Spatial and temporal variation in speech planning: Evidence from laterals.

Emily Gorman

Department of Linguistics and English Language
Lancaster University

A thesis submitted for the degree of
Doctor of Philosophy
September 2024

Spatial and temporal variation in speech planning:

Evidence from laterals.

Emily Gorman

Abstract

This thesis explores the relationship between spatial and temporal variation in English laterals in order to test predictions of the coupled oscillator model of speech timing. Electro-magnetic articulography data is used to examine spatial and temporal properties of laterals from two British English dialects which differ in degree of /l/ darkening, an Onset Lightening dialect: Standard Southern British English, and and Onset Darkening dialect: Lancashire English. This thesis explores how dialect-mediated spatial differences in laterals affect patterns of lateral cluster timing. This design allows for a systematic, within-language exploration of the effects of /l/ darkening on speech timing. Results find laterals to differ spatially across dialects, but not temporally in measures of /l/ cluster timing and inter-gestural timing. In this way, temporal stability in /l/ is found to be invariant to spatial differences. Further analyses reveal systematic dialectal differences in the velocity of the tongue body gesture of /l/ to facilitate this pattern. Implications of these results are discussed within the framework of Articulatory Phonology and the coupled oscillator model of speech timing. Specifically, results support a feed-forward model, whereby coupling relations are invariant to spatial properties.

Contents

List of Figures

List of Tables

Acknowledgements

Thank you to Sam Kirkham and Patrycja Strycharczuk for being the most wonderful and supportive supervisors throughout this journey. You are truly inspirational. Thank you to Danielle Turton and Marianne Pouplier for your thorough and extremely valuable feedback upon examining this thesis.

Thank you to all of my friends and colleagues in the Lancaster Linguistics department. It's been a privilege to be part of such a beautiful community.

Thank you to my lovely family who've always been right behind me in everything I do. And to my partner, thank you for being my rock, throughout the PhD and beyond.

A huge thank you to my participants for not being phased by EMA and for making this project possible. Finally, thank you to the Economic and Social Research Council for funding this research.

Declaration

I declare that this thesis is my own work and has not been submitted in substantially the same form for an award of a higher degree elsewhere.

Emily Gorman

Chapter 1

Introduction

The goal of this thesis is to investigate the relationship between spatial and temporal variation in speech production using the test case of British English laterals, where systematic spatio-temporal variation is known to occur. I focus in particular on the interaction between dialect-mediated spatial differences in lateral gestures and patterns of lateral cluster timing. The nature of this relationship, and spatio-temporal relations in speech production more broadly, has theoretical importance, for the extent to which spatial and temporal aspects of speech co-vary has consequences for models of phonology, which must account for these patterns.

One model of phonology which makes specific predictions about spatio-temporal variation is Articulatory Phonology. Articulatory Phonology (AP) proposes dynamical spatio-temporal representations to be the fundamental unit of phonological organisation (e.g., Browman and Goldstein, 1988; Browman and Goldstein, 1992). An assumption of the AP model is that patterns of temporal coordination between gestures are governed by coupling relations between planning oscillators associated with each gesture. Coupling relations between gestures differ with syllable position, such that C-V gestures are coupled in-phase in onsets, and V-C gestures are coupled anti-phase in codas. The model predicts that these coupling relations should occur irrespective of spatial variation, hence there should be no relationship between intrinsic gestural properties and gestural timing. However, empirical evidence is mixed on this matter. For example, Pastätter and Pouplier (2017) found the coarticulatory resistance of a consonant to influence C-V timing, while Shaw and Chen (2019) found that C-V timing was affected by the position of the tongue dorsum at the vowel onset, both suggesting covariance between spatial and temporal dimensions. Shaw and Chen's findings were not, however, replicated by Liu, Xu, and Hsieh (2022) who found CV onsets in Mandarin to begin synchronously in segmentally contrastive disyllabic pairs.

Evidence is also mixed regarding the predicted timing patterns of gestures in cluster contexts. The coupled oscillator model of AP (Browman and Goldstein, 2000) predicts gestures in a coda clusters (VCC) to show a local timing pattern through a series of anti-phase relationships, which results in a temporal stacking of gestures. Gestures in onset clusters (CCV) are rather predicted to exhibit a global timing pattern through a combination of in-phase coupling relations between each consonant and the vowel, and an anti-phase coupling between the consonants. This results in a displacement of consonants around the vowel. These timing patterns are reviewed in more detail in Section 2.2.3.

Contrary to these predictions, empirical evidence often find lateral clusters to show different timing patterns. For example, lateral codas show the predicted sequential timing pattern in German (Pouplier, 2012), but show an atypical timing pattern in English (e.g, Goldstein et al., 2009; Marin and Pouplier, 2010). These timing irregularities are possibly linked to the complex gestural make-up of laterals and variation in the phasing of the component gestures (Sproat and Fujimura, 1993), which varies across varieties (Gick et al., 2006).

Cross-linguistically, laterals exhibit considerable variation in lateral darkening (e.g, Gick et al., 2006; Recasens and Espinosa, 2005). One hypothesis is that lateral darkening mediates lateral cluster timing (Marin and Pouplier, 2014), suggesting a relationship between spatial variation in cluster components and the global timing patterns evidenced in those clusters. To date, it has however been difficult to test hypotheses around the effects of lateral darkening on cluster timing, as between-language comparisons are confounded by phonotactic and other segmental-prosodic differences between languages. Further, eliciting a wide range of variation from within a homogeneous group, also proves challenging.

Laterals in British English provide a compelling test case for addressing this relationship between spatial and temporal variation in lateral clusters. Systematic spatial and temporal variation in lateral darkening has been widely reported across dialects of British English (Turton, 2014; Wells, 1982). This within-language variation provides an interesting opportunity for measuring the effects of lateral darkening on cluster timing within a language, without the confounds of a between-language comparison.

This thesis explores the relationship between patterns of spatial phonetic variation and speech timing, using the test case of lateral clusters in two English dialects: an Onset Darkening dialect, where /l/ is dark in onset and coda positions, and an Onset Lightening dialect, where /l/ is clear or light in onset position and dark in coda position. Audio-synchronised electromagnetic articulography data of speakers of the two dialects are analysed. This allows for a comparison between two similar varieties of the same language, which differ in a crucial regard: /l/ darkening. The following sections provide a detailed breakdown of how the broad aim of the thesis will be approached. First, I establish how differences in /l/ darkening between Onset Lightening and Onset Darkening dialects manifest in articulation, before I explore the relationship between the dialect-mediated differences in /l/ darkening and patterns of lateral cluster timing.

Before detailing the studies of this thesis, I first wish to highlight the stance taken by the thesis regarding how /l/ is conceptualised at the gestural level, and the goals of lateral production. This thesis adopts the multi-gestural model of /l/ proposed by Sproat and Fujimura (1993), whereby /l/ is said to comprise a tongue dorsum retraction / tongue body lowering gesture and a tongue tip raising gesture. Others, however, have argued that /l/ may best be thought of as a single gesture. For example, Recasens and Espinosa (2005) suggest a single combined measure of post dorsum lowering and retraction to capture /l/ darkening. Debate is also to be had around the goals of lateral production. Browman and Goldstein (1995) suggest the goals of lateral production to be defined mid-sagittally, with lateralisation being a by-product of a primary mid-sagittal target. This is the stance assumed by this thesis which does not measure lateralisation. However, we must also consider the theoretical possibility that lateralisation, as oppose to tongue body lowering/dorsum retraction, is the primary target in /l/ production. This idea is raised by Sproat and Fujimura (1993) who suggest tongue dorsum to be a by-product of tongue lateralisation. Ying et al. (2021) supports the view, finding the timing of lateralisation to be stable in Australian English speakers across different morphosyntactic contexts. Thus, it should be acknowledged that, in the treatment of laterals, this thesis adopts one of several approaches.

1.1 Overview of Studies

1.1.1 Dialect Variation in /l/ Darkening

Before the interaction between lateral darkness and timing can be explored, I first establish the articulatory mechanisms responsible for the dialectal contrast in /l/ darkening. Commonly cited measures of /l/ darkening are applied to lateral data. Given the known effect of context on lateral darkening, (e.g., Lee-Kim, Davidson, and Hwang, 2013; Mackenzie et al., 2018; Sproat and Fujimura, 1993; Strycharczuk, Derrick, and Shaw, 2020; Turton, 2014), /l/ is examined across a range of morphosyntactic and vocalic contexts in order to better disentangle effects of dialect on /l/ darkening from effects of context. Throughout, I discuss methodological challenges posed by the task of arriving at a single measure of /l/ darkening.

1.1.2 Relationship Between /l/ Darkness and Clusters Timing

Chapter 5 explores the interaction between the dialect-mediated spatial patters of laterals, established in Chapter 4, and the timing of laterals in consonant clusters. As briefly outlined above, consonant cluster timing refers to the timing relationships between consonants and vowels in a consonant cluster context, in onset (CCV), or coda (VCC) position. Numerous studies have found systematic differences in the timing of consonant clusters in onsets compared to codas, which have been explained through the coupled oscillator model to result from differences in coupling relations, described above (e.g, Browman and Goldstein, 1995). While consonant clusters in onset position are typically observed to be organised in a C-centre timing pattern, a sequential timing pattern is observed for coda clusters (see Sections 2.2.4 and 2.2.5 for explanation of these terms).

For lateral coda clusters, atypical, non-sequential timing patterns have been reported for some languages but not others. Specifically, languages where /l/ is dark in coda position, such as in American English show such atypical patterns, while languages where coda /l/ is clear, such as Romanian do not (Katz, 2012; Marin and Pouplier, 2010, 2014). This suggests that the differences observed in cluster timing arise from differences in the articulatory properties of clear vs dark /l/s. This relationship between the internal articulatory dynamics of laterals and lateral timing in multi-gestural structures poses potentially interesting questions for an Articulatory Phonology framework regarding the nature of the relationship between spatial and temporal variation. Addressing this question, this study compares two dialects of British English where /l/ differs in systematic ways: An Onset Darkening dialect - Lancashire / Manchester English, where /l/ is dark in all positions, and an Onset Lightening dialect - Southern Standard British English, where /l/ is clear in onsets and dark in codas. Chapter 4 finds differences in /l/ darkening between dialects to manifest in the magnitude of tongue body lowering. This provides a context where the effect of dialect-mediated spatial differences in laterals on patterns of lateral cluster timing can be tested explicitly.

This study's synthesis of speech dynamics and language variation allows insights into important theoretical issues, including how the internal dynamics of a lateral segment affect global timing patterns within a consonant cluster context, and how this relationship is handled by an articulatory framework. In addition, the dialect comparison allows insights into patterns of micro variation of /l/ across two closely related varieties. This is important because much of what we currently know about articulatory variation in /l/ is drawn from meta-comparisons between different studies, which often employ non-comparable methodologies or are confounded by cross-linguistic differences.

1.1.3 Dialectal variation in the tongue body gesture of /l/

In the final study of this thesis in Chapter 6, I undertake a restricted temporal analysis of /l/ within a single segmental context - namely in the word pair *plick / lick*. This high-front vowel context provides a segmental environment whereby the two gestures of /l/ (assuming the Sproat and Fujimura model of /l/, see Section 2.3) can be easily identified. Identifying the gestures of /l/ proved challenging within the previous studies of this thesis, where wider segmental contexts of /l/ were considered. Here, I consider the timing of the tongue body lowering gesture of /l/ and the tongue tip raising gesture of /l/, both in relation to each other and surrounding segments. Further, this study investigates, explicitly, the relationship between the previously established spatial variation in the tongue body gesture of /l/ between dialects, and patterns of temporal variation.

1.2 Structure of Thesis

This thesis is structured as follows. Chapter 2 provides the theoretical and empirical background to the studies of this thesis. First, the theoretical framework of Articulatory Phonology is presented, and key theoretical concepts and areas of recent developments are outlined. Articulatory Phonology takes into account both spatial and temporal aspects of speech, making it physically relevant and testable. Importantly for this thesis, Articulatory Phonology makes specific predictions about the relationship between the spatial and temporal aspects of speech; these will be outlined in Section 2.2.7. Concepts from Articulatory Phonology will be a continuing thread throughout this thesis and will be drawn upon in conceptualising, measuring, and interpreting speech data. The latter half of Chapter 2 addresses laterals, the test case used within the thesis to examine spatio-temporal relations. Here, I provide an overview of patterns of lateral variation and consider how laterals may be conceptualised within an Articulatory Phonology framework. The chapter concludes by laying out the research questions of the thesis.

Chapter 3 provides an overview the experimental procedures used to obtain the audio-synchronised electromagnetic articulography (EMA) data for this study, speaker information, and stimuli. All data for this thesis were collected within single sessions. Each chapter of the thesis focuses on a different set of stimuli collected within this session. Methodological details which are specific to each study (i.e., temporal measures and statistical methods) are provided alongside the relevant studies in Chapters 4, 5, and 6.

Chapter 4 presents the first study of this thesis, "Dialect Variation in /l/ Darkening". The goal of this chapter is to establish the articulatory mechanisms driving the dialectal difference in /l/ darkening. Of particular importance is the nature of spatial articulatory differences in /l/ darkening between dialects, as Chapters 5 and 6, will build upon this finding in asking how spatial differences in /l/ between dialects affects patterns of /l/ cluster timing.

In Chapter 5, I present the second study, "Timing of /l/ Clusters in Onset Darkening and Onset Lightening dialects." Here, I explore the relationship between spatial differences in /l/ darkening between dialects, as found in Chapter 4, and patterns of /l/ cluster timing. Chapter 6 presents the final study of the thesis which performs a restricted analysis on /l/ within a single segmental context. This allows an explicit look into how spatial differences in the tongue body gesture of /l/ between dialects structure timing.

Chapter 7, draws together the themes which run throughout this thesis. The relationship between spatial and temporal variation in the context of /l/ is evaluated in relation to core assumptions of the Articulatory Phonology model. I discuss issues such as the validity of the C-centre measure as a heuristic for syllable structure, alongside other hypothesised explanations for the main findings of this thesis. In addition, I also discuss the contexts structuring variability in lateral darkening in the dialects considered here, the ways in which /l/ darkening manifests, and the implications of these for models of laterals.

Chapter 2

Literature Review

2.1 Introduction

In this chapter I review literature relevant to this thesis' investigation of the relationship between spatial and temporal variation in speech. The prediction of Articulatory Phonology for this relationship will be core to this exploration. I begin this chapter by outlining the fundamental concepts of Articulatory Phonology which underpin the models predictions for the nature of spatio-temporal relations. These include the notion of the gesture, task dynamic modelling, and coupling relations between gestures. This is followed by a discussion of how spatio-temporal variation is handled within the AP model.

In the second half of the chapter, I provide an overview of laterals. English laterals will provide a test case through which I examine the relationship between spatial and temporal variation. In these sections, I justify the use of laterals as a test case for examining the nature of spatio-temporal relations and outline the ways in which laterals vary along spatial and temporal dimensions. Drawing together the chapter, I discuss how laterals are handled within an AP framework, focusing in particular on the coupling relations between the multiple gestures of /l/. The research questions of the thesis are then presented.

2.2 Articulatory Phonology

This section outlines the fundamental concepts of Articulatory Phonology and theoretical issues that the model seeks to address. A major problem for theories of phonology is how to account for the discrete invariant and continuous variant aspects of speech. Articulatory Phonology accounts for this problem through adopting a dynamical approach to phonology, where variation emerges "naturally" from discrete spatio-temporal targets (Kelso, Saltzman, and Tuller, 1986, p. 30). Articulatory Phonology differs from symbolic and modular approaches to phonology in several ways. Firstly, there is no translation problem between the discrete abstract phonological representation and the continuous physical domain of phonetic implementation in AP; rather, discrete and continuous aspects of speech are considered the higher and lower dimensions of a single, dynamical system (Browman and Goldstein, 1992; Browman and Goldstein, 1995). This differs to symbolic or modular accounts, where cognitive and physical domains are highly distinct. Secondly, AP considers the gesture to be the fundamental unit of speech. Gestures are the forces which drive articulatory movements, and hence have both a spatial and temporal manifestation (Browman and Goldstein, 1992). In this way, AP shows time to be an emergent property of the system (Iskarous and Pouplier, 2022).

A central aspect of AP is the notion of "coordinative structures" (Kelso, Saltzman, and Tuller, 1986). Coordinative structures refer to the synergetic coming together of articulators to achieve a particular speech goal. How exactly the goal is achieved, i.e., which articulators contribute, and to what degree, may vary with context, hence leading to instances of motor equivalence. The point to emphasise here, is that it is the *goals* of speech, rather than the movements of individual articulators, which are of primary importance (Kelso, Saltzman, and Tuller, 1986), and variation within the contextual environment may yield variation in how the goal is achieved. In the sections which follow, I elaborate further on these ideas, beginning first with defining the notion of the gesture.

2.2.1 The Gesture

The precise definition of a gesture is that of an abstract system which specifies articulatory events: "abstract characterisations of articulatory events" (Browman and Goldstein, 1992)[pg., 155], or "a dynamical system specified with a characteristic set of parameter values" (Browman and Goldstein, 1995)[pg., 181]. That is, a gesture refers to the dynamical system which determines an articulatory movement, and not to the articulatory movement itself. However, within this thesis, the term "gesture" will often be used in the latter sense, to refer to a physical event. This is because, although gestures are abstract, they also manifest as physical changes to the vocal tract. One way to disambiguate the definition of the gesture is to define the AP notions of "tract variables" and "model articulators" (Browman and Goldstein, 1992; Saltzman and Munhall, 1989). The functional goals of speech, for example, the complete lip closure required for a bilabial plosive, are tract variables. These are achieved through the combined movements of multiple articulators; for example, lip closure may involve movements of the upper lip, the lower lip, and the jaw. There are multiple ways that the tract variable of lip closure could be achieved; the jaw may raise a little or a lot, and this has consequences for how much the lower lip has to move. Gestures are defined in terms of tract variables. Model articulators, on the other hand, refer to the functional synergies of articulators which can flexibly achieve a speech goal in different ways. The notion of tract variables will be elaborated upon further in the discussion of Task Dynamics below.

Periods of gestural activation can be visually represented using a gestural score. A gestural score depicts the intervals of time during which a gesture's activation is above an activation threshold, such that there is a meaningful modification to the physical state of the vocal tract (Tilsen, 2020b). Figure 2.1 depicts a gestural score for the word *tab*. The mobile articulators are listed on the y-axis and time on the x-axis. Within the boxes the nature of the constriction is specified (wide, or narrow), as well as the passive articulator where appropriate.

Gestures have further been categorised by the degree of constriction they specify. "Consonantal" gestures involve a considerable constriction (i.e., a close coming together of articulators) in the vocal tract, while "vocalic" gestures involve a comparatively open constriction (Sproat and Fujimura, 1993).

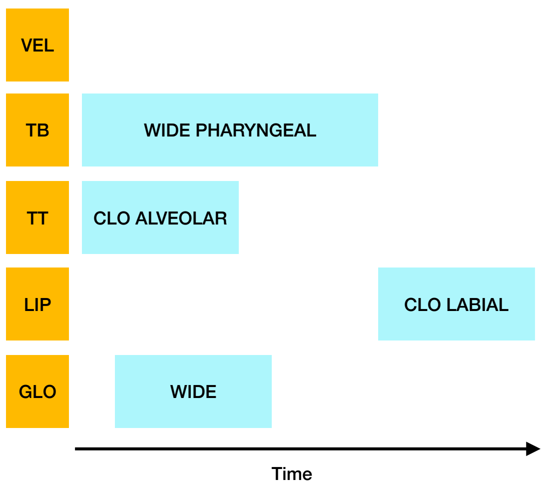

Figure 2.1: Gestural Score of the word *tab*. Orange boxes show the articulator and blue boxes show periods of activation for articulatory gestures.

Gestures are modelled as dynamical systems. Dynamical systems are common place within mathematical descriptions of the natural world. What is meant by a dynamical system is concisely summarised by Iskarous and Pouplier (2022): "A differential equation describing a dynamical system, specifies how a system will evolve over time based on its current state" [pg., 21]. We can consider gestures then as systems which specify how the vocal tract gets from its current state to another phonologically specified state. More specifically, gestures can be described as point attractors (Saltzman and Munhall, 1989). A point attractor in a dynamical system describes a point which attracts states to it, leading different trajectories towards the same point, i.e., a specified articulatory target, and do so independent of the starting position. To take a speech example, a velic speech target for /k/, can still be achieved, regardless of whether the preceding segment is a front vowel, requiring a fronted tongue body, or back vowel, requiring a posterior tongue body position.

2.2.2 Task Dynamics

Task Dynamics is a branch within Articulatory Phonology which allows gestural behaviour to be modelled and specified (Saltzman and Munhall, 1989). As discussed above, Articulatory Phonology considers gestures to be made up of tract variables: constrictions within the vocal tract, created from synergies of articulatory movements (Saltzman and Munhall, 1989). Tract variables include: lip protrusion, lip aperture, tongue tip constriction location and degree, tongue body constriction location and degree, velic aperture, and glottal aperture. A full list of tract variables and the articulators which they may recruit is shown in Table 2.1, taken from Browman and Goldstein (1992). For example, to achieve a lip aperture of 0, or perhaps a small negative number denoting lip compression, involves varying contributions of the upper lip, lower lip, and the jaw. Tract variables are phonologically specified for spatial target and stiffness through constant parameters in a dynamical equation. Stiffness determines the rate at which a gestural target is achieved (Saltzman and Munhall, 1989). For example, the constriction degree of a consonant will generally be larger than that of a vowel and will also differ in stiffness; a fricative will generally have a higher constriction degree than a vowel, but a lower constriction degree than the complete closure of a stop (Browman and Goldstein, 1992). Further, the model posits that when a gesture is not active, a neutral attractor acts to bring articulators towards a resting schwa position (Saltzman and Munhall, 1989). This shift in focus away from the actions of individual articulators, and towards the coordinated activity of multiple gestures (Browman and Goldstein, 1992) allows for a dynamic view of speech which is closer to its physical implementation and further away from the idea of speech as discrete abstract units.

Tract variables and Articulators	
Tract Variable	Articulators
LP - lip protrusion	Upper lip, lower lip, jaw
LA – lip aperture	Upper lip, lower lip, jaw
TTCL – tongue tip constriction location	Tongue tip, tongue body, jaw
TTCD – tongue tip constriction degree	Tongue tip, tongue body, jaw
TBCL – tongue body constriction location	Tongue body, jaw
TBCD – tongue body constriction degree	Tongue body, jaw
VEL – velic aperture	Velum
GLO – glottal aperture	Glottis

Table 2.1: Tract variables and articulators, from Browman and Goldstein, 1992

2.2.3 Gestural Coupling and Co-ordination

Another important aspect of Articulatory Phonology is the nature of gestural coordination, the temporal relationships between individual gestures, and the mechanisms which govern them. The mechanism responsible for gestural coordination in AP is the Coupled Oscillator Model (Browman and Goldstein, 2000).

The coupled oscillator model views the measurable timing patterns in speech to result from the coupling relations between planning oscillators with which gestures are associated (Nam and Saltzman, 2003). Affiliated with each gesture is a planning oscillator. A planning oscillator serves as a temporal clock, oscillating in 360° cycles. Gestures are coordinated with one another through the notion of coupling. Coupling occurs between planning oscillators of gestures and mediates relative gestural timing. Pairs of gestures may be coupled to one another in one of two primary coupling modes, "in-phase" and "anti-phase" (Browman and Goldstein, 2000). In-phase coupling refers to a synchronous coupling relationship between two planning oscillators, or a relative phase of 0° between oscillators. Anti-phase coupling refers to a sequential coupling relationship between two planning oscillators or a relative phase of 180° between oscillators. The coupled oscillator model gains appeal from its simplicity; gestural timing can be accounted for entirely by the in-phase and anti-phase coupling relations between planning oscillators.

Evidence for in-phase and anti-phase as the primary modes of movement coordination has been provided by coordination studies such as Kelso (1984), who found that participants, when asked to extend their index fingers on their left and right hands back and forth, could do so either synchronously, (in-phase) or in an alternating pattern (anti-phase) only. Further, these coupling modes have been regarded as "stable" or "intrinsic" modes since they are available without the need to be learned (Nam, Goldstein, and Saltzman, 2009) (see also Turvey, 1990 for a detailed overview of phasing relationships from a dynamic perspective). It is argued that these intrinsic in-phase and anti-phase modes are exploited in speech, since speech itself is a system which requires no explicit learning (Browman and Goldstein, 1988; Nam, Goldstein, and Saltzman, 2009), though compare Iskarous and Pouplier (2022) who discuss this issue from a broader perspective). While in-phase and anti-phase are the most readily available modes, more complex patterns of coordination are also possible with explicit learning, such as the coordination involved in the act of juggling for example (Turvey, 1990).

The most common illustrations of in-phase and anti-phase relationships within speech are the timing patterns of singleton onsets and codas. Articulatory evidence (e.g., Löfqvist and Gracco, 1999) has shown that the consonant and vowel within a CV onset sequence are produced concurrently, while in a VC coda sequence, the vowel and consonant are produced sequentially. The different modes of coupling between the consonant and vowel in an onset and coda sequence can be seen in the periods of gestural activation on gestural score for *tab* in Figure 2.1.

Of the two available coupling modes, in-phase is considered the most stable; observations of "phase transitions" – the rapid switch from one coupling mode to another (Turvey, 1990), offer evidence to support this. Returning again to the example from Kelso (1984), when participants were asked move their left and right index fingers in an alternating pattern in time with a speeding up metronome, participants switched to an in-phase mode after a certain rate. In-phase but not anti-phase coupling could be maintained at high rates, suggestive of greater stability within the in-phase mode (Kelso, 2009). This difference in stability between in-phase and anti-phase modes was subsequently modelled in the Haken-Kelso-Bunz model (HKB) (Haken, Kelso, and Bunz, 1985). Similar evidence has also been provided from induced speech error studies (Goldstein et al., 2007; Pouplier, 2003). For example, Goldstein et al. (2007) elicited speech errors by asking participants to repeat sequences of words with alternating consonants such as *"cop top, kip tip"* to a speeding up metronome. When observing the kinematic patterns of the tongue tip and tongue dorsum during the speech errors in the *"top cop"* sequence, the authors found that speakers were producing the /t/ and /k/ simultaneously. This was interpreted as showing speakers transitioning to a more stable and easier to maintain, in-phase mode. Further evidence corroborating in-phase coupling as the more stable mode is the finding of a shorter planning time for in-phase CV sequences compared to anti-phase VC sequences (Mooshammer et al., 2012) and the cross-linguistic preference for CV structures compared to VC structures (Nam, Goldstein, and Saltzman, 2009). Because in-phase coupling occurs within CV structures and anti-phase coupling occurs within VC structures, the model reframes the preference for CV structures as a preference for in-phase coupling, the more stable of the two modes (Nam, Goldstein, and Saltzman, 2009).

Coupling between planning oscillators mediates gestural timing in a very direct way. When two planning oscillators are coupled in-phase ($0°$), their associated gestures begin at the same time. When two planning oscillators are coupled anti-phase (180°), the oscillator of the first associated gesture must be at 180° before the second gesture begins. Figure 2.2 depicts in-phase and anti-phase coupling between pairs of planning oscillators; in-phase coupling is shown on the top and anti-phase coupling is shown on the bottom.

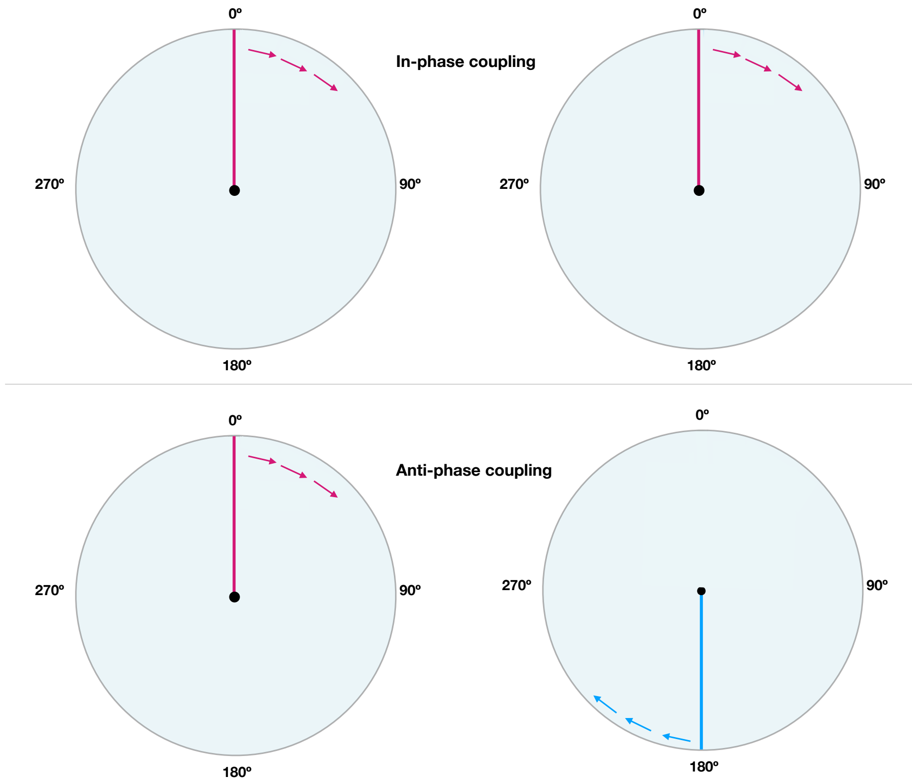

Figure 2.2: Phase circles depicting in-phase and anti-phase relationships between planning oscillators. The top two circles depict an in-phase coupling relationship, whereby pairs of oscillators begin at 0°. The bottom two circles depict an anti-phase coupling relationships, whereby the second oscillator begins at 180° in the oscillatory cycle.

In-phase and anti-phase coupling relationships between gestures can be graphically represented on a coupling graph as shown for the word *tab* in Figure 2.3. Solid green lines denote an in-phase relationship (0°), while the dashed purple line denotes an anti-phase (180°) relationship. The following sections will elaborate further on the interaction between in-phase and anti-phase relationships in relation to the more complex case of consonant clusters.

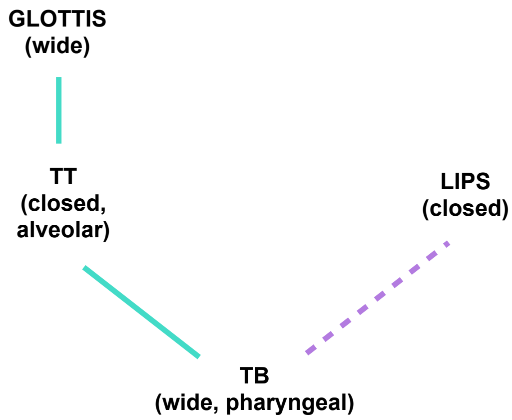

Figure 2.3: Coupling graph of the word *tab*. Solid green lines show an in-phase coupling relationship between gestures, while the purple dashed line denotes an anti-phase coupling relationship.

2.2.4 Consonant-Vowel Coupling Relations and Timing patterns in Onsets

Singleton Onsets CV

According to the coupled oscillator model, in a singleton CV onsets, (e.g., *ba* in "bat"), the consonant and vowel onsets are produced in-phase. This means that the planning oscillators associated with the consonant and vowel gestures are coupled in-phase with one another (at 0° in the oscillatory cycle). This, in turn, results in a synchronous onset of the consonant and vowel gestures. Synchronicity is permitted here for two reasons. First, the more open constriction of vowels allows for synchronous onset of the vowel with the consonant without interfering with the recoverability of either segment (Goldstein, Byrd, and Saltzman, 2006). Second, the vowel gesture takes longer to achieve its target, and is hence longer in duration than the consonant gesture. This difference in duration allows the sequence to be heard as a consonant followed by a vowel (Goldstein, Byrd, and Saltzman, 2006).

