2006. Phonological change in optimality theory. In *Encyclopedia of language and linguistics*, ed. Keith Brown, Vol. 9, 497–505. Oxford: Elsevier, 2 edition.

Bermúdez-Otero, Ricardo. 2007a. Diachronic phonology. In *The Cambridge handbook of phonology*, ed. Paul de Lacy, 497–517. Cambridge: Cambridge University Press.

Bermúdez-Otero, Ricardo. 2007b. Word-final prevocalic consonants in English: representation vs derivation. Paper presented at the Old World Conference in Phonology 4, Rhodes. Handout available at http://www.bermudez-otero.com/OCP4.pdf.

Bermúdez-Otero, Ricardo. 2010a. The history of English phrase-level syllabification: A sequel to Minkova (2003: ch. 4). Paper given at the inaugural event of the Historical Linguistics Research Cluster, Cambridge, 10 May 2010.

Bermúdez-Otero, Ricardo. 2010b. Morphologically conditioned phonetics? Not proven. Paper presented at On Linguistic Interfaces II, Belfast.

Bermúdez-Otero, Ricardo. 2011. Cyclicity. In *The blackwell companion to phonology*, eds. Marc van Oostendorp, Colin Ewen, Elizabeth Hume, and Keren Rice, Vol. 4, 2019–2048. Malden, MA: Wiley-Blackwell.

Bermúdez-Otero, Ricardo. 2013. Amphichronic explanation and the life cycle of phonological processes. In *The Oxford handbook of historical phonology*, eds. Patrick Honeybone and Joseph C. Salmons. Oxford: Oxford University Press.

Bermúdez-Otero, Ricardo. 2014. The diachronic rise and synchronic representation of phonological opacity: a case study. Paper given at the Rencontres d'Automne de Linguistique Formelle (RALFe) 2014, Université Paris 8, 9 October 2014.

Bermúdez-Otero, Ricardo, and Graeme Trousdale. 2012. Cycles and continua: on unidirectionality and gradualness in language change. In *The Oxford handbook of the history of English*, eds. Terttu Nevalainen and Elizabeth Closs Traugott, 691–720. Oxford: Oxford University Press.

Bladon, R. A., and A. Al-Bamerni. 1976. Coarticulation resistance of English /l/. *Journal of Phonetics* 4:135–150.

Boersma, Paul, and Bruce Hayes. 2001. Empirical tests of the gradual learning algorithm. *Linguistic Inquiry* 32:45–86.

Boersma, Paul, and David Weenink. 2010. Praat: doing phonetics by computer .

Borowsky, Toni. 2001. *The vocalisation of dark-l in Australian English.* Amsterdam/Philadelphia: John Benjamins.

Borowsky, Toni, and Barbara Horvath. 1997. L-vocalization in Australian English. In *Variation, change and phonological theory*, eds. Frans Hinskins, Roeland van Hout, and W. Leo Wetzels, 101–123. Amsterdam: John Benjamins.

Bressmann, T., C. Heng, and JC Irish. 2005. Applications of 2D and 3D ultrasound imaging in speech-language pathology. *Journal of Speech Language Pathology and Audiology* 29:158.

Britain, David. 2009. One foot in the grave. dialect death, dialect contact, and dialect birth in England. *International Journal of the Sociology of Language 2009* 121–155.

Browman, Carol, and Louis Goldstein. 1989. Articulatory gestures as phonological units. *Phonology* 6:201–251.

Browman, Carol, and Louis Goldstein. 1992. Articulatory phonology: An overview. *Phonetica* 49:155–180.

Browman, Catherine P. 1995. Assimilation as gestural overlap: comments on Holst and Nolan. *Phonology and phonetic evidence, papers in laboratory phonology I* 334.

Browman, Catherine P, and Louis M Goldstein. 1986. Towards an articulatory phonology. *Phonology* 3:219–252.

Campbell, Fiona, Bryan Gick, Ian Wilson, and Eric Vatikiotis-Bateson. 2010. Spatial and temporal properties of gestures in North American English /r/. *Language and speech* 53:49–69.

Carter, Paul. 2002. Structured variation in British English liquids. Doctoral Dissertation, University of York.

Carter, Paul. 2003. Extrinsic phonetic interpretation: spectral variation in English liquids. In *Phonetic interpretation: Papers in laboratory phonology vi*, eds. Richard Ogden John Local and Rosalind Temple, 237–252. Cambridge: Cambridge University Press.

Carter, Paul, and John Local. 2007. F2 variation in Newcastle and Leeds English liquid systems. *Journal of the International Phonetic Association* 37:183–199.

Chambers, Jack. 1973. Canadian raising. *Canadian Journal of Linguistics* 18:113–135.

Cheshire, Jenny, Paul Kerswill, Sue Fox, and Eivind Torgersen. 2011. Contact, the feature pool and the speech community: The emergence of Multicultural London English. *Journal of Sociolinguistics* 15:151–196.

Chládková, Kateřina, and Silke Hamann. 2011. High vowels in Southern British English:/u/-fronting does not result in merger. *Proceedings of the 17th International Congress of Phonetic Sciences (ICPhS 17)* 17–21.

Cho, Taehong, Yeomin Yoon, and Sahyang Kim. 2014. Effects of prosodic boundary and syllable structure on the temporal realization of CV gestures. *Journal of Phonetics* 44:96–109.

Chomsky, Noam, and Morris Halle. 1968. *The sound pattern of English.* New York: Harper & Row.

Cohn, Abigail. 1993. Nasalisation in English: phonology or phonetics. *Phonology* 10:43–81.

Crawley, Michael J. 2012. *The R book.* John Wiley & Sons.

Cruttenden, Alan, ed. 2008. *Gimson's pronunciation of English*. London: Hodder, 7th edition.

Davidson, L. 2006. Comparing tongue shapes from ultrasound imaging using smoothing spline analysis of variance. *The Journal of the Acoustical Society of America* 120:407.

Davidson, Lisa. 2012. Ultrasound as a tool for speech research. In *The Oxford handbook of laboratory phonology*, eds. Abigail Cohn, Cecile Fougeron, and Marie K. Huffman. Oxford: Oxford University Press.

Delattre, Pierre. 1965. *Comparing the phonetic features of English, French, German and Spanish : an interim report*. Heidelberg: Groos Verlag.

Docherty, Gerard J, and Paul Foulkes. 1999. Derby and Newcastle: instrumental phonetics and variationist studies. In *Urban voices: accent studies in the British Isles*, eds. Paul Foulkes and Gerard Docherty, 47–71. Arnold London.

Dodsworth, Robin. 2005. Attribute networking: A technique for modeling social perceptions1. *Journal of Sociolinguistics* 9:225–253.

Ellis, Lucy, and William J. Hardcastle. 2002. Categorical and gradient properties of assimilation in alveolar to velar sequences: evidence from EPG and EMA data. *Journal of Phonetics* 30:373–396.

Elphinston, James. 1765. *The principles of the English Language*. James Bettenham, and sold by P. Vaillant; L. Hawes, W. Clarke, and R. Collins; J. Dodsley; and W. Drummond, Edinburgh.

Erker, D. 2012. Of categories and continua: Relating discrete and gradient properties of sociophonetic variation. *University of Pennsylvania Working Papers in Linguistics* 18:3.

Espy-Wilson, Carol Y. 1992. Acoustic measures for linguistic features distinguishing the semivowels /w, j, r, l/ in American English. *The Journal of the Acoustical Society of America* 92:736–757.

Fabricius, Anne. 2000. T-glottalling between stigma and prestige: a sociolinguistic study of modern RP. *Unpublished PhD thesis, Copenhagen Business School* .

Fabricius, Anne. 2002. Ongoing change in modern RP. *English World-Wide* 23:115–136.

Fant, Gunnar. 1960. *Acoustic theory of speech production*. The Hague: Mouton.

Faraway, Julian J. 2004. *Linear models with R*. CRC Press.

Flynn, Nicholas. 2011. Comparing vowel formant normalization procedures. *York Working Papers in Linguistics (Series 2)* 11:1–28.

Fruehwald, Josef. 2013. The Philadelphia short-A system. Available online at: http://jofrhwld.github.io/assets/philly_short_a2.svg.

Gaitenby, Jane. 1965. The elastic word. *Haskins Laboratories Status Report on Speech Research* 2:3–1.

Garrett, Andrew, and Juliette Blevins. 2009. Analogical morphophonology. In *The nature of the word: Essays in honor of Paul Kiparsky*, 527–545. Cambridge, MA: Cambridge, MA: The MIT Press.

Gelman, Andrew, and Jennifer Hill. 2006. *Data analysis using regression and multilevel/hierarchical models.* Cambridge University Press.

Gibb, Charlotte. 2014. Generational change in /l/ vocalisation patterns: a mechanism of fool-full merger in Essex. BA dissertation, University of Manchester.

Gick, B., and S. Rahemtulla. 2004. Recent developments in quantitative analysis of ultrasound tongue data. In *Meeting of the Second Ultrasound Roundtable, Vancouver, BC*.

Gick, Bryan. 1999. A gesture-based account of intrusive consonants in English. *Phonology* 16:29–54.

Gick, Bryan. 2002. The American intrusive l. *American Speech* 77:167–183.

Gick, Bryan. 2003. Articulatory correlates of ambisyllabicity in English glides and liquids. In *Phonetic interpretation: papers in laboratory phonology vi*, eds. John Local, Richard Ogden, and Rosalind Temple. Cambridge: Cambridge University Press.

Gick, Bryan, Fiona Campbell, Sunyoung Oh, and Linda Tamburri-Watt. 2006. Toward universals in the gestural organization of syllables: A cross-linguistic study of liquids. *Journal of Phonetics* 34:49 – 72.

Gick, Bryan, and Ian Wilson. 2006. Excrescent schwa and vowel laxing: Cross-linguistic responses to conflicting articulatory targets. *Laboratory phonology* 8:635–659.

Giegerich, Heinz J. 1992. *English phonology: An introduction.* Cambridge University Press.

Gilbert, Richard. 2013. Diphthong allophony and durational enhancement: reconciling Canadian Raising, Southern monophthongisation and pre-fortis clipping. BA dissertation, University of Manchester.

Giles, S. B., and K. L. Moll. 1975. Cinefluorographic study of selected allophones of English /l/. *Phonetica* 31:206–227.

Gould, Stephen Jay. 1996. *The mismeasure of man.* WW Norton & Company.

Hall-Lew, Lauren. 2011. The completion of a sound change in California English. In *17th International Congress of Phonetic Sciences, Hong Kong.*.

Hall-Lew, Lauren, and Sonya Fix. 2012. Perceptual coding reliability of (l)-vocalisation in casual speech data. *Lingua* 122:794–809.

Halle, Morris, and K. P. Mohanan. 1985. Segmental phonology of Modern English. *Linguistic Inquiry* 16:57–116.

Hamann, Silke. 2014. Phonological change. In *Handbook of historical linguistics*, eds. Claire Bowern and Bethwyn Evans. Routledge.

Hardcastle, William, and William Barry. 1989. Articulatory and perceptual factors in /l/ vocalisations in English. *Journal of the International Phonetic Association* 15:3–17.

Harrington, Jonathan, Felicitas Kleber, and Ulrich Reubold. 2008. Compensation for coarticulation, /u/-fronting, and sound change in standard southern British: An acoustic and perceptual study. *The Journal of the Acoustical Society of America* 123:2825–2835.

Harris, James W. 1984. Autosegmental phonology, Lexical Phonology, and Spanish nasals. In *Language sound structure*, eds. Mark Aronoff and Richard Oehrle, 67–82. Cambridge: MIT Press.

Harris, John. 1994. *English sound structure*. Oxford: Blackwell.

Harris, John. 2003. Release the captive coda: the foot as a domain of phonetic interpretation. In *Phonetic interpretation: papers in laboratory phonology vi*, eds. John Local, Richard Ogden, and Rosalind Temple, 103–129. Cambridge: Cambridge University Press.

Harris, John. 2006. Wide-domain r-effects in English. *UCL Working Papers in Linguistics* 18.

Harris, John, and Jonathan Kaye. 1990. A tale of two cities: London glottalling and New York City tapping. *The Linguistic Review* 7:251–274.

Harshman, Richard, Peter Ladefoged, and Louis Goldstein. 1977. Factor analysis of tongue shapes. *The Journal of the Acoustical Society of America* 62:693.

Hartigan, John A, and PM Hartigan. 1985. The dip test of unimodality. *The Annals of Statistics* 70–84.

Hartmann, Dieter, and Sabine Zerbian. 2010. Rhoticity in Black South African English–a sociolinguistic study. *Southern African linguistics and applied language studies* 27:135–148.

Hawkins, S. 2004. Influence of syllable-coda voicing on the acoustic properties of syllable-onset /l/ in English. *Journal of Phonetics* 32:199 – 231.

Hawkins, Sarah, and Noël Nguyen. 2004. Influence of syllable-coda voicing on the acoustic properties of syllable-onset /l/ in English. *Journal of Phonetics* 32:199–231.

Hay, Jennifer, and Margaret MacLagan. 2010. Social and phonetic conditioners on the frequency and degree of intrusive /r/ in New Zealand English. In *A reader in sociophonetics*, eds. Dennis Preston and Nancy Niedzielski. Berlin: Mouton de Gruyter.

Hayes, Bruce. 2000. Gradient well-formedness in Optimality Theory. In *Optimality Theory: phonology, syntax, and acquisition*, eds. Joost Dekkers, Frank van der Leeuw, and Jeroen van de Weijer, 88–120. Oxford: Oxford University Press.

Hedrick, Wayne R, David L Hykes, and Dale E Starchman. 1995. *Ultrasound physics and instrumentation*. Mosby St. Louis.

Heselwood, Barry, and Louise McChrystal. 2000. Gender, accent features and voicing in Panjabi-English bilingual children. *Leeds working papers in Linguistics and Phonetics* 8:45–70.