Onset Clusters CCV

When moving from a singleton (CV) to a cluster (CCV) in onset position (e.g., the singleton onset of *sip* to the cluster onset of *skip*), consonants are competitively coupled. Within a CCV sequence, oscillators of both consonants are coupled in-

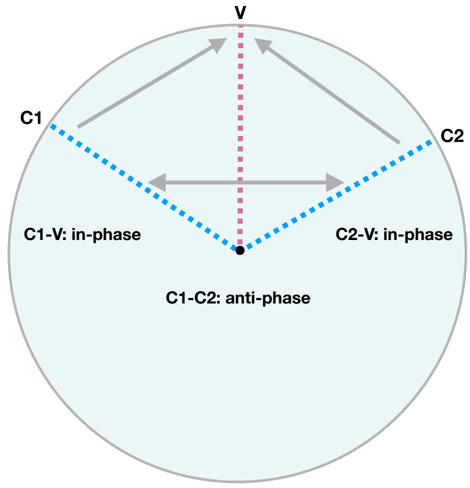

Figure 2.4: Schematic illustration of the consonant and vowel coupling relations in an onset cluster.

phase with the vowel, but anti-phase to one another, see Figure 2.4. Because all consonants cannot begin at the same time as the vowel without obscuring one another perceptually, a solution is sought whereby the centre of the consonant cluster (midpoint of C1 and C2) maintains a constant relationship with the centre of the vowel, a phenomenon described as the "C-centre Effect" (Browman and Goldstein, 1988). This effect entails a leftward shift of C1 away from the vowel, and a rightward shift of C2 towards the vowel, as shown in Figure 2.5. Through this pattern, the timing relationship between the centre of the consonant cluster and the vowel remains the same even as consonant cluster complexity increases (Marin and Pouplier, 2010, 2014). One condition required for the C-centre effect is that the anti-phase coupling between consonants (C1-C2) must have a higher coupling strength that the in-phase couplings between the consonants and vowel (C1-V; C2-V) (Browman and Goldstein, 2000). In addition, the model assumes that all consonants in the cluster are coupled equally to the vowel, i.e., coupling strengths are equal between C1-V and C2-V, and hence there are equal amounts of rightward and leftward shifting towards and away from the vowel (Mücke, Hermes, and Tilsen, 2020).

The coupled oscillator model predicts differences in consonant-vowel timing patterns between branching and non-branching onset languages. In accordance with this prediction, languages with branching onsets, such as English and German, are

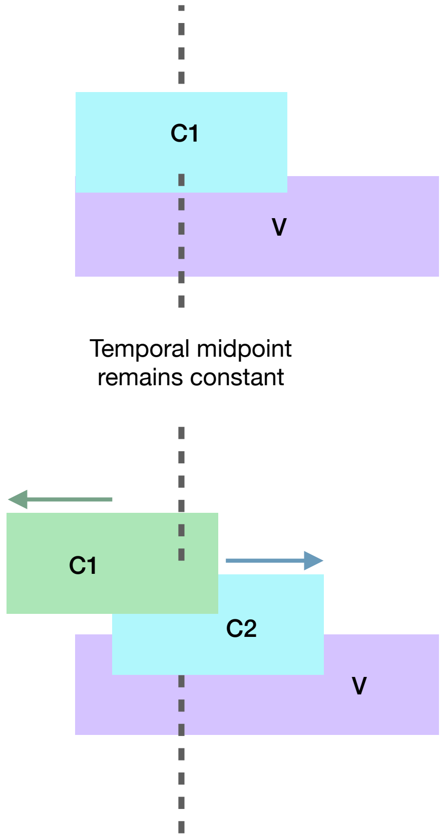

Figure 2.5: Schematic diagram of C-centre Effect, adapted from Marin and Pouplier (2010, p. 381) The top figure shows a singleton CV onset and the bottom figure shows a cluster CCV onset. Dashed vertical lines show a stable consonant centre (C-centre) across singleton and cluster contexts.

reported to show a C-centre effect. Languages with non-branching onsets, such as Moroccan Arabic show a different pattern. In non-branching onsets, there is no in-phase coupling relation between the vowel-remote consonant and the vowel (e.g. C1-V in C1-C2-V clusters). Coupling relationships are therefore non-competitive since consonants are coupled anti-phase to each other and only the oscillator of the vowel-adjacent consonant (C2) is coupled in-phase to the vowel (Mücke, Hermes, and Tilsen, 2020).

2.2.5 Consonant-Vowel Coupling Relations and Timing patterns in Codas

Singleton Codas VC

Unlike onsets, oscillators of consonant and vowel onsets in coda position are coupled anti-phase to each other, (at 180° in the oscillatory cycle). This means that the coda consonant begins after the onset of the vowel. Indeed, an anti-phase relationship here makes sense; if the vowel and the consonant in a coda VC sequence were to be coupled in-phase, the consonant and vowel would begin simultaneously, but because of the longer and slower execution of vowels (Goldstein, Byrd, and Saltzman, 2006), the vowel would outlast the consonant giving the percept of a CV sequence, i.e., an onset. This point shows that the anti-phase coupling relationship between the vowel and consonant in a coda sequence is not arbitrary, but rather a necessity of syllabic organisation.

Coda Clusters VCC

Coupling relationships within coda clusters are also non-competitive. As consonants are added (e.g., VC to VCC), each consonant is coupled in an anti-phase relationship to the preceding consonant and only the vowel-adjacent consonant (C1) is coupled anti-phase to the vowel. Unlike onsets, there is no need for weighted coupling strengths because there is no conflict between coupling modes (Browman and Goldstein, 2000). This means that added consonants are simply stacked in time without affecting the timing of preceding gestures, hence this timing pattern is referred to as a 'local' timing pattern (Marin and Pouplier, 2010) (see Figure 2.6 for illustration).

Sections 2.2.4 and 2.2.5 have outlined basic predictions of the coupled oscillator model for onset and coda coordination patterns: a global C-centre pattern for onsets, and a local sequential timing pattern for codas. However, as will be discussed in Chapter 5, such patterns are not always observed, particularly in the case of laterals (e.g., Marin and Pouplier, 2014), and hence require explanations beyond the basic assumptions of the model.

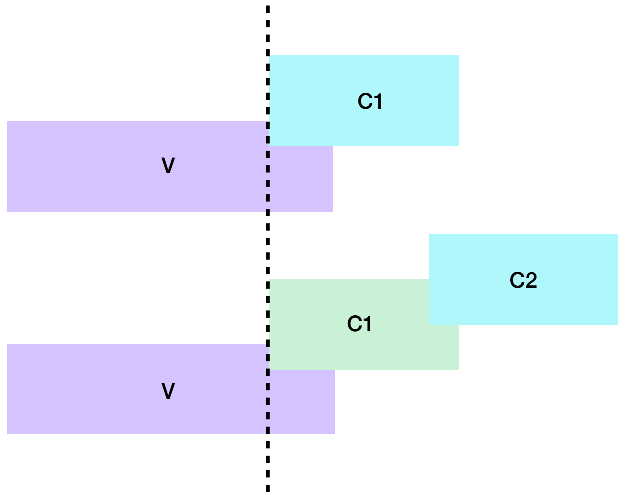

Figure 2.6: Schematic diagram of local coda timing pattern, adapted from Marin and Pouplier (2010, p. 381). A singleton VC context is shown on the top, and a cluster VCC context is show below. The dashed line shows a stable left edge across singleton and cluster contexts.

2.2.6 Developments in Articulatory Phonology

Despite its successes and explanatory power of empirical phenomena, there are many areas of the early model of Articulatory Phonology, (Browman and Goldstein, 2000), which require further refinement. One area for development is the need to integrate state based feedback. The need for a feedback mechanism is apparent from findings of perturbation studies, which show speakers to produce real time changes to the vocal tract in response to articulatory or auditory perturbation (Honda, Fujino, and Kaburagi, 2002; Munhall, Löfqvist, and Kelso, 1994; Munhall et al., 2009).

One recent development of the model which includes feedback is Tilsen's Selection-Coordination-Intention (SCI) model (Tilsen, 2018). Tilsen's SCI models expands the explanatory power of the AP model in such a way that it can better support empirical data. The SCI model not only includes a feedback loop, but assigns it the theoretically important role of duration manipulation (Tilsen, 2022). Feedback is responsible for suppressing the gestural activation period of a selected gesture, which in turn signals for a competitive gesture selection process to resume (see Tilsen (2018)). The sooner this feedback is received, the sooner the gesture activation interval will end and the sooner the gesture selection process will resume.

The SCI model distinguishes between two types of feedback, external feedback and internal feedback, which differ in regard to the time it takes for them to be received and thus take effect. External feedback is physical in nature and based on acoustic or somatosensory information, thus is only received upon execution of a gesture. Sole reliance on external feedback would thus result in a temporal lag between gestures, since the first gesture must be executed and fed-back, before the next cued gesture can be selected; this makes coarticulation impossible. Internal feedback, a predictive kind of feedback, resolves this need for the possibility of coarticulation. Internal feedback is available before a gesture has been executed and hence allows for quicker gestural suppression, shorter duration, and coarticulation. Duration then is modulated by varying degrees of reliance on internal and external feedback. Tilsen hypothesises that the well know durational difference between consonants and vowel can be explained in this framework. That is, because consonants tend to have a clearer, more tactile target, less time is needed for external feedback, while vowels, which have a more open articulation require greater external feedback time (Tilsen, 2022).

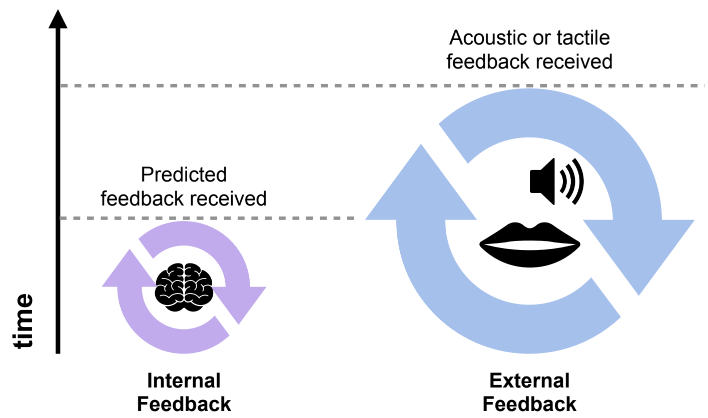

Figure 2.7: Graphical depiction of the difference between internal and external feedback as presented in Tilsen, 2022

The time at which gestural activation begins and ends in the standard AP model is a further area requiring refinement (Tilsen, 2022). This is also addressed by the

SCI model (Tilsen, 2018). In this model, gestures are defined as "systems which are continuously exerting force on the vocal tract" (Tilsen, 2020a)[pg., 2]. That is, gestures are always active; however, it is only when their activation reaches a certain threshold that they exert significant influence on the vocal tract, i.e., a movement of an articulator (Tilsen, 2020b). Within the model, gestures are competitively ordered in terms of their level of activation; activation continues to increase until a gesture reaches a selection threshold. Once a gesture has reached a selection threshold, that gesture is selected for execution and the activation growth of lower gestures is temporarily halted. The selected gesture is executed and feedback (either internal and/or external) cues deselection of the gesture. Once deselected, gestures continue the competitive queuing process, with the recently deselected gesture now assuming the position of least activation.

Also relevant to the lack of specificity of gestural activation time in AP is the argument for movement end points to be temporally specified (Turk and Shattuck-Hufnagel, 2020). The argument for privileging the timing of the gesture end point is made on the basis there is less variability in motor actions such as catching a ball or pressing a key on a keyboard at the end points than any other points. A possible counter-argument, voiced by Tilsen (2022), is that evidence for this is based on non-speech examples, which bare little resemblance to speech (see Tilsen, 2022 for further discussion on this matter).

The inability of the standard AP model to explain non-local timing patterns has also been identified as a challenge. For example, phenomena such as nasal consonant harmony, whereby distant segments are affected, while intervening segments are not (Tilsen, 2019). This cannot be accounted for by a standard AP account, i.e., through the spreading of gestures, since spreading would affect all gestures in between the source gesture and affected gesture. Tilsen (2019) proposes a modification to the model which can account for such patterns, whereby gestures can exert influence on the vocal tract even when gestural activation is below threshold.

Finally, AP must account for the possibility that the targets of speech are not restricted to a single modality, i.e., articulation. Developments in this area include the work of Burgdorf (2022), who shows how acoustic targets may incorporated into an AP model.

As a developing model, there are naturally areas in the AP model requiring further development , as reviewed above. However, for the purposes of this thesis, the standard AP model provides a useful framework of speech coordination and timing. As we have shown here, the model is continually evolving and developing in new directions, which will be embraced in this thesis.

2.2.7 Relationship between spatial and temporal variation from an AP perspective

Previous sections have described how spatial and temporal information is specified with the AP model. Against this backdrop, this section will discuss how the relationship between spatial and temporal variation is understood within an AP framework.

In contrast to extrinsic timing models such as Turk and Shattuck-Hufnagel (2020)'s 3C/XT model, there is no external time keeper in the AP model. This means that there is no explicit tracking of the spatial positions of articulators. Rather, information such as time of target achievement is only predicted based on a given onset time, as dictated by the gesture's phase coupling relations, as well as other factors such as stiffness (Saltzman and Munhall, 1989). The coupled oscillator model (Browman and Goldstein, 2000) suggests the nature of the coupling relations should apply to all segments which occur within the CV structure. In this respect, the model is considered to be a "feed-forward" model; the coordination structure is fed forward, and is unaffected by segmental variability (Shaw and Chen, 2019). This means that segmental information, such as the spatial positions of articulators, is not fed back to the coordination level.

However, several studies have challenged this idea by showing that the spatial properties of the consonant or vowel do in fact affect temporal coordination (e.g., Iskarous, 2010; Pastätter and Pouplier, 2017; Shaw and Chen, 2019). Shaw and Chen (2019) found a relationship between the position of the tongue dorsum at the start of a CV sequence, and the time of vowel initiation in Mandarin speakers. When the initial position of the tongue was closer to the vowel target position, tongue movement for the vowel was initiated later than when initial tongue position was further away from the vowel target. However, this finding was not replicated by Liu, Xu, and Hsieh (2022).

Pastätter and Pouplier (2017) reported a relationship between the coarticulation resistance of a consonant and consonant-vowel timing relations in Polish speakers. The degree to which a consonant is considered coarticulation resistant is determined by the relative involvement of the tongue body (Recasens and Rodríguez, 2016). The more competing demands placed on the tongue body, the greater the degree of coarticulation resistance. For example, labials /p, b, m/ which involve little to no active tongue body movement, are not considered to be coarticulation resistant, and so would be influenced more by surrounding vowels than a sibilant, for example, which involves greater tongue body control. Consonants may, therefore, be considered along a cline from most to least coarticulation resistant depending on the degree of tongue body involvement. Relatedly, Iskarous and Pouplier (2022) describe how it is mechanically impossible for a consonant-vowel sequence to be produced in-phase when the segments pose spatially conflicting demands, such as in /ta/, where the tongue is required to be simultaneously front for /t/ and back for /a/. They evidenced their claim by drawing upon Iskarous (2010), who found in-phase relationships for CV sequences which did not pose competing spatial demands, such as /pa/, but found sequential patterns for spatially conflicting segments, such as /ta/.

In summary, the assumption of the standard AP model is that there is no interaction between spatial and temporal variation in speech, positing a feed-forward control system. Empirical knowledge to date on the nature of the spatio-temporal relationship is, however, incomplete, hence it remains unclear how this relationship should be modelled within an Articulatory Phonology framework. There is some evidence to suggest the presence of an interaction between spatial, segmental variation and timing which cannot be straightforwardly accounted for within the standard coupled oscillator model. Various accounts have been proposed to explain how the coupled oscillator model may be compatible with such an interaction. One suggestion draws upon the role of the neutral attractor (Saltzman and Munhall, 1989). This was proposed by Shaw and Chen (2019) to explain their findings of an interaction between tongue body position and time of vowel initiation. They suggest that the neutral attractor comes into effect when the tongue is in an extreme, or spatially non-neutral position. This means that while the data may appear to show the vowel gesture being initiated earlier when the TB is further away from the vowel target, this early initiation is simply the neutral attractor bringing the articulators back to rest position, before the vowel gesture takes over (Shaw and Chen, 2019).

Another explanation proposed by Shaw and Chen (2019) to explain findings is that of "downstream targets". In the coupled oscillator model, the onset of gestural movement is typically considered to be the temporal point of a gesture which is coordinated with other gestures; though others have challenged this view, such as Turk and Shattuck-Hufnagel (2020) who argue for end point specification. The idea of downstream targets entertains the possibility that the coupled point may not be the initiation of movement, but rather a temporally later event, such as the gestural target, or release. The hypothesis then is that findings of an interaction between spatial position and gestural timing may be explained if the assumption of the onset of movement as being the privileged time point for coordination is relaxed. Within Shaw and Chen (2019)'s CV data, achievement of the vowel target was found to be most closely aligned with the offset of the consonant; this is in contrast to the pattern predicted by an in-phase timing pattern, i.e., alignment of C(onset) and V(onset). The authors further found a positive correlation between the C(offset) to V(target) alignment pattern and consonant duration, showing the timing of the vowel target to be affected by consonant duration. Such coupling between the vowel target and consonant offset means that the time of the consonant offset must be fed back for temporal alignment with the vowel (Shaw and Chen, 2019, p. 11); such a feedback mechanism is currently not implemented within the coupled oscillator model (see Tilsen's SCI model outlined in Section 2.2.6 for an example of how a feedback mechanism could be implemented into a gestural model).

2.3 Laterals

This thesis aims to develop our understanding of the nature of the relationship between spatial and temporal variation through an articulatory investigation of English laterals. British English laterals provide an interesting test case for exploring this relationship as this segment displays systematic spatial variation between dialects of the same language. This offers a scenario whereby the temporal effects of spatial variation in laterals can be explored without cross linguistic confounds. In addition to their systematic variability, laterals have also been reported to exhibit unusual timing patterns in clusters, which are not straightforwardly accounted for by the coupled oscillator model (e.g., Marin and Pouplier, 2014). This suggests that there may be a potentially interesting interaction between the spatial composition of laterals and their subsequent temporal patterning in consonant clusters. For these reasons, laterals will be the subject of empirical investigation into the relationship between spatial and temporal variability throughout this thesis.

Laterals belong to the class of liquids which in many ways combine articulatory elements of both consonants and vowels, owing to their multi gestural nature (Sproat and Fujimura, 1993). One consequence of multiple gestures is the greater scope for variability. Realisations of /l/ are commonly grouped into three perceptual categories: "clear" (or "light"), "dark", and "vocalised". Laterals have been widely reported to show allophonic variation, such that /l/ is clear in syllable onsets and dark in syllable codas. The empirical reality, is not always so clear-cut between clear and dark categories, with reports of gradient variation between clear and dark extremes, in both acoustic and articulatory domains (e.g., Recasens, 2004, 2012; Recasens and Espinosa, 2005; Sproat and Fujimura, 1993). As will be reviewed in Chapter 4, lateral darkening has widely been shown to interact with vocalic and morphosyntactic context, such that /l/ is typically darker before stronger morphosyntactic boundaries and in back vowel contexts (e.g., Lee-Kim, Davidson, and Hwang, 2013; Strycharczuk, Derrick, and Shaw, 2020; Strycharczuk and Scobbie, 2015; Turton, 2014).

2.3.1 Articulatory goals of lateral production

Current understandings of the gestural composition of laterals stem from the seminal work of Sproat and Fujimura (1993). Using X-ray Micro-beam data, Sproat and Fujimura (1993) identified two distinct gestures within American English laterals: a consonantal tongue tip raising gesture, where the tongue tip raises vertically towards the alveolar region of the palate, and a vowel-like dorsal raising and retraction gesture, which they noted to be highly correlated with the lowering of the tongue body. As their name implies, laterals also involve a degree of lateralisation. Lateralisation refers to a lowering of the sides of the tongue to create a channel for lateral airflow.

The articulatory goal of laterals is somewhat contested. In particular, there is a debate around whether lateralisation is the articulatory goal of laterals, or whether lateralisation is rather a by product of the mid sagittal stretching and narrowing of the tongue. The latter position, that lateralisation is a by-product, is adopted by Browman and Goldstein (1995). Browman and Goldstein (1995) show that laterals can be successfully simulated when only the mid-sagittal gestures are specified, hence lateralisation was not considered to be a primary articulatory target. On the other hand, Sproat and Fujimura (1993) argue that lateralisation is the goal, and that retraction of the tongue dorsum is a consequence of lateralisation. More recently, studies have sought to test this question empirically by measuring the dynamic movement of the tongue in both the mid sagittal and para sagittal planes. In an EMA study on Australian English speakers where sensors are affixed to mid sagittal and lateral locations on the tongue, Ying et al. (2021) show lateralisation to be stable across different vowel and syllable contexts, while the apical and dorsal movements of /l/ were not. From this finding for the stability of lateralisation, the authors inferred lateralisation to be the articulatory goal of laterals for Australian English. Strycharczuk, Derrick, and Shaw (2020) on the other hand offer an alternative account. They found that while there was the predicted correlation between tongue dorsum retraction and lateralisation, this was not maintained for dark /l/s. They go on to suggest that lateralisation may result in dorsal retraction initially, but this does not continue to be the case for darker /l/s.

2.3.2 Acoustic and articulatory correlates of clear, dark, and vocalised /l/

Systematic variation in laterals, leading to the labels such as "clear" and "dark" described above, occurs in both acoustics and articulation. Articulatorily, darker laterals are characterised by greater raising and retraction of the tongue dorsum (Giles and Moll, 1975; Sproat and Fujimura, 1993; Turton, 2015), and greater lowering of the tongue body (Lee-Kim, Davidson, and Hwang, 2013; Sproat and Fujimura, 1993; Strycharczuk, Derrick, and Shaw, 2020). The relative timing of gestures has also been found to pattern with /l/ darkening (Sproat and Fujimura, 1993). In dark coda /l/s in English, the dorsal gesture temporally precedes (or occurs simultaneously with) the apical gesture, while in a clear onset /l/, the apical gesture precedes the dorsal gesture, (see also Browman and Goldstein, 1995; Scobbie and Pouplier, 2010; Strycharczuk and Scobbie, 2015). Acoustically, darker /l/ is characterised by a relatively low F2 and high F1, meaning an increased proximity between F1 and F2 values (or higher F1−F2 value), while a clearer /l/ is characterised by a relatively low F1 and high F2, giving rise to a large F2−F1 distance (Kirkham et al., 2019; Nance, 2014; Recasens and Espinosa, 2005; Sproat and Fujimura, 1993). Similarly, F3-F2 has also been shown to correlate with variation in /l/ darkness; again, the larger the distance between formants, the clearer the /l/ (Kirkham et al., 2019).

Vocalisation of /l/ broadly refers to a more vocalic production of /l/, largely accompanied by reduction in the apical raising gesture of /l/. This can exist on a wide scale from reduced contact of the tongue tip on the alveolar ridge to altogether absence of a tongue tip gesture (e.g., Strycharczuk, Derrick, and Shaw, 2020). Possible intermediate stages include complete loss of tongue tip contact (e.g., Hardcastle and Barry, 1989), and reduced magnitude of the tongue tip gesture beyond a loss of contact. In addition to reduction in magnitude of the tongue tip gesture, Strycharczuk, Derrick, and Shaw (2020) found that the tongue tip gesture may also be temporally delayed in vocalised /l/, though this was found for only one speaker in their study. A further articulatory characteristic of vocalisation is lip rounding, as reported by Wells (1982) who found an increase in labialisation for vocalised tokens of /l/ in London speakers. Vocalisation has been increasingly observed in numerous varieties of English, including Southern British English, (Hardcastle and Barry, 1989; Trudgill, 1986; Turton, 2014) American English (Wrench and Scobbie, 2003), Scottish English (Scobbie and Pouplier, 2010), and New Zealand English (Strycharczuk, Derrick, and Shaw, 2020), which has led to the view of vocalisation as being a sound change in progress (Wells, 1982). Johnson and Britain (2007) argue vocalisation to be a "natural phenomenon" occurring syllable finally in varieties which have clear and dark allophones of /l/, or dark /l/s in all positions. However, this is not the case cross-linguistically; for example, /l/ vocalisation occurs in dialects of Bavarian, where /l/ is clear (Vollmann et al., 2017). Segmentally, reports of vocalisation are almost exclusively limited to word final pre consonantal position, though vocalisation has been reported to variably occur in word final pre vocalic position for some speakers (Turton, 2014). Vocalisation has also been shown to be more likely to occur when followed by a velar consonant (Hardcastle and Barry, 1989), or preceded by a long vowel (Johnson and Britain, 2007), and less likely following coronals (Johnson and Britain, 2007).

2.3.3 Cross linguistic variation in laterals

Considerable variation in /l/ darkening is reported across varieties. Cross-linguistically, languages and dialects have been categorised according to whether they have "intrinsic" or "extrinsic" allophones of /l/ (Recasens, 2012). Intrinsic allophones are defined as those where there is a positional effect on /l/ darkness, such that /l/ is clearer in onset and darker in coda position. Extrinsic allophones are those where the difference in darkness between variants of /l/ is too great to be explain by position alone. Recasens (2012) further proposes a non binary distinction between clear and dark /l/ across languages. In an acoustic study of 23 languages/dialects which examined F1, F2, and F3 of laterals in /a/ and /i/ contexts, Recasens (2012) suggested varieties with a clear /l/ to include RP English, Danish, Dutch, French, German, and Hungarian. Varieties with dark /l/s included Russian, American English, and Portuguese, while varieties with intermediate clear or dark /l/ included Czech and Eastern Catalan respectively.

Wide variability has also been observed in the articulatory properties of laterals across languages and dialects. Gick et al. (2006) show cross-linguistic differences in both gestural timing between apical and dorsal gestures of /l/, and gestural magnitude. While Gick et al. (2006) observe varieties including West Canadian English and Squamish Salish to show inter-gestural timing patterns (between dorsal and apical gestures of /l/) which are similar to those reported for English above (e.g., Browman and Goldstein, 1995; Sproat and Fujimura, 1993), different patterns were observed in the laterals of Korean and Serbo-Croatian speakers. Korean laterals were found to have no observable tongue body gesture in pre-vocalic and intervocalic positions, and negative gestural lag in post vocalic position, whereby the tongue tip gesture preceded the tongue body gesture. This latter observation runs counter to the predictions of Sproat and Fujimura (1993) for the tongue body gesture to be temporally closer to the vowel. In Serbo Croatian non palatalized /l/, Gick et al. (2006) found apical and dorsal gestures of /l/ to be simultaneous in pre-vocalic, post-vocalic and intervocalic positions, and the magnitude of both gestures was observed to be consistent across all positions. Such findings highlight the diversity of lateral gestural timing patterns, beyond those reported for English.

2.3.4 Laterals in SSBE and Lancashire / Manchester English

The present study examines laterals from two broad dialect regions of British English. The Onset Darkening dialect includes speakers of Standard Southern British English (SSBE), and the Onset Darkening dialect comprises Lancashire and Manchester English speakers. These dialect regions were selected because speakers in these regions are reported to show systematic differences in /l/ darkness. Moreover, study of these dialects facilitates a comparative analysis between two systematically distinct but related varieties.

SSBE is reported to exhibit clear /l/s in syllable onset position, and dark or vocalised /l/s in syllable coda (e.g., Turton, 2014; Wells, 1982). Articulatorily, this has been shown to be realised through more retracted tongue dorsum and reduced tongue tip gesture in codas compared to onsets (e.g, Scobbie and Pouplier, 2010).

Further, Turton (2014) found /l/ in word final pre consonantal position to show markedly greater tongue root retraction, relative to 5 other morphosyntactic positions for one RP speaker. Speakers of Lancashire and Manchester English, on the other hand, are reported to produce a dark /l/ in all syllable positions, as well as vocalised /l/ in coda position (Beal, 2008; Hughes, Trudgill, and Watt, 2012). While /l/ is reported to be dark in all positions for these speakers, increased darkening in coda position is still expected. For example, Turton (2014) found /l/ in Manchester English to exhibit tongue tongue root retraction in all positions, however, observed slightly greater retraction for /l/ in word final pre consonantal position. An interaction between /l/ darkening and social class has also been observed for speakers of Manchester English. Turton (2014) observed middle class speakers to produce clearer /l/s than those of working class speakers.

Vocalisation has been widely documented within Southern British English (e.g., Scobbie and Pouplier, 2010; Strycharczuk et al., 2020; Tolfree, 1999; Wells, 1982). Tolfree (1999) reports on vocalisation in the South East of England in both regional and standard dialects in word final pre consonantal, word final pre pausal, and word final intervocalic positions. Greater instances of vocalisation were found for younger speakers than older speakers, and only younger speakers vocalised in word final intervocalic position, suggesting /l/ vocalisation to be a change in progress (see also Wells, 1982). Unlike Wells (1982), Tolfree (1999) did not find vocalisation to be associated with an increase in labialisation. Gradience in vocalisation has also been reported in Southern British English speakers, such that the tongue tip raising gesture is reduced to varying degrees across speakers (Scobbie and Pouplier, 2010; Strycharczuk et al., 2020). While vocalisation is most strongly linked to a change spreading from Southern British English, historically, vocalisation has been reported much more widely (Johnson and Britain, 2007). Illustrating this point, Johnson and Britain (2007) point to the dropping of /l/ in all dialects of British English, in words such as *calm* and *talk*.

2.4 Laterals within Articulatory Phonology

I here draw together the elements of the previous sections by considering laterals from an Articulatory Phonology framework. An AP account of laterals begins with Sproat and Fujimura (1993)'s description of laterals as comprising two gestures. These include an apical, consonant-like gesture, and a dorsal vowel-like gesture, with the relative timing and magnitude of these gestures mediating the clearness or darkness of the /l/ (Sproat and Fujimura, 1993). The multi gestural nature of /l/ offers scope for interpretation in terms of how it may be conceptualised within a gestural model. In particular, there are a number of possibilities for the type and strength of gestural coupling relations in /l/.

One possibility is that different coupling relations, of varying strengths exist for the composite gestures of /l/ within different morphological environments (Lee-Kim, Davidson, and Hwang, 2013; Scobbie and Pouplier, 2010). Lee-Kim, Davidson, and Hwang (2013) propose context mediated coupling relations for /l/ as an explanation for gradient patterns of increasing /l/ darkening with morphological boundary strength. In their data, /l/ darkening, captured by increasing tongue body lowering, was found to proceed from most to least dark for word final stem position (e.g., *cool*), word medial pre boundary (e.g., *coolest*), and word medial post boundary (e.g., *coupless*). Further, the following coupling relations were proposed to capture the varying levels of darkness with morphological context (see also Table 2.2 for a summary of these relationships). For /l/ in stem final contexts such as *"cool"*, the TT and TD gestures of /l/ are coupled to each other, and the TD gesture is also coupled to the preceding vowel, hence pushing the TD gesture towards the vowel to temporally precede the TT gesture, as is typically found for coda /l/s. For post-boundary contexts such as *"coupless"* where there is only a following vowel, the TT and TD gestures of /l/ are coupled to one another, and the TT gesture is coupled to the following vowel. This results in an onset-like pattern whereby the TT precedes the TB gesture. For a pre-boundary context, such as *"coolest"*, the TT and TD gestures of /l/ are coupled to each other, the TD gesture is coupled to the preceding vowel, and the TT gesture is coupled to the following vowel, resulting in an onset and coda like configuration. The idea pursued here, that /l/ darkening is mediated by varying coupling relations between the gestures of /l/ has important implications for the nature of the relationship between spatial and temporal variation in /l/.

/l/ context	Example	TT+TD	TD+pre V	TT+post V
Word final stem	*cool*	Y	Y	N
Word medial pre boundary	*coolest*	Y	Y	Y
Word medial post boundary	*coupless*	Y	N	Y

Table 2.2: Context mediated coupling relationships of gestures of /l/, summarised from Lee-Kim, Davidson, and Hwang (2013).

Another interpretation of /l/ within an AP framework is provided in the lateral cluster timing analyses of Goldstein et al. (2009). In this study, the authors found a comparatively smaller rightward shift of /l/ towards the vowel in lateral onset clusters, compared to that of segments in non-lateral onset clusters, hence, /l/ clusters showed a non C-centre pattern. To explain such findings within an AP framework, the authors suggested that both the apical and dorsal gestures of /l/ (as C2 in an onset cluster) were coupled to the following vowel. This was in contrast to non lateral segments measured which were made up of a single gesture and had only a single coupling with the vowel. The result of /l/'s multiple couplings to the vowel was a higher overall coupling between the lateral and the vowel than the non lateral C1. Greater coupling increases the stability between the lateral and the vowel, hence resulting in comparatively reduced movement of the lateral towards the vowel. Using computational modelling, the authors found that the empirical data was well replicated when the coupling strength parameter between /l/ and the vowel was increased.

Finally, in discussing laterals within a gestural account of phonology, it is useful to also draw parallels with other multi-gestural segments. As reviewed extensively by Krakow (1999), laterals and nasals show similar patterns of gestural timing and magnitude in syllable onsets and codas (Browman and Goldstein, 1995; Krakow, 1989; Sproat and Fujimura, 1993). While /l/ is reported to show tip delay in coda position (Browman and Goldstein, 1995; Sproat and Fujimura, 1993), the lip and velum gestures in /m/ are reported to show a similar pattern. Krakow (1989) shows how, relative to the lip gesture, the velum gesture of /m/ is timed earlier and has a greater magnitude in coda position compared to in onset position. Such patterning at the inter gestural level is analogous to the widely reported syllable level patterns of coordination, namely the in-phase timing pattern in CV onsets, and the anti-phase pattern in VC codas. In both cases, the more consonantal gesture (the TT raising in /l/ and the lip closure in /m/) occurs relatively earlier in onsets and relatively later in codas. This is suggestive of a broader pattern of coordination which operates at multiple structural levels (Nam, 2007). From this perspective, it may be reasonable to hypothesise that the coordination patterns of multi-gesture segments may be predicted from broader patterns of syllable coordination.

2.5 Summary and Research Questions

This chapter has outlined the core principles of Articulatory Phonology, including the notion of the gesture – where spatial specifications of articulatory constrictions are made, and the coupled oscillator model – where patterns of temporal coordination are modelled. The relationship between spatial and temporal variation in the AP model was also discussed. While the model predicts there to be no relationship between spatial and temporal variation (for example, between the constriction location of a gesture and the timing of that gesture), empirical evidence for this is mixed. This thesis aims to provide further clarity to this matter. I proposed laterals as a suitable means of testing the nature of this relationship given that laterals exhibit systematic variation across language varieties and contexts. Specifically, by comparing the /l/-cluster timing patterns of two English dialects which exhibit systematic differences in /l/ darkening, this thesis will assess the robustness of the predictions of a feed-forward model. Before addressing this question, some methodological groundwork is first required to determine the articulatory nature of differences in /l/ darkening between dialects. The research questions of each Chapter are below presented in turn.

RQs: Chapter 4

- RQ1: *What is the nature of the articulatory differences in /l/ darkening between Onset Darkening and Onset Lightening dialects?*
- Sub RQ: *How does the vowel and morphosyntactic context of /l/ interact with dialectal effects on /l/ darkening?*

Evidence shows an acoustic difference in /l/ darkening between Onset Lightening and Onset Darkening dialects. Articulatorily, darker /l/ is typically characterised by greater tongue body lowering, relative to lighter /l/, and a reduced or later tongue tip gesture in coda position. What is less clear from cross dialectal studies is the articulatory mechanism driving the dialect contrast in /l/ darkening, be it the magnitude of tongue body lowering, or differences in the relative timing between gestures. Chapter 4 seeks to address this question through a dialect comparison of /l/ darkening across a range of vowel and morphosyntactic contexts. A particular focus is placed on establishing the nature of spatial articulatory differences in /l/ darkening between dialects, focussing particularly on tongue body lowering, which has been found to be a robust correlate of /l/ darkening (e.g., Sproat and Fujimura, 1993), across multiple vowel contexts (Lee-Kim, Davidson, and Hwang, 2013). Identifying the nature of spatial differences will serve subsequent chapters which examine the relationship between spatial differences in /l/ darkening and patterns of /l/ cluster timing.

A further sub research question is posed within Chapter 4 to address the interaction between /l/ darkening and morphosyntactic / vocalic context. This question is motivated by previous findings which show interactions between /l/ darkening and vocalic/morphosyntactic context (e.g., Lee-Kim, Davidson, and Hwang, 2013; Strycharczuk, Derrick, and Shaw, 2020; Strycharczuk and Scobbie, 2015; Turton, 2014) (See Chapter 4 for a more detailed review). In light of this interaction, /l/ is examined in a range of morphosyntactic and vocalic environments, with the aim to tease apart the effects of dialect on /l/ darkening from effects of vowel and morphosyntactic context. The specific vowel and morphosyntactic contexts examined here were based on the stimuli of Strycharczuk, Derrick, and Shaw (2020), which included three vowel contexts, FLEECE, KIT and THOUGHT, and five morphosyntactic contexts: word initial, word medial morpheme internal, word medial morpheme final, word final pre vocalic, and word final pre consonantal. These contexts were found to elicit a wide range of variation in the laterals of the New Zealand English speakers within Strycharczuk, Derrick, and Shaw (2020)'s study, including clear, dark and vocalised realisations, hence were considered suitable for the purpose of this investigation.

RQs: Chapter 5

- RQ2: *How do spatial articulatory differences in /l/ darkening between dialects affect patterns of /l/ cluster timing in onset and coda clusters?*

Once the nature of the spatial articulatory differences in /l/ darkening between dialects has been established, Chapter 5 asks how spatial articulatory differences in /l/ darkening interact with patterns of /l/ cluster timing. This analysis will examine /l/ clusters in front and back vowel contexts, allowing any potential interactions between dialect and vocalic context to be exposed.

This question has important implications for how spatio-temporal variability is modelled within the coupled oscillator model of speech timing. As reviewed above, the AP model predicts patterns of temporal coordination to be unaffected by spatial variability. The prediction from the standard AP model for these data then is that articulatory differences in /l/ darkening between dialects, such as the spatial position of the tongue should not result in different patterns of timing in /l/ clusters. On the other hand, there is some evidence that this relationship is not so clear cut, as will be reviewed in Chapter 5. Most salient to this study is the finding for cross-linguistic differences in /l/ cluster timing patterns across varieties where /l/ differs in darkness (Marin and Pouplier, 2014). This finding suggests that there may in fact be a relationship between the spatial properties of /l/ which condition its darkness and patterns of cluster timing.

RQs: Chapter 6

- RQ4: *How can an analysis of both the tongue tip and tongue body gestures of /l/ enhance our understanding of the relationship between /l/ darkening and /l/ cluster timing?*
- Sub RQ: *How do patterns of inter-gestural timing in /l/ differ between dialects and singleton vs cluster contexts?*

Chapter 6 undertakes a restricted analysis of cluster timing on a subset of the data where both the tongue tip and tongue body gestures of /l/ can be reliably identified. Literature reviewed here have shown the timing and magnitude of the tongue body gesture of /l/ to be important correlates of /l/ darkening (e.g., Sproat and Fujimura, 1993), and to have potential implications for the nature of the inter and intra gestural coupling relations (Lee-Kim, Davidson, and Hwang, 2013). Considering /l/ darkening from this gestural perspective (Lee-Kim, Davidson, and Hwang, 2013), it is reasonable to speculate that differences in /l/ darkening between the dialects of this study may result from differences in coupling relations between gestures of /l/, which may have subsequent implications for the broader patterns of lateral cluster timing between dialects.

This chapter thus seeks to resolve important methodological issues encountered in previous chapters which prevented both the tongue tip and tongue body gesture from being included in measures of lateral timing and lateral cluster timing. The broad research question for this analysis asks how considering both the tongue body and tongue tip gestures of /l/ can further our understanding of the relationship between spatial differences in /l/ darkening, and timing patterns in clusters. Within this analysis, patterns of inter gestural timing are compared across dialects and singleton versus cluster contexts, offering insight into the gestural composition of /l/ across varieties and contexts.

Chapter 3

Methods

All data presented in this thesis are from a single experiment which lasted between 1 to 1.5 hours per speaker. Speakers were ask to read two sets of stimuli within the experiment. The first set is used in Chapter 4, the second set is used for Chapter 5, and a subset of the second stimuli set is used in Chapter 6.

3.1 Speakers

This thesis presents data from 16 speakers, aged between 18 and 35 years. All speakers were recruited from Lancaster University via internally distributed emails, poster advertisements around campus, and word-of-mouth. Twenty one speakers were recorded in total, however, three speakers were used for pilot data, and two speakers were not included due to post processing issues. All speakers were native monolingual English speakers, and lived in Lancashire at the time of the experiment. Speakers were evenly distributed between the broad dialectal regions of Lancashire / Manchester English and Southern Standard British English (SSBE), 8 speakers per dialect. Dialectal identity was self reported by the speaker; however, all speakers were required to have been native to the dialect region. Specific regions of speakers who broadly identified as having a Lancashire / Manchester English dialect included Blackpool, Stockport, Oldham and Burnley; specific regions of speakers who broadly identified as having a SSBE dialect, included Norwich, Bristol, and Buckinghamshire. Data on biological sex was not collected, however an estimate of vocal tract size based on formant data is calculated for each speaker and is used within statistical models to account for anatomical differences. All speakers were reimbursed for their participation.

3.1.1 OLD and ODD speakers

Throughout this thesis, the two broad dialect regions, SSBE and Lancashire / Manchester English, are referred to as either the Onset Lightening dialect, or the Onset Darkening dialect respectively. The Onset Lightening dialect will be used to refer to Southern Standard British English speakers, since /l/ is reported to be clear in onset position for these speakers (Turton, 2014). The Onset Darkening dialect will be used to refer to Lancashire / Manchester English, since /l/ is reported to be dark in onset position for these speakers (Hughes, Trudgill, and Watt, 2012; Turton, 2014). The acronyms also give a faint nod to /l/ darkening being a change in progress, whereby OLD speakers may be considered the old, non darkened variety.

3.1.2 VT length estimation

Estimates of vocal tract length were calculated for each speaker using a method of average formant spacing (Johnson, 2020). Vocal tract length (VT length) estimates were created in light of absent information on speakers' biological sex. VT length is reported to differ with speaker sex, with males having, on average, longer vocal tracts than females, with differences occurring at puberty (Barbier et al., 2015). Differences in VT length between males and females have been shown to correlate with differences in gestural magnitude and velocity, such that the larger VT length of males results in larger and faster lingual movements (Simpson, 2001). Given that this study's main goal is explore spatio-temporal relations in speech, gestural magnitude and velocity are both key variables for which we wish to control. An estimate of VT length was thus considered an appropriate method for capturing a potentially important confound for this investigation within the speaker sample.

To derive an estimate of VT length for each speaker, I used the following calculation from Johnson (2020) (specifically, formula 7 from Johnson, 2020) which measures the average spacing of mean F1, F2 and F3 values across vowel tokens.

Coefficients are weighted across formants according to their contribution to the VT length estimate.

$$\text{deltaF} = 0.6667 * \text{mean (f1)} + 0.222 * \text{mean (f2)} + 0.133 * \text{mean (f3)}$$
$$\text{vtl} = 34000 \ / \ (2 * \text{deltaF})$$

Average formant spacing equation from Johnson (2020)

The input to the formula consisted of midpoint frequencies of F1, F2, and F3 from a total of 17,516 vowel tokens from (average of 1094.75 tokens per speaker). Vowels were extracted from all the entire spoken stimuli of each participant. This included both the carrier phrases and target words detailed in Section 4.3.2, plus some additional sentences not included within this thesis. As a result, a wide range of vowels were included which was important to the validity of the calculation. Segmentation was performed by Montreal Forced Aligner (McAuliffe et al., 2017) and formant midpoints were extracted in Praat (Boersma and Weenink, 2011) using the FastTrack plugin (Barreda, 2021).

Tables 3.1 and 3.2 show VT length estimates for speakers of the Onset Lightening and Onset Darkening dialects respectively. To get a sense of how estimated VT lengths are distributed across dialects, Figure 3.1 shows a plot of the estimated VT length of each speaker across the two dialects; dialect is shown by colour. Figure 3.1 shows, across the two dialects, VT lengths largely cluster into two groups, a high VT length group - above 16cm, and a low VT length group - below 15cm. There is one speaker, S08, who falls mid way between these groups with a VT length of 15.3cm. In the higher range (16-17cm), there is a larger proportion of Onset Darkening speakers (4 OD speakers vs 1 OL speaker), while in the lower range, (below 15cm), there is a larger proportion of Onset Lightening speakers (6 OL speakers vs 4 OD speakers). Further, VT length estimates are not balanced across dialect groups, hence VT length will be included as a fixed effect within all statistical models.

Speaker code	Dialect	est. vtl (cm)
S01	OLD	16.2
S02	OLD	14.0
S03	OLD	13.3
S04	OLD	13.4
S05	OLD	14.5
S06	OLD	14.2
S07	OLD	14.1
S08	OLD	15.3

Table 3.1: Table showing the speaker code and estimated vocal tract length of each Onset Lightening speaker.

Speaker code	Dialect	est. vtl (cm)
L01	ODD	14.3
L02	ODD	14.1
L03	ODD	14.1
L04	ODD	14.6
L05	ODD	17.0
L06	ODD	16.0
L07	ODD	16.3
L08	ODD	16.8

Table 3.2: Table showing the speaker code and estimated vocal tract length of each Onset Darkening speaker.

3.1.3 Stimuli

Two sets of stimuli were included within the experiment, Set 1 is used within the analysis in Chapter 4, and Set 2 is used within the analysis in Chapter 5, a subset of which is also used in Chapter 6.

Set 1

Stimuli Set 1 (Table 3.3) contains target /l/ words in 5 morphosyntactic contexts (initial, mono morphemic intervocalic, pre-boundary intervocalic, word-final preceding a vowel, and word-final preceding a consonant), and 3 vowel contexts (FLEECE, KIT, THOUGHT). Target words were embedded within a carrier phrase *"Say the (target word) again"* Each sentence was repeated 4 times.

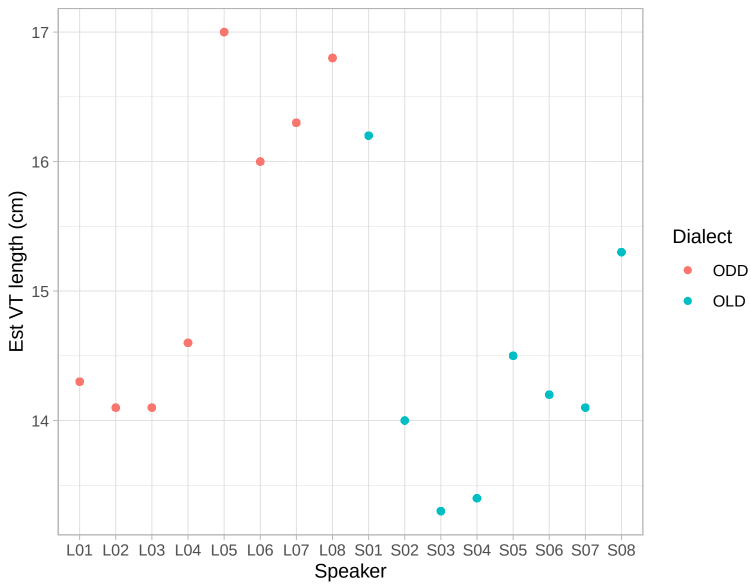

Figure 3.1: Estimated VT lengths plotted for each speaker. Dialect is further indicated by colour: Onset Darkening speakers are in red, and Onset Lightening speakers are in blue.

Table 3.3: Stimuli Set 1

	Initial	Mono-morphemic intervocalic	Pre-boundary intervocalic	Word final pre-vocalic	Word final pre-consonantal
FLEECE	leap	helix	healing	heal it	heal mick
KIT	lip	fillet	filling	fill it	fill mick
THOUGHT	law	paula	mauling	maul it	maul mick

Set 2

Stimuli Set 2 (Table 3.4) contained a series of singleton-cluster /l/ pairs (e.g., *clip* - *lip*), within the carrier phrase *say the xxx*, or *say xxx again*. Target words included both front and back vowels. Cluster contexts included /pl-/ and /kl-/ contexts for onset tokens, and /-lp/ and /-lk/ contexts for coda tokens. All tokens were repeated 4 times.

Table 3.4: Stimuli Set 2

Onset singleton	Onset cluster	Coda singleton	Coda cluster
tea lip	tea clip	mill in	milk in
tea lug	tea club	gull it	gulp it
tea lick	tea plick	fill it	philp it
tea lug	tea plug		

3.2 Data collection

3.2.1 Electromagnetic articulography

Electromagnetic articulography is the articulatory measure used throughout this thesis. Electromagnetic articulography (EMA) is point tracking method which tracks the positions of sensors glued onto articulators during the production of speech. Positions of sensors are calculated using electromagnetic principles. A weak electromagnetic field is generated around the subject's head, and movements of the sensors, which are made from small metal coils, create detectable disturbances to the field. For the AG501 model used here, sensor positions are recorded in three dimensions (x, y, and z), plus two angular coordinates, phi and theta, which provide information about the directionality of sensor movement. Positions were recorded at a frame rate of 1250Hz, which was subsequently downsampled to 250Hz, sufficient for capturing fast tongue movements. Sensor positions are visible in real time using the AG501 *cs view* real time display, and can be exported for further processing and analysis.

3.2.2 Experimental procedure

Audio synchronised electromagnetic articulography data were collected in the Lancaster University EMA Lab. Data for all studies were recorded within a single session. Participants were briefed about the nature of the experiment before the experiment took place, but were naive to the specific purpose of the experiment.

Prior to the experiment, each participant completed a consent form and a questionnaire to check for latex allergies. Latex was used to coat the sensors and latex gloves were worn by the experimenter. In addition, participants were asked whether they used a pacemaker, given the potential for interference here (Rebernik et al., 2021) – see Appendix 1, Section 9.1 for the relevant section of the consent form provided to participants. Acoustic and articulatory recordings were made using the Carstens AG501 electro-magnetic articulograph, and a DPA 4006A microphone. During the experiment participants were seated under the articulograph, (see Figure **??**). The height of the articulograph was adjusted to ensure the subject's head was within an appropriate range. Participants were asked to read sentences from a computer monitor at a normal pace while their speech and articulatory movements were recorded. The monitor was positioned at approximately eye-level, 1 meter in front of the participant. Each sentence was shown on an individual Power Point slide, and each constituted a single (EMA and audio) recording. Sentences were written in a clear font (pt. size 44) on a white background. Because all sentences were similar, the text colour of sentences alternated between black and dark grey between slides to differentiate sentences and reduce eye fatigue. The sentence presentation and recordings were both controlled manually by the experimenter, the former from a desktop computer, and the latter from a Dell laptop. Each sentence contained a target word embedded within a carrier phrase e.g., *Say tea xxx again / Say the xxx it again*. Recording sessions, including sensor attachment and removal, took between 1 and 1.5 hours per speaker. Opportunities for breaks and water were provided throughout the experiment.

3.2.3 Sensor preparation and attachment

To prepare sensors for use in the experiment, sensors were first calibrated using the AG501 calibration software. Calibrated sensors were then labelled from 1-16 according to their position number in the magazines. To protect the sensors, each sensor was covered in liquid latex, as recommended by the articulograph manufacturers, and left to dry overnight prior to the experiment. Between participants, the sensors and bite plate were sterilised in a sterilising solution and rinsed.

Sensors were glued to articulators using EpiGlu, a dental glue also recommended by the articulograph manufacturers. For the most part, the glue secured sensors in place for the duration of the experiment, but could be removed with relative ease after the experiment. On occasions where a sensor did fall off during the experiment, the time at which it fell off was noted down, and the sensor was reattached using the dried glue residue as a guide for the place of reattachment.

Lingual sensors were first attached to three mid sagittal points on the tongue. The tongue sensors were considered the most likely of the articulatory sensors to fall off, hence attaching them first increased the chance that they would fall off before the experiment began, rather than during the experiment. To begin, the dorsal sensor was attached and positioned as far back as was comfortable for the participant. The tongue tip sensor was then glued approximately 1cm behind the anatomical tongue tip when the tongue was in a stretched position. A sensor could not be attached directly to the tongue tip due to the substantial impediment to speech this would cause. The tongue body sensor was then glued equidistant between the tongue tip and dorsum sensors. To ensure that the sensors stuck to the tongue surface, the tongue was first dried by the participant using paper napkins and a cool hairdryer immediately prior to the sensor being glued in place. While the participant was drying the area, the back of the relevant sensor was covered in glue. Once the area had been thoroughly dried, the sensor was carefully placed onto the tongue and secured into place using by applying a little pressure onto the sensor using a tongue depressor. Participants were then asked to check that the sensors felt comfortable and secure.

Sensors were also glued mid sagittally to the gingival surfaces above the upper and below the lower incisors. This was more or less difficult depending on the size of gingival surface, which was found to be highly variable between speakers. Prior to attachment, speakers were asked to stretch out their lip to make the gum accessible, and dry the area with a paper napkin. To capture lip movement, sensors were also affixed mid sagittally to the vermillion boarder of the upper and lower lips. Since this was an external sensor, no drying was required. Caution was taken not to

(a) Visible sensor positions. (b) A single sensor.

Figure 3.2: Sensors

get glue on the lips themselves, as this may have caused damage to the lips and discomfort to the participant upon removal.

Additional reference sensors were also attached to safety goggles worn by the participant during the experiment. Reference sensors are sensors which are attached to non-mobile structures of the head, allowing sensor positions to be corrected for head movement. Reference sensor locations included the gingival surface above the upper incisor described above, and three sensors taped to a pair of safety glasses at the nasion and the left and right mastoids. Glasses were secured into position via a ribbon tied around the participant's head to avoid movement of the glasses during the experiment. The reason for attaching the external reference sensors to safety glasses was that in previous experiments, I found that reference sensors were generally more secure and moved less when attached to glasses compared to when affixed directly to the participant's skin, (for another example of this, see Thompson and Kim, 2019).

3.2.4 Other steps

During the experiment, participants wore a grounding clamp around their ankle to prevent interference (Medizinelektronik, 2014). Loose wires were taped to participants' clothing, while ensuring that there was enough slack in the sensor wires for minor movements. A bite plate recording was made in order to rotate sensor positions to the speakers' occlusal plane. Three sensors were taped to the Carstens bite plate and a recording was made while the participant bit down on the plate. Sensor positions were rotated to the bite plane during post experiment processing. To obtain the shape of the speaker's palate, each participant was asked to tape a sensor to their thumb and trace their thumb along the mid sagittal plane of their palate while a recording was made.

3.3 Processing

Articulatory data was recorded at a frame rate of 1250Hz on the AG501, and subsequently downsampled to 250Hz. Initial data processing was performed using the AG501 software. Positions were first calculated, before head correction was applied and data was rotated to the speakers' occlusal plane. Data files were then converted into text files for further analysis. Speech sensors were filtered using a Kaiser-windowed low pass filter at 40 – 50Hz, and reference sensors were filtered at 5Hz. Audio recordings were made using a DPA 4006A microphone, which was connected to a Alesis io2 express audio interface and recorded onto the EMA laptop at 44.1kHz.

Acoustic data was annotated in Praat (Boersma and Weenink, 2011). For each sound file, which contained a single token, a TextGrid was created and the sentence was manually transcribed. Segmentation was performed using Montreal Forced Aligner (McAuliffe et al., 2017) and hand corrected and relabelled where necessary. Where /l/ was segmented from the following vowel (Chapter 4), segmental boundaries were placed at the point of an overall change in spectral energy. Depending on the data set, a different annotation scheme was implemented (for example, for the cluster timing data, an alphanumeric system was used).

Articulatory data from each speaker underwent further processing in R, using the *tadaR* (Kirkham, 2024a) and *tardis* (Kirkham, 2024b) packages. A Butterworth filter was applied to sensor data with a filter order of 5 and filter cut off of 20Hz. A "time" column was created for each token which began at 0 and proceeded in increments of 1/250 (where 250Hz was the frame rate). A derived measure of lip aperture was calculated, from the x, y, z coordinates of upper and lower lip sensor, as well as the tangential velocities of tongue sensors (in xz, and xyz dimensions). Audio files were read into R, and for each speaker, EMA files and corresponding audio files were combined into a single data frame and subjected to various visual and statistical analyses, which are outlined within the methodology of the relevant studies.

3.3.1 Normalisation of articulatory data

Despite efforts to standardise sensor placement, there is necessarily a degree of speaker-variation in the location of sensors due to differences in tongue size and participant comfort. Steps were thus taken to normalise articulatory data. All articulatory data were rotated to the speaker's occlusal plane, as outlined above. Articulatory displacement data were then further z-scored (scaled, and centred) by speaker unless otherwise specified, such as in Chapter 5, Figure 5.23, where data is centred by speaker – but not scaled – so as to preserve the magnitude of displacement. To control for anatomical differences in vocal tract size, an estimate of vocal tract length described above was included as a fixed effect in all statistical models.

3.4 Evaluation of EMA

This section evaluates the appropriateness of Electromagnetic Articulography (EMA) as a method of obtaining acoustic-articulatory data for the purposes of this thesis. EMA is a point tracking technique which tracks the dynamic movement of sensors attached to key locations on articulators involved in speech, such as the tongue and the lips. Other point tracking techniques include optical tracking, which tracks points on external articulatory structures, and x-ray micro-beam, which is now widely considered to pose too great a health risk for use within a research context (e.g., Kochetov, 2020a,b). An alternative to point tracking techniques is articulatory imaging, where images can be recorded of the tongue surface, as with ultrasound imagining (UTI), or of entire articulators, as with magnetic resonance imaging (MRI). Further articulatory techniques are also available, including electropalatography (EPG) which records lingual-palatal contact using electrodes on an artificial palate. In a survey of articulatory methods used in articles published in well regarded phonetic journals between 2000 and 2019, EMA was used in 29% of cases - more than any other articulatory technique (Kochetov, 2020a). This is not to say that EMA is always the most useful method for the study at hand. This section proceeds under the premise that the utility of a method can only be evaluated in light of the research questions it serves to address (Léger et al., 2024). In the following sections, I discuss the extent to which the methodological requirements of this investigation are met by EMA. I focus in particular on the requirements for (i) measuring the articulatory correlates of /l/ darkening, and (ii) measuring /l/ cluster timing.

Measuring /l/ darkening using EMA

In order to establish a difference in darkening between dialects, the method must be able to precisely measure the timing and displacement of tongue tip, tongue body, and tongue dorsum gestures. This requires: (i) access to the tongue tip, tongue body and tongue dorsum, and (ii) a high temporal resolution, given that measures of tip delay are often very small, for example, a tip delay of 0.005 seconds was reported for one speaker in Sproat and Fujimura (1993).

Relative to MRI and ultrasound imaging, EMA has relatively limited access to the tongue surface. Given that sensors must be glued directly onto the articulators, positioning of sensors is largely dictated by participant comfort (Rebernik et al., 2021). For example, gag reflexes mean that sensors can only realistically be affixed the anterior portion of the tongue. In addition, the impediment to speech caused by the sensors (Meenakshi et al., 2014) also means that sensors cannot be glued to the very apex of the tongue. Instead, sensors are typically glued approximately 1cm behind the tongue tip (e.g., Marin and Pouplier, 2010; Rebernik et al., 2021). There is also a limit on the proximity of sensors, with a minimum distance of 1cm required between sensors, if interference between sensors is to be avoided (Medizinelektronik, 2014). Addressing the requirement for a high temporal resolution, EMA has a high frame at 250Hz for the Carstens AG501 relative to other methods such as ultrasound imaging, which has a frame rate of around 87 frames per second (for the Articulate Instruments' Echo B system). EMA also permits simultaneous recording of high quality audio data, which enables acoustic and articulatory correlates of lateral darkening to be compared.

Evaluating EMA, in relation to the methodological demands of measuring /l/ darkening between dialects, there is a clear spatial-temporal tradeoff. While EMA has the advantage of a high temporal precision it is considerably limited in the positioning of sensors and its ability to capture movements of only small fixed points. Ultrasound imaging is somewhat more informative in this respect, offering data on the movement profiles of the tongue surface contour, though shadows may obscure visability of edges.

Measuring /l/ cluster timing using EMA

To measure lateral cluster timing, simultaneous capture of the high speed movement of multiple gestures from multiple articulators is required, including those of the tongue and the lips. As a point tracking method, EMA affords the capture of simultaneous spatial and temporal information from multiple articulators with high temporal precision (250Hz). Unlike Ultrasound, which captures the tongue surface only, EMA sensors can be attached to both internal and external structures, so long as sensors are within range of the electromagnetic field (compare for example motion detection techniques, where tracked points must be visible) (Rebernik et al., 2021). In this sense, EMA is highly appropriate for the study of gestural timing.

Overall evaluation

Given the centrality of timing in this thesis, EMA's high temporal resolution, along with the ability to simultaneously track the positions of multiple articulators makes EMA a suitable method for the purposes of this study. The disadvantages of EMA should be acknowledged, for example, the perceptually detectable effect of sensors on speech (Meenakshi et al., 2014), and that EMA informs only on the positions of singular fixed sensors which are limited in their placement by participant comfort (Rebernik et al., 2021). However, in light of the goals of this thesis, these limitations are considered to be outweighed by EMA's temporal precision which is here crucial to the capture of small timing differences between gestures.

Chapter 4

Dialect Variation in /l/ Darkening

The aim this chapter is to establish the articulatory mechanism underlying acoustic dialectal differences in /l/ darkening. Previous literature shows clear acoustic differences in /l/ darkening between Onset Lightening and Onset Darkening dialects. Relative to lighter /l/s, darker /l/s typically show greater tongue body lowering, as well a later apical raising gesture in coda position. What is less clear, is which articulatory mechanism drives the /l/ darkening contrast between dialects.

Establishing the articulatory nature of dialectal differences in /l/ darkening is fundamental to the subsequent studies of the thesis, which ask how articulatory differences in /l/ darkening between dialects affects patterns of /l/ cluster timing. In pursuit of the thesis' goal: to investigate the nature of the spatio-temporal relationship in laterals, a particular focus will here be placed on establishing the nature of spatial differences in /l/ darkening between dialects. This will allow subsequent chapters to ask how dialect-mediated spatial differences in /l/ affect patterns of /l/ cluster timing, hence allowing an explicit investigation into the spatio-temporal relationship of laterals. Temporal measures will also be discussed as it is also relevant to the spatio-temporal focus of the thesis to consider lateral-internal spatio-temporal relations.

One difficulty in establishing articulatory differences in /l/ darkening between dialects is that lateral darkening is affected by factors such as the strength of the following boundary (e.g., Sproat and Fujimura, 1993), and vocalic context (e.g., Mackenzie et al., 2018; Strycharczuk, Derrick, and Shaw, 2020). This makes it difficult to make overarching claims about the darkness of /l/ in a given variety. In order to disentangle these effects of vowel, morphosyntactic context and dialect, this chapter examines /l/ in a range of morphosyntactic and vocalic positions, designed to elicit a wide range of variation.

The chapter will begin by orienting the analysis within the broader literature. Here, I provide a review of the nature and scope of variability in /l/ darkening, including the effects of vowel and morphosyntactic context on /l/ darkening. I then discuss the practical matter of how to measure /l/ darkening, where I cover both acoustic and articulatory measures, some of which will be implemented within this analysis.