Hickey, Raymond. 2005. Dublin English: current changes and their motivation. In *Urban voices: Accent studies in the British Isles*, eds. Paul Foulkes and Gerard Docherty. London: Arnold.

van Hofwegen, Janneke. 2010. Apparent-time evolution of /l/ in one African American community. *Language Variation and Change* 22:373–396.

Holst, Tara, and Francis Nolan. 1995. The influence of syntactic structure on [s] to [ʃ] assimilation. *Papers in Laboratory Phonology, IV* 315–3333.

Honeybone, Patrick. 2001. Lenition inhibition in Liverpool English. *English Language and Linguistics* 5:213–249.

Honeybone, Patrick. 2004. Influences in Liverpool English koineisation. In *Symposium on the Influence of the Languages of Ireland and Scotland on Linguistic Varieties in Northern England*. University of Aberdeen.

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

Honorof, Douglas N, Jeffrey Weihing, and Carol A Fowler. 2011. Articulatory events are imitated under rapid shadowing. *Journal of phonetics* 39:18–38.

Hoole, Philip. 1999. On the lingual organization of the German vowel system. *The Journal of the Acoustical Society of America* 106:1020–1032.

Horvath, Barbara M, and Ronald J Horvath. 2002. The geolinguistics of /l/-vocalization in Australia and New Zealand. *Journal of Sociolinguistics* 6:319–346.

Hudson, Richard, and AF Holloway. 1977. *Variation in London English.* Final Report to the Social Sciences Research Council. Department of Phonetics & Linguistics, University College London.

Huffman, M. K. 1997. Phonetic variation in intervocalic onset /l/s in English. *Journal of Phonetics* 25:115 – 141.

Hughes, Arthur, Peter Trudgill, and Dominic Watt. 2012. *English accents and dialects: an introduction to social and regional varieties of English in the British Isles.* Routledge.

Huwaldt, JA. 2010. Plot digitizer. Available online at plotdigitizer.sourceforge.net.

Hwang, Sangjin, Sang-Im Lee, and Lisa Davidson. 2010. Morphological effects on the articulation of English intervocalic /l/. Paper presented at NELS 41, University of Pennsylvania.

Jackson, Michel TT. 1988. Analysis of tongue positions: Language-specific and cross-linguistic models. *The Journal of the Acoustical Society of America* 84:124.

Jakobson, Roman, Gunnar Fant, and Morris Halle. 1952. *Preliminaries to speech analysis.* Cambridge, MA: MIT Press.

Jansen, Sandra. 2012. Language variation and change in the Cumbrian City Dialect of Carlisle. Doctoral Dissertation, University of Duisburg-Essen.

Jensen, John T. 1993. *English phonology*. Amsterdam: John Benjamins.

Jensen, John T. 2000. Against ambisyllabicity. *Phonology* 17:187–235.

Johnson, Keith. 2011. *Quantitative methods in linguistics.* John Wiley & Sons.

Johnson, Wyn, and David Britain. 2007. L-vocalisation as a natural phenomenon: Explorations in sociophonology. *Language Sciences* 29:294–315.

Jones, Daniel. 1966. *The pronunciation of English.* Cambridge University Press, 4th edition.

Kahn, Daniel. 1976. Syllable-based generalizations in English phonology. Doctoral dissertation, MIT.

Kawahara, Shigeto. 2011. Experimental approaches in theoretical phonology. In *The Blackwell companion to phonology*, eds. Marc van Oostendorp, Colin J. Ewen, Elizabeth Hume, and Keren Rice, Vol. 4. Malden, MA: Wiley-Blackwell.

Kelly, J., and J.L. Local. 1986. Long-domain resonance patterns in English. In *Institute of Electronic Engineers: Proceedings of the International Conference on Speech Input/Output*, 304–308.

Kenstowicz, Michael. 1996. Base-identity and uniform exponence: alternatives to cyclicity. In *Currents trends in phonology: models and methods*, eds. Jacques Durand and Bernard Laks, Vol. 1, 363–393. Salford: European Studies Research Institute, University of Salford.

Kerswill, Paul, and Susan Wright. 1990. The validity of phonetic transcription: Limitations of a sociolinguistic research tool. *Language Variation and Change* 2:255–275.

Kerswill, Paul E. 1985. A sociophonetic study of connected speech processes in Cambridge English: an outline and some results. *Cambridge papers in phonetics and experimental linguistics* 4:25–49.

Kiparsky, Paul. 1979. Metrical structure assignment is cyclic. *Linguistic Inquiry* 10:421–441.

Kiparsky, Paul. 1988. Phonological change. In *Linguistics: the Cambridge survey*, ed. Frederick J. Newmeyer, Vol. 1, Linguistic theory: foundations, 363–415. Cambridge: Cambridge University Press.

Kiparsky, Paul. 2006. The amphichronic program vs. evolutionary phonology. *Theoretical Linguistics* 32:217–236.

Kirkham, Sam. 2014. Sociophonetic variation in Sheffield English liquids. Paper presented at British Association of Academic Phoneticians Conference, University of Oxford, UK, 7-9 April 2014.

Knowles, Gerald O. 1973. Scouse: the urban dialect of Liverpool. Doctoral Dissertation, University of Leeds.

Labov, William. 1972. *Sociolinguistic patterns.* Oxford: Blackwell.

Labov, William. 1994. *Principles of linguistic change: internal factors.* Oxford, U.K.: Blackwell.

Ladefoged, Peter. 2001. *Vowels and consonants: An introduction to the sounds of languages.* Malden, MA & Oxford: Blackwell.

Ladefoged, Peter, and Keith Johnson. 2014. *A course in phonetics.* Cengage learning.

Ladefoged, Peter, and Ian Maddieson. 1996. *The sounds of the world's languages.* Oxford: Blackwell.

Lawson, Eleanor, James M Scobbie, and Jane Stuart-Smith. 2011. The social stratification of tongue shape for postvocalic /r/ in Scottish English. *Journal of Sociolinguistics* 15:256–268.

Lawson, Eleanor, Jane Stuart-Smith, and James M. Scobbie. 2008. Articulatory insights into language variation and change: Preliminary findings from an ultrasound study of derhoticization in Scottish English. *University of Pennsylvania Working Papers in Linguistics* 14:1524–1549.

Lawson, Eleanor, Jane Stuart-Smith, James M. Scobbie, Malcah Yaeger-Dror, and Margaret MacLagan. 2010. Liquids. In *Sociophonetics: A student's guide*, eds. Malcah Yaeger-Dror and Marianna Di Paolo, 72–86. London: Routledge.

Lee-Kim, Sang-Im, Lisa Davidson, and Sangjin Hwang. 2013. Morphological effects on the darkness of English intervocalic /l/. *Laboratory Phonology* 4:475–511.

Lehiste, Ilse. 1964. Acoustical characteristics of selected English consonants. *Indiana research centre in anthropology, folklore and linguistics* 34:10–50.

Lehiste, Ilse. 1980. Phonetic manifestation of syntactic structure in English. *Annual Bulletin of the Research Institue of Logopedics and Phoniatrics.* 14:1–27.

Lehiste, Ilse, Joseph P Olive, and Lynn A Streeter. 1976. Role of duration in disambiguating syntactically ambiguous sentences. *The Journal of the Acoustical Society of America* 60:1199–1202.

Lewis-Beck, Michael S. 1980. *Applied regression: An introduction*, Vol. 22. Sage.

Li, M., C. Kambhamettu, and M. Stone. 2005. Automatic contour tracking in ultrasound images. *Clinical Linguistics & Phonetics* 19:545–554.

Lignos, Constantine. 2012. Modeling words in the mind. Doctoral Dissertation, University of Pennsylvania.

Lin, Susan, Patrice Speeter Beddor, and Andries W. Coetzee. 2014. Gestural reduction, lexical frequency, and sound change: a study of post-vocalic /l/. *Laboratory Phonology* 5:9–26.

Lynch, Erika. 2009. The role of syllable structure in connected speech: an ultrasound study of Anglo-English /l/ allophony. BA thesis, Queen Margaret University College, Edinburgh.

Maechler, Martin. 2013. *diptest: Hartigan's dip test statistic for unimodality - corrected code.* URL http://CRAN.R-project.org/package=diptest, r package version 0.75-5.

Maeda, Shinji. 1990. Compensatory articulation during speech: Evidence from the analysis and synthesis of vocal-tract shapes using an articulatory model. In *Speech production and speech modelling*, 131–149. Springer.

Mathisen, Anne. 1999. Sandwell, West Midlands: ambiguous perspectives on gender patterns and models of change. In *Urban voices: Accent studies in the British Isles*, eds. Paul Foulkes and Gerald Docherty, Vol. 107–123. London: Arnold.

Maye, Jessica, Janet F Werker, and LouAnn Gerken. 2002. Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition* 82:B101–B111.

McCafferty, Kevin. 1999. Derry: between Ulster and local speech–class, ethnicity and language change. In *Urban voices: accent studies in the British Isles*, eds. Paul Foulkes and Gerard Docherty, 246–264. London: Arnold.

McCarthy, John. 1993. Containment, consistency, and alignment. In *Rutgers Optimality Workshop I*. New Brunswick, NJ.

Mees, Inger M, and Beverley Collins. 1999. Cardiff: a real-time study of glottalisation. In *Urban voices: Accent studies in the British Isles*, eds. Paul Foulkes and Gerard Docherty, 185–202. London: Arnold.

Milroy, James, Lesley Milroy, Sue Hartley, and David Walshaw. 1994. Glottal stops and Tyneside glottalization: Competing patterns of variation and change in British English. *Language Variation and Change* 6:327–357.

Minkova, Donka. 2003. *Alliteration and sound change in early English*. Cambridge Studies in Linguistics 101. Cambridge: Cambridge University Press.

Mohanan, K. P. 1982. Lexical Phonology. Doctoral dissertation, MIT.

Moisik, Scott R. 2013. The epilarynx in speech. Doctoral Dissertation, University of Victoria.

Moisik, Scott R. 2014. A study of laryngeal gestures in mandarin citation tones using simultaneous laryngoscopy and laryngeal ultrasound (SLLUS). *Journal of the International Phonetic Association* 44:21–58.

Morris, Jonathan. 2013. Sociolinguistic variation and regional minority language bilingualism: an investigation of Welsh-English bilinguals in North Wales. Doctoral Dissertation, University of Manchester.

Müller, Daniela. 2011. Developments of the lateral in Occitan dialects and their Romance and cross-linguistic context. Doctoral Dissertation, Université Toulouse le Mirail-Toulouse II.

Mullooly, Richard. 2004. An electromagnetic articulograph study of alternating [ɹ] and the effects of stress on rhotic consonants. Doctoral Dissertation, Queen Margaret University College, Edinburgh.

Nance, Claire. 2013. Phonetic variation, sound change, and identity in Scottish Gaelic. Doctoral Dissertation, University of Glasgow.

Narayanan, Shrikanth S, Abeer A Alwan, and Katherine Haker. 1997. Toward articulatory-acoustic models for liquid approximants based on MRI and EPG data. Part i. The laterals. *The Journal of the Acoustical Society of America* 101:1064–1077.

Naseem, Tahira. 2002. Acoustic analysis of Punjabi phonemes l and n. *CRULP Annual Student Report 2002-3* 3.

Newton, David E. 1993. Clear and dark: Resonance and its perception in English. Master's thesis, University of York.

Newton, David E. 1994. Resonance in English: principles and investigations. In *Proceedings of the Edinburgh Linguistics Department Conference*, 167–172.

Nolan, F., T. Holst, and B. Kühnert. 1996. Modelling [s] to [š] accommodation in English. *JPh* 24:113–137.

Nolan, Francis, ed. 1983. *The phonetic bases of speaker recognition.* Cambridge: Cambridge University Press.

Nolan, Francis. 1995. Part 1: Introduction to the IPA. *Journal of the International Phonetic Association* 25:3–33.

Nolan, Francis, and Paul E Kerswill. 1990. The description of connected speech processes. In *Studies in the pronunciation of English: a commemorative volume in honour of A.C. Gimson*, ed. Susan Ramsaran, 295–316. London and New York: Routledge.

Ohala, J. J. 1990. There is no interface between phonetics and phonology. a personal view. *JPh* 18:153–171.

Ohala, John J. 1981. The listener as a source of sound change. In *Papers from the parasession on language and behavior*, eds. Carrie Masek, Roberta Hendrick, and Mary Frances Miller, 178–203. Chicago: Chicago Linguistics Society.

Olive, Joseph P., Alice Greenwood, and John Coleman. 1993. *Acoustics of American English speech: a dynamic approach.* New York: Springer Verlag.

Oller, D Kimbrough. 1973. The effect of position in utterance on speech segment duration in English. *The Journal of the Acoustical Society of America* 54:1235–1247.

Oller, D Kimbrough, and Bruce L Smith. 1977. Effect of final-syllable position on vowel duration in infant babbling. *The Journal of the Acoustical Society of America* 62:994–997.

Orton, H., S. Sanderson, and J. Widdowson, eds. 1978. *The linguistic atlas of England*. London: Croom Helm.

Orton, Harold. 1933. *The phonology of a south Durham dialect: descriptive, historical, and comparative.* London: Keegan Paul, Trench and Trübner.

Pederson, Lee. 2001. Dialects. In *Cambridge history of the English Language: English in North America*, ed. J. Algeo, Vol. 6. Cambridge: Cambridge University Press.

Pouplier, Marianne, and Philip Hoole. 2013. Comparing principal component analysis of ultrasound images with contour analyses in a study of tongue body control during german coronals. Paper presented at Ultrafest 6, QMU, Edinburgh.

Proctor, Michael. 2009. Gestural characterization of a phonological class: the liquids. Doctoral Dissertation, Yale University.

Przedlacka, Joanna. 2001. Estuary English and RP: Some recent findings. *Studia Anglica Posnaniensia* 36:35–50.

R Core Team. 2014. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

Ramsammy, Michael. 2011. The realisation of coda nasals in Spanish. Doctoral Dissertation, The University of Manchester.

Ramsammy, Michael. forthcoming. The life cycle of phonological processes: accounting for dialectal microtypologies. *Language and Linguistics Compass* .