4.1 A review of lateral darkening: Measures, and variability

4.1.1 The articulation and acoustics of laterals

In discussing variation in lateral darkening, we must first understand how laterals are structured in articulation and the consequences of this for acoustics. English laterals are produced using a raised tongue tip which makes contact with the alveolar ridge, a lateral channel, which is formed through narrowing of one or both sides of the tongue blade, and a retracted and lowered tongue body (Stevens, 1998, p. 543). There are two airways described for this configuration: a main airway, which is the lateral channel formed through the narrowing of the tongue blade edge, and a side branch – the airway formed in the mid-sagittal region below the palate (Stevens, 1998, p. 545); (Johnson, 2012, p. 196); (Fant, 1971). The side branch results in anti-formants around the F3 frequency region (Johnson, 2012, p. 198). F2 is considered the resonant frequency of the region behind the apical constriction (Stevens, 1998, p. 554), and hence corresponds to the degree of tongue retraction; greater tongue retraction, characteristic of a darker /l/, increases the length of this cavity, resulting in a lowering of F2. The frequency of F3 is mostly mediated by the resonant frequency of the frontal cavity, anterior to the constriction at the tongue apex (Stevens, 1998, p. 555).

Lateral vocalisation

Lateral vocalisation occurs when the apical gesture is reduced or altogether absent from the lateral. Vocalisation in London speakers has also been associated with increased labialisation (Wells, 1982) and has been observed to occur in increasing levels in varieties of English such as SSBE (Hardcastle and Barry, 1989; Trudgill, 1986; Turton, 2014) and American English (Wrench and Scobbie, 2003). One view is that vocalisation is an extreme form of gestural undershoot of the apical gesture in /l/ darkening (Strycharczuk, Derrick, and Shaw, 2020). From this perspective, it is possible that vocalisation is an extreme form of /l/ darkening, which follows from the relationship between gestural timing and gestural magnitude. In support of this hypothesis, Strycharczuk, Derrick, and Shaw (2020) highlight that lateralisation of /l/, which is reduced in dark /l/ compared to clear /l/, is reduced even further in vocalised /l/, suggesting a trajectory of darkness. Further support for the hypothesis that vocalisation is an extreme stage of darkening comes from the observation that vocalisation only occurs in varieties with a dark /l/, and never occurs in clear /l/ varieties (Johnson and Britain, 2007); however, this is not the case cross-linguistically (e.g., Vollmann et al., 2017). Methodologically, vocalisation warrants a different approach to non-vocalised laterals given the reduction or absence of the tongue tip gesture.

4.1.2 Contexts of variability in lateral darkening

Within a given variety, lateral darkening can be mediated by morphosyntactic context (e.g., Sproat and Fujimura, 1993), and vocalic context (e.g., Mackenzie et al., 2018; Strycharczuk, Derrick, and Shaw, 2020). This section outlines findings of variability in /l/ darkening as mediated by contextual factors of morphosyntactic position and vocalic context.

Effects of morphosyntactic context on /l/ darkening

In the seminal X-ray Micro-beam study, Sproat and Fujimura (1993) found a relationship between lateral darkening and prosodic boundary strength. In their study of four American English, and one British English speaker, variation in lateral darkening was elicited by positioning /l/ within a range of morphosyntactic contexts that varied in boundary strength. Greater tongue body lowering and retraction was reported for /l/ at stronger boundaries – i.e., before an intonational boundary – than at weaker boundaries, such as in intervocalic position. Intermediate degrees of tongue body lowering and retraction were found at intermediate boundaries, such as in word final prevocalic position. The presence of increasing /l/ darkening with increasing boundary strength lay the foundations for the argument that lateral darkening is a gradient rather than categorical phenomenon. In this account, clear and dark /l/ are not distinct allophones, but contextual variants that arise due to coarticulation, prosody, and syllable structure.

Sproat and Fujimura (1993) further argued that the patterning of /l/ darkening with boundary strength could be accounted for by duration. This was explained with reference to the findings of Lehiste (1980) for rime duration to be longer at stronger phrase boundaries. At a stronger phrase boundary, /l/ is longer in duration and hence has sufficient time for the dorsal gesture to be fully realised; this in turn, results in a darker /l/. Conversely, at weaker phrasal boundaries, /l/ is shorter in duration and so has less time to achieve its dorsal target; the result of this is gestural undershoot and a clearer realisation of /l/. See Figure 4.1 for a visual summary of the relationship between /l/ darkness and phrase boundary strength. Other studies have however reported similar correlations between /l/ darkness and duration, but only for dark /l/s (Yuan and Liberman, 2009), which has been suggested as evidence for a more complex interplay between both gradience in /l/ darkness and categorical allophony (Turton, 2014).

Lee-Kim, Davidson, and Hwang (2013) challenge the conclusion that duration conditions the darkness of /l/ by pointing to patterns within Sproat and Fujimura's data which could not be explained by duration alone. Specifically, Lee-Kim, Davidson, and Hwang (2013) showed that the differences in /l/ darkening observed between productive and non-productive morpheme boundaries in Sproat and Fujimura's data were not be explained by a durational difference, hence must have been a product of morphology.

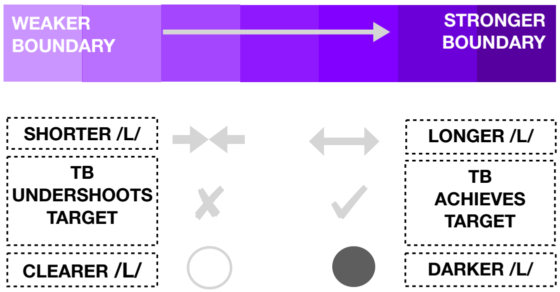

Figure 4.1: Visual summary of the relationship between /l/ darkness and strength of prosodic boundary discussed in Sproat and Fujimura (1993).

Patterns of gradient variation in lateral darkening have been reported across a range of English dialects (e.g., Lee-Kim, Davidson, and Hwang, 2013; Strycharczuk and Scobbie, 2015; Turton, 2014). In an ultrasound study on American English speakers, Lee-Kim, Davidson, and Hwang (2013) found gradient effects of morphological boundary on tongue body lowering / tongue dorsum retraction. Three contexts were examined: stem-final pre boundary *(tall-est)*, post boundary *(flaw-less)* and word final *(tall)*. Greatest tongue body lowering / tongue dorsum retraction was found for the word final context, which was followed by the stem final pre boundary context; least tongue body lowering / tongue dorsum retraction was found for /l/ in the post boundary position.

Similarly, gradient variation in /l/ darkening has been reported for SSBE speakers in Strycharczuk and Scobbie (2015). Within their study, degree of /l/ darkening was determined by a temporal measure of Tip Delay - the lag between the apical and the dorsal gesture of /l/. Incremental increases in tip delay, corresponding to a darker /l/, were observed across morphemic boundaries of increasing strength: word medial, word final intervocalic, and word final pre-consonantal contexts.

Turton (2014) reported both gradient and categorical patterns of lateral darkening across 10 morphosyntactic contexts for speakers of different English dialects.

Turton's acoustic and ultrasound study was one of the first to expose the extent of articulatory dialectal variability in /l/, and to offer evidence for both categorical and gradient processes of darkening. For example, while the RP speaker within her study showed clear categorical patterning of /l/ darkening in all acoustic and articulatory measures, the Manchester working class speaker showed categorical variability in some aspects, such as duration, but gradient variability in others, such as acoustics. A big contribution of Turton (2014)'s extensive analysis was that it highlighted the importance of examining /l/ within a full range of morphosyntactic contexts to establishing the range of /l/ variance.

Effects of vocalic context on /l/ darkening

Other studies have found that it is not just the morphosyntactic boundary strength that mediates lateral darkening, but also vocalic context. Lee-Kim, Davidson, and Hwang (2013), examined ultrasound data of American English laterals in five post vocalic contexts /u o ai au a/. The authors suggested that the back vowel contexts rendered the widely established measure of tongue dorsum retraction less useful, given that there was little room for further retraction from the already posterior position of tongue during the production of the vowel. This is not the case for a front vowel context where the tongue has ample room for retraction. From this, it is evident that the suitability of the measure of /l/ darkening is contingent upon the vocalic context.

Further studies have found interactions between vocalic context and morphosyntactic/prosodic context in patterns of lateral darkening. Such an interaction has been reported for American English (Mackenzie et al., 2018), SSBE (Strycharczuk and Scobbie, 2015) and New Zealand English (Strycharczuk, Derrick, and Shaw, 2020). In Strycharczuk, Derrick, and Shaw (2020), a greater range of lateral darkening was found for laterals in a FLEECE context, compared to that of KIT or THOUGHT contexts. Laterals in the FLEECE context were more sensitive to changes in the morphosyntactic environment than laterals in any other vowel context. Explaining this interaction Strycharczuk, Derrick, and Shaw (2020) note that the effect of boundary strength on lateral darkness is exaggerated in high front vowel contexts. This is because the coarticulatory effects of high front vowels are such that in word initial position, /l/ is particularly clear, while in word final position, coarticulatory effects of the preceding vowel are minimal. This means that relative to back vowels, the scale of darkening is greater for front vowel contexts because of their "lightening" effect on word-initial laterals. A greater scale of darkening means that differences in boundary strength have stronger effects.

The considerable variability in /l/ darkening, whether gradient or categorical, thus presents a challenge for comparing the degree of darkening between varieties. It highlights the importance of considering /l/ within a broad range of contexts in order to capture a sufficient range of variability – examining /l/ darkening within only front vowels, for example, offers insights of a very limited scope. The implication for this analysis is that, if the effects of dialect on /l/ darkening are to be established, they must be disentangled from effects of context. To do this, it is necessary to look across a multiple contexts. Further, this also carries the methodological implication that the measure of /l/ darkening must then be robust to the context induced variation.

4.1.3 Measures of lateral darkening

The following sections review established measures of /l/ darkening including acoustic measures and articulatory measures. Such will inform the measures of /l/ darkening implemented within the present analysis.

Acoustic measures

Acoustically, /l/ darkening is typically measured by the frequency of lower formant values, including F1, F2 and F3, as a consequence of their relationship with lingual configuration and cavity size/shape described above. Darker laterals are characterised by a higher F1 and lower F2, such that they occupy a closer spatial relationship on the spectrogram. Many studies have captured this relative relationship through a measure of F2−F1 (e.g., Kirkham, 2017; Sproat and Fujimura, 1993; Turton, 2014), while others have gained insights into lateral darkening from individual formants, such as F2 (e.g., Carter and Local, 2007), which broadly corresponds to the degree of tongue retraction (Stevens, 1998). Variability between studies is also observed in the place at which the measure is taken; for example, some studies measure the formant frequency from a single point in time, such as the lateral midpoint (e.g., Turton, 2014) or F2 minimum (e.g., Lee-Kim, Davidson, and Hwang, 2013), while others have taken a dynamic approach whereby multiple measures are taken across the duration of the lateral (e.g., Carter and Local, 2007).

Articulatory measures

Articulatory measures typically focus on either temporal or spatial aspects of lateral articulation. Temporal measures concern the relative timing of gestural events, and spatial measures concern the specific values of gestural displacement and magnitude. A common temporal measure of /l/ darkening is tongue tip delay, hereafter 'Tip Delay'. Sproat and Fujimura (1993) show systematic differences in relative gestural timing between the apical and dorsal gestures of laterals according to syllable position (initial vs final). Tip Delay is defined as the time lag between the maximal fronting and raising of the tongue tip and the maximal lowering of the tongue body. Sproat and Fujimura (1993) find Tip Delay to be negative for syllable final /l/ and positive for syllable initial /l/. This means that in syllable initial position, the apical gesture precedes the tongue body gesture, while in syllable final position, the tongue body gesture precedes the apical gesture. Explaining these results, the authors argued for general rules of attraction for consonantal and vocalic gestures. The apical gesture of /l/ is considered a consonantal gesture, since a substantial obstruction is made in the vocal tract, while the dorsal gesture is considered vocalic. Vocalic gestures are attracted to syllable nuclei, while consonantal gestures are attracted to syllable margins. This means that in syllable initial position, the dorsal gesture is attracted towards the following vowel and the apical gesture is attracted towards the left syllable margin. As a result, the apical gesture precedes the dorsal gesture syllable initially. In syllable final, post vocalic position, the dorsal gesture is attracted to the preceding vowel and the apical gesture to the right syllable margin, hence the dorsal gesture precedes the apical gesture. Through this line of reasoning, Sproat and Fujimura (1993) demonstrate a systematic relationship between gestural timing and lateral darkening (see also, Browman and Goldstein 1995; Lee-Kim, Davidson, and Hwang 2013).

Measures of Tip Delay have also been operationalised within ultrasound studies (e.g., Strycharczuk and Scobbie, 2015). In Strycharczuk and Scobbie (2015), Tip Delay was measured based on gestural minima, which were defined using displacement of the tongue contour along a vector specified in the relevant region of interest. Results revealed an interaction between morphosyntactic and vocalic context.

Spatial measures of /l/ darkening typically include dorsal raising and retraction (Giles and Moll, 1975; Sproat and Fujimura, 1993; Turton, 2015). Other related measures have also been used, such as a combined measure of post-dorsum lowering and pre-dorsum retraction (Recasens and Espinosa, 2005), and the magnitude of tongue body displacement (Lee-Kim, Davidson, and Hwang, 2013). Sproat and Fujimura (1993) have shown that, due to the hydrostatic properties of the tongue, tongue body lowering can be used as a robust correlate for the tongue dorsum retraction gesture in laterals. They note the tongue body lowering measure to be more reliable, with more easily identifiable extremums than dorsum retraction (Sproat and Fujimura, 1993). Similar comments on the robustness of tongue body lowering in measuring lateral darkening have also reported in Lee-Kim, Davidson, and Hwang (2013), Proctor et al. (2019), and Strycharczuk, Derrick, and Shaw (2020). Lee-Kim, Davidson, and Hwang (2013) show that vocalic context considerably affects the suitability of the measure used, meaning that tongue dorsum retraction is largely unsuitable for measuring /l/ darkening in non front vowel contexts (Proctor and Walker, 2012), while tongue body lowering is a more robust measure across vowel contexts.

Following from the similar findings of Sproat and Fujimura (1993), Strycharczuk, Derrick, and Shaw (2020) suggest an interaction between timing measures such as Tip Delay, and spatial measures, such as the magnitude of tongue body lowering. Within a fixed temporal window, a gesture which occurs later within that window has less time to achieve its target and is more likely to be undershot. Given the finding for the dorsal gesture to preceded the apical gesture in dark coda /l/s, and apical gesture to preceded the dorsal gesture in a clear onset /l/ (e.g., Sproat and Fujimura, 1993), the following can be predicted. When the dorsal gesture precedes the apical gesture, as is the case for dark coda /l/, the dorsal gesture has plenty of time to fully achieve its target, while the apical gesture has relatively little time. This results in a full dorsal gesture, and a reduced apical gesture. The reverse is predicted for clear onset /l/s.

4.1.4 Review summary

This section has reviewed literature showing lateral darkening to exhibit complex patterns of variation, interacting with vocalic and morphosyntactic context in seemingly gradient and categorical ways. This makes it difficult to made specific predictions about how variation in /l/ will manifest across varieties. The goal of this chapter is to establish the articulatory differences in /l/ darkening between dialects. It is therefore important for this study that dialectal differences in /l/ darkening are disentangled from the darkening effects of vocalic and morphosyntactic context. This will be achieved by conducting a dialect comparison of /l/ across range of morphosyntactic and vocalic contexts known to mediate /l/ darkening.

This chapter will consider both spatial and temporal measures of /l/ darkening reviewed above in order to understand the internal dynamics of laterals across varieties. Understanding the internal dynamics of laterals between dialects will allow us to better understand dialectal differences in lateral timing patterns at the consonant cluster level, investigated in subsequent chapters. However, given the overarching goal of this thesis - to explore the relationship between spatial and temporal variation in laterals - a specific focus will here be placed on establishing the nature of spatial differences in /l/ between dialects. This will facilitate an explicit investigation into the the relationship between spatial variation in laterals, and temporal variation in lateral clusters across dialects. Spatial correlates of /l/ darkening reviewed here include tongue body lowering, tongue dorsum retraction, and post dorsum lowering and retraction. Importantly for this study, the measure of tongue body lowering was widely reported to be robust across contexts and more easily identifiable than dorsal retraction. For these reasons, tongue body lowering will be a focal measure within this analysis.

4.2 A study of lateral darkening

The aim of study analysis is to establish the nature of articulatory differences in lateral darkening between dialects. The findings of this analysis will feed directly into subsequent chapters which ask how (spatial) articulatory differences between dialects affect patterns of timing in lateral clusters. The research questions for this analysis are here reiterated:

Broad RQ: *What is the nature of (spatial) articulatory differences in darkening between dialects?*

Sub RQ: *How does the vowel and morphosyntactic context of /l/ interact with dialectal effects on /l/ darkening?*

The analysis will be structured as followed. I first present the methods used within the analysis, including the stimuli used and the steps taken to process the acoustic and articulatory data. I begin the analysis by considering /l/ in onset position, where dialect differences in /l/ darkening have been reported. From here, I consider the remaining contexts (including mono-morphemic intervocalic, pre-boundary intervocalic, word final pre-vocalic, and word final pre-consonantal contexts) allowing for an investigation into the effects of dialect, morphosyntactic and vocalic context. Onset and coda contexts are considered separately given that the onset /l/s are preceded by an unstressed schwa e.g., *the leap*, where as coda /l/s are preceded by a stressed vowel, which makes it difficult to make visual comparisons.

4.3 Methods

4.3.1 Study overview

Audio-synchronised electromagnetic articulography data were collected from 8 Onset Lightening speakers, and 8 Onset Darkening speakers using the Carstens AG501 and a DPA 4006A microphone. Participants were instructed to read sentences aloud from a facing computer monitor. Further details on the experimental setup can be found in Chapter 3.

4.3.2 Stimuli

The stimuli list was based on the stimuli employed in Strycharczuk, Derrick, and Shaw (2020) and was designed to extract maximal variation in /l/ production. This was here necessary in order to tease apart the effects of dialect on /l/ darkening from effects of vocalic and morphosyntactic context. Tokens included /l/ within 5 morphosyntactic contexts (initial, mono-morphemic intervocalic, pre-boundary intervocalic, word final preceding a vowel, and word final preceding a consonant), and three vowel contexts: FLEECE, KIT, THOUGHT. See Table 4.1 below for a full list of target words. Target words were embedded within the carrier phrase: *Say the (target word) again*, with the exception of word final pre vocalic tokens, which were embedded in the carrier phrase *Say the (target word) quick*. Each unique sentence were repeated 4 times over the course of the experiment.

	Initial	Mono-morphemic intervocalic	Pre-boundary intervocalic	Word final pre-vocalic	Word final pre-consonantal
FLEECE	leap	helix	healing	heal it	heal mick
KIT	lip	fillet	filling	fill it	fill mick
THOUGHT	law	paula	mauling	maul it	maul mick

Table 4.1: Study 1 stimuli list

The onset analysis included a total of 187 tokens; this included /l/ in word initial position in the FLEECE, KIT and THOUGHT vowel contexts. Of these, 92 tokens were from Onset Darkening speakers, and 95 tokens were from Onset Lightening speakers. 5 tokens were excluded due to audio errors.

The coda analysis included a total of 739 tokens; this included /l/ within mono-morphemic intervocalic, pre-boundary intervocalic, word final pre-vocalic, and word-final pre-consonantal contexts, for FLEECE, KIT and THOUGHT vowel contexts. Of these, 365 tokens were from Onset Darkening speakers, and 371 were from Onset Lightening speakers. 29 tokens were were excluded due to audio errors / mispronunciations.

4.3.3 Acoustic data processing

For each audio file containing a single sentence, (e.g., *"say the healing again"*) a TextGrid was created in Praat (Boersma and Weenink, 2011). Sentences were labelled manually on each TextGrid, before acoustic segmentation was performed using Montreal Forced Aligner (McAuliffe et al., 2017) and hand corrected where necessary. Where /l/ was segmented from the following vowel, segmental boundaries were placed at the point of an overall change in spectral energy. The FastTrack plugin (Barreda, 2021) was then used in Praat to extract formants from the lateral (plus vowel) segments of target words. FastTrack offers a means of automatically performing several formant analyses, and then choosing the best one (Barreda, 2021). A requirement of FastTrack is that sound files contain continuous formants, with no gaps. Formants are estimated using linear predictive coding, and the selected formant track is that which has the least residuals (i.e., is smooth), and falls within set parameters (Barreda, 2021). F1 and F2 values were extracted at time steps of 0.004 seconds using the Burg tracking method, with 5 frequency bins between the range of of 5000 - 7000 Hz. Outputs were visually checked for each speaker to ensure that the correct formant band had been tracked. The output csv files were then imported into R, where a measure of F2−F1 was calculated for each time step, and further analysis and visualisation took place.

4.3.4 Articulatory data processing

For each sentence, TextGrids were imported into R alongside the corresponding EMA files, where position and derived velocity data were calculated using the *tardis* and *tadaR* R packages (Kirkham, 2024a,b), in accordance to the process outlines in the general methods section (Chapter 3). Where the point of maximum displacement or velocity minima was taken along the trajectory of an articulatory dimension, the *findpeaks* function from the *pracma* R package (Borchers, 2022) was used. Speakers L03 and L04 are removed from all tongue body analyses due to an absent TB sensor.

The analysis will focus primarily on the spatial measure of vertical tongue body displacement (TBz). This measure has been reported to be a robust measure of /l/ darkening and has been shown to also correlate with dorsal retraction (e.g., Sproat and Fujimura, 1993). This is particularly important for this data set, since we are unable to measure dorsal retraction directly due to the relatively anterior placement of the sensors. The analysis also considers the tongue tip sensor in the vertical dimension (TTz) in seeking to calculate measures of Tip Delay. However, this measure is problematised by difficulties in identifying the time of gestural achievement for tongue body lowering, which is a necessary component of this measure.

To quantify effects of vowel, dialect, and contexts on acoustic and articulatory measures of /l/ darkening, linear mixed effects models are performed in R using the *lme4* package (Bates et al., 2015), and post-hoc tests are performed on significant interaction terms using the *emmeans* R package (Lenth, 2025). Specific details of the model structures used are reported within the relevant sections.

4.4 Onset analysis

This section examines the nature of articulatory differences in onset /l/ between dialects. Previous literature reports onset position to be the context where dialects differ the most in /l/ darkening, hence it is important to the rationale of the thesis to address this context. I begin by verifying the presence of an acoustic difference in onset /l/ between dialects, specifically looking at minimum F2−F1 values. In Section 4.4.2, I then visualise tongue body lowering profiles of /l/ to establish broad patterns of tongue body lowering with vowel and dialect, which I then explore further through a functional principal component analysis in Section 4.4.3.

4.4.1 Minimum F1−F2

Many studies of /l/ darkening have focused exclusively on acoustics (e.g., Carter and Local, 2007). The acoustic measure explored here is F2 minus F1 at its minimum, (e.g., as used in Lee-Kim, Davidson, and Hwang (2013) for F2). F2−F1 is a widely used measure of /l/ darkening (e.g., Kirkham, 2017; Recasens and Espinosa, 2005; Sproat and Fujimura, 1993; Turton, 2014). A clear /l/ is typically characterised by a low F1 and a high F2 which means there is a larger difference between the formants, and hence a higher value of F2−F1 compared to that of a darker /l/ where F1 and

F2 are typically closer together. It is thus predicted that Onset Darkening speakers will show lower F2−F1 values than Onset Lightening speakers across vowel contexts.

Figure 4.2 shows minimum F2−F1 values across the schwa + /l/ + vowel window for onset /l/ in FLEECE, KIT, and THOUGHT contexts. Clear differences can be observed across all vowel contexts, with Onset Darkening speakers showing a lower minimum F2−F1 value, characteristic of a darker /l/. The lower F2−F1 values of Onset Darkening speakers are also accompanied by greater variation. An effect of vowel can also be observed for both dialects, with both dialects showing lower F2−F1 values within the back THOUGHT vowel context. Onset Darkening speakers also show a difference between FLEECE and KIT contexts, with higher F2−F1 values in the FLEECE context; however, the high variability means there considerable overlap between these contexts.

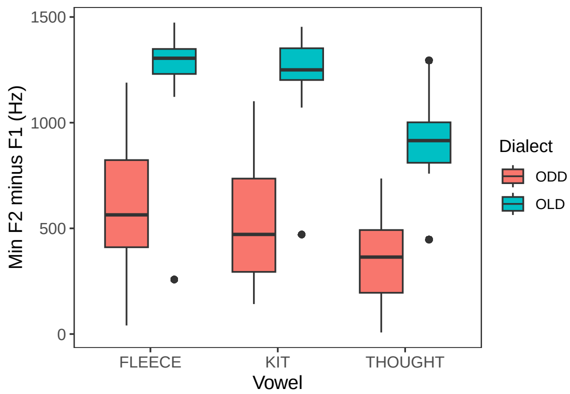

Figure 4.2: Minimum F2−F1 over schwa + /l/ + V window. Onset Lightening Dialect (OLD) and Onset Darkening Dialect (ODD) are compared across three vowel contexts, FLEECE - *leap*, KIT - *lip*, and THOUGHT - *law*.

Given the patterns observed between dialect and vowel and minimum F2−F1 values in Figure 4.2, model comparisons were performed to quantify these effects. Models were run in R using the *lme4* package (Bates et al., 2015). A full model was first performed; the full model contained fixed effects of dialect, vowel and estimated vocal tract length, an interaction term for a dialect by vowel interaction, and a random intercept for speaker. The procedure for model comparisons was as follows:

1. First, an effect of the dialect by vowel interaction was tested for by comparing

the full model to a model where the interaction term had been excluded.

2. If the interaction was found to be significant at $p < .05$, no further comparisons on fixed effects were made and a p value was reported.

3. If the interaction was found to be non-significant at $p > .05$, further comparisons were made on the fixed effects of each of the terms of the interaction. The full model was compared to a model where both the relevant fixed effect and interaction term had been removed, and a p value and model summary were reported.

Effect Tested	χ^2	df	p-value
dialect x vowel	4.72	2	0.09441
dialect	29.67	3	< **.001**
vowel	110.36	4	< **.001**

Table 4.2: F2−F1 model comparisons. Significant effects ($p < .05$) are shown in bold.

Results of the model comparison (Table 4.2) found a non-significant effect of the dialect by vowel interaction on minimum F2−F1 for onset /l/ at $p > 0.05$. Because the interaction was non-significant, model comparisons were performed for fixed effects of dialect and vowel, both of which were found to be significant at $p < .001$. This means that minimum F1−F2 is significantly affected by both vowel (FLEECE, KIT and THOUGHT), and dialect, but not an interaction between the two.

4.4.2 Vertical tongue body displacement of onset /l/

The previous section showed the predicted differences in acoustics between dialects, finding a significant effect of dialect on minimum F2−F1 values. This section seeks to establish the articulatory mechanisms underpinning these acoustic differences. I here focus on tongue body lowering given that this measure has been shown to be a robust correlate of /l/ darkening, with greater lowering being observed for darker /l/s (Sproat and Fujimura, 1993; Strycharczuk, Derrick, and Shaw, 2020). In light of this, Onset Darkening speakers are here predicted to exhibit great tongue body lowering than Onset Lightening speakers.

Vertical tongue body displacement trajectories are here presented for onset /l/ in FLEECE, KIT and THOUGHT contexts, i.e., *leap, lip* and *law*. Figure 4.3 shows z-scored tongue body trajectories in the vertical dimension plotted over the V + /l/ sequences for the three vowel contexts. Visual inspection of TBz trajectories showed variation between dialects in the post lateral vowel within the word initial context; for this reason, the post-lateral vowel was here excluded. In the high front vowel contexts, *leap* and *lip*, Onset Lightening speakers show a pattern of initial tongue body raising, most notably in the context of *lip*, where raising is seen throughout the interval. Conversely, Onset Darkening speakers show a consistent pattern of initial tongue body lowering across vowel contexts. This suggests that Onset Darkening speakers resist the lightening effect of the following front vowel, while Onset Lightening speakers do not. Dialect differences are considerably reduced in context of *law*, where tongue body height is similar across dialects; however, differences can be seen in the shape of trajectory, with Onset Lightening speakers maintaining a small amount of initial tongue body raising before lowering.

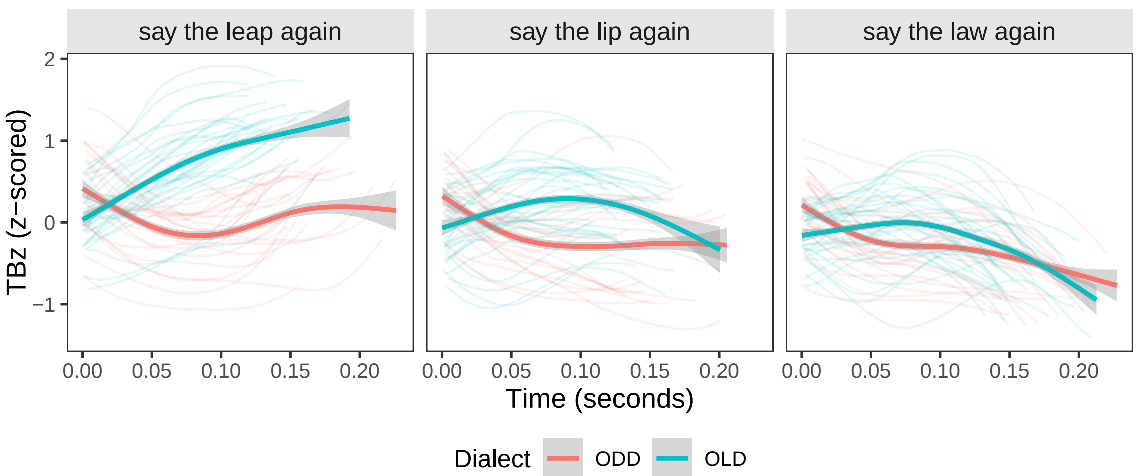

Figure 4.3: Z-scored vertical tongue body displacement trajectories over the schwa + /l/ window for *leap*, *lip* and *law*. Thick opaque lines show smoothed trajectories by dialect, while faint lines show individual displacement trajectories.

4.4.3 By-vowel Function Principal Component Analysis on vertical tongue body displacement in onset /l/

Figure 4.3 showed visual differences in tongue body vertical displacement for front vowel contexts for word initial /l/ in Figure 4.3. This section seeks to further home in on these dialectal differences in tongue body displacement through a Functional Principal Component Analysis (fPCA).

fPCA provides a useful means of extracting the orthogonal variance in ranked order from a set of functional data, hence has become increasingly prevalent in phonetic research, which is often concerned with the patterning of phonetic parameters over time (e.g., Cole and Strycharczuk, 2024; Gubian, Torreira, and Boves, 2015). An fPCA is here performed on z-scored TBz values across the vowel + /l/ interval. The analysis was conducted in R using the *fdapace* package (Wang, Chiou, and Müller, 2016). fPCAs were performed separately for each of the three preceding vowel contexts FLEECE, KIT, THOUGHT. Dialectal differences are the major point of interest for this analysis, thus looking at vowel context in turn allows the variance which results from dialectal differences to be better exposed. In the subsections which follow, I discuss fPCA results of word initial laterals within FLEECE, KIT and THOUGHT vowel contexts in turn.

Before presenting the results, I will first orient the structure of this analysis. For each vowel context, three figures are presented. The first is a perturbation plot (Figures 4.4, 4.7, and 4.10). The perturbation plots show a white midline which represents mean z-scored TBz displacement over the /l/ + V sequence. Coloured lines either side of the mean, represents +/− the PC from the mean, multiplied by 2 SDs (Gubian, Torreira, and Boves, 2015). From this, the effect of the PC on the shape of the mean TBz trajectory can be observed for increasingly positive or negative PC values. The colour coding illustrates how differences in PC scores affect the shape of the trajectory, blue indicates positive, and red indicates negative. For each vowel context, the first 3 PCs are presented in this way.

The second figure shows box plots of PC scores from PCs 1 to 3, plotted by dialect (Figures 4.5, 4.8, and 4.11). This allows us to see how each PC patterns with dialect. Since this analysis is primarily concerned with the dialectal differences in TBz, linear mixed effects models are performed to statistically quantify the effect of dialect on PC values in each vowel context. For each PC within a given vowel context, a full model is compared to a partial model whereby the fixed effect of dialect is excluded. Full models include fixed effects of dialect and estimated vocal tract length, and a random intercept for speaker. An effect of dialect is considered significant for p-values of $< .05$.