Recasens, Daniel. 2004. Darkness in [l] as a scalar phonetic property: implications for phonology and articulatory control. *Clinical linguistics & phonetics* 18:593–603.

Recasens, Daniel, and Aina Espinosa. 2005. Articulatory, positional and coarticulatory characteristics for clear /l/ and dark /l/: evidence from two Catalan dialects. *Journal of the International Phonetic Association* 35:1–25.

Recasens, Daniel, and Aina Espinosa. 2009. An articulatory investigation of lingual coarticulatory resistance and aggressiveness for consonants and vowels in Catalan. *The Journal of the Acoustical Society of America* 125:2288–2298.

Recasens, Daniel, Jordi Fontdevila, and Maria Dolors Pallarés. 1996. Linguopalatal coarticulation and alveolar-palatal correlations for velarized and non-velarized /l/. *Journal of Phonetics* 24:165–185.

Roach, Peter, Daniel Jones, and James Hartman. 2006. *Cambridge English pronouncing dictionary*. Cambridge University Press,, 16th edition.

Rosen, Kristin M. 2005. Analysis of speech segment duration with the lognormal distribution: A basis for unification and comparison. *Journal of Phonetics* 33:411–426.

Rubach, Jerzy. 1996. Shortening and ambisyllabicity in English. *Phonology* 13:197–237.

Schilling, Mark F, Ann E Watkins, and William Watkins. 2002. Is human height bimodal? *The American Statistician* 56:223–229.

Schuchardt, Hugo. 1885. *Über die lautgesetze. gegen die junggrammatiker*. Berlin.

Scobbie, James M. 2005. The phonetics phonology overlap. *QMU Speech Science Research Centre Working Papers, WP-1* .

Scobbie, James M. 2007. Interface and overlap in phonetics and phonology. In *The oxford handbook of linguistic interfaces*, eds. Gillian Ramchand and Charles Reiss. Oxford University Press.

Scobbie, James M. 2009. Ultrasound-based tongue root imaging and measurement. Oral paper at the Workshop on Pharyngeals and Pharyngealisations, Newcastle Upon Tyne.

Scobbie, James M., Lawson E., S. Cowen, J. Cleland, and Alan A. Wrench. 2011. A common co-ordinate system for mid-sagittal articulatory measurement. Proceedings of Interspeech, Florence, 2011.

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

Scobbie, James M, Alan A Wrench, and Marietta van der Linden. 2008. Head-probe stabilisation in ultrasound tongue imaging using a headset to permit natural head movement. In *Proceedings of the 8th International Seminar on Speech Production*, 373–376.

Scobbie, J.M., and M. Pouplier. 2010. The role of syllable structure in external sandhi: An EPG study of vocalisation and retraction in word-final English /l/. *Journal of Phonetics* 38:240–259.

Sharma, Devyani. 2012. Stylistic activation in ethnolinguistic repertoires. *University of Pennsylvania Working Papers in Linguistics* 18:15.

Shorrocks, Graham. 1999. *A grammar of the dialect of the Bolton area.* Frankfurt am Main: Peter Lang.

Slud, Eric, Maureen Stone, Paul J Smith, and Moise Goldstein Jr. 2005. Principal components representation of the two-dimensional coronal tongue surface. *Phonetica* 59:108–133.

Speitel, Hans-Henning. 1983. *A sociolinguistic investigation of Edinburgh speech.* Economic and Social Research Council end of grant report.

Sproat, Richard, and Osamu Fujimura. 1993. Allophonic variation in English /l/ and its implications for phonetic implementation. *Journal of Phonetics* 21:291–311.

Stoddart, Jana, Clive Upton, and J. Widdowson. 1999. Sheffield dialect in the 1990s: Revisiting the concept of NORMs. In *Urban voices: Accent studies in the British Isles*, eds. Paul Foulkes and Gerald Docherty, 72–89. London: Arnold.

Stone, Maureen. 1990. A three-dimensional model of tongue movement based on ultrasound and x-ray microbeam data. *The Journal of the Acoustical Society of America* 87:2207–2217.

Stone, Maureen. 2005. A guide to analysing tongue motion from ultrasound images. *Clinical linguistics & phonetics* 19:455–501.

Streeter, Lynn A. 1978. Acoustic determinants of phrase boundary perception. *The Journal of the Acoustical Society of America* 64:1582–1592.

Strycharczuk, Patrycja. 2012. Sonorant transparency and the complexity of voicing in Polish. *Journal of Phonetics* 40:655–671.

Strycharczuk, Patrycja, and Ellen Simon. 2013. Obstruent voicing before sonorants. the case of west-flemish. *Natural Language & Linguistic Theory* 31:563–588.

Stuart-Smith, Jane. 1999. Glasgow: Accent and voice quality. In *Urban voices: Accent studies in the British Isles*, eds. Paul Foulkes and Gerard Docherty. London: Arnold.

Stuart-Smith, Jane, Lawson Eleanor, James M. Scobbie, Alice E. Turk, Mercedes Durham, Claire Timmins, David Beavan, William Simon Barras, and Flora Edmonds. 2013. Seeing speech: An online ultrasound tongue imaging resource for phonetics, linguistics, and speech therapy teaching at scottish universities. Carnegie Trust, University of Glasgow. URL http://www.seeingspeech.arts.gla.ac.uk/uti/.

Stuart-Smith, Jane, Claire Timmins, and Farhana Alam. 2011. Hybridity and ethnic accents. *Selected Papers ICLaVE Amsterdam: Benjamins* 5:43–57.

Syrdal, Ann K, and Hundrai S Gopal. 1986. A perceptual model of vowel recognition based on the auditory representation of American English vowels. *The Journal of the Acoustical Society of America* 79:1086–1100.

Tollfree, L.F. 1999. South-east London English: discrete versus continuous modelling of consonantal reduction. In *Urban voices: Accent studies in the British Isles*, eds. Paul Foulkes and Gerald Docherty, 163–184. London: Arnold.

Trask, R. L. 1996. *A dictionary of phonetics and phonology*. London and New York: Routledge.

Traunmüller, Hartmut. 1990. Analytical expressions for the tonotopic sensory scale. *The Journal of the Acoustical Society of America* 88:97–100.

Trudgill, Peter. 1999. Norwich: endogenous and exogenous linguistic change. In *Urban voices: Accent studies in the British Isles*, eds. Paul Foulkes and Gerard Docherty, 124–140. London: Arnold.

Turk, Alice E, and Stefanie Shattuck-Hufnagel. 2007. Multiple targets of phrase-final lengthening in american english words. *Journal of Phonetics* 35:445–472.

Turton, Danielle. 2010. The sociophonology of linking and intrusive /r/ in Manchester English. Masters thesis, University of Manchester.

Turton, Danielle. 2012. The darkening of English /l/: A Stochastic Stratal OT analysis. Unpublished manuscript, University of Manchester. Available at http://ling.auf.net/lingbuzz/001524.

Turton, Danielle, and Maciej Baranowski. 2014. t[ʉː] c[ʉːɫ] for sch[ʉːɫ]: the interaction of /u/-fronting and /l/-darkening in Manchester. Paper presented at New Ways of Analysing Variation (NWAV) 2014 conference. 23rd-25th Oct 2014. Chicago.

Turton, Danielle, and Michael Ramsammy. 2012. /ɪ, ə/-lowering in manchest[ʌ]: Contextual patterns of gradient variability. Paper given at the 20th Manchester Phonology Meeting. Manchester, 24th-26th May 2012.

Uffmann, Christian. 2010. The non-trivialness of segmental representations. Paper presented at the Old World Conference in Phonology (OCP) 7, Nice.

Umeda, N. 1977. Consonant duration in American English. *The Journal of the Acoustical Society of America* 61:846–858.

Vaissière, Jacqueline. 1983. Language-independent prosodic features. In *Prosody: Models and measurements*, eds. Anne Cutler and D. Robert Ladd, 53–66. Berlin: Springer Verlag.

Venneman, Theo. 1972. Rule inversion. *Lingua* 29:209–242.

Wahba, G. 1990. *Spline models for observational data*, Vol. 59. Society for Industrial Mathematics.

Wallace, Kate Elizabeth. 2007. Social variation in the English of the Southampton area. Doctoral Dissertation, University of Leeds.

Watson, Kevin. 2007. Liverpool English. *Journal of the International Phonetic Association* 37:351–360.

Watt, Dominic, and Lesley Milroy. 1999. Variation in three Tyneside vowels: is this dialect levelling?'. In *Urban voices: Accent studies in the British Isles*, eds. Paul Foulkes and Gerard Docherty, 25–46. London: Arnold.

Weinreich, Uriel, William Labov, and Marvin Herzog. 1968. Empirical foundations for a theory of language change .

Wells, J. C. 1982. *Accents of English.* Cambridge: Cambridge University Press.

Wells, J. C. 1990. Syllabification and allophony. In *Studies in the pronunciation of english: a commemorative volume in honour of a.c. gimson*, ed. Susan Ramsaran, 76–86. London: Routledge.

Wells, John C. 2008. *Longman pronunciation dictionary.* Harlow: Longman, 3rd edition.

Wickham, Hadley. 2009. *ggplot2: elegant graphics for data analysis.* Springer Publishing Company, Incorporated.

Williams, Ann, and Paul Kerswill. 1999. Dialect levelling: change and continuity in Milton Keynes, Reading and Hull. In *Urban voices: Accent studies in the British Isles*, eds. Paul Foulkes and Gerald Docherty, 141–162. London: Arnold.

Wrench, A.A. 2007. Articulate Assistant Advanced user guide: Version 2.07. *Edinburgh: Articulate Instruments Ltd* .

Wrench, Alan A., and James M. Scobbie. 2003. Categorising vocalisation of English /l/ using EPG, EMA and ultrasound. In *Proceedings of the 6th International Seminar on Speech Production*, 314–319.

Wrench, Alan A., and James M. Scobbie. 2008. High-speed cineloop ultrasound vs. video ultrasound tongue imaging: Comparison of front and back lingual gesture location and relative timing. In *Proceedings of the Eighth International Seminar on Speech Production (ISSP).*

Wright, Joseph. 1905. *The English dialect grammar.* Oxford: Clarendon Press.

Wright, Susan. 1987. The interaction of sociolinguistic and phonetically-condtioned CSPs in Cambridge English: auditory and electropalatographic evidence. *Cambridge papers in phonetics and experimental linguistics* 5.

Wright, Susan. 1989. The effects of style and speaking rate on /l/-vocalisation in Local Cambridge English. *York Papers in Linguistics* 13:355–365.

Wright, Susan, and Paul Kerswill. 1989. Electropalatography in the analysis of connected speech processes. *Clinical linguistics & phonetics* 3:49–57.

Yuan, Jiahong, and Mark Liberman. 2009. Investigating /l/ variation in English through forced alignment. In *INTERSPEECH 2009: 10th Annual Conference of the International Speech Communication Association, Brighton, UK, 6-10 September 2009.*, 2215–2218.

Yuan, Jiahong, and Mark Liberman. 2011. /l/ variation in American English: A corpus approach. *Journal of Speech Sciences* 1:35–46.

Zsiga, Elizabeth C. 1995. An acoustic and electropalatographic study of lexical and postlexical palatalization in American English. In *Phonology and phonetic evidence: papers in laboratory phonology IV*, eds. Bruce Connell and Amalia Arvaniti, 282–302. Cambridge: Cambridge University Press.

Zwicky, Arnold. 1972. Note on a phonological hierarchy in English. *Linguistic change and generative theory* 275:301.

Appendix A

Consent forms

MANCHESTER
1824

The University of Manchester

Variation in English Liquids: Ultrasound Experiment
CONSENT FORM

Principle Investigator: Danielle Turton

If you are happy to participate please complete and sign the consent form below

	Please Initial Box
1. I confirm that I have read the attached information sheet on the above project and have had the opportunity to consider the information and ask questions and had these answered satisfactorily.	[redacted]
2. I understand that my participation in the study is voluntary and that I am free to withdraw at any time without giving a reason and without detriment to any treatment/service	[redacted]
3. I understand that the experiment will be audio/video-recorded	[redacted]
4. I have received £15 for participating.	[redacted]

I agree to take part in the above project

[redacted]	10/12/2013	[redacted]
Name of participant	Date	Signature
______________________	______________	______________________

Appendix B

Experimental stimuli

Experiment 1 stimuli and number of tokens per participant

		Number of tokens for:				
Context	Sentence	RP	Mcr WC	Middlesbrough	American Eng	Essex
word-initial	leap	3	5	5	5	5
stem-medial posttonic	helium	5	5	5	5	5
stem-final presuffixal	Is it healing?	5	5	5	5	5
word-final prevocalic	Can you heal it?	5	5	5	5	5
utterance final	heal	5	5	5	5	5
word-final pre-/h/	Can you peel heals?	0	0	0	5	5
word-final pre-/b/	Can you heal beasts?	0	0	5	5	5
Total number of splines		23	25	30	35	35

Experiment 2 stimuli and number of tokens per participant

Context	*Sentence*	*Number of tokens for:*							
		RP	London F	London M	Mcr WC	Mcr MC	Newcastle	Belfast	Liverpool
word-initial	Mr. Bee Linoleum's from Manchester	5	3	4	5	5	0	1	5
	Mr. B. Leoni's from Colchester	5	6	6	0	0	5	5	6
	Mr B. Leoni's fat belly*	0	0	0	5	0	0	0	0
stem-medial pretonic	Don't believe the hype	5	5	4	4	5	5	5	3
	Don't believe the hive	5	5	5	5	5	5	1	4
suffix-initial	The swarm moves freely	5	5	0	5	6	5	5	5
	The hive is beeless	5	5	3	5	6	0	4	4
stem-medial posttonic	The helix is curved	5	5	5	5	6	5	6	3
	And Felix is served	5	5	5	5	5	5	5	5
stem-final presuffixal	He's peeling grapes	5	5	5	5	4	5	5	5
	She's healing apes	5	5	5	5	5	5	3	5
compound boundary	The peel-instrument's useful	5	5	5	5	5	5	5	4
	The seal-index is accurate	5	5	5	5	5	5	5	5
word-final phrase-medial	I sent Neil interesting emails	5	5	5	5	5	5	6	5
	I sent Neil innocuous pleas	5	5	5	5	5	5	5	5
phrase-final	Seal, enjoy the ride	5	5	5	5	5	5	3	6
	Neil, ingest the pie	5	5	5	5	5	5	5	5
utterance final	Mr Peale	5	5	5	5	4	5	5	4
	Mr Neil	5	5	5	4	4	4	5	4
word-final pre-consonantal	Mr. Peale Banana's from Manchester	5	4	5	5	5	5	5	5
	Mr. Neil Banana's from Manchester	5	3	5	5	6	5	5	5
Total number of splines		100	96	97	98	96	89	89	93