Visually, it is difficult to map the values from the box plots of PC scores onto the corresponding contour of the perturbation plots. At best, we may be able to make vague estimates about what the corresponding trajectory shape would look like for a given PC value. In order to better visualise the modifying effect of PC scores on the shape of the mean TBz trajectory, trajectories are reconstructed based on variation in the relevant PC (i.e., the PCs for which a significant effect of dialect is reported). Reconstructed trajectories are shown in Figures 4.6, 4.9, and 4.12. The following sections will outline the results of the fPCA analysis on each vowel context in turn.

4.4.4 Functional Principal Component Analysis: word initial FLEECE context

Figure 4.4 shows perturbation plots for principal components (PCs) on z-scored TBz trajectories in the FLEECE context. PC1 captures 81.2% of variation, PC2 captures 12.5%, and PC3 captures 6.1%.

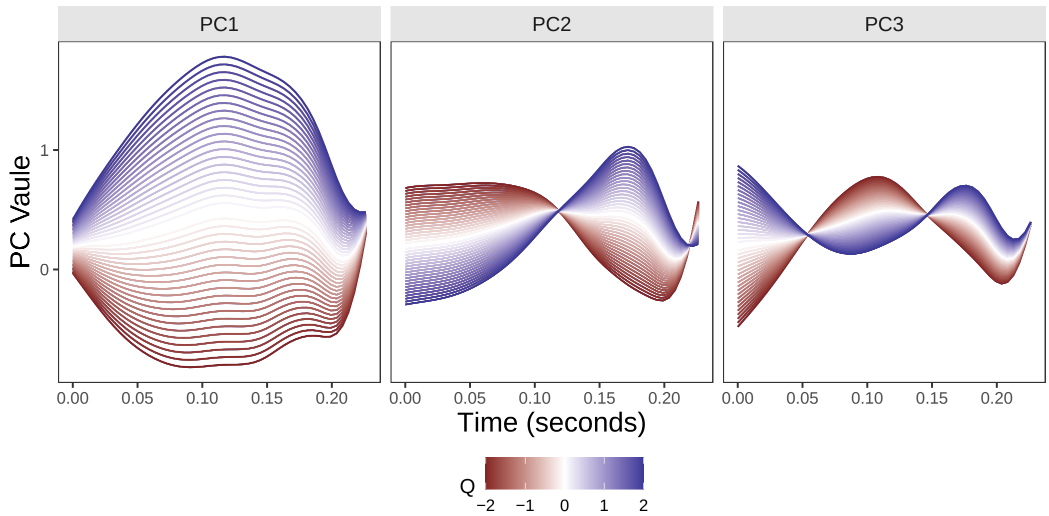

Figure 4.4: Average vertical tongue body displacement over the V + /l/ window of *leap* is shown by the white line. Blue and red lines show positive (blue) and negative (red) PC values for +/- 2 SDs. PC1 captures 81.2% of variation, PC2 captures 12.5%, and PC3 captures 6.1%.

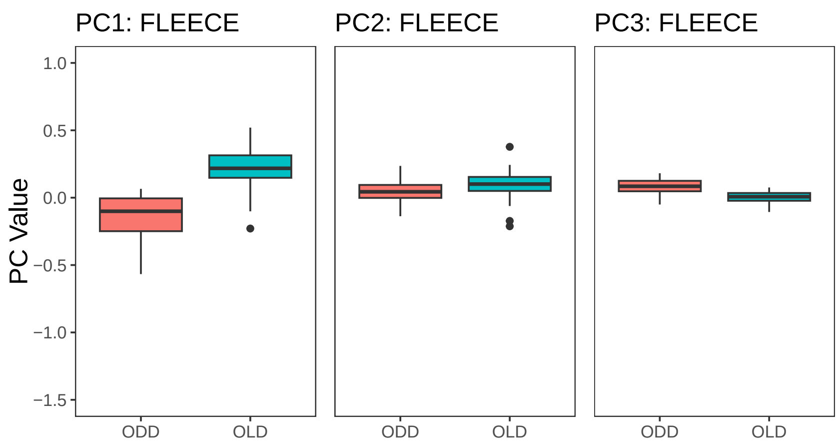

Figure 4.5: PC values from the FLEECE context plotted by dialect, where ODD = Onset Darkening Dialect, OLD = Onset Lightening Dialect. PC1 captures 81.2% of variation, PC2 captures 12.5%, and PC3 captures 6.1%.

Figure 4.5 shows box plots of PC1, PC2, and PC3 by dialect. The clearest dialectal differences can be observed for PC1, whereby Onset Lightening speakers have higher PC values than Onset Darkening speakers. Smaller dialect differences can also be observed in PC3, whereby Onset Lightening speakers have relatively lower PC values. By referring back to Figure 4.4, we can see that a higher PC1, here seen for Onset Lightening speakers, reflects a higher tongue body position, while a lower PC1, seen for Onset Darkening speakers, reflects a lower tongue body position. Higher values of PC3, as seen for Onset Darkening speakers, correspond to a trajectory of tongue body lowering then raising, while the reverse is true for lower values of PC3.

To determine the effect of dialect on each PC, results of model comparisons which test for the effect of dialect are reported for each PC in Table 4.3. Dialect has a significant effect on PC1 ($p = < .001$) and PC3 ($p = 0.0125$), and a non-significant effect on PC2 ($p = 0.939$).

Effect Tested	χ^2	df	p-value
PC1	15.016	1	**< .001**
PC2	0.006	1	0.939
PC3	6.235	1	**0.0125**

Table 4.3: Model comparisons for an effect of dialect on PCs 1-3 within the FLEECE context. Significant effects ($p < .05$) are shown in bold.

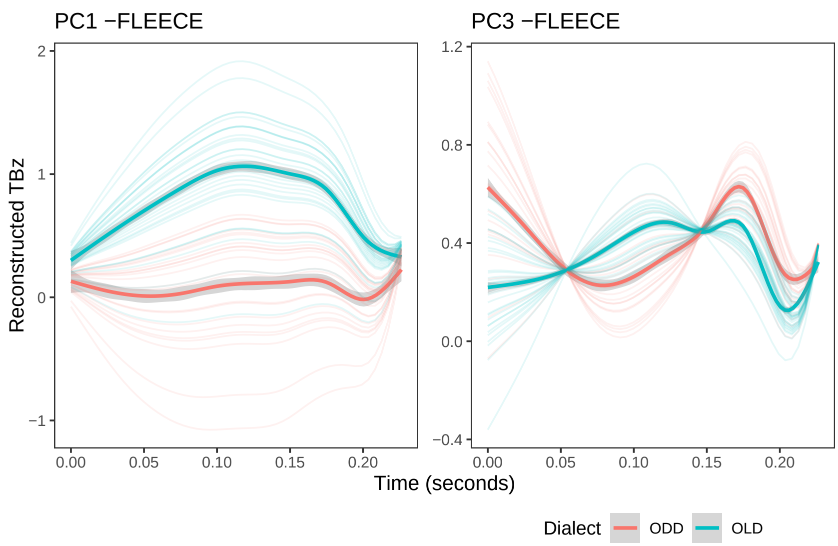

Figure 4.6: Reconstructed changes to the mean TBz trajectory based on variation on PC1 and PC3 for word initial /l/ across V + /l/ window in *leap*.

To visualise these dialect differences in PC1 and PC3 more clearly, 4.6 shows a reconstruction to the changes in the TBz displacement trajectory based on variation in PC1 and PC3. Reconstructions based on variation in PC1 show an overall lower tongue body position for Onset Darkening speakers relative to Onset Lightening speakers. For reconstructions based on variation in PC3, Onset Darkening speakers show a pattern of tongue body lowering followed by raising, while Onset Lightening speakers show a pattern raising followed by lowering.

4.4.5 Functional Principal Component Analysis: word initial KIT context

Figure 4.7 shows perturbation plots for principal components (PCs) on z-scored TBz trajectories in the KIT context. PC1 captures 80.5% of variation, PC2 captures 11.1%, and PC3 captures 8.5% of variation.

Figure 4.8 shows box plots of PC1, PC2, and PC3 for each dialect. Dialectal differences can be observed for PC1 and PC3. Onset Darkening speakers have higher PC1 values, but lower PC3 values. The clearest dialectal difference can be seen in PC1. By referring back to Figure 4.7, a higher PC1, here seen for Onset Darkening speakers, can be seen to reflect an overall lower tongue body. A higher PC3, here

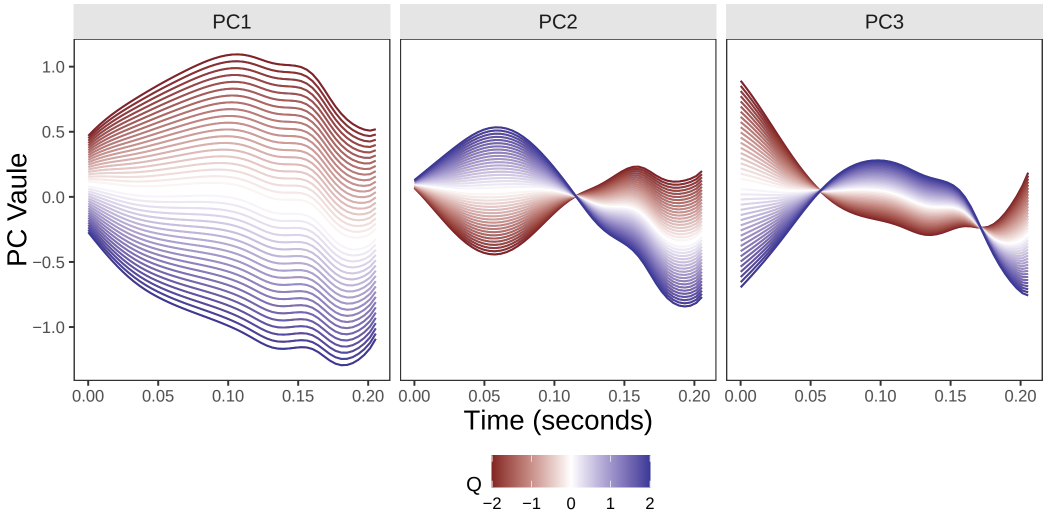

Figure 4.7: Average vertical tongue body displacement over the V + /l/ window of *lip* is shown by the white line. Blue and red lines show positive (blue) and negative (red) PC values for +/- 2 SDs. PC1 captures 80.5% of variation, PC2 captures 11.1%, and PC3 captures 8.5%.

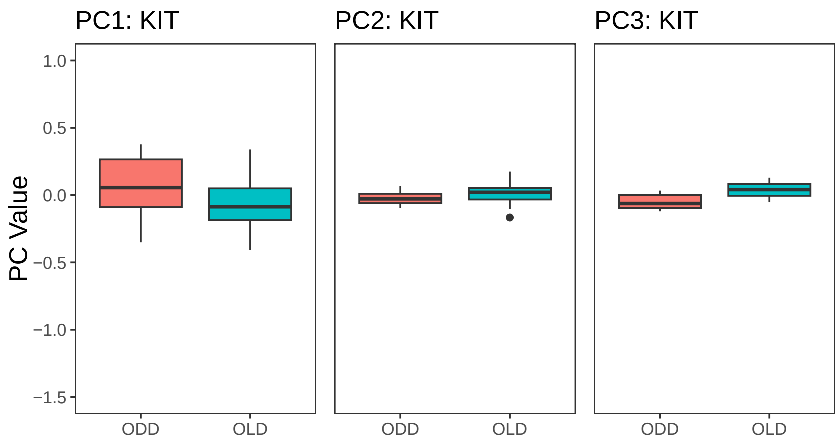

Figure 4.8: PC values from the KIT context plotted by dialect, where ODD = Onset Darkening Dialect, OLD = Onset Lightening Dialect. PC1 captures 80.5% of variation, PC2 captures 11.1%, and PC3 captures 8.5%.

seen for Onset Lightening speakers, appears to reflect an initially low tongue body position which raises and then lowers, while a lower PC3 reflects an initially high tongue body, which lowers and then raises.

Table 4.4 reports the results of model comparisons which test for the effect of dialect on each PC. Dialect has a significant effect on PC2 (p = 0.0435) and PC3 (p = 0.0156), and a non-significant effect on PC1 (p = 0.12). The lack of a statistically significant effect of dialect on PC1 is somewhat surprising given visual differences observed in Figure 4.8 - likely owing to the larger range of overlapping values for this PC.

Effect Tested	χ^2	df	p-value
PC1	2.416	1	0.12
PC2	4.075	1	**0.0435**
PC3	5.843	1	**0.0156**

Table 4.4: Model comparisons for an effect of dialect on PCs 1-3 within the KIT context. Significant effects (p <.05) are shown in bold.

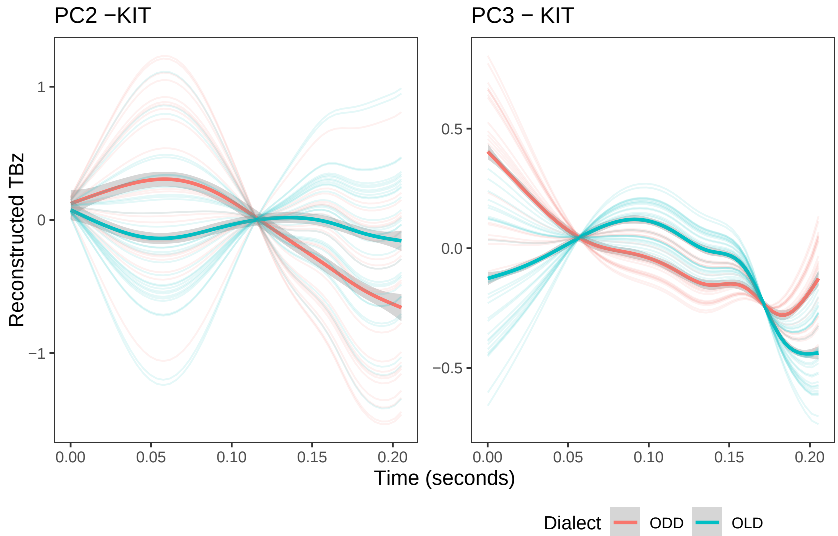

Figure 4.9: Reconstructed changes to the mean TBz trajectory based on variation on PC2 and PC3 for word initial /l/ across V + /l/ window in *lip*.

To see how dialects pattern with PC2 and PC3 more clearly, Figure 4.9 shows a reconstruction to the changes in the TBz displacement trajectory based on variation in PC2 and PC3. For reconstructions based on variation in PC2, differences between dialects manifest in differences in the slope of the TBz trajectory. Onset Darkening speakers are shown to have a higher tongue body position initially, which then lowers quite steeply. Onset Lightening speakers on the other hand, show a relatively flatter trajectory. For PC3, where the strongest effect of dialect was observed, Onset Darkening speakers show a continual lowering of the tongue body, while Onset Lightening speakers exhibit a trajectory of tongue body raising then lowering.

4.4.6 Functional Principal Component Analysis: word initial THOUGHT context

Figure 4.10 shows perturbation plots for principal components (PCs) on z-scored TBz trajectories in the THOUGHT context. PC1 captures 80.4% of variation, PC2 captures 10.7%, and PC3 captures 5.9% of variation.

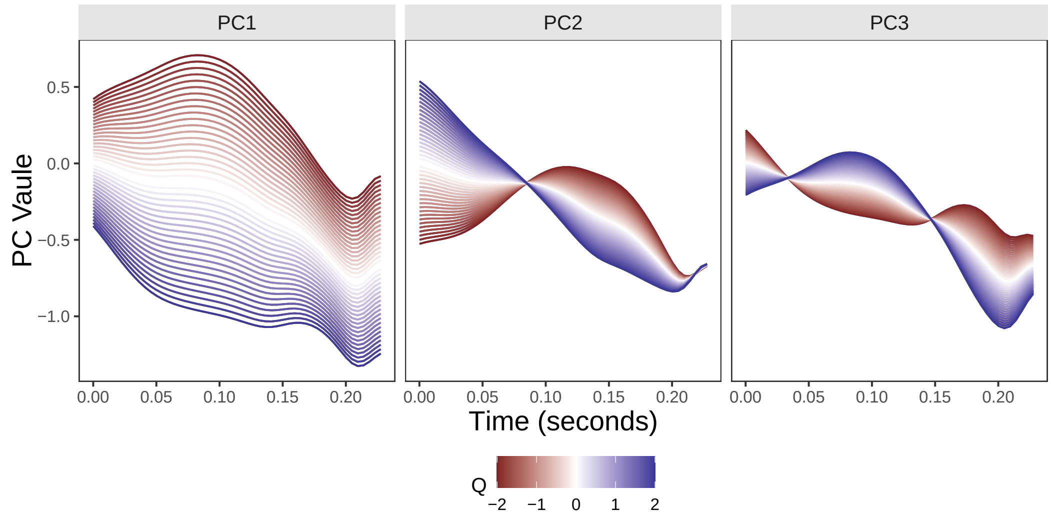

Figure 4.10: The average vertical tongue body displacement over the V + /l/ of *law* is shown by the white line. Blue lines show positive PC values and red lines show negative PC values for +/- 2 SDs. PC1 captures 80.4% of variation, PC2 captures 10.7%, and PC3 captures 5.9%.

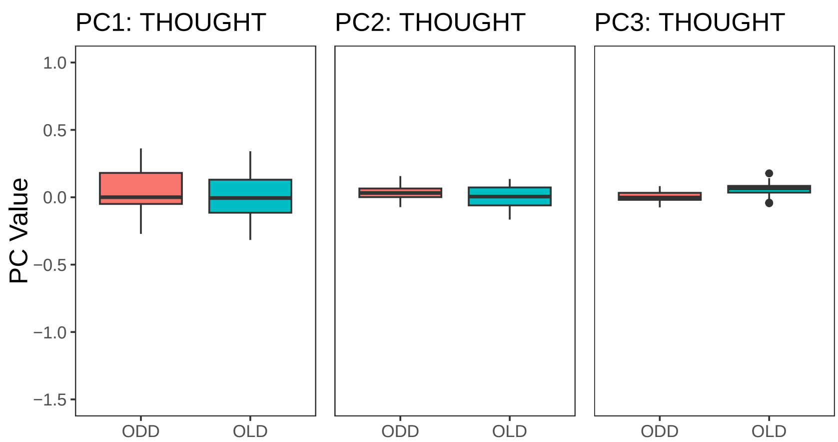

Figure 4.11: PC values from the THOUGHT context plotted by dialect, where ODD = Onset Darkening Dialect, OLD = Onset Lightening Dialect. PC1 captures 80.4% of variation, PC2 captures 10.7%, and PC3 captures 5.9%.

Figure 4.11 shows box plots of PC1, PC2 and PC3 for each dialect. Dialectal differences between PCs are minimal; this is expected given the near-floor effect observed for the THOUGHT context in Figure 4.3. A small difference can be observed between dialects in PC1, with Onset Darkening speakers showing higher PC values than Onset Lightening speakers. By referring back to Figure 4.10, a higher PC1, here observed for Onset Darkening speakers, is indicative of a lower tongue body position. A small dialectal difference is also seen in PC3, whereby relatively higher values are observed for Onset Lightening speakers. Again referring back to Figure 4.10, higher PC3 values, here seen for Onset Lightening speakers reflect a trajectory of initial tongue body raising followed by lowering. Conversely, lower PC3 values, seen for Onset Darkening speakers, rather reflects a pattern of initial tongue body lowering followed by raising.

Table 4.5 reports the results of model comparisons which test for the effect of dialect on each PC. Dialect has a significant effect on PC3 (p = <.001) only, and a non-significant effect on PC1 (p = 0.219) and PC2 (p = 0.865).

Effect Tested	χ^2	df	p-value
PC1	1.512	1	0.219
PC2	0.0291	1	0.865
PC3	17.207	1	**<.001**

Table 4.5: Model comparisons for an effect of dialect on PCs 1-3 within the THOUGHT context. Significant effects (p <.05) are shown in bold.

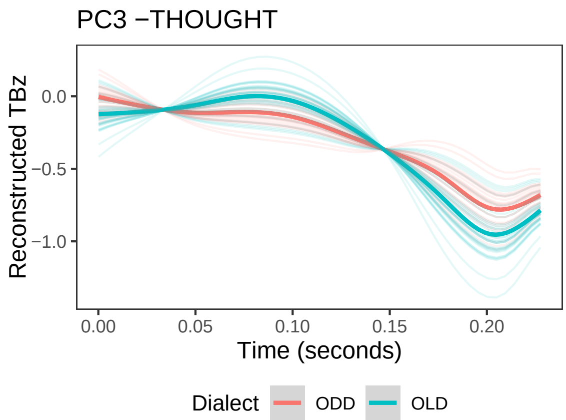

Figure 4.12: Reconstructed changes to the mean TBz trajectory based on variation on PC3 for word initial /l/ across V + /l/ window in *law*.

To better understand the effect of dialect on PC3, Figure 4.12 shows a reconstruction to the changes in the TBz displacement trajectory based on variation in PC3. Differences between dialects can be seen in the shape of the trajectory. While Onset Lightening speakers exhibit a trajectory of initial tongue body raising followed by a steep lowering, Onset Darkening speakers show a more gradual patterns of lowering across the V + /l/ window.

4.4.7 Onset analysis summary

This section has compared onset /l/ in three vowel contexts across Onset Darkening and Onset Lightening dialects. The aim of this analysis was to explore the articulatory mechanism driving the /l/ darkening contrast between dialects. I first began by verifying a dialectal difference in acoustics, where lower minimum F2−F1 values were observed for Onset Darkening speakers across all vowel contexts.

The articulatory analysis focussed on tongue body lowering given that this measure has been shown to be a robust correlate of /l/ darkening (e.g., Sproat and Fujimura, 1993; Strycharczuk, Derrick, and Shaw, 2020). First, I presented dynamic trajectories of the tongue body sensor in the vertical dimension of the V + /l/ window. This revealed spatial differences between dialects in the high front vowel contexts. For Onset Lightening speakers, a clear effect of vowel on tongue body lowering was observed, while Onset Darkening speakers exhibited a consistently low tongue body position across vowel contexts, and hence resisted the lightening effect of the high vowel context. In the THOUGHT context, there appeared to be a floor effect whereby dialects exhibited similar amounts of tongue body lowering.

Findings of the by-vowel fPCA largely echoed these initial observations. fPCA results in the FLEECE context found dialect differences in TBz to manifest in an overall greater tongue body lowering in Onset Darkening speakers. For the KIT context, differences were seen in the magnitude of tongue lowering, which was greater for Onset Darkening speakers. Consistent with the earlier observed floor effect, fPCA results for the THOUGHT context captured only a small difference in trajectory shape between dialects, whereby Onset Lightening speakers showed a pattern of initial tongue body raising followed by lowering, while onset Darkening speakers showed continual lowering.

4.5 Coda /l/ analysis

The remainder of the analysis will look at /l/ across mono-morphemic intervocalic, pre-boundary intervocalic, word final pre-vocalic, and word final pre-consonantal contexts. While the previous literature reports a difference in /l/ darkening between dialects in onset position, it is less clear whether the dialects examined here also differ in coda position. For this reason, acoustic measures are first used to determine whether there is an acoustic difference in /l/ darkening between dialects for /l/ in coda position.

4.5.1 Minimum F2−F1

Figure 4.13 shows the minimum F2−F1 values across the V+/l/ window for each vowel context; F2−F1 is shown on the vertical axis, and morphosyntactic context is shown on the horizontal axis. Dialect is indicated by colour. Figure 4.13 shows a clear effect of preceding vowel on the minimum F2−F1 values for both dialects, where higher values are observed for high-front vowels (FLEECE and KIT), and lower values are observed for the back vowel context (THOUGHT). An effect of context can also be observed. There is a contrast between the word final pre consonantal context and all other contexts in Onset Lightening and Onset Darkening dialects when /l/ is preceded by FLEECE and KIT, but not THOUGHT. Regarding the effect of dialect, Onset Darkening speakers show lower minimum F2−F1 values than Onset Lightening speakers when /l/ is preceded by FLEECE and KIT, but not when preceded by THOUGHT; here a floor effect can again be observed.

Given the clear visual effects of dialect, vowel and context seen for F2−F1 at its minimum, linear mixed effects models were performed to formally quantify the interaction between variables. A full model was first performed; the full model contained fixed effects of dialect, vowel, context, and estimated vocal tract length. Interaction terms were included for a dialect by vowel interaction, and a dialect by context interaction, along with a random intercept for speaker. To test for the significance of each effect, model comparisons were performed according to the procedure reported in Section 4.4.1, whereby effects of interaction terms are first

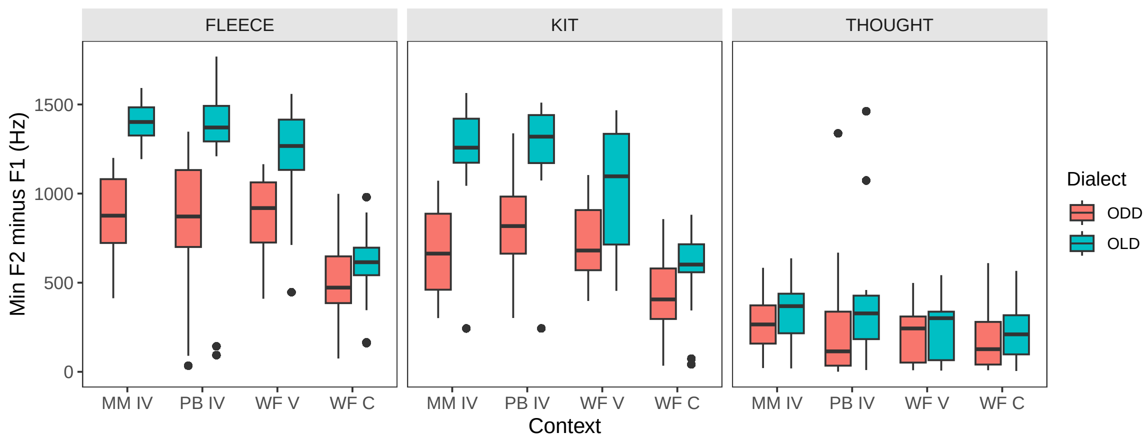

Figure 4.13: Minimum F2−F1 over V + /l/ window: Dialect comparison of Onset Lightening Dialect (OLD) and Onset Darkening Dialect (ODD). The horizontal axis shows positional context, where MM IV = mono-morphemic intervocalic, PB IV = pre boundary intervocalic, WF V = word final pre-vocalic, WF C = word final pre-consonantal.

tested, and effects of fixed effects are only tested for if interaction terms are found to be non significant at p > .05. Significant effects (p <.05) are shown in bold.

Effect Tested	χ^2	df	p-value
dialect x vowel	87.157	2	<**0.001**
dialect x context	49.774	3	<**0.001**

Table 4.6: F2−F1 model comparisons.

Dialect-pairwise	Vowel-pairwise	Estimate	SE	p-value
ODD - OLD	FLEECE - KIT	25.992	39.121	1.0
ODD - OLD	FLEECE - THOUGHT	-307.285	39.157	<**0.001**
ODD - OLD	KIT - THOUGHT	-333.277	38.649	<**0.001**

Table 4.7: Pairwise comparisons on dialect by vowel interaction for F2−F1 in coda position.

Results of the model comparisons showed both the dialect by vowel and the dialect by context interactions to be significant to the minimum value of F2−F1 (p < 0.001) over the vowel plus /l/ window. Because interaction terms were found to be significant, model comparisons were not performed for individual fixed effects.

To interpret the significant effect of the interaction terms, pairwise comparisons were performed on the dialect by vowel interaction (Table 4.7) and the dialect by

Dialect-pairwise	Context-pairwise	Estimate	SE	p-value
ODD - OLD	PB IV - WFC	-259.045	45.373	**<0.001**
ODD - OLD	PB IV - WFV	-98.478	44.725	0.168
ODD - OLD	PB IV - MM IV	37.042	44.716	1.000
ODD - OLD	WFC - WFV	160.567	45.294	**0.0025**
ODD - OLD	WFC - MM IV	296.087	45.295	**<0.001**
ODD - OLD	WFV - MM IV	135.520	44.646	**0.0149**

Table 4.8: Pairwise comparisons on dialect by context interaction for F2−F1 in coda position.

context interaction (Table 4.8). Table 4.7 shows pairwise comparisons for the dialect by vowel interaction to be significant for KIT - THOUGHT and FLEECE - THOUGHT pairings at $p < 0.001$, but not significant for the FLEECE - KIT pairing. This suggests a significant interaction between vowel and dialect for pairing between front and back vowels, but not for the pairing between front vowels.

For the dialect by context interaction, pairwise comparisons showed a significant effect for all pairings except for pairings between the two intervocalic contexts (pre-boundary intervocalic and mono-morphemic intervocalic), and between the pre-boundary intervocalic and word final pre vocalic contexts. A consistently significant interaction between dialect and context was observed for all pairings with the word final pre consonantal context - the context with the strongest boundary strength.

4.5.1.1 Acoustic measures summary

This acoustic analyses has shown the measure of F2−F1 at its minimum proved robust in capturing variation conditioned by vowel, morphosyntactic context and dialect. Predictably, a smaller minimum F2−F1 value, corresponding to a darker lateral, was found for the Onset Darkening Dialect speakers, compared to the Onset Lightening Dialect in front vowel contexts. The back vowel context of THOUGHT had a lowering effect on F2−F1 for both dialects, to such an extent that the dialectal and morphosyntactic variation observed for the front vowel contexts, was not observed for this context.

4.5.2 Articulatory analysis

In this section, I seek to establish the how the interactions found in acoustics between dialect and vowel, and dialect and morphosyntactic context manifest in articulation. I begin by identifying the apical gesture of /l/ as a first step towards calculating the widely used measure of lateral darkening, Tip Delay (Sproat and Fujimura, 1993; Strycharczuk, Derrick, and Shaw, 2020). It is predicted that the apical gesture of /l/ will more readily identifiable than the dorsal gesture, given that the tongue tip, unlike the tongue dorsum, is relatively independent from the more anterior movements recruited for the production of neighbouring vowels.

4.5.2.1 Identifying the Tongue Tip Gesture

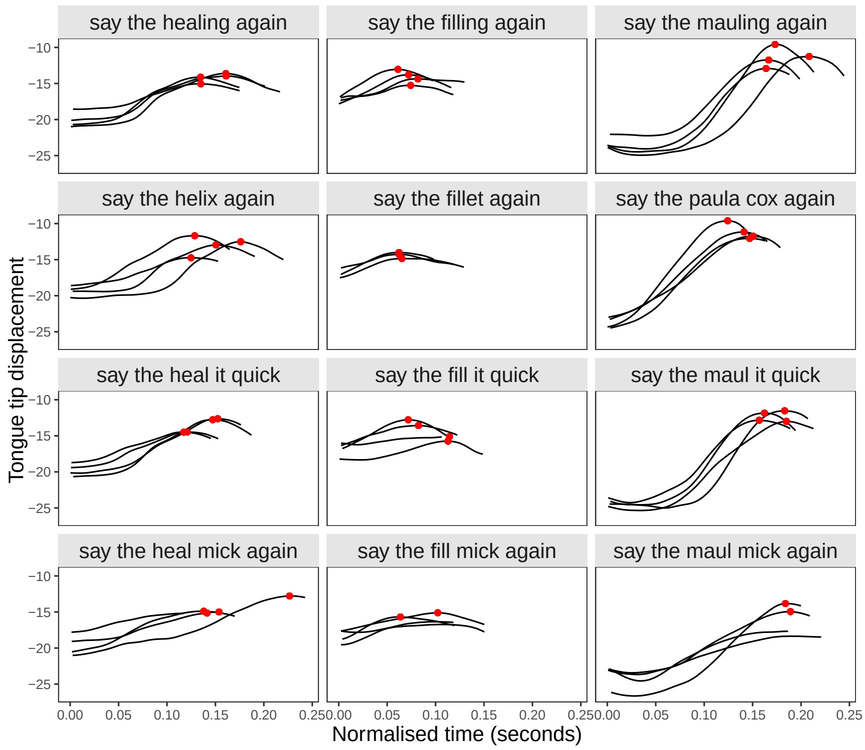

Figure 4.14: L04 tongue tip trajectories and identified gestures. Trajectories show the [V+ /l/ + V (or C)].