Distracters: Experiment 1	
Who are you fooling?	Who are you drawing?
trilby	The war is nigh
The heel is high	The fool is right
roar	pouring
pool	poo
hooligan	Can you draw it?
boot	pet
who?	Pete
pat	hoot
the law is wrong	pit
pot	pull

Distracters: Experiment 2	
The fool is happy	when fooling about
Happy in Paris	Ship
Street	Boot
boo!	spook
Draw in between the lines	He's drawing Paul
Rich man	The holly and the ivy
Swimming pool	
Belfast only	
The killer has a gun	A pillar of the community
It really doesn't matter	The husband is much fatter
He owes me a tenner	Pavarotti is a tenor

Appendix C

Tukey HSD tests

Experiment 1

Acoustics

	diff	lwr	upr	p.adj
helix-leap	537.15	-78.32	1152.63	0.105
healing-leap	493.23	-122.25	1108.71	0.154
heal it-leap	201.86	-413.62	817.34	0.856
heal-leap	-1115.92	-1731.40	-500.44	0.000
healing-helix	-43.92	-576.94	489.10	0.999
heal it-helix	-335.29	-868.31	197.73	0.351
heal-helix	-1653.07	-2186.09	-1120.05	0.000
heal it-healing	-291.37	-824.39	241.65	0.485
heal-healing	-1609.15	-2142.17	-1076.13	0.000
heal-heal it	-1317.78	-1850.80	-784.76	0.000

Table C.1: Tukey HSD comparisons of F2-F1 values for RP speaker in Experiment 1

	diff	lwr	upr	p.adj
helix-leap	-179.10	-399.68	41.48	0.148
healing-leap	-261.86	-482.44	-41.28	0.015
heal it-leap	-12.93	-233.51	207.65	1.000
heal-leap	-391.71	-612.28	-171.13	0.000
healing-helix	-82.76	-303.34	137.82	0.793
heal it-helix	166.17	-54.41	386.75	0.201
heal-helix	-212.61	-433.19	7.97	0.062
heal it-healing	248.93	28.35	469.51	0.022
heal-healing	-129.85	-350.43	90.73	0.422
heal-heal it	-378.78	-599.35	-158.20	0.000

Table C.2: Tukey HSD comparisons of F2-F1 values for Manchester WC speaker in Experiment 1

	diff	lwr	upr	p.adj
helix-leap	541.79	113.72	969.85	0.008
healing-leap	402.67	-25.40	830.73	0.074
heal it-leap	306.23	-121.84	734.29	0.269
heal-leap	-225.47	-653.54	202.59	0.589
heal_C-leap	-244.78	-672.84	183.28	0.503
healing-helix	-139.12	-567.19	288.94	0.912
heal it-helix	-235.56	-663.62	192.51	0.544
heal-helix	-767.26	-1195.32	-339.19	0.000
heal_C-helix	-786.57	-1214.63	-358.50	0.000
heal it-healing	-96.44	-524.50	331.63	0.981
heal-healing	-628.14	-1056.20	-200.07	0.002
heal_C-healing	-647.45	-1075.51	-219.38	0.001
heal-heal it	-531.70	-959.76	-103.63	0.009
heal_C-heal it	-551.01	-979.07	-122.94	0.007
heal_C-heal	-19.31	-447.37	408.76	1.000

Table C.3: Tukey HSD comparisons of F2-F1 values for Middlesbrough speaker

	diff	lwr	upr	p.adj
helix-leap	-59.07	-171.70	53.55	0.604
healing-leap	-181.98	-294.61	-69.35	0.000
heal it-leap	-236.18	-348.80	-123.55	0.000
heal-leap	-284.87	-397.50	-172.24	0.000
heal_C-leap	-246.78	-346.10	-147.45	0.000
healing-helix	-122.91	-229.09	-16.72	0.016
heal it-helix	-177.10	-283.29	-70.91	0.000
heal-helix	-225.80	-331.99	-119.61	0.000
heal_C-helix	-187.70	-279.66	-95.74	0.000
heal it-healing	-54.19	-160.38	51.99	0.630
heal-healing	-102.89	-209.08	3.30	0.062
heal_C-healing	-64.79	-156.76	27.17	0.291
heal-heal it	-48.70	-154.89	57.49	0.726
heal_C-heal it	-10.60	-102.56	81.36	0.999
heal_C-heal	38.10	-53.86	130.06	0.800

Table C.4: Tukey HSD comparisons of F2-F1 values for American English speaker

	diff	lwr	upr	p.adj
helix-leap	239.78	-189.18	668.75	0.472
healing-leap	-129.38	-558.35	299.59	0.893
heal it-leap	41.62	-387.35	470.58	0.998
heal-leap	-1088.45	-1517.41	-659.48	0.000
healing-helix	-369.17	-798.13	59.80	0.113
heal it-helix	-198.17	-627.14	230.80	0.645
heal-helix	-1328.23	-1757.20	-899.26	0.000
heal it-healing	171.00	-257.97	599.97	0.755
heal-healing	-959.06	-1388.03	-530.10	0.000
heal-heal it	-1130.06	-1559.03	-701.09	0.000

Table C.5: Tukey HSD comparisons of F2-F1 values for Essex speaker

Experiment 2

PC1

	diff	lwr	upr	p.adj
believe-leap	-0.0290	-2.4039	2.3458	1.0000
free-ly-leap	-0.7471	-3.1220	1.6277	0.9902
helix-leap	-1.3942	-3.7691	0.9806	0.6659
peel-ing-leap	0.1669	-2.2080	2.5418	1.0000
peel-index-leap	-6.8608	-9.2357	-4.4859	0.0000
heal#V-leap	-7.4129	-9.7877	-5.0380	0.0000
heal]#V-leap	-8.7020	-11.0769	-6.3271	0.0000
peel-leap	-9.2920	-11.6669	-6.9172	0.0000
peel bananas-leap	-8.3583	-10.7332	-5.9834	0.0000
free-ly-believe	-0.7181	-3.0930	1.6568	0.9927
helix-believe	-1.3652	-3.7401	1.0096	0.6920
peel-ing-believe	0.1959	-2.1790	2.5708	1.0000
peel-index-believe	-6.8318	-9.2066	-4.4569	0.0000
heal#V-believe	-7.3839	-9.7587	-5.0090	0.0000
heal]#V-believe	-8.6730	-11.0479	-6.2981	0.0000
peel-believe	-9.2630	-11.6379	-6.8882	0.0000
peel bananas-believe	-8.3293	-10.7041	-5.9544	0.0000
helix-free-ly	-0.6471	-3.0220	1.7278	0.9966
peel-ing-free-ly	0.9140	-1.4608	3.2889	0.9620
peel-index-free-ly	-6.1137	-8.4885	-3.7388	0.0000
heal#V-free-ly	-6.6658	-9.0406	-4.2909	0.0000
heal]#V-free-ly	-7.9549	-10.3297	-5.5800	0.0000
peel-free-ly	-8.5449	-10.9198	-6.1701	0.0000
peel bananas-free-ly	-7.6112	-9.9860	-5.2363	0.0000
peel-ing-helix	1.5611	-0.8137	3.9360	0.5100
peel-index-helix	-5.4666	-7.8414	-3.0917	0.0000
heal#V-helix	-6.0186	-8.3935	-3.6438	0.0000
heal]#V-helix	-7.3078	-9.6826	-4.9329	0.0000
peel-helix	-7.8978	-10.2727	-5.5229	0.0000
peel bananas-helix	-6.9641	-9.3389	-4.5892	0.0000
peel-index-peel-ing	-7.0277	-9.4025	-4.6528	0.0000
heal#V-peel-ing	-7.5798	-9.9546	-5.2049	0.0000
heal]#V-peel-ing	-8.8689	-11.2438	-6.4940	0.0000
peel-peel-ing	-9.4589	-11.8338	-7.0841	0.0000
peel bananas-peel-ing	-8.5252	-10.9000	-6.1503	0.0000
heal#V-peel-index	-0.5521	-2.9270	1.8228	0.9990
heal]#V-peel-index	-1.8412	-4.2161	0.5336	0.2747
peel-peel-index	-2.4312	-4.8061	-0.0564	0.0403
peel bananas-peel-index	-1.4975	-3.8724	0.8774	0.5698
heal]#V-heal#V	-1.2891	-3.6640	1.0857	0.7571
peel-heal#V	-1.8792	-4.2540	0.4957	0.2487
peel bananas-heal#V	-0.9454	-3.3203	1.4295	0.9531
peel-heal]#V	-0.5900	-2.9649	1.7848	0.9983
peel bananas-heal]#V	0.3437	-2.0311	2.7186	1.0000
peel bananas-peel	0.9338	-1.4411	3.3086	0.9566

Table C.6: Tukey HSD comparison of PC1 values across context for RP

	diff	lwr	upr	p.adj
believe-leap	1.0766	-0.4426	2.5958	0.3986
free-ly-leap	0.3692	-1.1500	1.8884	0.9986
helix-leap	1.4465	-0.0728	2.9657	0.0756
peel-ing-leap	1.3668	-0.1524	2.8860	0.1153
peel-index-leap	-1.9665	-3.4857	-0.4473	0.0025
heal#V-leap	-3.6703	-5.1895	-2.1511	0.0000
heal]#V-leap	-3.8247	-5.3439	-2.3055	0.0000
peel-leap	-4.8852	-6.4044	-3.3660	0.0000
peel bananas-leap	-2.8378	-4.5041	-1.1715	0.0000
free-ly-believe	-0.7074	-2.1861	0.7713	0.8659
helix-believe	0.3698	-1.1089	1.8485	0.9982
peel-ing-believe	0.2902	-1.1885	1.7688	0.9997
peel-index-believe	-3.0431	-4.5218	-1.5644	0.0000
heal#V-believe	-4.7469	-6.2256	-3.2683	0.0000
heal]#V-believe	-4.9013	-6.3800	-3.4227	0.0000
peel-believe	-5.9618	-7.4405	-4.4831	0.0000
peel bananas-believe	-3.9145	-5.5439	-2.2850	0.0000
helix-free-ly	1.0772	-0.4015	2.5559	0.3589
peel-ing-free-ly	0.9976	-0.4811	2.4762	0.4706
peel-index-free-ly	-2.3357	-3.8144	-0.8570	0.0001
heal#V-free-ly	-4.0395	-5.5182	-2.5608	0.0000
heal]#V-free-ly	-4.1939	-5.6726	-2.7153	0.0000
peel-free-ly	-5.2544	-6.7331	-3.7757	0.0000
peel bananas-free-ly	-3.2071	-4.8365	-1.5776	0.0000
peel-ing-helix	-0.0797	-1.5584	1.3990	1.0000
peel-index-helix	-3.4129	-4.8916	-1.9342	0.0000
heal#V-helix	-5.1168	-6.5955	-3.6381	0.0000
heal]#V-helix	-5.2712	-6.7499	-3.7925	0.0000
peel-helix	-6.3317	-7.8104	-4.8530	0.0000
peel bananas-helix	-4.2843	-5.9137	-2.6548	0.0000
peel-index-peel-ing	-3.3332	-4.8119	-1.8546	0.0000
heal#V-peel-ing	-5.0371	-6.5158	-3.5584	0.0000
heal]#V-peel-ing	-5.1915	-6.6702	-3.7128	0.0000
peel-peel-ing	-6.2520	-7.7307	-4.7733	0.0000
peel bananas-peel-ing	-4.2046	-5.8340	-2.5752	0.0000
heal#V-peel-index	-1.7039	-3.1825	-0.2252	0.0115
heal]#V-peel-index	-1.8583	-3.3369	-0.3796	0.0038
peel-peel-index	-2.9187	-4.3974	-1.4401	0.0000
peel bananas-peel-index	-0.8714	-2.5008	0.7581	0.7713
heal]#V-heal#V	-0.1544	-1.6331	1.3243	1.0000
peel-heal#V	-1.2149	-2.6936	0.2638	0.2032
peel bananas-heal#V	0.8325	-0.7969	2.4619	0.8144
peel-heal]#V	-1.0605	-2.5392	0.4182	0.3813
peel bananas-heal]#V	0.9869	-0.6425	2.6163	0.6237
peel bananas-peel	2.0474	0.4179	3.6768	0.0038

Table C.7: Tukey HSD comparison of PC1 values across context for London female

	diff	lwr	upr	p.adj
believe-leap	-4.9902	-8.5566	-1.4238	0.0007
free-ly-leap	-4.9642	-8.6461	-1.2823	0.0013
helix-leap	-3.9876	-7.4589	-0.5163	0.0120
peel-ing-leap	-1.4100	-4.8813	2.0613	0.9464
peel-index-leap	-12.7095	-16.1809	-9.2382	0.0000
heal#V-leap	-15.4100	-18.9765	-11.8436	0.0000
heal]#V-leap	-15.7961	-19.2675	-12.3248	0.0000
peel-leap	-16.0002	-19.4715	-12.5289	0.0000
peel bananas-leap	-14.5457	-18.0170	-11.0744	0.0000
free-ly-believe	0.0260	-3.7457	3.7977	1.0000
helix-believe	1.0026	-2.5639	4.5690	0.9957
peel-ing-believe	3.5802	0.0137	7.1466	0.0483
peel-index-believe	-7.7193	-11.2858	-4.1529	0.0000
heal#V-believe	-10.4198	-14.0789	-6.7608	0.0000
heal]#V-believe	-10.8060	-14.3724	-7.2395	0.0000
peel-believe	-11.0100	-14.5764	-7.4435	0.0000
peel bananas-believe	-9.5555	-13.1219	-5.9890	0.0000
helix-free-ly	0.9766	-2.7053	4.6585	0.9972
peel-ing-free-ly	3.5542	-0.1277	7.2361	0.0676
peel-index-free-ly	-7.7453	-11.4272	-4.0634	0.0000
heal#V-free-ly	-10.4458	-14.2175	-6.6741	0.0000
heal]#V-free-ly	-10.8319	-14.5138	-7.1500	0.0000
peel-free-ly	-11.0359	-14.7178	-7.3541	0.0000
peel bananas-free-ly	-9.5814	-13.2633	-5.8996	0.0000
peel-ing-helix	2.5776	-0.8937	6.0489	0.3320
peel-index-helix	-8.7219	-12.1932	-5.2506	0.0000
heal#V-helix	-11.4224	-14.9888	-7.8560	0.0000
heal]#V-helix	-11.8085	-15.2798	-8.3372	0.0000
peel-helix	-12.0125	-15.4838	-8.5412	0.0000
peel bananas-helix	-10.5580	-14.0293	-7.0867	0.0000
peel-index-peel-ing	-11.2995	-14.7708	-7.8282	0.0000
heal#V-peel-ing	-14.0000	-17.5665	-10.4336	0.0000
heal]#V-peel-ing	-14.3861	-17.8574	-10.9148	0.0000
peel-peel-ing	-14.5902	-18.0615	-11.1188	0.0000
peel bananas-peel-ing	-13.1357	-16.6070	-9.6643	0.0000
heal#V-peel-index	-2.7005	-6.2669	0.8659	0.3052
heal]#V-peel-index	-3.0866	-6.5579	0.3847	0.1250
peel-peel-index	-3.2906	-6.7619	0.1807	0.0782
peel bananas-peel-index	-1.8361	-5.3074	1.6352	0.7822
heal]#V-heal#V	-0.3861	-3.9525	3.1803	1.0000
peel-heal#V	-0.5901	-4.1566	2.9763	0.9999
peel bananas-heal#V	0.8644	-2.7021	4.4308	0.9986
peel-heal]#V	-0.2040	-3.6753	3.2673	1.0000
peel bananas-heal]#V	1.2505	-2.2208	4.7218	0.9750
peel bananas-peel	1.4545	-2.0168	4.9258	0.9355

Table C.8: Tukey HSD comparison of PC1 values across context for London male

	diff	lwr	upr	p.adj
believe-leap	-1.2743	-3.3123	0.7638	0.5809
free-ly-leap	-5.0696	-7.0533	-3.0859	0.0000
helix-leap	-2.3191	-4.3028	-0.3354	0.0097
peel-ing-leap	-0.3108	-2.2945	1.6729	1.0000
peel-index-leap	-2.3166	-4.3003	-0.3329	0.0098
heal#V-leap	-3.8491	-5.8328	-1.8654	0.0000
heal]#V-leap	-3.8401	-5.8238	-1.8564	0.0000
peel-leap	-6.5182	-8.5563	-4.4802	0.0000
peel bananas-leap	-3.9764	-5.9601	-1.9927	0.0000
free-ly-believe	-3.7953	-5.8334	-1.7572	0.0000
helix-believe	-1.0449	-3.0829	0.9932	0.8120
peel-ing-believe	0.9635	-1.0745	3.0016	0.8743
peel-index-believe	-1.0423	-3.0804	0.9958	0.8142
heal#V-believe	-2.5749	-4.6129	-0.5368	0.0035
heal]#V-believe	-2.5658	-4.6039	-0.5278	0.0036
peel-believe	-5.2439	-7.3350	-3.1529	0.0000
peel bananas-believe	-2.7021	-4.7402	-0.6641	0.0017
helix-free-ly	2.7504	0.7667	4.7341	0.0008
peel-ing-free-ly	4.7588	2.7751	6.7425	0.0000
peel-index-free-ly	2.7530	0.7693	4.7367	0.0008
heal#V-free-ly	1.2204	-0.7633	3.2041	0.6033
heal]#V-free-ly	1.2295	-0.7542	3.2132	0.5931
peel-free-ly	-1.4487	-3.4867	0.5894	0.3948
peel bananas-free-ly	1.0932	-0.8905	3.0769	0.7402
peel-ing-helix	2.0084	0.0247	3.9921	0.0447
peel-index-helix	0.0026	-1.9811	1.9863	1.0000
heal#V-helix	-1.5300	-3.5137	0.4537	0.2810
heal]#V-helix	-1.5210	-3.5047	0.4627	0.2889
peel-helix	-4.1991	-6.2371	-2.1610	0.0000
peel bananas-helix	-1.6573	-3.6410	0.3264	0.1851
peel-index-peel-ing	-2.0058	-3.9895	-0.0221	0.0453
heal#V-peel-ing	-3.5384	-5.5221	-1.5547	0.0000
heal]#V-peel-ing	-3.5294	-5.5130	-1.5457	0.0000
peel-peel-ing	-6.2075	-8.2455	-4.1694	0.0000
peel bananas-peel-ing	-3.6656	-5.6493	-1.6819	0.0000
heal#V-peel-index	-1.5326	-3.5163	0.4511	0.2789
heal]#V-peel-index	-1.5235	-3.5072	0.4602	0.2866
peel-peel-index	-4.2016	-6.2397	-2.1636	0.0000
peel bananas-peel-index	-1.6598	-3.6435	0.3239	0.1835
heal]#V-heal#V	0.0090	-1.9747	1.9927	1.0000
peel-heal#V	-2.6691	-4.7071	-0.6310	0.0021
peel bananas-heal#V	-0.1273	-2.1110	1.8564	1.0000
peel-heal]#V	-2.6781	-4.7162	-0.6401	0.0019
peel bananas-heal]#V	-0.1363	-2.1200	1.8474	1.0000
peel bananas-peel	2.5418	0.5038	4.5799	0.0042

Table C.9: Tukey HSD comparison of PC1 values across context for Manchester WC

	diff	lwr	upr	p.adj
believe-leap	-2.3522	-6.3382	1.6338	0.6596
free-ly-leap	-9.3640	-13.2377	-5.4903	0.0000
helix-leap	-9.1435	-13.0687	-5.2184	0.0000
peel-ing-leap	-6.2174	-10.2765	-2.1583	0.0001
peel-index-leap	-12.9155	-16.7892	-9.0418	0.0000
heal#V-leap	-13.1272	-16.9568	-9.2976	0.0000
heal]#V-leap	-12.2538	-16.2398	-8.2678	0.0000
peel-leap	-16.9516	-21.1004	-12.8029	0.0000
peel bananas-leap	-14.2119	-18.1370	-10.2867	0.0000
free-ly-believe	-7.0118	-10.1278	-3.8958	0.0000
helix-believe	-6.7913	-9.9710	-3.6116	0.0000
peel-ing-believe	-3.8652	-7.2089	-0.5214	0.0110
peel-index-believe	-10.5632	-13.6792	-7.4473	0.0000
heal#V-believe	-10.7750	-13.8360	-7.7139	0.0000
heal]#V-believe	-9.9016	-13.1561	-6.6470	0.0000
peel-believe	-14.5994	-18.0513	-11.1474	0.0000
peel bananas-believe	-11.8596	-15.0393	-8.6799	0.0000
helix-free-ly	0.2205	-2.8173	3.2582	1.0000
peel-ing-free-ly	3.1466	-0.0624	6.3556	0.0593
peel-index-free-ly	-3.5515	-6.5225	-0.5805	0.0073
heal#V-free-ly	-3.7632	-6.6765	-0.8499	0.0025
heal]#V-free-ly	-2.8898	-6.0058	0.2262	0.0928
peel-free-ly	-7.5876	-10.9093	-4.2659	0.0000
peel bananas-free-ly	-4.8478	-7.8856	-1.8101	0.0001
peel-ing-helix	2.9261	-0.3448	6.1971	0.1203
peel-index-helix	-3.7719	-6.8097	-0.7342	0.0044
heal#V-helix	-3.9837	-6.9650	-1.0023	0.0015
heal]#V-helix	-3.1103	-6.2900	0.0695	0.0606
peel-helix	-7.8081	-11.1896	-4.4266	0.0000
peel bananas-helix	-5.0683	-8.1714	-1.9652	0.0000
peel-index-peel-ing	-6.6981	-9.9071	-3.4890	0.0000
heal#V-peel-ing	-6.9098	-10.0655	-3.7541	0.0000
heal]#V-peel-ing	-6.0364	-9.3801	-2.6926	0.0000
peel-peel-ing	-10.7342	-14.2704	-7.1980	0.0000
peel bananas-peel-ing	-7.9944	-11.2654	-4.7235	0.0000
heal#V-peel-index	-0.2117	-3.1250	2.7016	1.0000
heal]#V-peel-index	0.6617	-2.4543	3.7777	0.9995
peel-peel-index	-4.0361	-7.3578	-0.7145	0.0059
peel bananas-peel-index	-1.2964	-4.3341	1.7414	0.9290
heal]#V-heal#V	0.8734	-2.1876	3.9345	0.9952
peel-heal#V	-3.8244	-7.0946	-0.5542	0.0096
peel bananas-heal#V	-1.0846	-4.0660	1.8967	0.9736
peel-heal]#V	-4.6978	-8.1498	-1.2458	0.0011
peel bananas-heal]#V	-1.9581	-5.1378	1.2217	0.6031
peel bananas-peel	2.7398	-0.6418	6.1213	0.2201

Table C.10: Tukey HSD comparison of PC1 values across context for Manchester MC

	diff	lwr	upr	p.adj
believe-leap	0.6168	-2.1114	3.3450	0.9992
free-ly-leap	0.0772	-2.6510	2.8055	1.0000
helix-leap	-0.1423	-2.8706	2.5859	1.0000
peel-ing-leap	-0.5607	-3.2889	2.1676	0.9996
peel-index-leap	-2.0497	-4.7780	0.6785	0.3166
heal#V-leap	-3.9043	-6.6326	-1.1761	0.0005
heal]#V-leap	-5.7919	-8.5202	-3.0637	0.0000
peel-leap	-5.5396	-8.2678	-2.8113	0.0000
peel bananas-leap	-5.6381	-8.3663	-2.9098	0.0000
free-ly-believe	-0.5396	-3.1950	2.1159	0.9997
helix-believe	-0.7591	-3.4146	1.8963	0.9951
peel-ing-believe	-1.1775	-3.8329	1.4780	0.9113
peel-index-believe	-2.6665	-5.3220	-0.0111	0.0482
heal#V-believe	-4.5211	-7.1766	-1.8656	0.0000
heal]#V-believe	-6.4087	-9.0642	-3.7533	0.0000
peel-believe	-6.1564	-8.8118	-3.5009	0.0000
peel bananas-believe	-6.2549	-8.9103	-3.5994	0.0000
helix-free-ly	-0.2196	-2.8750	2.4359	1.0000
peel-ing-free-ly	-0.6379	-3.2934	2.0176	0.9987
peel-index-free-ly	-2.1270	-4.7824	0.5285	0.2335
heal#V-free-ly	-3.9815	-6.6370	-1.3261	0.0002
heal]#V-free-ly	-5.8692	-8.5246	-3.2137	0.0000
peel-free-ly	-5.6168	-8.2723	-2.9613	0.0000
peel bananas-free-ly	-5.7153	-8.3708	-3.0598	0.0000
peel-ing-helix	-0.4184	-3.0738	2.2371	1.0000
peel-index-helix	-1.9074	-4.5629	0.7481	0.3800
heal#V-helix	-3.7620	-6.4175	-1.1065	0.0006
heal]#V-helix	-5.6496	-8.3051	-2.9941	0.0000
peel-helix	-5.3972	-8.0527	-2.7418	0.0000
peel bananas-helix	-5.4958	-8.1512	-2.8403	0.0000
peel-index-peel-ing	-1.4890	-4.1445	1.1664	0.7210
heal#V-peel-ing	-3.3436	-5.9991	-0.6882	0.0036
heal]#V-peel-ing	-5.2313	-7.8867	-2.5758	0.0000
peel-peel-ing	-4.9789	-7.6344	-2.3234	0.0000
peel bananas-peel-ing	-5.0774	-7.7329	-2.4219	0.0000
heal#V-peel-index	-1.8546	-4.5101	0.8009	0.4208
heal]#V-peel-index	-3.7422	-6.3977	-1.0867	0.0006
peel-peel-index	-3.4898	-6.1453	-0.8344	0.0019
peel bananas-peel-index	-3.5883	-6.2438	-0.9329	0.0013
heal]#V-heal#V	-1.8876	-4.5431	0.7679	0.3951
peel-heal#V	-1.6353	-4.2907	1.0202	0.6023
peel bananas-heal#V	-1.7338	-4.3892	0.9217	0.5195
peel-heal]#V	0.2524	-2.4031	2.9078	1.0000
peel bananas-heal]#V	0.1539	-2.5016	2.8093	1.0000
peel bananas-peel	-0.0985	-2.7540	2.5570	1.0000