To identify the point of maximum tongue tip raising, tongue tip vertical (TTz) displacement trajectories were first plotted over a temporal window which the spanned the V+/L/+/SEG/ of each target word to check for the presence of a visible tongue tip raising gesture. The point of maximum tongue tip raising was then identified over the same temporal window using the *findpeaks* function from the *pracma* R package (Borchers, 2022). The larger temporal window was necessary for the suc-

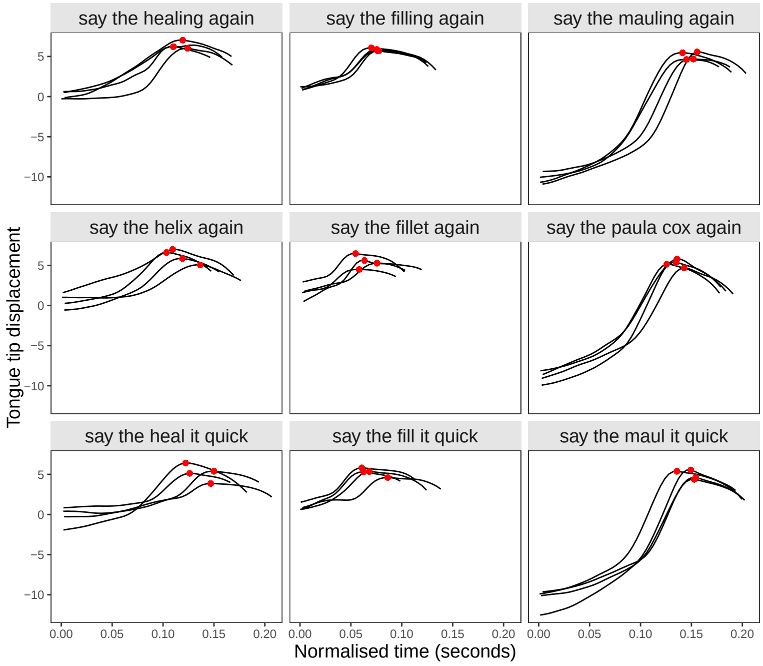

Figure 4.15: S03 tongue tip trajectories and identified gestures. Trajectories show the [V+ /l/ + V (or C)].

cessful identification of a displacement peak. To optimise performance, minimum peak height values and the temporal search window were manually specified for each speaker and context, and each token was manually checked for accuracy.

For some speakers, this method proved a very robust means of identifying the point of maximal tongue tip raising. This can be seen in Figures 4.14 and 4.15 which show the vertical tongue tip displacement trajectories of two speakers, one from each dialect region; an Onset Darkening speaker in Figure 4.14, and an Onset Lightening speaker in Figure 4.15. Red points indicate the point of maximal tongue tip vertical displacement, as identified by the *findpeaks* function. Despite the success of this method in locating the time of maximum tongue tip raising, identification of the tongue tip gesture was severely hindered by the frequent presence of vocalisation.

To better contextualise the tongue tip raising gesture, Figures 4.16 and 4.17 show non normalised tongue body vertical displacement and tongue tip vertical displacement for one token of *helix, healing, heal it* and *heal mick* per dialect. Figure 4.16 shows tokens from Onset Darkening speaker L07, and Figure 4.17 shows tokens from Onset Lightening speaker S08. Clear tongue tip raising can be seen for both speakers in the *helix, healing* and *heal it* contexts. For the *heal mick* context, speaker L07 (Figure 4.16) shows reduced tongue tip raising, accompanied by greater tongue body lowering relative to the other contexts, while speaker S08 (Figure 4.17) shows no tongue tip raising in the word final pre consonantal context, which is likewise accompanied by greater tongue body lowering.

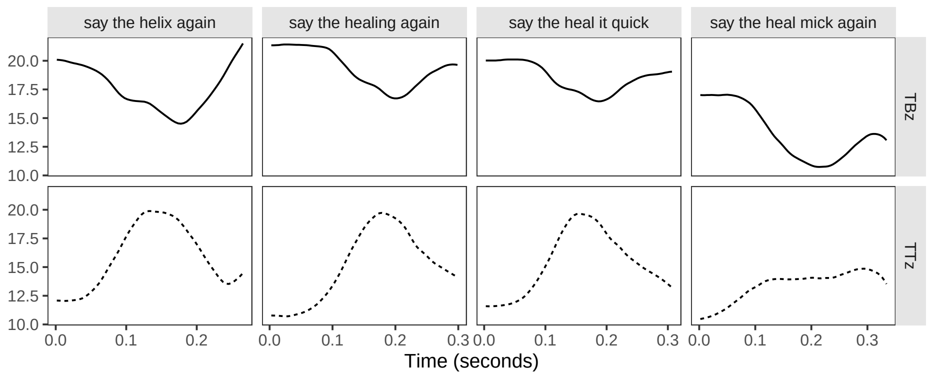

Figure 4.16: Time aligned (non normalised) tongue body vertical displacement and tongue tip vertical displacement for *helix*, *healing*, *heal it*, and *heal mick* contexts. Each context shows trajectories of one token by Onset Darkening speaker L07. The time interval spans the acoustically defined V + /L/ + seg window.

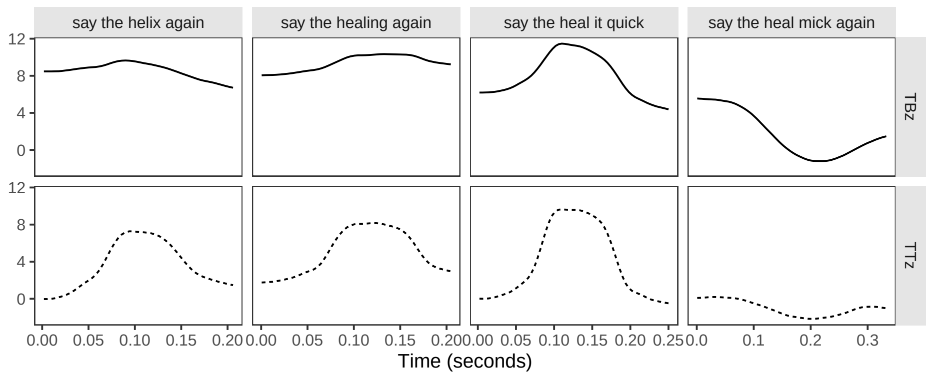

Figure 4.17: Time aligned (non-normalised) tongue body vertical displacement and tongue tip vertical displacement for *helix*, *healing*, *heal it*, and *heal mick* contexts. Each context shows trajectories of one token by Onset Lightening speaker S08. The time interval spans the acoustically defined V + /L/ + seg window.

Vocalisation

The high frequency of vocalisation in speakers of both dialects within this data set presents a challenge for measures of /l/ darkening which rely upon spatial or temporal information about the point of tongue tip gesture achievement. Vocalisation, suggested to be an extreme form of /l/ darkening (Strycharczuk, Derrick, and Shaw, 2020), involves a significant reduction or absence of the TT gesture. Illustrative examples of vocalised tokens are shown in Figure 4.18, while a complete list of instances of partial and full vocalisation are provided Tables 4.9 and 4.10. In addition, the vertical tongue tip displacement trajectories for each speaker in all vowel and morphosyntactic contexts are provided in Appendix 2 (Section 9.3).

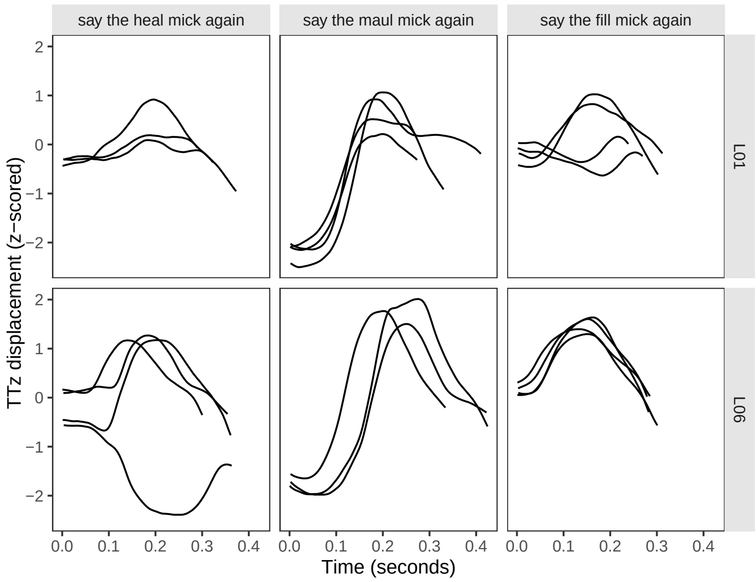

Figure 4.18: Figure showing tongue tip raising for speakers L01 and L06 across the V + /l/ + seg widow in word final pre consonantal position. From left to right panels show the FLEECE context - *heal mick*, the THOUGHT context - *maul mick*, and the KIT context - *fill mick*.

Figure 4.18 shows z-scored tongue tip vertical displacement over time for two Onset Darkening speakers, (L01 and L06). The figure shows /l/ in the word final pre consonantal context, preceding FLEECE, THOUGHT, and KIT. Clear instances of vocalisation can be seen for L01 in the KIT context (top right), and for L06 in the FLEECE context (bottom left), as shown by a fall-rise TTz trajectory, which contrast to the rise-fall trajectories of non vocalised tokens. In addition to full tongue tip raising, and the absence of, Figure 4.18 also shows an intermediate stage of vocalisation whereby there is a reduction in, but not an altogether absence of tongue tip raising. Such partial reduction in tongue tip raising can be seen in the lower tongue tip trajectories of L01 in the FLEECE context (top left), and to an extent in KIT context (top middle). Instances of partial reduction in tongue tip raising are also recorded in Table 4.9 and Table 4.10 for each speaker. These results suggest that there is not a binary distinction between vocalised and non-vocalised tokens for speakers of both dialects, across and within contexts. Tables 4.9 and 4.10 report comparable levels of vocalisation across dialects: a total of 29 (partially and fully) vocalised tokens for Onset Darkening speakers, and a total of 35 for Onset Lightening speakers. However, while all Onset Darkening speakers show some degree of vocalisation, 3 of the 8 Onset Lightening speakers do not vocalise at all. All instances of full vocalisation occur in word final pre consonantal, front vowel contexts. Instances of partial vocalisation varies across dialects. For Onset Lightening speakers, partially vocalised tokens occur in word final pre consonantal front vowel contexts, with the exception of speaker S03 who shows partial vocalisation in the word final pre consonantal THOUGHT context. For Onset Darkening speakers, partially vocalised tokens occur in word final pre consonantal front vowel contexts for speakers L01, L02 and L08, the word final pre consonantal THOUGHT context for speakers L04 and L08, the word final pre vocalic KIT context for speaker L04, and in the intervocalic FLEECE contexts for speaker L05. These results show that vocalisation is highly variable across speakers and contexts, with Onset Darkening speakers in particular showing considerable variation in the contexts of partial vocalisation.

Speaker	Full vocalisation	Partial vocalisation	Total
L01	KIT WFC x2	FLEECE WFC x2	4
L02	FLEECE WFC x3	KIT WFC x3	6
L03	FLEECE WFC x3		3
L04		THOUGHT WFC x2 KIT WFV x2	4
L05		FLEECE PB IV x1 FLEECE MM IV x1	2
L06	FLEECE WFC x1		1
L07	FLEECE WFC x1		1
L08	FLEECE WFC x1 KIT WFC x3	FLEECE WFC x1 THOUGHT WFC x3	8

Table 4.9: Individual instances of partial and full vocalisation for Onset Darkening speakers. WFC = word final pre consonantal; WFV = word final pre vocalic; MM IV = mono-morphemic intervocalic; PB IV = pre boundary intervocalic context. Total of (partially and fully) vocalised tokens = 29.

While interesting to the nature of lateral variation, vocalisation poses a methodological challenge for the measures of /l/ darkening which rely upon identifying the time of maximum tongue tip rasing. Given the high number vocalised tokens in these data, and hence the number of missing data points of maximum tongue tip

Speaker	Full vocalisation	Partial vocalisation	Total
S01	KIT WFC x1 FLEECE WFV x1	KIT WFC x3 FLEECE WFC x3	8
S02	FLEECE WFC x1 KIT WFC x1	KIT WFC x3	5
S03		FLEECE WFC x4 KIT WFC x4 THOUGHT WFC x4	12
S04	FLEECE WFC x1	FLEECE WFC x2	3
S05			0
S06			0
S07	FLEECE WFC x4 KIT WFC x1	KIT WFC x2	7
S08			0

Table 4.10: Individual instances of partial and full vocalisation for Onset Lightening speakers. WFC = word final pre consonantal; WFV = word final pre vocalic context. Total of (partially and fully) vocalised tokens = 35.

raising, measures of lateral darkening are not derived from time of maximum tongue tip raising here.

4.5.2.2 Identifying the Tongue Body Gesture

Since a measure of Tip Delay could not be calculated due to instances of vocalisation, measures of /l/ darkening which do not rely upon the tongue tip gesture are here explored. In particular, I examine the measure of tongue body lowering, reported to be a particularly robust correlate of /l/ darkening (Proctor et al., 2019; Sproat and Fujimura, 1993; Strycharczuk, Derrick, and Shaw, 2020).

Figure 4.19 shows tongue body vertical (TBz) displacement trajectories over a V+/l/ window, z-scored by speaker. Relative to the point of maximum tongue tip raising (Figures 4.14 and 4.15), the point of maximum tongue body lowering is much more difficult to identify. To gain a sense of how problematic it would be to locate the point of maximum tongue body lowering in my data, I performed a restricted analysis on 4 speakers, two Onset Lightening speakers, and two Onset Darkening speakers using the same procedure as used to identify the point of maximum tongue

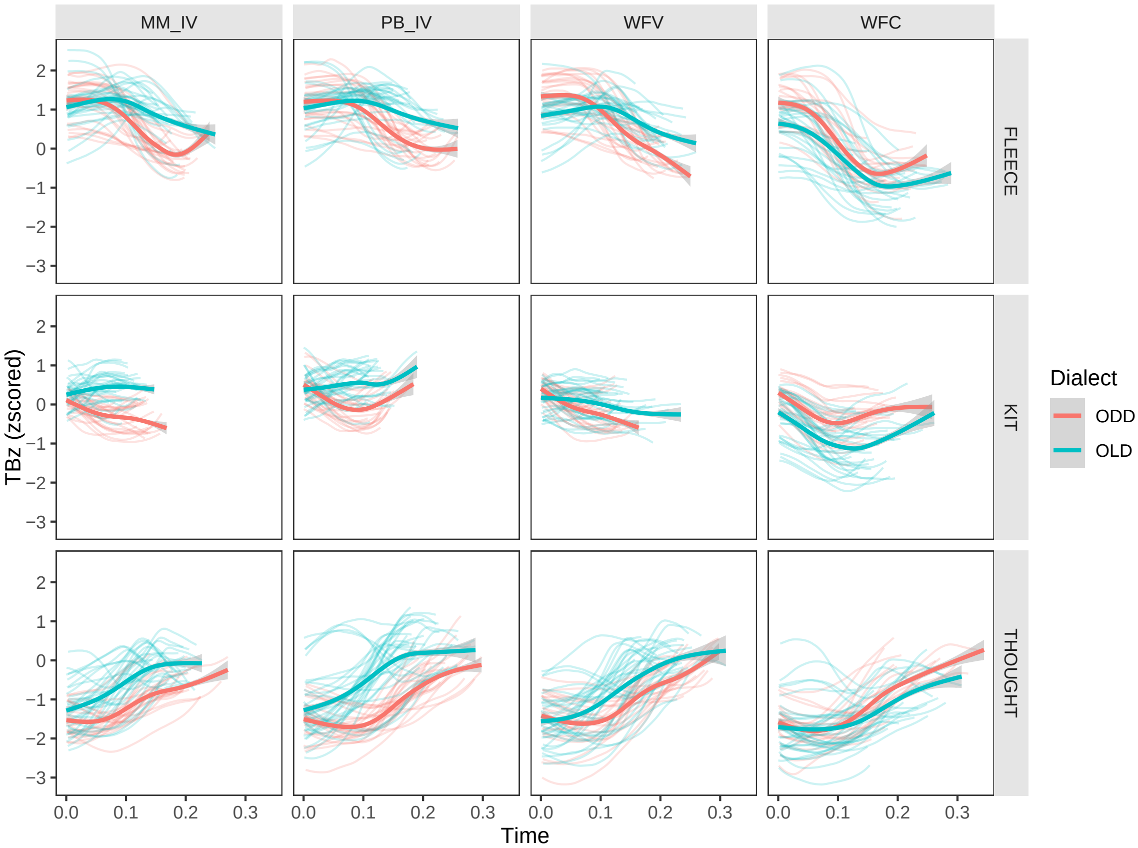

Figure 4.19: Z-scored TBz displacement trajectories across the V+/l/ window for Onset Lightening Dialect (OLD) and Onset Darkening Dialect (ODD). Solid lines show gam-smoothed trajectories by dialect; faint lines show individual trajectories per token. MM IV = mono-morphemic intervocalic; PB IV = pre-boundary intervocalic; WFV = word final pre vocalic; WFC = word final pre consonantal.

tip raising in Section 4.5.2.1. The frequency with which the *findpeaks* function identified the tongue body lowering gesture from the TBz velocity trajectories of a V + /L/ + SEG was reported. Tokens included two contrasting morphosyntactic environments (pre boundary intervocalic and word final pre consonant), which differ considerably in boundary strength.

Results of this small scale analysis are shown in Table 4.11 for each speaker and vocalic and morphosyntactic environment. Results showed that across the 4 speakers, there were four contexts where the tongue body lowering gesture was not identified (shown in bold). Further manual exploration of the TBz displacement and velocity trajectories also revealed considerable challenges in distinguish the lateral tongue body gesture from that of the neighbouring vowels, where in many cases, they seemed to be one in the same. Even in cases where a tongue body gesture could be found automatically, its independence from the vowel gesture was unclear. Use of the point of maximal tongue body lowering as a measure of /l/ darkness thus presents challenges with regard to its practical implementation. In addition, this poses problems for measures of Tip Delay which not only relies on the time of the tongue tip gesture, which may be hindered by vocalisation, but is also contingent on being able to identify the time of achievement of the tongue body gesture.

Speaker	Context	Time Gestures Identified (/4)
L05	PRE BOUNDARY IV	**F: 0** - K: 3 - T: 4
	WORD-FINAL C	F: 3 - K: 4 - T: 2
L01	PRE BOUNDARY IV	F: 4 - K: 4 - T: 3
	WORD-FINAL C	F: 4 - K: 4 - T: 2
S01	PRE BOUNDARY IV	**F: 0** - K: 3 - T: 4
	WORD-FINAL C	F: 4 - K: 4 - T: 4
S02	PRE BOUNDARY IV	**F: 0** - **K: 0** - T: 4
	WORD-FINAL C	F: 4 - K: 4 - T: 4

Table 4.11: Rate of identification of tongue body lowering gesture (i.e., the velocity minima of TBz) for each vocalic and morphosyntactic context, where F = FLEECE, K = KIT, and T = THOUGHT. The maximum number of times a gesture could be identified for each context is 4 (Each prompt was repeated 4 times during the experiment).

4.5.2.3 Minimum TBz Position

Given the difficulties incurred in identifying the point at which the tongue body lowering gesture was achieved (here defined as velocity minimum of tongue body lowering), an alternative time point was considered, namely, the minimum tongue body vertical position across the V + /L/ window. Figures 4.20 show box plots of the minimum tongue body position for each context/preceding vowel combination, compared across dialects. While methodologically simpler to implement, this approach also carries the limitation that an assumption must be made about precisely what the positional minimum of the TBz refers to within the vowel plus lateral window.

Figure 4.20 shows an effect of vowel for both dialects, with a lower minimum TBz

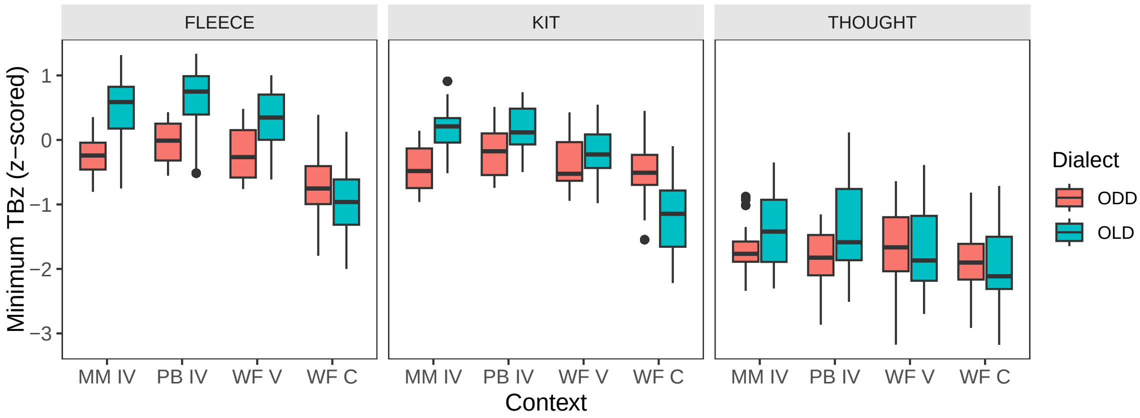

Figure 4.20: Minimum TBz (z-scored) over V + /L/ window: Dialect comparison of Onset Lightening Dialect (OLD) and Onset Darkening Dialect (ODD). The horizontal axis shows positional context, where MM IV = mono-morphemic intervocalic; PB IV = pre-boundary intervocalic; WFV = word final pre vocalic; WFC = word final pre consonantal.

position observed when /l/ is preceded by THOUGHT compared to when high vowels FLEECE and KIT are the preceding vowel. Regarding the effects of context, lower TBz values are observed for the word final pre consonantal context compared to all other contexts for both dialects when /l/ is preceded by FLEECE, and for Onset Lightening speakers when /l/ is preceded by KIT. Effects of dialect are also observed; Onset Darkening speakers show lower minimum TBz positions when /l/ follows FLEECE and KIT in non word final pre consonantal contexts (compare to Figure 4.13). A dialect difference is not observed when /l/ is preceded by THOUGHT, where, similar to Figure 4.13, there is a floor effect.

To quantify interactions between dialect and preceding vowel, and dialect and context, mixed effects models were performed on values of TBz minima. Linear mixed effects models were performed using the same procedure described for the acoustic models in Section 4.5.1, with the F2−F1 minima being here substituted with the TBz minima. The full model included fixed effects for dialect, vowel, context, and estimated vocal tract length. Also included were interaction terms for the dialect by vowel and dialect by context interactions, a random intercept for speaker, and a random slope of speaker for the effect of context. To test the significance of effects, a full model was compared to a partial model where the relevant effect had been excluded. Effects of interaction terms were first tested by comparing the full model to a partial model where an interaction term had been removed. If interaction terms were found to be significant ($p < .05$), no further model comparisons of individual fixed effects were performed. If interaction terms were found to be non significant ($p > .05$), further model comparisons tested for the significance of the individual fixed effects terms of the interaction. Results of models comparisons are reported in Table 4.12. Both dialect by vowel and dialect by context interactions were found to have a significant effect on the TBz minima ($p < .05$); for this reason, further comparisons were not made for the for fixed effects.

To unpack the significant effect of the dialect by vowel and dialect by context interaction on the TBz minima, pairwise tests were performed for the dialect by vowel interaction (Table 4.13), and the dialect by context interaction (Table 4.14). Table 4.13 shows the dialect by vowel interaction to be significant for FLEECE pairs (FLEECE - KIT and FLEECE - THOUGHT), but not for the KIT - THOUGHT pairing. Table 4.14 shows the dialect by context interaction to be significant for all context pairings except for the pairing between intervocalic contexts (PB IV - MM IV), and between pre boundary intervocalic and word final pre vocalic contexts (PB IV - WFV), as was also found for minimum F2−F1 in Table 4.8.

Effect Tested	χ^2	df	p-value
dialect - vowel interaction	11.093	2	**0.004**
dialect - context interaction	73.232	3	<**0.001**

Table 4.12: TBz minimum model comparisons.

Dialect-pairwise	Vowel-pairwise	Estimate	SE	p-value
ODD - OLD	FLEECE - KIT	-0.305	0.096	**0.005**
ODD - OLD	FLEECE - THOUGHT	-0.241	0.096	**0.038**
ODD - OLD	KIT - THOUGHT	0.063	0.095	1.0

Table 4.13: Pairwise comparisons on dialect by vowel interaction for minimum TBz.

This section has shown an interaction between dialect and morphosyntactic context, and between dialect and preceding vowel on the TBz minimum, revealing similar patterns of variation to those reported for the F2−F1 minimum. However, given that the TBz minimum captures only a single point in the lingual trajectory,

Dialect-pairwise	Context-pairwise	Estimate	SE	p-value
ODD - OLD	PB IV - MM IV	0.077	0.111	1.0
ODD - OLD	PB IV - WFV	-0.257	0.110	0.123
ODD - OLD	PB IV - WFC	-0.804	0.112	**<.001**
ODD - OLD	MM IV - WFV	0.333	0.11	**0.015**
ODD - OLD	MM IV - WFC	0.881	0.111	**<.001**
ODD - OLD	WFV - WFC	0.547	0.111	**<.001**

Table 4.14: Pairwise comparisons on dialect by context interaction for minimum TBz. MM IV = mono-morphemic intervocalic; PB IV = pre-boundary intervocalic; WFV = word final pre vocalic; WFC = word final pre consonantal.

it is possible that this measure does not provide a complete picture of the dialectal variation in /l/ darkness. This is addressed in the subsequent section which implements a dynamic approach to measuring /l/ darkening.

4.5.2.4 By-vowel Functional Principal Component Analysis on TBz displacement

The previous section discussed the practical difficulties in consistently identifying a tongue body lowering gesture in /l/, and alluded to potential limitations of using a static measure taken at a single time point. Here, variation in lateral darkening is investigated from a dynamic perspective using functional principal component analysis (fPCA) (Gubian, Torreira, and Boves, 2015). One advantage of fPCA is that it does not privilege a specific point in time, such as the point of maximal tongue body lowering, which, as demonstrated can be both difficult to identify and sensitive to error. fPCA is here performed on z-scored TBz values across the vowel plus /l/ interval. Use of the TBz dimension in this analysis enables a direct comparison to be drawn with the previous single time point analysis of TBz displacement. As with Section 4.4.3, the analysis was conducted in R using the *fdapace* package (Wang, Chiou, and Müller, 2016). fPCAs were performed separately for each of the three preceding vowel contexts FLEECE, KIT, THOUGHT. Again, dialectal differences are the major point of interest for this analysis, thus running an fPCA on multiple vowel contexts would dilute the effect of dialect given that the preceding vowel has a considerable effect on the lateral. The structure of the analysis here mirrors that used in Section 4.4.3; fPCA results of laterals within FLEECE, KIT and THOUGHT vowel contexts will be addressed in turn. For clarity, I will here reiterate the structure of the fPCA analysis and the explanation of the various plots presented, however, the reader may skip to Section 4.5.2.5 if a refresh of this is not required.

For each vowel context, three figures are presented. The first is a perturbation plot (Figures 4.21, 4.24, and 4.27). The perturbation plots show a white midline which represents mean z-scored TBz displacement over the V + /L/ sequence. Coloured lines either side of the mean, represents +/− the PC from the mean, multiplied by 2 SDs (Gubian, Torreira, and Boves, 2015). From this, the effect of the PC on the shape of the mean TBz trajectory can be observed for increasingly positive or negative PC values. The colour coding illustrates how differences in PC scores affect the shape of the trajectory, blue indicates positive, and red indicates negative. PCs accounting for $> 5\%$ of variance are presented in this way; PCs accounting for $< 5\%$ variance are excluded from the analysis. The second figure shows box plots of PC scores, plotted by dialect and morphosyntactic context (Figures 4.22, 4.25, and 4.28). This allows us to see how each PC patterns with dialect and context.

To quantify the effect of interactions between dialect and context on each PC, model comparisons of linear mixed effects models are performed using the same procedure described for previous sections. Full models include fixed effects for dialect, context, and estimated vocal tract length, an interaction term for dialect by context, and a random intercept for speaker. To test the significance of effects, a full model is compared to a partial model where the relevant effect is excluded. Effects of interaction terms are first tested by comparing the full model to a partial model where the interaction term had been removed. If the interaction is found to be significant ($p < .05$), no further model comparisons of individual fixed effects are performed and results of post-hoc tests are reported.

Since it is visually difficult to map PC values from the box plots onto the corresponding contour of the perturbation plots, mean TBz trajectories are reconstructed based on variation in the PCs which account for the largest amount of variance and show a significant effect of dialect, or the dialect by context interaction. Reconstructed trajectories are shown in Figures 4.23, 4.26, and 4.29.

4.5.2.5 Functional Principal Component Analysis: FLEECE + /L/

Results are here presented for an fPCA performed on z-scored TBz trajectories for /l/ following FLEECE. Figure 4.21 shows perturbation plots for PCs 1-3. The percentage of variance explained by each PC is as follows: PC1 accounts for 68.2%, PC2 accounts for 24.7% and PC3 accounts for 6.7%.

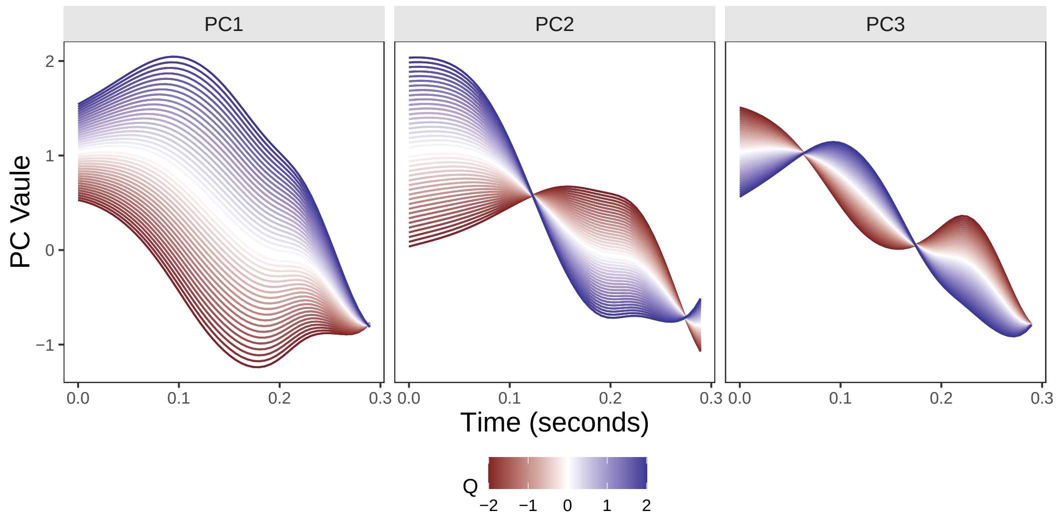

Figure 4.21: The average vertical tongue body displacement over the FLEECE + /L/ window is shown by the white line. Blue and red lines show positive (blue) and negative (red) PC values for +/- 2 SDs.

Figure 4.22 shows box plots of PCs 1-3 plotted by dialect and context. For PC1, there is a notable difference between the word final pre consonantal context and all other contexts for both dialects. Relative to other contexts, the word final pre consonantal context is lower in PC1; referring to Figure 4.21, a lower value of PC1 corresponds to a lower tongue body. In all non word final pre consonantal contexts, dialectal differences can be observed, with Onset Darkening speakers showing comparatively lower PC1 values, indicative of a relatively lower tongue body position.