Table C.11: Tukey HSD comparison of PC1 values across context for Newcastle

	diff	lwr	upr	p.adj
believe-leap	-2.1217	-5.2648	1.0215	0.4643
free-ly-leap	0.3076	-2.5296	3.1449	1.0000
helix-leap	0.5964	-2.1936	3.3863	0.9995
peel-ing-leap	2.1462	-0.7231	5.0155	0.3178
peel-index-leap	0.8225	-1.9888	3.6338	0.9938
heal#V-leap	0.9403	-1.8496	3.7302	0.9832
heal]#V-leap	-0.4212	-3.2905	2.4481	1.0000
peel-leap	-2.2918	-5.1031	0.5195	0.2096
peel bananas-leap	0.3922	-2.4192	3.2035	1.0000
free-ly-believe	2.4293	0.2483	4.6103	0.0173
helix-believe	2.7180	0.5989	4.8371	0.0030
peel-ing-believe	4.2679	2.0453	6.4904	0.0000
peel-index-believe	2.9442	0.7970	5.0914	0.0011
heal#V-believe	3.0619	0.9428	5.1811	0.0005
heal]#V-believe	1.7005	-0.5221	3.9230	0.2876
peel-believe	-0.1702	-2.3173	1.9770	1.0000
peel bananas-believe	2.5138	0.3666	4.6610	0.0098
helix-free-ly	0.2888	-1.3425	1.9200	0.9999
peel-ing-free-ly	1.8386	0.0750	3.6021	0.0341
peel-index-free-ly	0.5149	-1.1527	2.1825	0.9909
heal#V-free-ly	0.6327	-0.9986	2.2640	0.9581
heal]#V-free-ly	-0.7288	-2.4924	1.0348	0.9383
peel-free-ly	-2.5994	-4.2670	-0.9319	0.0001
peel bananas-free-ly	0.0845	-1.5831	1.7521	1.0000
peel-ing-helix	1.5498	-0.1366	3.2363	0.0983
peel-index-helix	0.2261	-1.3597	1.8119	1.0000
heal#V-helix	0.3439	-1.2037	1.8915	0.9993
heal]#V-helix	-1.0176	-2.7040	0.6689	0.6227
peel-helix	-2.8882	-4.4740	-1.3024	0.0000
peel bananas-helix	-0.2042	-1.7900	1.3816	1.0000
peel-index-peel-ing	-1.3237	-3.0453	0.3979	0.2811
heal#V-peel-ing	-1.2059	-2.8923	0.4805	0.3808
heal]#V-peel-ing	-2.5674	-4.3821	-0.7527	0.0007
peel-peel-ing	-4.4380	-6.1596	-2.7164	0.0000
peel bananas-peel-ing	-1.7540	-3.4756	-0.0325	0.0423
heal#V-peel-index	0.1178	-1.4680	1.7036	1.0000
heal]#V-peel-index	-1.2437	-2.9653	0.4779	0.3662
peel-peel-index	-3.1143	-4.7374	-1.4912	0.0000
peel bananas-peel-index	-0.4304	-2.0535	1.1928	0.9970
heal]#V-heal#V	-1.3615	-3.0479	0.3250	0.2208
peel-heal#V	-3.2321	-4.8179	-1.6463	0.0000
peel bananas-heal#V	-0.5481	-2.1339	1.0377	0.9801
peel-heal]#V	-1.8706	-3.5922	-0.1491	0.0226
peel bananas-heal]#V	0.8134	-0.9082	2.5349	0.8705
peel bananas-peel	2.6840	1.0609	4.3071	0.0000

Table C.12: Tukey HSD comparison of PC1 values across context for Belfast

	diff	lwr	upr	p.adj
believe-leap	0.6615	-1.2179	2.5409	0.9786
free-ly-leap	0.2278	-1.5194	1.9749	1.0000
helix-leap	-0.2546	-2.0608	1.5516	1.0000
peel-ing-leap	0.7742	-0.9242	2.4727	0.8963
peel-index-leap	-2.1259	-3.8243	-0.4275	0.0040
heal#V-leap	-6.7083	-8.4067	-5.0099	0.0000
heal]#V-leap	-8.3076	-9.9650	-6.6501	0.0000
peel-leap	-7.5779	-9.3841	-5.7717	0.0000
peel bananas-leap	-6.7898	-8.3823	-5.1974	0.0000
free-ly-believe	-0.4337	-2.3927	1.5252	0.9993
helix-believe	-0.9162	-2.9279	1.0956	0.8969
peel-ing-believe	0.1127	-1.8029	2.0283	1.0000
peel-index-believe	-2.7874	-4.7030	-0.8718	0.0004
heal#V-believe	-7.3698	-9.2854	-5.4542	0.0000
heal]#V-believe	-8.9691	-10.8485	-7.0897	0.0000
peel-believe	-8.2395	-10.2512	-6.2277	0.0000
peel bananas-believe	-7.4514	-9.2737	-5.6291	0.0000
helix-free-ly	-0.4824	-2.3712	1.4064	0.9979
peel-ing-free-ly	0.5465	-1.2395	2.3325	0.9919
peel-index-free-ly	-2.3537	-4.1397	-0.5677	0.0019
heal#V-free-ly	-6.9361	-8.7221	-5.1501	0.0000
heal]#V-free-ly	-8.5353	-10.2825	-6.7882	0.0000
peel-free-ly	-7.8057	-9.6945	-5.9169	0.0000
peel bananas-free-ly	-7.0176	-8.7032	-5.3320	0.0000
peel-ing-helix	1.0289	-0.8149	2.8727	0.7259
peel-index-helix	-1.8713	-3.7151	-0.0274	0.0438
heal#V-helix	-6.4537	-8.2975	-4.6098	0.0000
heal]#V-helix	-8.0529	-9.8591	-6.2467	0.0000
peel-helix	-7.3233	-9.2669	-5.3797	0.0000
peel bananas-helix	-6.5352	-8.2819	-4.7885	0.0000
peel-index-peel-ing	-2.9002	-4.6385	-1.1618	0.0000
heal#V-peel-ing	-7.4825	-9.2209	-5.7442	0.0000
heal]#V-peel-ing	-9.0818	-10.7802	-7.3834	0.0000
peel-peel-ing	-8.3522	-10.1960	-6.5084	0.0000
peel bananas-peel-ing	-7.5641	-9.1991	-5.9291	0.0000
heal#V-peel-index	-4.5824	-6.3208	-2.8440	0.0000
heal]#V-peel-index	-6.1817	-7.8801	-4.4832	0.0000
peel-peel-index	-5.4520	-7.2959	-3.6082	0.0000
peel bananas-peel-index	-4.6639	-6.2989	-3.0289	0.0000
heal]#V-heal#V	-1.5993	-3.2977	0.0991	0.0825
peel-heal#V	-0.8696	-2.7135	0.9742	0.8756
peel bananas-heal#V	-0.0815	-1.7166	1.5535	1.0000
peel-heal]#V	0.7296	-1.0766	2.5358	0.9483
peel bananas-heal]#V	1.5177	-0.0747	3.1102	0.0750
peel bananas-peel	0.7881	-0.9586	2.5348	0.9020

Table C.13: Tukey HSD comparison of PC1 values across context for Liverpool

F2-F1

	diff	lwr	upr	p.adj
believe-leap	-171.0220	-532.4886	190.4445	0.8743
free-ly-leap	-131.3649	-492.8314	230.1016	0.9737
helix-leap	4.1979	-357.2687	365.6644	1.0000
peel-ing-leap	154.8846	-206.5819	516.3512	0.9272
peel-index-leap	-1078.7182	-1440.1848	-717.2517	0.0000
heal#V-leap	-1205.5179	-1566.9845	-844.0514	0.0000
heal]#V-leap	-1301.7790	-1663.2456	-940.3125	0.0000
peel-leap	-1125.0243	-1486.4908	-763.5577	0.0000
peel bananas-leap	-1263.5765	-1625.0430	-902.1100	0.0000
free-ly-believe	39.6571	-321.8094	401.1237	1.0000
helix-believe	175.2199	-186.2466	536.6864	0.8576
peel-ing-believe	325.9067	-35.5599	687.3732	0.1138
peel-index-believe	-907.6962	-1269.1627	-546.2297	0.0000
heal#V-believe	-1034.4959	-1395.9624	-673.0294	0.0000
heal]#V-believe	-1130.7570	-1492.2235	-769.2904	0.0000
peel-believe	-954.0022	-1315.4688	-592.5357	0.0000
peel bananas-believe	-1092.5545	-1454.0210	-731.0879	0.0000
helix-free-ly	135.5628	-225.9038	497.0293	0.9678
peel-ing-free-ly	286.2495	-75.2170	647.7161	0.2477
peel-index-free-ly	-947.3533	-1308.8199	-585.8868	0.0000
heal#V-free-ly	-1074.1530	-1435.6196	-712.6865	0.0000
heal]#V-free-ly	-1170.4141	-1531.8807	-808.9476	0.0000
peel-free-ly	-993.6594	-1355.1259	-632.1928	0.0000
peel bananas-free-ly	-1132.2116	-1493.6781	-770.7451	0.0000
peel-ing-helix	150.6868	-210.7798	512.1533	0.9380
peel-index-helix	-1082.9161	-1444.3826	-721.4496	0.0000
heal#V-helix	-1209.7158	-1571.1823	-848.2493	0.0000
heal]#V-helix	-1305.9769	-1667.4434	-944.5103	0.0000
peel-helix	-1129.2221	-1490.6887	-767.7556	0.0000
peel bananas-helix	-1267.7744	-1629.2409	-906.3078	0.0000
peel-index-peel-ing	-1233.6029	-1595.0694	-872.1363	0.0000
heal#V-peel-ing	-1360.4026	-1721.8691	-998.9360	0.0000
heal]#V-peel-ing	-1456.6637	-1818.1302	-1095.1971	0.0000
peel-peel-ing	-1279.9089	-1641.3754	-918.4424	0.0000
peel bananas-peel-ing	-1418.4611	-1779.9277	-1056.9946	0.0000
heal#V-peel-index	-126.7997	-488.2662	234.6668	0.9792
heal]#V-peel-index	-223.0608	-584.5273	138.4058	0.5998
peel-peel-index	-46.3060	-407.7726	315.1605	1.0000
peel bananas-peel-index	-184.8583	-546.3248	176.6083	0.8147
heal]#V-heal#V	-96.2611	-457.7276	265.2055	0.9971
peel-heal#V	80.4937	-280.9729	441.9602	0.9993
peel bananas-heal#V	-58.0586	-419.5251	303.4080	1.0000
peel-heal]#V	176.7548	-184.7118	538.2213	0.8512
peel bananas-heal]#V	38.2025	-323.2640	399.6691	1.0000
peel bananas-peel	-138.5522	-500.0188	222.9143	0.9630

Table C.14: Tukey HSD comparison of F2-F1 values across context for RP

	diff	lwr	upr	p.adj
believe-leap	48.5997	-218.3367	315.5360	0.9999
free-ly-leap	162.8951	-104.0413	429.8315	0.6136
helix-leap	47.8006	-219.1358	314.7370	0.9999
peel-ing-leap	213.2191	-53.7173	480.1555	0.2363
peel-index-leap	-585.8918	-852.8282	-318.9554	0.0000
heal#V-leap	-764.7511	-1031.6875	-497.8147	0.0000
heal]#V-leap	-732.5516	-999.4879	-465.6152	0.0000
peel-leap	-852.3438	-1119.2801	-585.4074	0.0000
peel bananas-leap	-540.7205	-833.5008	-247.9401	0.0000
free-ly-believe	114.2954	-145.5214	374.1122	0.9147
helix-believe	-0.7991	-260.6159	259.0177	1.0000
peel-ing-believe	164.6194	-95.1974	424.4362	0.5615
peel-index-believe	-634.4915	-894.3083	-374.6747	0.0000
heal#V-believe	-813.3507	-1073.1675	-553.5339	0.0000
heal]#V-believe	-781.1512	-1040.9680	-521.3344	0.0000
peel-believe	-900.9434	-1160.7602	-641.1266	0.0000
peel bananas-believe	-589.3201	-875.6243	-303.0160	0.0000
helix-free-ly	-115.0945	-374.9113	144.7223	0.9113
peel-ing-free-ly	50.3240	-209.4928	310.1408	0.9998
peel-index-free-ly	-748.7869	-1008.6037	-488.9701	0.0000
heal#V-free-ly	-927.6462	-1187.4630	-667.8294	0.0000
heal]#V-free-ly	-895.4466	-1155.2634	-635.6298	0.0000
peel-free-ly	-1015.2388	-1275.0556	-755.4220	0.0000
peel bananas-free-ly	-703.6155	-989.9197	-417.3114	0.0000
peel-ing-helix	165.4185	-94.3983	425.2353	0.5546
peel-index-helix	-633.6924	-893.5092	-373.8756	0.0000
heal#V-helix	-812.5517	-1072.3685	-552.7349	0.0000
heal]#V-helix	-780.3522	-1040.1690	-520.5354	0.0000
peel-helix	-900.1444	-1159.9612	-640.3276	0.0000
peel bananas-helix	-588.5211	-874.8252	-302.2169	0.0000
peel-index-peel-ing	-799.1109	-1058.9277	-539.2941	0.0000
heal#V-peel-ing	-977.9702	-1237.7870	-718.1534	0.0000
heal]#V-peel-ing	-945.7707	-1205.5875	-685.9539	0.0000
peel-peel-ing	-1065.5628	-1325.3796	-805.7461	0.0000
peel bananas-peel-ing	-753.9396	-1040.2437	-467.6354	0.0000
heal#V-peel-index	-178.8593	-438.6761	80.9575	0.4409
heal]#V-peel-index	-146.6598	-406.4766	113.1570	0.7123
peel-peel-index	-266.4520	-526.2688	-6.6352	0.0397
peel bananas-peel-index	45.1713	-241.1328	331.4755	1.0000
heal]#V-heal#V	32.1995	-227.6173	292.0163	1.0000
peel-heal#V	-87.5927	-347.4095	172.2241	0.9840
peel bananas-heal#V	224.0306	-62.2735	510.3348	0.2619
peel-heal]#V	-119.7922	-379.6090	140.0246	0.8895
peel bananas-heal]#V	191.8311	-94.4731	478.1353	0.4807
peel bananas-peel	311.6233	25.3191	597.9275	0.0219

Table C.15: Tukey HSD comparison of F2-F1 values across context for London female

	diff	lwr	upr	p.adj
believe-leap	-154.5988	-413.1840	103.9864	0.6416
free-ly-leap	95.6158	-171.7105	362.9422	0.9762
helix-leap	2.6626	-248.7109	254.0362	1.0000
peel-ing-leap	58.8569	-192.5166	310.2304	0.9989
peel-index-leap	-671.4039	-922.7774	-420.0304	0.0000
heal#V-leap	-740.9236	-992.2971	-489.5501	0.0000
heal]#V-leap	-573.5509	-824.9244	-322.1774	0.0000
peel-leap	-770.9040	-1022.2775	-519.5304	0.0000
peel bananas-leap	-741.4619	-992.8354	-490.0884	0.0000
free-ly-believe	250.2146	-29.3385	529.7678	0.1196
helix-believe	157.2614	-107.0779	421.6008	0.6481
peel-ing-believe	213.4557	-50.8836	477.7950	0.2236
peel-index-believe	-516.8051	-781.1444	-252.4658	0.0000
heal#V-believe	-586.3248	-850.6642	-321.9855	0.0000
heal]#V-believe	-418.9521	-683.2915	-154.6128	0.0001
peel-believe	-616.3052	-880.6445	-351.9658	0.0000
peel bananas-believe	-586.8631	-851.2025	-322.5238	0.0000
helix-free-ly	-92.9532	-365.8494	179.9430	0.9829
peel-ing-free-ly	-36.7590	-309.6552	236.1372	1.0000
peel-index-free-ly	-767.0197	-1039.9160	-494.1235	0.0000
heal#V-free-ly	-836.5395	-1109.4357	-563.6433	0.0000
heal]#V-free-ly	-669.1668	-942.0630	-396.2706	0.0000
peel-free-ly	-866.5198	-1139.4160	-593.6236	0.0000
peel bananas-free-ly	-837.0778	-1109.9740	-564.1816	0.0000
peel-ing-helix	56.1942	-201.0948	313.4833	0.9994
peel-index-helix	-674.0665	-931.3555	-416.7775	0.0000
heal#V-helix	-743.5863	-1000.8753	-486.2972	0.0000
heal]#V-helix	-576.2136	-833.5026	-318.9245	0.0000
peel-helix	-773.5666	-1030.8556	-516.2776	0.0000
peel bananas-helix	-744.1246	-1001.4136	-486.8356	0.0000
peel-index-peel-ing	-730.2608	-987.5498	-472.9718	0.0000
heal#V-peel-ing	-799.7805	-1057.0695	-542.4915	0.0000
heal]#V-peel-ing	-632.4078	-889.6968	-375.1188	0.0000
peel-peel-ing	-829.7608	-1087.0499	-572.4718	0.0000
peel bananas-peel-ing	-800.3188	-1057.6078	-543.0298	0.0000
heal#V-peel-index	-69.5197	-326.8087	187.7693	0.9968
heal]#V-peel-index	97.8530	-159.4360	355.1420	0.9646
peel-peel-index	-99.5001	-356.7891	157.7890	0.9606
peel bananas-peel-index	-70.0580	-327.3470	187.2310	0.9966
heal]#V-heal#V	167.3727	-89.9163	424.6617	0.5244
peel-heal#V	-29.9803	-287.2693	227.3087	1.0000
peel bananas-heal#V	-0.5383	-257.8273	256.7507	1.0000
peel-heal]#V	-197.3530	-454.6420	59.9360	0.2883
peel bananas-heal]#V	-167.9110	-425.2000	89.3780	0.5198
peel bananas-peel	29.4420	-227.8470	286.7310	1.0000

Table C.16: Tukey HSD comparison of F2-F1 values across context for London male

	diff	lwr	upr	p.adj
believe-leap	-338.3523	-499.8964	-176.8083	0.0000
free-ly-leap	-214.5593	-376.1034	-53.0153	0.0016
helix-leap	-37.8263	-199.3704	123.7177	0.9990
peel-ing-leap	35.7940	-125.7500	197.3381	0.9993
peel-index-leap	-156.1562	-317.7003	5.3879	0.0669
heal#V-leap	-266.9016	-428.4457	-105.3576	0.0000
heal]#V-leap	-320.6697	-482.2137	-159.1256	0.0000
peel-leap	-329.2997	-490.8438	-167.7557	0.0000
peel bananas-leap	-224.2623	-385.8063	-62.7182	0.0008
free-ly-believe	123.7930	-37.7510	285.3371	0.2901
helix-believe	300.5260	138.9819	462.0701	0.0000
peel-ing-believe	374.1464	212.6023	535.6904	0.0000
peel-index-believe	182.1961	20.6521	343.7402	0.0148
heal#V-believe	71.4507	-90.0933	232.9948	0.9128
heal]#V-believe	17.6827	-143.8614	179.2267	1.0000
peel-believe	9.0526	-152.4914	170.5967	1.0000
peel bananas-believe	114.0901	-47.4540	275.6342	0.4048
helix-free-ly	176.7330	15.1889	338.2770	0.0207
peel-ing-free-ly	250.3533	88.8093	411.8974	0.0001
peel-index-free-ly	58.4031	-103.1409	219.9472	0.9746
heal#V-free-ly	-52.3423	-213.8863	109.2018	0.9880
heal]#V-free-ly	-106.1103	-267.6544	55.4337	0.5111
peel-free-ly	-114.7404	-276.2845	46.8036	0.3965
peel bananas-free-ly	-9.7029	-171.2470	151.8411	1.0000
peel-ing-helix	73.6204	-87.9237	235.1644	0.8971
peel-index-helix	-118.3299	-279.8739	43.2142	0.3522
heal#V-helix	-229.0753	-390.6193	-67.5312	0.0006
heal]#V-helix	-282.8433	-444.3874	-121.2993	0.0000
peel-helix	-291.4734	-453.0174	-129.9293	0.0000
peel bananas-helix	-186.4359	-347.9800	-24.8918	0.0113
peel-index-peel-ing	-191.9502	-353.4943	-30.4062	0.0079
heal#V-peel-ing	-302.6956	-464.2397	-141.1516	0.0000
heal]#V-peel-ing	-356.4637	-518.0077	-194.9196	0.0000
peel-peel-ing	-365.0937	-526.6378	-203.5497	0.0000
peel bananas-peel-ing	-260.0563	-421.6003	-98.5122	0.0000
heal#V-peel-index	-110.7454	-272.2895	50.7986	0.4484
heal]#V-peel-index	-164.5135	-326.0575	-2.9694	0.0424
peel-peel-index	-173.1435	-334.6876	-11.5995	0.0257
peel bananas-peel-index	-68.1060	-229.6501	93.4380	0.9338
heal]#V-heal#V	-53.7680	-215.3121	107.7760	0.9855
peel-heal#V	-62.3981	-223.9422	99.1459	0.9611
peel bananas-heal#V	42.6394	-118.9047	204.1834	0.9973
peel-heal]#V	-8.6301	-170.1741	152.9140	1.0000
peel bananas-heal]#V	96.4074	-65.1366	257.9515	0.6447
peel bananas-peel	105.0375	-56.5066	266.5815	0.5259

Table C.17: Tukey HSD comparison of F2-F1 values across context for Manchester WC

	diff	lwr	upr	p.adj
believe-leap	-224.3434	-552.8140	104.1272	0.4542
free-ly-leap	-506.3362	-825.5522	-187.1201	0.0001
helix-leap	-608.8727	-932.3283	-285.4172	0.0000
peel-ing-leap	-437.7819	-772.2800	-103.2838	0.0020
peel-index-leap	-1058.5022	-1377.7183	-739.2862	0.0000
heal#V-leap	-1143.1087	-1458.6931	-827.5243	0.0000
heal]#V-leap	-1127.9168	-1456.3874	-799.4462	0.0000
peel-leap	-1252.9594	-1594.8425	-911.0763	0.0000
peel bananas-leap	-1119.3869	-1442.8424	-795.9313	0.0000
free-ly-believe	-281.9928	-538.7701	-25.2155	0.0198
helix-believe	-384.5294	-646.5583	-122.5004	0.0003
peel-ing-believe	-213.4385	-488.9828	62.1058	0.2761
peel-index-believe	-834.1589	-1090.9362	-577.3816	0.0000
heal#V-believe	-918.7653	-1171.0136	-666.5171	0.0000
heal]#V-believe	-903.5734	-1171.7686	-635.3783	0.0000
peel-believe	-1028.6160	-1313.0799	-744.1521	0.0000
peel bananas-believe	-895.0435	-1157.0724	-633.0146	0.0000
helix-free-ly	-102.5366	-352.8666	147.7934	0.9443
peel-ing-free-ly	68.5542	-195.8897	332.9982	0.9977
peel-index-free-ly	-552.1661	-796.9936	-307.3385	0.0000
heal#V-free-ly	-636.7725	-876.8457	-396.6994	0.0000
heal]#V-free-ly	-621.5806	-878.3579	-364.8033	0.0000
peel-free-ly	-746.6232	-1020.3488	-472.8977	0.0000
peel bananas-free-ly	-613.0507	-863.3807	-362.7207	0.0000
peel-ing-helix	171.0908	-98.4554	440.6371	0.5608
peel-index-helix	-449.6295	-699.9595	-199.2995	0.0000
heal#V-helix	-534.2360	-779.9181	-288.5539	0.0000
heal]#V-helix	-519.0441	-781.0730	-257.0151	0.0000
peel-helix	-644.0867	-922.7446	-365.4287	0.0000
peel bananas-helix	-510.5141	-766.2282	-254.8001	0.0000
peel-index-peel-ing	-620.7203	-885.1643	-356.2763	0.0000
heal#V-peel-ing	-705.3268	-965.3753	-445.2783	0.0000
heal]#V-peel-ing	-690.1349	-965.6792	-414.5906	0.0000
peel-peel-ing	-815.1775	-1106.5806	-523.7744	0.0000
peel bananas-peel-ing	-681.6049	-951.1512	-412.0587	0.0000
heal#V-peel-index	-84.6065	-324.6796	155.4667	0.9786
heal]#V-peel-index	-69.4146	-326.1919	187.3627	0.9968
peel-peel-index	-194.4572	-468.1827	79.2684	0.3965
peel bananas-peel-index	-60.8846	-311.2146	189.4454	0.9986
heal]#V-heal#V	15.1919	-237.0563	267.4402	1.0000
peel-heal#V	-109.8507	-379.3322	159.6308	0.9459
peel bananas-heal#V	23.7218	-221.9603	269.4039	1.0000
peel-heal]#V	-125.0426	-409.5065	159.4213	0.9159
peel bananas-heal]#V	8.5299	-253.4990	270.5588	1.0000
peel bananas-peel	133.5725	-145.0854	412.2304	0.8657

Table C.18: Tukey HSD comparison of F2-F1 values across context for Manchester MC

	diff	lwr	upr	p.adj
believe-leap	-62.1302	-235.7946	111.5342	0.9763
free-ly-leap	-41.2712	-214.9355	132.3932	0.9988
helix-leap	-84.3514	-258.0157	89.3130	0.8559
peel-ing-leap	-207.5878	-381.2522	-33.9235	0.0073
peel-index-leap	-266.4159	-440.0802	-92.7515	0.0001
heal#V-leap	-433.8169	-607.4813	-260.1526	0.0000
heal]#V-leap	-553.4838	-727.1482	-379.8195	0.0000
peel-leap	-474.0674	-647.7318	-300.4030	0.0000
peel bananas-leap	-506.3521	-680.0165	-332.6877	0.0000
free-ly-believe	20.8590	-148.1734	189.8915	1.0000
helix-believe	-22.2212	-191.2537	146.8113	1.0000
peel-ing-believe	-145.4576	-314.4901	23.5749	0.1550
peel-index-believe	-204.2857	-373.3182	-35.2532	0.0063
heal#V-believe	-371.6867	-540.7192	-202.6542	0.0000
heal]#V-believe	-491.3536	-660.3861	-322.3212	0.0000
peel-believe	-411.9372	-580.9697	-242.9047	0.0000
peel bananas-believe	-444.2219	-613.2544	-275.1894	0.0000
helix-free-ly	-43.0802	-212.1127	125.9523	0.9980
peel-ing-free-ly	-166.3167	-335.3492	2.7158	0.0576
peel-index-free-ly	-225.1447	-394.1772	-56.1122	0.0016
heal#V-free-ly	-392.5458	-561.5783	-223.5133	0.0000
heal]#V-free-ly	-512.2127	-681.2452	-343.1802	0.0000
peel-free-ly	-432.7962	-601.8287	-263.7638	0.0000
peel bananas-free-ly	-465.0809	-634.1134	-296.0485	0.0000
peel-ing-helix	-123.2365	-292.2690	45.7960	0.3586
peel-index-helix	-182.0645	-351.0970	-13.0320	0.0244
heal#V-helix	-349.4656	-518.4980	-180.4331	0.0000
heal]#V-helix	-469.1325	-638.1650	-300.1000	0.0000
peel-helix	-389.7160	-558.7485	-220.6835	0.0000
peel bananas-helix	-422.0007	-591.0332	-252.9682	0.0000
peel-index-peel-ing	-58.8280	-227.8605	110.2045	0.9803
heal#V-peel-ing	-226.2291	-395.2616	-57.1966	0.0015
heal]#V-peel-ing	-345.8960	-514.9285	-176.8635	0.0000
peel-peel-ing	-266.4796	-435.5121	-97.4471	0.0001
peel bananas-peel-ing	-298.7643	-467.7967	-129.7318	0.0000
heal#V-peel-index	-167.4011	-336.4335	1.6314	0.0545
heal]#V-peel-index	-287.0680	-456.1005	-118.0355	0.0000
peel-peel-index	-207.6515	-376.6840	-38.6190	0.0051
peel bananas-peel-index	-239.9362	-408.9687	-70.9037	0.0006
heal]#V-heal#V	-119.6669	-288.6994	49.3656	0.4010
peel-heal#V	-40.2505	-209.2830	128.7820	0.9988
peel bananas-heal#V	-72.5352	-241.5677	96.4973	0.9264
peel-heal]#V	79.4164	-89.6160	248.4489	0.8786
peel bananas-heal]#V	47.1317	-121.9007	216.1642	0.9960
peel bananas-peel	-32.2847	-201.3172	136.7478	0.9998