For PC2, Onset Darkening speakers show slightly higher and more variable values, which, referring to Figure 4.21, reflects a trajectory of steeper tongue body lowering. No clear patterns can be observed for PC3.

To quantify the effects of dialect and context on each PC, model comparisons were performed. First, a comparison is made to test for the effect of a dialect by context interaction. If the interaction is found to be non-significant, further com-

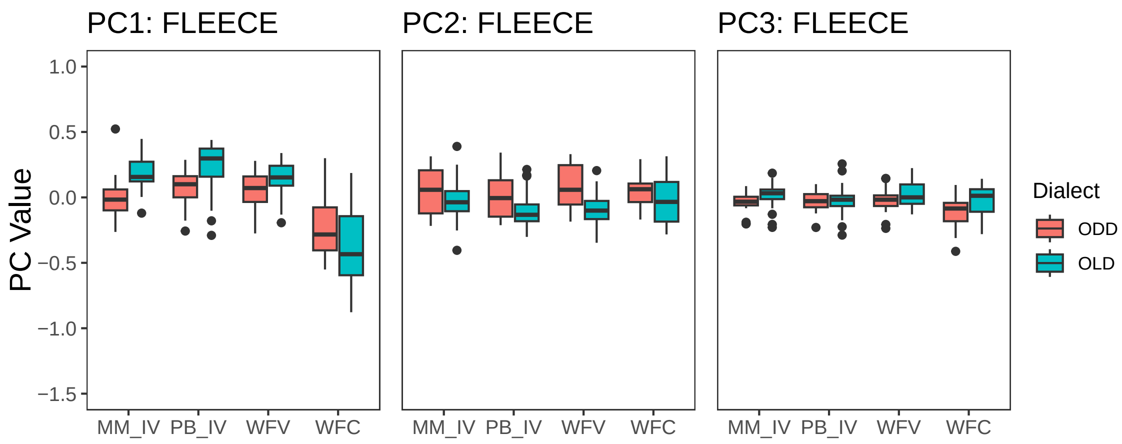

Figure 4.22: PC values from the FLEECE context plotted by dialect and context. Onset Darkening Dialect is shown in red, and Onset Lightening Dialect is shown in blue. The horizontal axis shows positional context, where MM IV = monomorphemic intervocalic; PB IV = pre-boundary intervocalic; WFV = word final pre vocalic; WFC = word final pre consonantal.

parisons are made to test for the effect of dialect and context. Results are reported in Table 4.15. Results show the dialect by context interaction to be significant for PC1 and PC2, and non-significant for PC3. For PC3, subsequent comparisons find a significant effect of context, and a non-significant effect of dialect.

PC	Effect Tested	χ^2	df	p-value
PC1	dialect*context	31.699	3	< **.001**
PC2	dialect*context	10.048	3	**0.0182**
PC3	dialect*context	4.871	3	0.182
PC3	dialect	5.096	4	0.278
PC3	context	21.012	6	**0.0018**

Table 4.15: Model comparisons for an effect of dialect and context on PCs 1-3 within the coda FLEECE context.

To better understand the effects of the dialect by context interactions on PC1 and PC2, results of post-hoc tests are here reported. Table 4.16 reports results of post-hoc tests for PC1, showing the dialect by context interaction to be significant for pairings with the word final pre consonantal context only. Table 4.17 reports results of post-hoc tests for PC2, showing the dialect by context interaction to be significant for the word final pairing only (i.e., WFV - WFC).

To visualise how variation in PC1 and PC2 interact with morphosyntactic con-

Dialect-pairwise	Context-pairwise	Estimate	SE	p-value
ODD - OLD	MM IV - PB IV	-0.027	0.062	1.0
ODD - OLD	MM IV - WFV	-0.095	0.062	0.767
ODD - OLD	MM IV - WFC	-0.335	0.064	**<.001**
ODD - OLD	PB IV - WFV	-0.067	0.062	1.0
ODD - OLD	PB IV - WFC	-0.308	0.064	**<.001**
ODD - OLD	WFV - WFC	-0.240	0.064	**0.0014**

Table 4.16: Pairwise comparisons on dialect by context interaction for PC1 in the FLEECE context. MM IV = mono-morphemic intervocalic; PB IV = pre-boundary intervocalic; WFV = word final pre vocalic; WFC = word final pre consonantal.

Dialect-pairwise	Context-pairwise	Estimate	SE	p-value
ODD - OLD	MM IV - PB IV	-0.054	0.037	0.866
ODD - OLD	MM IV - WFV	-0.083	0.036	0.138
ODD - OLD	MM IV - WFC	0.022	0.038	1.0
ODD - OLD	PB IV - WFV	-0.03	0.037	1.0
ODD - OLD	PB IV - WFC	0.075	0.038	0.29
ODD - OLD	WFV - WFC	0.105	0.038	**0.035**

Table 4.17: Pairwise comparisons on dialect by context interaction for PC2 in the FLEECE context. MM IV = mono-morphemic intervocalic; PB IV = pre-boundary intervocalic; WFV = word final pre vocalic; WFC = word final pre consonantal.

text and dialect, Figure 4.23 shows reconstructions to the TBz trajectory based on variance in PC1 (left) and PC2 (right). Reconstructions based on PC1 show an overall lower tongue body position for Onset Darkening speakers within the intervocalic and word final pre vocalic context, though differences in the word final pre vocalic context are minimal. In the word final pre consonantal context, Onset Lightening speakers exhibit greater tongue body lowering than Onset Darkening speakers. This shows that Onset Lightening speakers show considerable variation in the magnitude of tongue body lowering across morphosyntactic contexts.

For reconstructions based on PC2 (right-most plot), greater tongue body lowering can be observed for Onset Darkening speakers, whereby the tongue body starts high for the vowel, and then lowers considerably, similar to trajectory course of the dark blue positive line in Figure 4.21. Onset Lightening speakers, on the other hand, have a flatter trajectory, reflecting tongue body lowering of a reduced magnitude. In this way, Onset Darkening speakers can be considered to resist the lightening

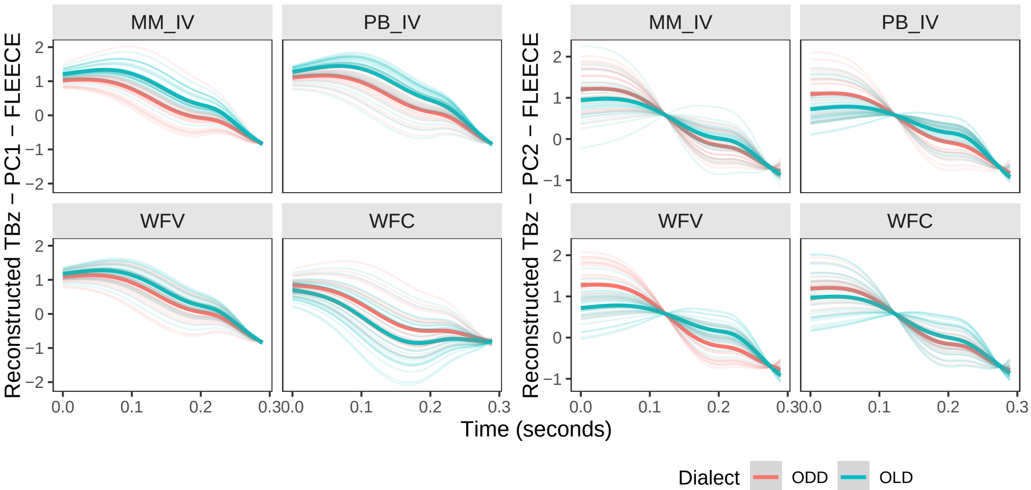

Figure 4.23: Reconstructed TBz trajectories over the FLEECE + /l/ window based on variance in PC1 (left) and PC2 (right). Faint lines show reconstructions for each token.

effect of the high front vowel on the lateral, while the flatter trajectory of Onset Lightening speakers suggests that there has been a lightening effect. The largest difference between dialects can be observed for word final pre vocalic context, while the smallest difference can be seen for the word final pre consonantal context. This echoes the significant interaction between context and dialect for word final context pairing reported in Table 4.17.

4.5.2.6 Functional Principal Component Analysis: KIT + /L/

Results are here presented for an fPCA performed on z-scored TBz trajectories for /l/ following KIT. The percentage of variance explained by each PC is a follows: PC1 accounts for 81.9% of variance, PC2 accounts for 9.7%, and PC3 accounts for 3.6%. Because PC3 accounts for $<$ 5% variance, it is excluded from this analysis. Figure 4.24 shows perturbation plots for PCs 1-2. For each of the PCs, the white midline shows the average trajectory of TBz displacement over the V+/L/ window. Increasingly darker blue lines indicate higher PC values, while the increasingly darker red lines show lower PC values.

Figure 4.25 shows box plots of PC1 and PC2, plotted for dialect and context. For PC1, Onset Lightening speakers show a clear effect of context, whereby PC1 is higher at stronger morphosyntactic boundaries. From Figure 4.24, we can see that

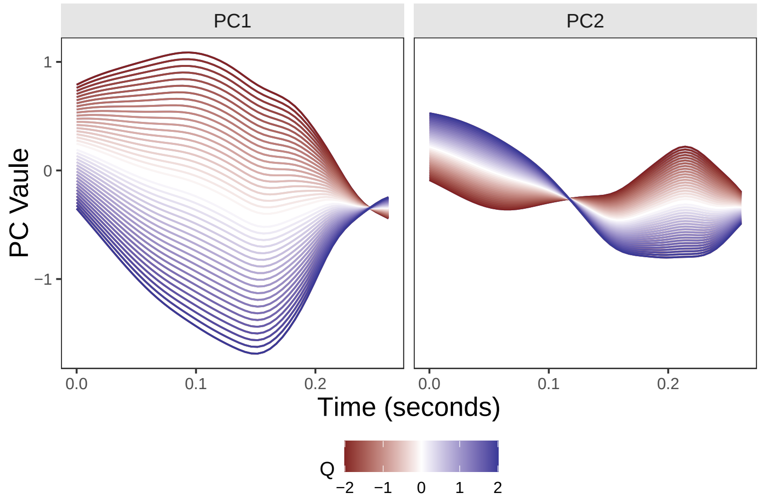

Figure 4.24: The average vertical tongue body displacement over the KIT + /L/ window is shown by the white line. Blue and red lines show positive (blue) and negative (red) PC values for +/- 2 SDs.

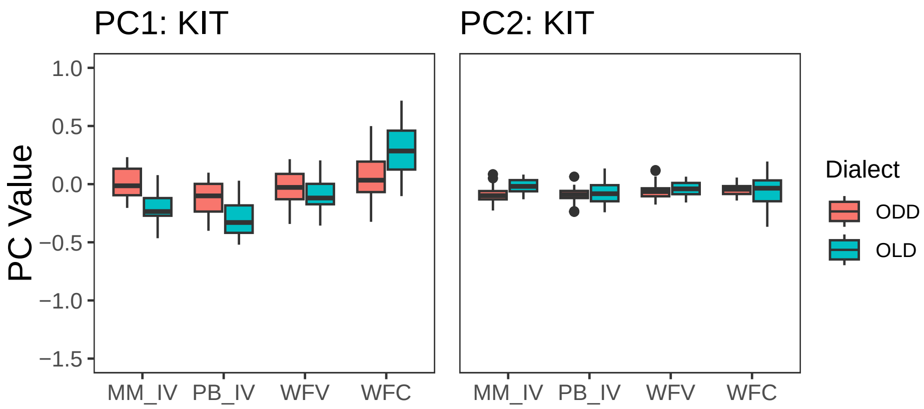

Figure 4.25: PC values from the KIT context plotted by dialect and context. The horizontal axis shows positional context, where MM IV = mono-morphemic intervocalic; PB IV = pre-boundary intervocalic; WFV = word final pre vocalic; WFC = word final pre consonantal.

higher PC1 values correspond to a lower tongue body, suggesting Onset Lightening speakers to show increased tongue body lowering at stronger boundaries. Regarding dialectal differences, Onset Darkening speakers show higher PC1 values across all contexts except the word final pre consonantal context, indicative of an overall lower tongue body position. For PC2, no clear patterns can be observed.

To quantify the effects of dialect and morphosyntactic context on each PC, model comparisons were performed. First, a comparison was made to test for the effect of a dialect by context interaction. If the interaction is found to be non-significant, further comparisons are made to test for the effect of dialect and context. Results are reported in Table 4.18. Results show the dialect by context interaction to be significant for PC1 and PC2.

To unpack the significant effect of the dialect by context interactions on PC1 and PC2, results of further post hoc tests are here reported. Table 4.19 reports results for post-hoc tests for PC1, finding the dialect by context interaction to be significant for all context pairings with the word final pre consonantal context, as well as the mono-morphemic intervocalic and word final pre vocalic context pairing. Table 4.20 reports results for post-hoc tests for PC2. Results find the dialect by context interaction to be significant for the mono-morphemic intervocalic and word final pre consonantal context pairing only, hence showing an interaction between dialect and context at the weakest and strongest morphosyntactic boundaries.

PC	Effect Tested	χ^2	df	p-value
PC1	dialect*context	73.936	3	< **.001**
PC2	dialect*context	8.55	3	**0.036**

Table 4.18: Model comparisons for an effect of dialect and context on PC1 and PC2 within the coda KIT context.

Dialect-pairwise	Context-pairwise	Estimate	SE	p-value
ODD - OLD	MM IV - PB IV	0.060	0.055	1.0
ODD - OLD	MM IV - WFV	0.172	0.054	**0.01**
ODD - OLD	MM IV - WFC	0.463	0.054	<**.0001**
ODD - OLD	PB IV - WFV	0.112	0.055	0.252
ODD - OLD	PB IV - WFC	0.402	0.055	<**.0001**
ODD - OLD	WFV - WFC	0.291	0.054	<**.0001**

Table 4.19: Pairwise comparisons on dialect by context interaction for PC1 in the KIT context. MM IV = mono-morphemic intervocalic; PB IV = pre-boundary intervocalic; WFV = word final pre vocalic; WFC = word final pre consonantal.

To better visualise how variance captured by the PCs interacts with dialect and context, Figure 4.26 shows a reconstruction of changes to the TBz trajectory based on variation in PC1 (left) and PC2 (right). Reconstructions based on PC1 (left) show Onset Lightening speakers to have a higher tongue body position within intervocalic contexts relative to Onset Darkening speakers. Dialectal differences in

Dialect-pairwise	Context-pairwise	Estimate	SE	p-value
ODD - OLD	MM IV - PB IV	-0.051	0.029	0.48
ODD - OLD	MM IV - WFV	-0.059	0.029	0.242
ODD - OLD	MM IV - WFC	-0.081	0.029	**0.034**
ODD - OLD	PB IV - WFV	-0.008	0.029	1.0
ODD - OLD	PB IV - WFC	-0.03	0.029	1.0
ODD - OLD	WFV - WFC	-0.022	0.029	1.0

Table 4.20: Pairwise comparisons on dialect by context interaction for PC2 in the KIT context. MM IV = mono-morphemic intervocalic; PB IV = pre-boundary intervocalic; WFV = word final pre vocalic; WFC = word final pre consonantal.

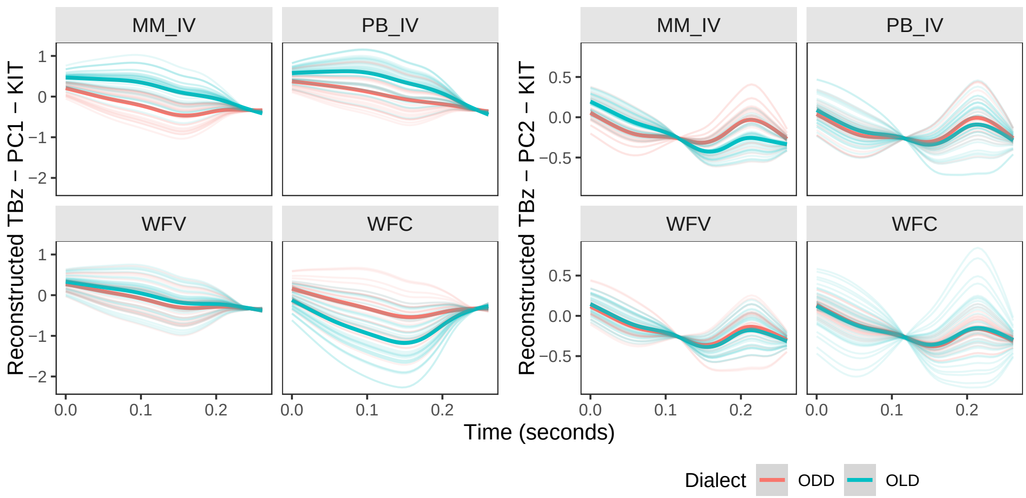

Figure 4.26: Reconstructed TBz trajectories over the KIT + /l/ window based on variance in PC1 (left) and PC2 (right). Faint lines show reconstructions for each token.

tongue body height are minimal in the word final pre vocalic context, and in the word final pre consonantal position, Onset Lightening speakers have a considerably lower tongue body position than Onset Darkening speakers. These results largely echo the patterns observed for the FLEECE context, and suggest that Onset Lightening speakers are using a larger range of tongue body positions across contexts, while Onset Darkening speakers are relatively stable across contexts. Reconstructions based on PC2 (right), show a dialectal difference in trajectory shape for the mono-morphemic intervocalic context only. Here, Onset Lightening speakers show a larger dynamic range of tongue body displacement, beginning relatively higher during the vowel, before lowering, while Onset Darkening speakers exhibit a relatively flatter trajectory.

4.5.2.7 Functional Principal Component Analysis: THOUGHT + /L/

This section presents results from an fPCA performed on z-scored TBz trajectories across the THOUGHT plus /l/ window. Figure 4.27 shows perturbation plots for PCs 1-3. The percentage of variance explained by each PC is a follows: PC1 accounts for 66.2% of variance, PC2 accounts for 24.1%, PC3 accounts for 7.1%. PC4 was excluded from the analysis since it explained < 5% of variance (2.4%).

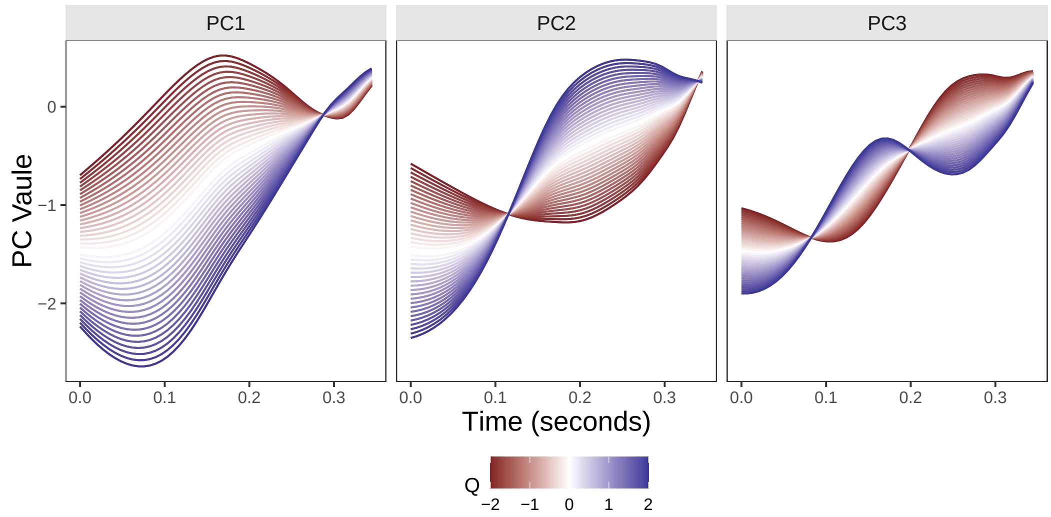

Figure 4.27: The average vertical tongue body displacement over the THOUGHT + /L/ window is shown by the white line. Blue and red lines show positive (blue) and negative (red) PC values for +/- 2 SDs.

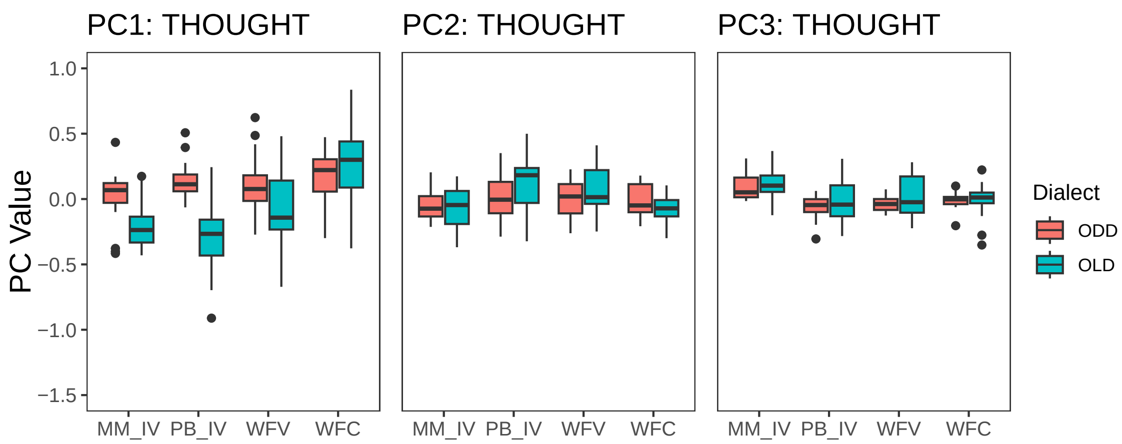

Figure 4.28: PC values from the THOUGHT context plotted by dialect and context. Onset Darkening Dialect is shown in red, and Onset Lightening Dialect is shown in blue. The horizontal axis shows positional context, where MM IV = monomorphemic intervocalic; PB IV = pre-boundary intervocalic; WFV = word final pre vocalic; WFC = word final pre consonantal.

Figure 4.28 shows box plots of PCs 1-3, plotted by dialect and context. PC1 shows patterning with dialect, with Onset Darkening speakers having higher PC1 values than Onset Lightening speakers in all but the word final pre consonantal context. Referring to Figure 4.27, a higher PC1 value, here observed for Onset Darkening speakers, corresponds to a relatively lower tongue body. In addition, PC1 also patterns with morphosyntactic context, but for Onset Lightening speakers only, with higher PC1 values, and hence a lower tongue body observed at stronger boundaries. For PC2, higher values can be seen for both dialects in pre boundary intervocalic and word final pre vocalic contexts. From Figure 4.27, we can see that higher values of PC2 reflect a greater magnitude of tongue body raising from the low vowel position. For PC3, both dialects show higher values in the mono-morphemic intervocalic position relative to other contexts, reflecting a greater magnitude of tongue body raising, followed by subsequent lowering for this context.

To quantify the effect of dialect and context on each PC within the THOUGHT context, model comparisons were performed. A comparison was first made to test for the effect of a dialect by context interaction. If the interaction is found to be non-significant, further comparisons are made to test for the effects of dialect and context. Results are reported in Table 4.21. Results find the dialect by context interaction to be significant for PC1 and PC2, and non-significant for PC3. Further model comparisons of fixed effect terms show a significant effect of context, and a non-significant effect of dialect on PC3.

Results of further post hoc tests on the dialect by context interaction for PC1 and PC2 are here reported. Table 4.22 reports results of post-hoc test for PC1, finding the dialect by context interaction to be significant for all context pairs except the intervocalic pairing (MM IV - PB IV), and the pairing between mono-morphemic intervocalic and word final pre vocalic contexts. Table 4.23 reports results of post-hoc test for PC2, showing the dialect by context interaction to be significant for the pairing between intervocalic contexts, the pre-boundary intervocalic and word final pre-consonantal pairing, and for the word final pre-vocalic and word final pre-consonantal pairing.

PC	Effect Tested	χ^2	df	p-value
PC1	dialect*context	43.22	3	< **.001**
PC2	dialect*context	16.83	3	< **.001**
PC3	dialect*context	3.92	3	0.27
PC3	dialect	4.25	4	0.373
PC3	context	66.18	6	< **.001**

Table 4.21: Model comparisons for an effect of dialect and context on PC1 and PC2 within the coda THOUGHT context.

Dialect-pairwise	Context-pairwise	Estimate	SE	p-value
ODD - OLD	MM IV - PB IV	-0.193	0.073	0.056
ODD - OLD	MM IV - WFV	0.076	0.073	1.0
ODD - OLD	MM IV - WFC	0.299	0.073	**0.0004**
ODD - OLD	PB IV - WFV	0.268	0.073	**0.0017**
ODD - OLD	PB IV - WFC	0.492	0.073	<**.0001**
ODD - OLD	WFV - WFC	0.223	0.073	**0.0143**

Table 4.22: Pairwise comparisons on dialect by context interaction for PC1 in the THOUGHT context. MM IV = mono-morphemic intervocalic; PB IV = pre-boundary intervocalic; WFV = word final pre vocalic; WFC = word final pre consonantal.

Dialect-pairwise	Context-pairwise	Estimate	SE	p-value
ODD - OLD	MM IV - PB IV	0.118	0.043	**0.036**
ODD - OLD	MM IV - WFV	0.088	0.042	0.241
ODD - OLD	MM IV - WFC	-0.033	0.043	1.0
ODD - OLD	PB IV - WFV	-0.031	0.042	1.0
ODD - OLD	PB IV - WFC	-0.151	0.042	**0.0027**
ODD - OLD	WFV - WFC	-0.120	0.042	**0.0285**

Table 4.23: Pairwise comparisons on dialect by context interaction for PC2 in the THOUGHT context. MM IV = mono-morphemic intervocalic; PB IV = pre-boundary intervocalic; WFV = word final pre vocalic; WFC = word final pre consonantal.

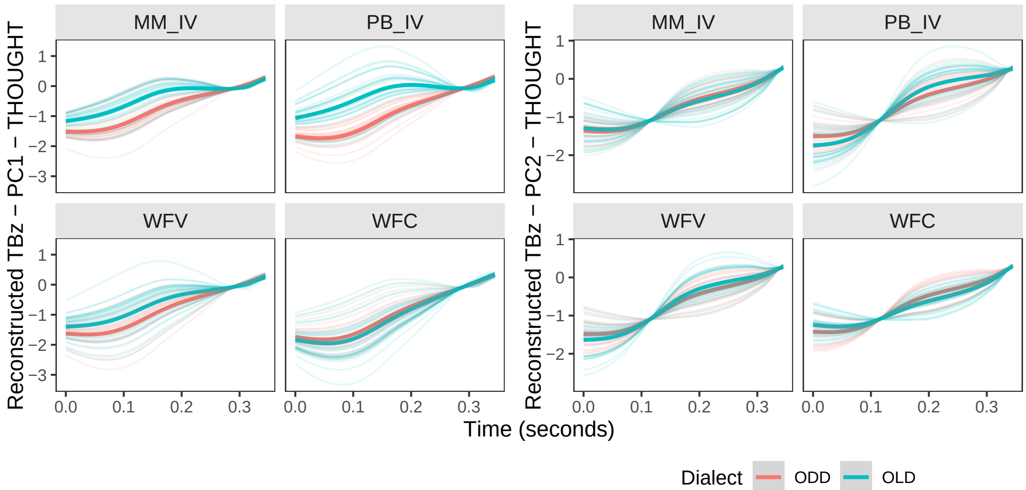

Figure 4.29: Reconstructed TBz trajectories over the THOUGHT + /l/ window based on variance in PC1 (left) and PC2 (right). Faint lines show reconstructions for each token.

To better visualise how variation in PC1 and PC2 interact with morphosyntactic context and dialect, Figure 4.29 shows a reconstruction of changes to the TBz trajectory in the THOUGHT context based on variation in PC1 (left) and PC2 (right). For reconstructions based on PC1, dialectal differences are observed in all but the word final pre consonantal context. In non word final pre consonantal positions, Onset Lightening speakers show an overall higher tongue body position relative to Onset Darkening speakers; this is particularly prevalent in the pre-boundary intervocalic position, where the largest difference in tongue body height is seen. In the word final pre consonantal context, Onset Lightening speakers can be seen to pattern with Onset Darkening speakers, i.e., showing tongue body lowering during the vowel before rising gradually over the lateral. This context then appears to elicit a floor effect which levels out dialectal differences in TBz displacement. For reconstructions based on PC2, dialectal differences are minimal. A small difference can be observed in the pre-boundary intervocalic position, whereby Onset Lightening speakers show a slightly greater magnitude of tongue body raising. For both dialects, the word final pre consonantal context elicits a relatively flatter trajectory of tongue body raising.

4.5.3 Summary of TBz fPCA results

fPCA results show dialectal differences and context driven variation in tongue body vertical displacement. For the FLEECE and KIT contexts, similar patterns of dialect variation were observed, whereby Onset Darkening exhibited a consistently low tongue body across contexts, while Onset Lightening speakers showed variation in tongue body height with context. In intervocalic positions, Onset Lightening speakers had a higher tongue body than Onset Darkening speakers; in word final pre vocalic contexts, dialect differences were minimal, and in word final pre consonantal positions, Onset Lightening speakers had a lower tongue body than Onset Darkening speakers. In addition, differences in the magnitude of tongue body lowering were also seen in the FLEECE context. The greater magnitude of lowering in Onset Darkening speakers suggests a resistance to the lightening effect of the high front vowel context. Dialectal differences in the THOUGHT context manifested in an overall difference in tongue body height within non word final pre consonantal contexts, where Onset Darkening speakers had a relatively lower tongue body. In word final pre consonantal contexts, a floor effect was observed whereby both dialects exhibit considerable lowering, with only small differences between the two.

4.6 Summary and discussion

This chapter set out to answer the following questions: (1) What is the nature of (spatial) articulatory differences in /l/ darkening between dialects?; (2) How does the vowel and morphosyntactic context of /l/ interact with dialectal effects on /l/ darkening? In addressing these questions, I first compared onset /l/s across Onset Darkening and Onset Lightening dialects in three vocalic contexts, before looking across a wider range of coda contexts. A summary of findings from the onset and coda analysis are presented in turn.

4.6.1 Summary of findings

4.6.2 Onset /l/

Dialects were first compared across onset /l/ in FLEECE, KIT, and THOUGHT vowel contexts. A dialect difference in onset /l/ was first verified in acoustics using a measure of F2−F1 at its minimum, where Onset Darkening speakers were observed to have consistently lower minimum F2−F1 values across FLEECE, KIT, and THOUGHT contexts, and values were lower for both dialects in the THOUGHT context. Subsequent articulatory analysis revealed a difference between dialects in tongue body lowering in front FLEECE and KIT vowel contexts, whereby Onset Darkening speakers showed greater initial tongue body lowering, and an overall lower tongue body position. Within the THOUGHT context, a floor effect in tongue body lowering was observed across dialects, and minimal differences were observed.

4.6.3 Coda /l/

The coda analysis compared /l/ across dialects in four morphosyntactic contexts (mono morphemic intervocalic, pre boundary intervocalic, word final pre vocalic, and word final pre consonantal), and FLEECE, KIT, and THOUGHT vowel contexts. The acoustic measure of minimum F2−F1 showed an effect of an interaction between dialect and morphosyntactic context, and dialect and vowel; however, effects of the dialect by vowel interaction were significant only between front and back vowel pairings. Visually, lower F2−F1 values were observed for the Onset Darkening dialect compared to the Onset Lightening dialect within front, but not back vowel contexts. For both dialects, a relatively low F2−F1 value was observed for the word final pre consonantal context compared to the other contexts in front vowel contexts, while in the THOUGHT context, a floor effect was observed with variation between dialect or morphosyntactic context.