Table C.19: Tukey HSD comparison of F2-F1 values across context for Newcastle

	diff	lwr	upr	p.adj
believe-leap	-23.7186	-200.1448	152.7076	1.0000
free-ly-leap	123.1572	-36.0976	282.4120	0.2737
helix-leap	140.4422	-16.1582	297.0426	0.1169
peel-ing-leap	174.8395	13.7852	335.8939	0.0228
peel-index-leap	97.0595	-60.7409	254.8599	0.5968
heal#V-leap	42.8522	-113.7482	199.4526	0.9962
heal]#V-leap	51.4800	-109.5743	212.5344	0.9884
peel-leap	47.5807	-110.2197	205.3811	0.9923
peel bananas-leap	65.1838	-92.6166	222.9842	0.9385
free-ly-believe	146.8758	24.4557	269.2960	0.0072
helix-believe	164.1608	45.2143	283.1074	0.0010
peel-ing-believe	198.5581	73.8060	323.3103	0.0001
peel-index-believe	120.7781	0.2560	241.3002	0.0491
heal#V-believe	66.5708	-52.3757	185.5173	0.7165
heal]#V-believe	75.1987	-49.5535	199.9508	0.6241
peel-believe	71.2993	-49.2227	191.8214	0.6486
peel bananas-believe	88.9024	-31.6197	209.4245	0.3370
helix-free-ly	17.2850	-74.2801	108.8501	0.9998
peel-ing-free-ly	51.6823	-47.3075	150.6722	0.7897
peel-index-free-ly	-26.0977	-119.7003	67.5049	0.9957
heal#V-free-ly	-80.3050	-171.8701	11.2601	0.1357
heal]#V-free-ly	-71.6772	-170.6670	27.3127	0.3630
peel-free-ly	-75.5765	-169.1791	18.0261	0.2207
peel bananas-free-ly	-57.9734	-151.5760	35.6292	0.5873
peel-ing-helix	34.3973	-60.2630	129.0576	0.9723
peel-index-helix	-43.3827	-132.3942	45.6287	0.8484
heal#V-helix	-97.5900	-184.4563	-10.7238	0.0158
heal]#V-helix	-88.9622	-183.6225	5.6981	0.0832
peel-helix	-92.8615	-181.8729	-3.8501	0.0339
peel bananas-helix	-75.2584	-164.2699	13.7530	0.1704
peel-index-peel-ing	-77.7800	-174.4127	18.8526	0.2243
heal#V-peel-ing	-131.9873	-226.6477	-37.3270	0.0008
heal]#V-peel-ing	-123.3595	-225.2192	-21.4998	0.0064
peel-peel-ing	-127.2588	-223.8914	-30.6262	0.0020
peel bananas-peel-ing	-109.6557	-206.2884	-13.0231	0.0141
heal#V-peel-index	-54.2073	-143.2187	34.8041	0.6103
heal]#V-peel-index	-45.5794	-142.2121	51.0532	0.8716
peel-peel-index	-49.4788	-140.5849	41.6273	0.7499
peel bananas-peel-index	-31.8757	-122.9818	59.2304	0.9784
heal]#V-heal#V	8.6278	-86.0325	103.2882	1.0000
peel-heal#V	4.7285	-84.2829	93.7400	1.0000
peel bananas-heal#V	22.3316	-66.6798	111.3430	0.9981
peel-heal]#V	-3.8993	-100.5319	92.7333	1.0000
peel bananas-heal]#V	13.7037	-82.9289	110.3364	1.0000
peel bananas-peel	17.6031	-73.5030	108.7092	0.9998

Table C.20: Tukey HSD comparison of F2-F1 values across context for Belfast

	diff	lwr	upr	p.adj
believe-leap	99.8176	-72.1369	271.7720	0.6804
free-ly-leap	-95.0550	-254.4865	64.3764	0.6464
helix-leap	-63.5140	-218.3233	91.2952	0.9438
peel-ing-leap	-85.2378	-240.0470	69.5715	0.7422
peel-index-leap	-252.2221	-396.9606	-107.4836	0.0000
heal#V-leap	-441.6252	-596.4344	-286.8159	0.0000
heal]#V-leap	-413.0763	-563.9985	-262.1541	0.0000
peel-leap	-339.5984	-504.6256	-174.5711	0.0000
peel bananas-leap	-328.7035	-479.6257	-177.7813	0.0000
free-ly-believe	-194.8726	-377.0800	-12.6652	0.0262
helix-believe	-163.3316	-341.5086	14.8454	0.1011
peel-ing-believe	-185.0553	-363.2324	-6.8783	0.0352
peel-index-believe	-352.0397	-521.5400	-182.5393	0.0000
heal#V-believe	-541.4427	-719.6198	-363.2657	0.0000
heal]#V-believe	-512.8939	-687.7042	-338.0835	0.0000
peel-believe	-439.4159	-626.5392	-252.2926	0.0000
peel bananas-believe	-428.5211	-603.3314	-253.7107	0.0000
helix-free-ly	31.5410	-134.5828	197.6648	0.9998
peel-ing-free-ly	9.8173	-156.3065	175.9410	1.0000
peel-index-free-ly	-157.1671	-313.9485	-0.3856	0.0489
heal#V-free-ly	-346.5701	-512.6939	-180.4464	0.0000
heal]#V-free-ly	-318.0213	-480.5289	-155.5137	0.0000
peel-free-ly	-244.5433	-420.2283	-68.8584	0.0008
peel bananas-free-ly	-233.6485	-396.1561	-71.1409	0.0004
peel-ing-helix	-21.7237	-183.4167	139.9693	1.0000
peel-index-helix	-188.7081	-340.7868	-36.6293	0.0044
heal#V-helix	-378.1111	-539.8041	-216.4181	0.0000
heal]#V-helix	-349.5623	-507.5377	-191.5868	0.0000
peel-helix	-276.0843	-447.5857	-104.5830	0.0000
peel bananas-helix	-265.1895	-423.1649	-107.2140	0.0000
peel-index-peel-ing	-166.9843	-319.0631	-14.9056	0.0199
heal#V-peel-ing	-356.3874	-518.0804	-194.6944	0.0000
heal]#V-peel-ing	-327.8385	-485.8140	-169.8631	0.0000
peel-peel-ing	-254.3606	-425.8619	-82.8593	0.0002
peel bananas-peel-ing	-243.4657	-401.4412	-85.4903	0.0001
heal#V-peel-index	-189.4031	-341.4818	-37.3243	0.0042
heal]#V-peel-index	-160.8542	-308.9743	-12.7341	0.0224
peel-peel-index	-87.3763	-249.8448	75.0923	0.7671
peel bananas-peel-index	-76.4814	-224.6015	71.6387	0.8065
heal]#V-heal#V	28.5489	-129.4266	186.5243	0.9999
peel-heal#V	102.0268	-69.4745	273.5281	0.6492
peel bananas-heal#V	112.9217	-45.0538	270.8971	0.3876
peel-heal]#V	73.4779	-94.5230	241.4789	0.9183
peel bananas-heal]#V	84.3728	-69.7954	238.5410	0.7487
peel bananas-peel	10.8949	-157.1061	178.8958	1.0000

Table C.21: Tukey HSD comparison of F2-F1 values across context for London female

Appendix D

ANOVA comparison tables for Experiment 2

RP

	Res.Df	RSS	Df	Sum of Sq	Pr(>Chi)
1	97	294.98			
2	96	267.53	1	27.45	0.0017

Table D.1: Comparing models for 2 and 3 for RP

	Res.Df	RSS	Df	Sum of Sq	Pr(>Chi)
1	96	267.53			
2	95	267.53	1	0.01	0.9590

Table D.2: Comparing models for 3 and 4 for RP

London Female

	Res.Df	RSS	Df	Sum of Sq	Pr(>Chi)
1	92	149.25			
2	91	110.21	1	39.04	0.0000

Table D.3: Comparing models for 2 and 3 for London F

	Res.Df	RSS	Df	Sum of Sq	Pr(>Chi)
1	91	110.21			
2	90	98.90	1	11.32	0.0013

Table D.4: Comparing models for 3 and 4 for London F

London Male

	Res.Df	RSS	Df	Sum of Sq	Pr(>Chi)
1	94	754.21			
2	93	745.47	1	8.75	0.2962

Table D.5: Comparing models 2 and 3 for London M

	Res.Df	RSS	Df	Sum of Sq	Pr(>Chi)
1	93	745.47			
2	92	721.19	1	24.28	0.0784

Table D.6: Comparing models 3 and 4 for London M

	Res.Df	RSS	Df	Sum of Sq	Pr(>Chi)
1	94	754.21			
2	92	721.19	2	33.03	0.1216

Table D.7: Comparing models 2 and 4 for London M

Manchester WC

	Res.Df	RSS	Df	Sum of Sq	Pr(>Chi)
1	96	416.73			
2	95	413.79	1	2.94	0.4113

Table D.8: Comparing models for 2 and 3 for Manchester WC

	Res.Df	RSS	Df	Sum of Sq	Pr(>Chi)
1	85	250.22			
2	84	244.94	1	5.29	0.1782

Table D.9: Comparing models for 3 and 6 for Manchester WC

	Res.Df	RSS	Df	Sum of Sq	Pr(>Chi)
1	96	416.73			
2	94	411.87	2	4.86	0.5740

Table D.10: Comparing models for 2 and 4 for Manchester WC

	Res.Df	RSS	Df	Sum of Sq	Pr(>Chi)
1	85	250.56			
2	84	250.22	1	0.34	0.7371

Table D.11: Comparing models for 5 and 6 for Manchester WC

	Res.Df	RSS	Df	Sum of Sq	Pr(>Chi)
1	84	250.22			
2	82	236.51	2	13.71	0.0929

Table D.12: Comparing models for 6 and 7 for Manchester WC

	Res.Df	RSS	Df	Sum of Sq	Pr(>Chi)
1	85	250.56			
2	82	236.51	3	14.04	0.1817

Table D.13: Comparing models for 5 and 7 for Manchester WC

Manchester MC

	Res.Df	RSS	Df	Sum of Sq	Pr(>Chi)
1	99	1133.80			
2	98	950.82	1	182.98	0.0000

Table D.14: Comparing models for 2 and 3 for Manchester MC

	Res.Df	RSS	Df	Sum of Sq	Pr(>Chi)
1	98	950.82			
2	97	930.18	1	20.64	0.1424

Table D.15: Comparing models for 3 and 4 for Manchester MC

	Res.Df	RSS	Df	Sum of Sq	Pr(>Chi)
1	97	557.74			
2	96	550.41	1	7.33	0.2582

Table D.16: Comparing models 5 and 6 for Manchester MC

Newcastle

	Res.Df	RSS	Df	Sum of Sq	Pr(>Chi)
1	97	363.33			
2	96	336.45	1	26.88	0.0056

Table D.17: Comparing models for 2 and 3 for Newcastle

	Res.Df	RSS	Df	Sum of Sq	Pr(>Chi)
1	96	336.45			
2	95	335.99	1	0.46	0.7195

Table D.18: Comparing models for 3 and 4 for Newcastle

Belfast

	Res.Df	RSS	Df	Sum of Sq	Pr(>Chi)
1	81	212.04			
2	80	212.01	1	0.03	0.9108

Table D.19: Comparing models for 2 and 3 for Belfast

	Res.Df	RSS	Df	Sum of Sq	Pr(>Chi)
1	80	212.01			
2	79	205.64	1	6.37	0.1177

Table D.20: Comparing models for 3 and 4 for Belfast

Liverpool

	Res.Df	RSS	Df	Sum of Sq	Pr(>Chi)
1	95	197.13			
2	94	183.68	1	13.45	0.0087

Table D.21: Comparing models for 2 and 3 for Liverpool

	Res.Df	RSS	Df	Sum of Sq	Pr(>Chi)
1	94	183.68			
2	93	182.01	1	1.67	0.3556

Table D.22: Comparing models for 3 and 4 for Liverpool

	Res.Df	RSS	Df	Sum of Sq	Pr(>Chi)
1	94	149.96			
2	93	143.22	1	6.74	0.0365

Table D.23: Comparing models for 5 and 6 for Liverpool

	Res.Df	RSS	Df	Sum of Sq	Pr(>Chi)
1	93	143.22			
2	91	142.57	2	0.65	0.8127

Table D.24: Comparing models for 6 and 7 for Liverpool

Mixed-effects models

	Df	AIC	BIC	logLik	deviance	Chisq	Chi Df	Pr(>Chisq)
mod1	4	2508.16	2526.82	-1250.08	2500			
mod3	5	2488.32	2511.63	-1239.16	2478	21.84	1	0.0000

Table D.25: Comparing models for 1 and 3 in acoustics

	Df	AIC	BIC	logLik	deviance	Chisq	Chi Df	Pr(>Chisq)
mod2	4	3017.63	3036.28	-1504.82	3010			
mod3	5	2488.32	2511.63	-1239.16	2478	531.31	1	0.0000

Table D.26: Comparing models for 2 and 3 in acoustics

	Df	AIC	BIC	logLik	deviance	Chisq	Chi Df	Pr(>Chisq)
mod3	5	2488.32	2511.63	-1239.16	2478			
mod4	6	2431.16	2459.13	-1209.58	2419	59.16	1	0.0000

Table D.27: Comparing models for 3 and 4 in acoustics

	Df	AIC	BIC	logLik	deviance	Chisq	Chi Df	Pr(>Chisq)
mod4	6	2431.16	2459.13	-1209.58	2419			
mod5	6	2103.62	2131.58	-1045.81	2092	327.55	0	0.0000

Table D.28: Comparing models for 4 and 5 in acoustics

	Df	AIC	BIC	logLik	deviance	Chisq	Chi Df	Pr(>Chisq)
mod5	6	2103.62	2131.58	-1045.81	2092			
mod6	7	2094.76	2127.39	-1040.38	2081	10.85	1	0.0010

Table D.29: Comparing models for 5 and 6 in acoustics

	Df	AIC	BIC	logLik	deviance	Chisq	Chi Df	Pr(>Chisq)
mod6	7	2094.76	2127.39	-1040.38	2081			
mod7	8	2075.46	2112.74	-1029.73	2059	21.31	1	0.0000

Table D.30: Comparing models for 6 and 7 in acoustics