The articulatory analysis began by identifying the tongue tip raising gesture of /l/; this was done with the view of deriving a measure of Tip Delay (Sproat and Fujimura, 1993; Strycharczuk and Scobbie, 2015). However, no measures of lateral darkening were derived from this measure given the high frequency of vocalised tokens in the data which meant that the tongue tip gesture could not always be identified. This highlights a limitation of the measure of Tip Delay, which is commonly used temporal measure of lateral darkening, namely that it is only implementable only for non-vocalised variants of /l/. A question left open then is whether we should be treating vocalised variants as categorically separate in terms of how we measure /l/. If so, this raises the further question of where the line for vocalisation should be drawn: at the absence of any TT raising, or at loss of apical contact with alveolar ridge. Within this analysis, I decided to not consider vocalised and non vocalised variants separately, and rather to use a measure that could be operationalised across all realisations of /l/. In this way, a vocalised /l/ was here considered simply an extreme form of /l/ darkening (Strycharczuk, Derrick, and Shaw, 2020).

As with the onset analysis, the tongue body lowering gesture of /l/ was also considered as a measure of /l/ darkening. Given the considerable coarticulatory effects of the preceding vowel, the time of achievement of the tongue body lowering gesture of /l/ could not be reliably identified, hence the maximum tongue body lowering value (TBz) was compared across dialects, vowel contexts and morphosyntactic contexts. Note, the inability to locate the time of the tongue body lowering gesture meant that a measure of tip delay could not be calculated on the subset of non-vocalised tokens. Visually, the tongue body lowering maximum showed an effect of morphosyntactic context and dialect, however, in front vowel contexts only. Within these contexts, Onset Darkening speakers had an overall lower tongue body position than Onset Lightening speakers in all but the word final pre consonantal position. Lower tongue body positions were observed in word final pre consonant position compared to any other morphosyntactic position in the FLEECE context for both dialects, and in the KIT context for Onset Lightening speakers. While a three-way interaction between dialect, vowel, and context was not explicitly tested for, interactions between dialect and vowel, and dialect and morphosyntactic context were found to be significant. A functional principal component analysis was performed on the dynamic trajectories of tongue body vertical displacement. Unlike the other measures, this approach did not privilege a discrete time point but considered entire trajectories of displacement across the lateral and preceding vowel. fPCA results showed dialectal differences in TB vertical displacement to be a function of the interaction between vocalic and morphosyntactic context. In the high front vowel context, Onset Lightening speakers showed variation in tongue body height with morphosyntactic context, showing increasing lowering at stronger morphosyntactic boundaries. Onset Lightening speakers also showed a lightening effect of the high vowel context through a reduced magnitude of lowering, while Onset Darkening speakers showed a resistance to this effect. Within the THOUGHT context, dialectal differences in tongue body height were observed in the expected direction in all but the word final pre consonantal context; here a floor effect was observed, where dialects showed little variation in tongue body height and Onset Lightening speakers were seen to adopt the pattern of initial lowering used in Onset Darkening speakers.

4.6.4 Contexts of variability

Results of this analysis have shown dialectal differences in lateral darkening, conditioned by interactions with the preceding vowel and the morphosyntactic context. While three way interactions were not measured statistically, patterns were visually observed between vowel and morphosyntactic context. Such patterns were observed to be measure specific. Within the onset context, dialectal differences in tongue body lowering were observed in front vowel contexts, while only minimal differences in trajectory shape were observed for the THOUGHT context. In the coda contexts, minimum F2−F1 and minimum TBz also showed dialectal differences in front vowel contexts only. Within back vowel contexts, a floor effect was observed where F2−F1/TBz values were low for all coda contexts, with little variation between them. Such findings are compatible with results of Strycharczuk and Scobbie (2015), who found greater morphosyntactic variability within front, compared to back vowel contexts. Their explanation for the findings can also be applied here; because a front vowel context has a lightening effect on the lateral, there is a larger possible range of variation of darkening compared to back vowel contexts. However, dynamic analyses of the tongue body lowering across THOUGHT contexts did in fact show dialectal differences in tongue body height in non word final pre consonantal contexts.

4.6.5 Incidental findings

Vocalisation

In pursuit of deriving a measure of tip delay, the coda analysis revealed patterns of varying degrees of vocalisation for speakers of both dialects, including full tongue tip raising, absence of tongue tip raising, and partially reduction in raising. The varying degrees of tongue reduction captured here suggest vocalisation to be a non binary phenomenon for these dialects. This finding is consistent with previous observations for gradient reduction in the spatial magnitude of the tongue tip raising gesture of /l/ in Southern British English (Strycharczuk et al., 2020), and Australian English (Szalay et al., 2022), however, contrasts with the categorical pattern of vocalisation found in New Zealand English (Strycharczuk, Derrick, and Shaw, 2020). Further, multiple stages of vocalisation were observed within the same speaker, and for the same token. Such within speaker variability in vocalisation suggests a degree of instability, typical of an ongoing change. The degrees of vocalisation which were captured here was made possible by the articulatory method used. EMA tracks the position of sensors which are glued directly onto the tongue surface, which more readily lends itself to capturing gradient variation. This contrasts with electropalatography, which captures only the presence or absence of lingual-palatal contact, and not reduction in tongue tip raising where no palatal contact is made (e.g., Scobbie and Pouplier, 2010).

Regarding the contexts of vocalisation, all instances of (full and partial) vocalisation were observed exclusively in word final pre consonantal contexts for Onset Lightening speakers, consistent with findings of Strycharczuk et al. (2020). For Onset Darkening speakers, instances of full vocalisation were also observed in word final pre consonantal contexts only, while partial vocalisation spanned a wider range of contexts, including word final pre vocalic and intervocalic contexts. For both dialects, partial vocalisation was found predominantly in front vowel contexts, but was also observed in the THOUGHT context for 1 Onset Lightening speaker, and 2 Onset Darkening speakers. Full vocalisation was observed in front vowel contexts only, consistent with findings of Hardcastle and Barry (1989), who suggest vocalisation to favour front vowel contexts because front vowels remain perceptually recoverable from the adjacent lateral in absence of a tongue tip gesture.

Categorical or gradient?

This analysis examined /l/ within multiple morphosyntactic and vocalic contexts in order disentangle darkening effects of context from dialectal effects. Thus, while is was not the aim of this chapter to resolve questions regarding the categoricity or gradience of /l/ darkening, findings of this analysis are inevitably relevant to this issue. Within the coda analysis, where multiple morphosyntactic positions were considered, evidence lending support for both categorical and gradient variation were observed, depending on the combination of the measure used, the context (morphosyntactic and vocalic), and dialect. Categorical patterns are defined as those where there is a clear binary split between morphosyntactic contexts along a given measure of /l/ darkening, for example between the word final pre consonantal context and all other contexts. Gradient patterns are defined as those where differences proceed incrementally across contexts. In using the term "gradient", there is of course a large caveat; because not all the possible morphosyntactic contexts were examined in this study, whether or not a truly gradient pattern of variation was observed, cannot be determined (Turton, 2014). Given this, I make only tentative comments on the patterns of /l/ darkening observed within this analysis. The above section made comment on how the degrees of tongue tip reduction in /l/ vocalisation may lend support for a gradient view of /l/ darkening. Beyond this, the measure of F2−F1 at the minimum saw that, in front vowel contexts, the word final pre consonantal context behaved markedly different to all other morphosyntactic contexts across both dialects. Seemingly gradient patterns of variation were observed in maximum tongue body lowering for the Onset Lightening Dialect in front vowel contexts, and for the Onset Darkening dialect in the FLEECE context only. However, for the Onset Lightening dialect, the tongue body position for the word final pre consonantal context, i.e., the strongest boundary, was considerably lower than a strictly gradient pattern would predict. The fPCA results revealed seemingly categorical results within the FLEECE context between the word final pre consonantal context and other contexts, however, variation with boundary strength was seen within KIT and THOUGHT contexts, but for Onset Lightening speakers only. That different measures vary in the seemingly categorical or gradient patterns of variation they elicit was also found in Turton (2014), likely a result of differing sensitivities of different measures, and the different domains to which they pertain.

These results also suggest that the categorical or gradient nature of /l/ darkening may be partially conditioned by vocalic context. One hypothesis is that the greater the lightening effect of the vowel on /l/, the more categorical the data is going to seem since the there is likely going to be a bigger contrast with the strongest morphological boundary (in this case, the word final pre consonantal context), and the weakest (in this case, the mono-morphemic intervocalic context). This was not fully confirmed by these data, however, the front vowel contexts did see a wider range of morphosyntactic variation than the back vowel context. To test this hypothesis more rigorously would require examining /l/ darkening within a wider range of vowel and morphosyntactic contexts.

4.7 Conclusion

In this chapter, I presented an analysis of lateral darkening across onset and coda contexts whereby Onset Darkening and Onset Lightening dialects were compared. Both onset and coda analyses revealed dialectal differences in the degree of tongue body lowering and F2−F1. Measure specific interactions between preceding vowel and morphosyntactic context were also observed. These findings have particular relevance to the next chapter of this thesis which relies upon the existence of an articulatory difference in lateral darkening between dialects. Findings for interactions between dialect and vocalic context, and dialect and morphosyntactic context further highlights to the need to study /l/ darkening within a range of morphosyntactic and vocalic contexts in order to better understand the nature of this interaction.

Chapter 5

Timing of /l/ Clusters in Onset Lightening and Onset Darkening Dialects

In the previous chapter, I examined the articulatory correlates of lateral darkening. Differences in lateral darkness between Onset Lightening and Onset Darkening dialects manifested in differences in vertical tongue body displacement. I here build upon this finding by exploring the possible interaction between lateral darkness and cluster timing. The chapter will begin with an overview of the lateral cluster timing literature which problematises previously discussed relationships of consonant cluster timing. I then turn to the study of this chapter, which explores the interaction between lateral darkening and cluster timing empirically, asking whether the darkness of a lateral influences timing of lateral constant clusters.

5.1 Lateral Cluster Timing: Literature

In Section 2.2, a neat relationship between syllable structure complexity and consonant-vowel timing was depicted; namely, a global C-centre timing pattern for onset clusters, and a local, sequential timing pattern for coda clusters. A C-centre pattern describes the stable relationship maintained between the centre on the consonant onset and the vowel across both singleton and cluster onsets. A sequential timing pattern describes the serial stacking of coda consonants within a coda cluster. In practice, surface timing patterns regularly deviate from these neat ideals of gestural timing. Clusters containing laterals are particularly susceptible to variable and deviant timing patterns (e.g., Goldstein et al., 2009; Marin and Pouplier, 2010; Mücke, Hermes, and Tilsen, 2020). One suggestion is that the darkness of the /l/ may mediate the timing of lateral clusters because of their spatial gestural differences (Marin and Pouplier, 2014). This hypothesis provides the rationale for the study in the latter half of this chapter. For now, I present findings from studies which have measured the timing of lateral clusters, first in onsets and then in codas. English and non-English examples will be provided to enable comparisons to be made between lateral variants from different systems.

5.1.1 Lateral Onsets

Variable timing patterns have been reported for lateral onset clusters. To recap, the typical onset timing pattern is the C-centre pattern, whereby the distance between the consonant centre and a fixed anchor point remains the same across singleton and cluster contexts. Such typical C-centre timing patterns have been reported for lateral onset clusters for English speakers, as was shown in Browman and Goldstein (1988) in their study of one American English speaker; Honorof and Browman (1995), in their X-ray Micro-beam study on four American English speakers, and in Marin and Pouplier (2010), in their EMA study on seven American English speakers. For each of these studies, the C-centre was found to be the most stable interval across singleton-cluster word pairs.

Non-C-centre patterns have been reported for lateral onset clusters in English (Goldstein et al., 2009) and German (Brunner et al., 2014; Mücke, Hermes, and Tilsen, 2020; Pouplier, 2012). For English speakers, Goldstein et al. (2009) reports on an X-ray Micro-beam study which found a non-symmetrical shifting of /p/ and /l/ consonants within /p l + V/ structures. Specifically, /l/ was observed to shift rightward towards the following vowel to a lesser degree than /p/ was observed to shift leftwards. This runs counter to predictions for onset cluster timing, in which equal degrees of leftward shift of C1 and rightward shift of C2 are required to maintain a stable relationship between the vowel and consonant cluster centre, i.e., the 'C-centre Effect'. An asymmetrical shifting pattern on the other hand, means that a stable relationship between the anchor and consonant cluster centre is no longer maintained and there is no C-centre Effect. Mücke, Hermes, and Tilsen (2020) report on an EMA study of /pl/ onset clusters in younger and older German speakers. Neither group showed the expected symmetrical shift patterns; younger speakers showed no rightward shift of /l/ towards the vowel, and the older speakers showed only a little rightward shift of /l/ which was less than the leftward shift of /p/. These differences were explained by the increased tendency of older speakers to hyper-articulate, thus making a stronger anti-phase coupling between C1 and C2 more likely in older speakers than younger speakers. The authors note these patterns, in particular those of younger speakers, to be superficially more compatible with the timing patterns expected for non-branching languages such as Moroccan Arabic. Figure 5.1 illustrates symmetrical and asymmetrical shifting patterns for onset clusters. The lack of a sufficient rightward shift of C2 in the bottom diagram of Figure 5.1 results in a leftward shift of the C-centre compared the singleton context.

To explain these deviant timing patterns of /l/ clusters, a number of accounts have been proposed. For the asymmetrical shifts of /pl/ onset clusters in English, Goldstein et al. (2009) hypothesise about the multi-gestured nature of /l/. Because /l/ has multiple gestures - specifically, an apical and a dorsal gesture - coupling both of these gestures to the vowel results in more coupling links and a tighter, more stable overall coupling between /l/ and the vowel. This results in less overlap with the vowel when moving from a singleton to a cluster context. They further verified this using computational modelling; the model which best fit their findings reduced the coupling strength of /p/ to the vowel to 0.74 (down from the default of 1) and coupled both the apical and dorsal gestures of /l/ to the vowel. Coupling just the apical gesture of /l/ to the vowel failed to show results which matched their findings.

Mücke, Hermes, and Tilsen (2020) offer a similar account for the asymmetrical shifts in /pl/ clusters observed for younger and older German speakers. They

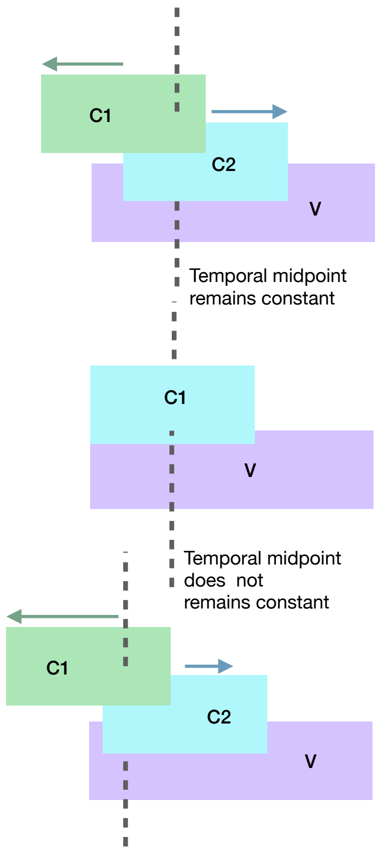

Figure 5.1: Illustration of an asymmetrical timing pattern for onset clusters. Diagram shows how a lack of rightward shift of C2 results in a shift of the C-centre.

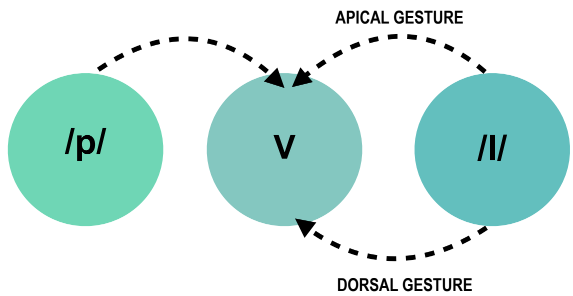

Figure 5.2: Diagram illustrates multiple in-phase couplings of /l/ - V (apical and dorsal gestures), and single in-phase coupling from /p/ - V, as part of Goldstein et al. (2009)'s explanation for non-C-centre organisation

suggest, like Goldstein et al. (2009) that the non-C-centre pattern may still be compatible with the coupled oscillator model and with the branching structure of German onsets. They argue that if non-theoretically motivated constraints within the coupled oscillator model are relaxed then the anomalous patterns can be explained, achieved by relaxing the assumption of balanced coupling between C1 and V, and C2 and V. If coupling strengths are allowed to be unbalanced (e.g., C2 may be more strongly coupled to the vowel than C1 is to the vowel), then asymmetrical shifts of C1 and C2 away from and towards the vowel are coherent with the coupled oscillator model and a branching onset structure. The authors seek to make a broader point with their paper, the point that seemingly incongruent results, such as a non C-centre timing pattern in branching onsets, do not necessitate radical changes to, or rejection of the theory. This thesis will proceed with a mindful appreciation of the importance of such a stance.

Another explanation for the non-C-centre timing patterns found in German onsets is proposed by Brunner et al. (2014) who discusses the role of segmental properties, including vowel height and the coarticulatory resistance of the consonants within the cluster. Brunner et al. (2014) argue that segmental factors influence C-centre measures to such an extent that the C-centre effect is an ineffective measure of the underlying coupling relations. To take an example of how measures can be affected by segmental properties, the authors consider the lag between the vowel and the pre-vocalic consonant in a CV versus a CCV onset as part of a coarticulation analysis, where they look at the relationship between this lag measure and the degree of vowel compression. Citing Lehiste (1970) the authors show how this lag measure can be influenced by the coarticulatory properties of the consonants involved. Where C1 in a CCV cluster is "coarticulatory aggressive" it has a considerable coarticulatory effect on C2 and hence changes its articulation. This in turn changes the degree of tongue displacement and hence the time required to reach the following vowel. The height of the vowel also affects the amount of tongue displacement and hence the consonant-to-vowel lag in the same way. Pastätter and Pouplier (2017) also observe onset timing patterns to be conditioned, at least in part, by the coarticulation resistance of the vowel adjacent consonant. "Coarticulation resistance" (Bladon and Al-Bamerni, 1976) describes the resistance of a consonant to the coarticulatory influence of its surrounding context. This is closely related to other terms such as coarticulatory dominance or aggressiveness mentioned above, which are used to refer to the degree to which a consonant both resists and exerts coarticulatory influence (Proctor et al., 2019; Recasens and Rodríguez, 2016, 2017). Relatedly, evidence has been found for a relationship between the starting position of the tongue dorsum and the onset of gestural movement. For example, Shaw and Chen (2019) report evidence from Mandarin speakers which found the starting position of the tongue dorsum to mediate the onset time of a following vowel gesture. Specifically, the study found vowel gestures to begin earlier when the distance between the starting position of the tongue dorsum was further away from the vowel target. Shaw and Chen (2019) suggest these result to be incompatible with a feed-forward model of speech timing as implemented within the articulatory phonology framework (see Section 2.2.7 for a more detailed discussion on this.) As such, the authors propose that results may be explained by the presence of a "neutral attractor" or "downstream targets" (Shaw and Chen, 2019).

5.1.2 Lateral Codas

Both typical and atypical timing patterns have also been observed for lateral clusters in coda position. In an EMA study of American English speakers, Marin and Pouplier (2010), found that laterals did not show the expected sequential timing pattern for complex coda clusters. Rather, when moving from a singleton coda (vowel + lateral) to a complex coda cluster (vowel + lateral + stop consonant), /l/ showed a leftward shift towards the vowel, similar to that of the C-centre pattern predicted for onsets (see Figure 5.3). This was also confirmed by an acoustic study of American English speakers by Katz (2012) who found that when moving from a singleton coda to a complex coda cluster, for /l/ and /r/, the vowel was reduced in duration, i.e., there was a leftwards shift of /l/ and /r/ causing greater overlap with the preceding vowel. Notably, this shift did not occur for obstruents which follow the expected sequential timing pattern. Articulatorily, this may correspond to an earlier onset of one or both of the lateral's tongue tip and tongue body gestures, during the vocalic interval.

Subsequent studies have found such deviant patterns of /l/ coda clusters to be language specific. Pouplier (2012) found clear /l/ coda clusters of German speakers to exhibit the predicted local organisation, contrary to the deviant timing pattern found for the dark /l/ coda clusters of American English speakers (Marin and Pouplier, 2010). Further, the differences in coda cluster timing observed between American English and German coda /l/ clusters was hypothesised to arise from the articulatory differences between clear and dark /l/ between languages (Marin and Pouplier, 2014; Pouplier, 2012). To test this hypothesis, Marin and Pouplier (2014) compared the timing patterns of /l/ and /r/ coda clusters in Romanian speakers, where /l/ and /r/ are polarised in terms of their articulatory realisations. Specifically, Romanian /l/ is clear, similar to German, while the Romanian alveolar trill has a similar retracted tongue dorsum position to American English dark /l/ (e.g., (Recasens, 2012)). They found /l/ coda clusters in Romanian speakers to show a local organisation, similar to German, while /r/ coda clusters showed an atypical leftward shift in the cluster context compared to the singleton context, similar to coda /l/ in American English (Marin and Pouplier, 2014). Moreover, by comparing liquids with distinct articulatory configurations within a language, Marin and Pouplier (2014) provided more evidence that the cross-linguistic differences in /l/ coda cluster timing patterns had an articulatory explanation.

One explanation for the difference in coda cluster timing of dark liquids suggested by Marin and Pouplier (2014) is that the dark /l/s of American English speakers exhibit greater tongue body reduction in a cluster compared to a singleton context. Two premises underlie their hypothesis. The first is that consonants are shorter in duration when in a cluster compared to a singleton context (Haggard, 1973, cited by Marin and Pouplier, 2014). The second is that the tongue body lowering/retraction gesture precedes the tongue tip raising gesture of American English coda /l/ (e.g., Sproat and Fujimura, 1993). Since lateral timing was measured from the tongue tip gesture of /l/ in Marin and Pouplier (2014)'s study, a tongue body gesture of /l/ which is reduced due to the constraints of a cluster context will result in a shorter lateral to anchor lag relative to a singleton context. Figure 5.4 provides a schematic illustration of where the tongue body gesture is expected to occur relative to the tongue tip gesture for a dark coda /l/. From this diagram, we can see how a reduction in the tongue body gesture, shown by the red dashed line, could cause a leftward shift of the tongue tip gesture, and hence give the appearance of increased overlap with the vowel or vowel reduction. While a plausible suggestion in light of the available evidence, the prediction of a reduced tongue body gesture is somewhat speculative given that only the tongue tip gesture was measured in Marin and Pouplier (2014). This is a common methodological issue that arises due to the difficulty in identifying the tongue body gesture of /l/ in vocalic contexts (see also Marin and Pouplier, 2010).

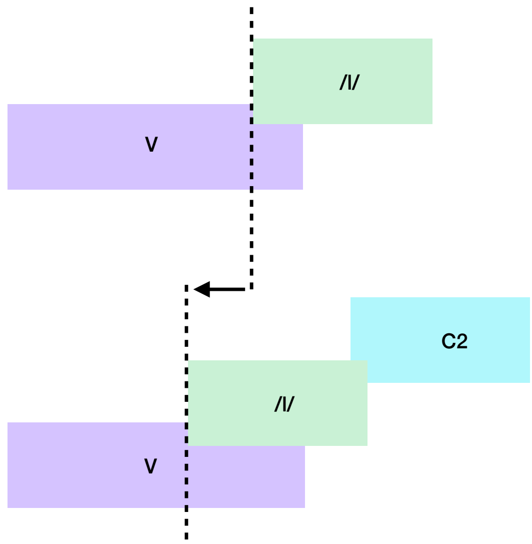

Figure 5.3: Illustration of atypical coda /l/ timing pattern, adapted from Marin and Pouplier (2014, p. 25). Arrow shows how /l/ shifts leftwards in a coda cluster relative to when in a singleton coda. A stable left edge timing pattern is not maintained across singleton and cluster contexts.

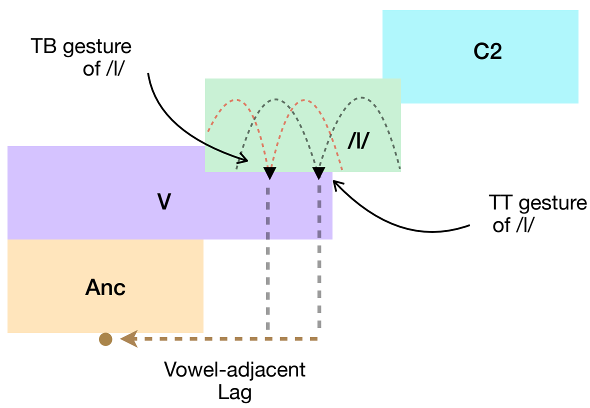

Figure 5.4: Location of the tongue body gesture relative to the tongue tip gesture in Marin and Pouplier (2014) hypothesised explanation for the atypical timing pattern of dark liquids in a coda cluster context.

5.1.3 Summary of Lateral Timing Patterns

The above sections have reviewed numerous studies which have found both typical and atypical timing patterns of lateral clusters, the latter seemingly conflicting with the predictions put forth by the coupled oscillator model. However, as we have also seen, these unexpected timing patterns may be accounted for by: (1) relaxing non-theoretically motivated constraints of the coupled oscillator model, and (2) taking into consideration segmental properties, coarticulatory resistance, and the articulatory distance between adjacent sounds.

Of the studies reported on here, evidence is presented from speakers of German (e.g., Mücke, Hermes, and Tilsen, 2020), American English (e.g., Marin and Pouplier, 2010), and Romanian (e.g., Marin and Pouplier, 2014). One difficulty with which we are faced when interpreting these results is the cross-linguistic confound presented by comparing data collected from speakers of different languages. For example, while there appears to be converging evidence for an articulatory explanation for the differences in lateral coda cluster timing across languages, this is difficult to verify. The aim of this study is to resolve this problem by examining lateral clusters within two dialects of the same system. While sharing the general phonotactic constraints of the language, the dialects differ in a crucial regard: lateral darkening. Through this design, it is hoped that this study will gain clear insights into the interaction between lateral darkening and cluster timing.

5.2 Articulatory Study of Dialect Variation in Lateral Cluster Timing

5.2.1 Overview

The remainder of this chapter details an articulatory study into the lateral cluster timing patterns of Onset Darkening and Onset Lightening dialects, and in so doing, explores the relationship between lateral darkening and cluster timing. The specific research question posed is:

RQ: *How do spatial differences in /l/ darkening between dialects affect patterns of /l/ cluster timing?*

Before moving onto the details of this study, it is important to draw attention to assumptions which underlie the study design. The first point to make is that this study is not testing the syllabic organisation of the language. Rather, I take for granted that the British English speakers used within this study have a complex syllabic organisation, whether or not onset clusters show a strict C-centre pattern. Recent studies have shown variability in C-centre measures to result from segmental confounds within stimuli, such as spatio-temporal differences between segments (Sotiropoulou and Gafos, 2022). Sotiropoulou and Gafos (2022) argue that it is rather *how* different syllable organisations respond to variability in confounding parameters that we should really interested in. Specifically, Sotiropoulou and Gafos (2022) have shown that complex versus simplex syllable organisations respond differently to these variations (or scaling) of parameters, and it is this differential response which is really diagnostic of syllable organisation.

The second assumption regards the finding of the previous chapter for the tongue body vertical gesture to best capture differences in lateral darkness between dialects. Within the study of this chapter, however, lateral timing will be measured using the velocity minimum of the tongue tip vertical gesture. While it may appear counter intuitive to focus on the tongue tip when the tongue body was found to be the most salient discriminator of lateral darkness, for reasons discussed in the previous chapter, it was not possible to pinpoint the time of the achievement of the tongue body lowering gesture across the range of contexts examined here. However, given the findings presented in Marin and Pouplier (2014) which suggest a relationship between the timing and magnitude of the tongue body gesture within liquid clusters and lateral cluster timing, when timing is measured using the tongue tip gesture of /l/, it is reasonable to hypothesise that a similar relationship will be found here. This hypothesis is grounded in the findings of Chapter 4 for spatial differences in the tongue body lowering gesture of /l/ found to mediate /l/ darkening between Onset Lightening and Onset Darkening dialects.

5.2.2 Methods

Audio-synchronised electromagnetic articulography data were obtained from Onset Darkening and Onset Lightening dialect speakers. Further details of the general experimental method for this study, including the speakers used and the nature of the experiment can be found in Chapter 3. Below, I provide details of the methods used which are specific to this study. Most of the methodological decisions are based upon the works of Marin and Pouplier (2010) and Marin and Pouplier (2014), which were central to this study.

Stimuli

Stimuli consisted of a series of singleton-cluster pairs, embedded within the carrier phrase *say the xxx*, or *say xxx again*. The stimuli list was loosely based upon the stimuli used in Marin and Pouplier (2010). Target words contained /pl-/, /kl- /, /-lp/, /-lk/ clusters, for example, *plug* and *gulp*, and each cluster occurred within a front and back vowel context, for example *clip* and *club*. Each unique token was repeated 4 times, though some tokens were excluded due to audio errors and mispronunciations. Tables of the stimuli tokens can be found below. Note, one coda context which contained a back vowel in the /-lk/ context was excluded from the study due to substantial durational differences in the vowel between the cluster and singleton tokens. For all onset and coda analyses, speakers L03 and L08 were excluded due to a missing TB sensor. For the coda *fill - philp* context, speaker L07 was excluded due to a missing lower teeth sensor, as were speakers S05 and S08

- though coda results are not presented for Onset Lightening speakers due to the large amount of vocalised tokens. Applying these exclusions, a total of 376 tokens were used in the onset analysis, 213 for Onset Lightening speakers and 163 for Onset Darkening speakers. For the coda analysis, there were a total of 299 tokens, 167 for Onset Lightening speakers, and 132 for Onset Darkening speakers.

Table 5.1: Onset cluster - singleton pairs

tea clip	tea lip
tea club	tea lug
tea plick	tea lick
tea plug	tea lug

Table 5.2: Coda cluster - singleton pairs

milk in	mill in
gulp in	gull it
philp it	fill it

Measurements

Extracting Time of Gesture Achievement

The time of gestural achievement was extracted for each consonantal segment in the target word. Gesture achievement of a given segment was defined as the point in time when the absolute velocity of the relevant sensor(s), in the relevant dimension, reached its minimum value, see Figure 5.7. The time of velocity minimum was found to be consistently measurable within the data and hence was considered appropriate for this purpose. Further, the decision was made not to define the time of gesture achievement as the time at which articulator velocity dropped to 20% of the preceding velocity peak, as in Marin and Pouplier (2014). This was because velocity peak height was observed to be impressionistically variable between speakers and dialects. Hence, 20% of a high peak versus 20% of a low peak may have resulted in large differences in timing. For these reasons, the time of the velocity minimum was considered the most sensible measure of time of gesture achievement.

For bilabial consonants, /p/, /b/, /m/, a measure of lip aperture was calculated from the euclidean distance of the upper and lower lip sensors in the vertical and horizontal planes (z/x dimensions). The achievement of a bilabial gesture then was defined as the point in time when the absolute velocity of the lip aperture (LAxz) reached its minima. Note, due to difficulties with the lip sensors during recording, three speakers of the Onset Darkening Dialect were recorded without lip sensors. For one of the three speakers, the velocity minima of the lower jaw in the x/z plane was used to generate the time of bilabial target achievement, since the movement of the jaw is coupled to that of the lower lip. For the remaining two speakers, the lower jaw sensor was unavailable, so these speakers were excluded from this analysis. For /t/, /d/, /s/, /l/ segments, the relevant sensor was the tongue tip sensor in the vertical plane, while for /k/, /g/ segments, the tongue dorsum sensor, also in the vertical plane, was used. As with lip aperture, gestural achievement was defined as the point in time when the absolute velocity of the relevant sensor reached its minima.

For each audio file (each containing a single sentence, e.g., *say tea plug again*), a TextGrid was created in Praat (Boersma and Weenink, 2011) and sentences were manually labelled. Acoustic segmentation was then performed using Montreal Forced Aligner (McAuliffe et al., 2017), henceforth MFA. Segment intervals were manually edited according to an alphanumeric annotation system which specified a segments' position within the target word. For anchor consonants (e.g., /g/ in *plug*), a "0" was added to the consonant (e.g., phone label = "G0"). In cluster tokens, such as *plug*, the first consonant in a cluster was suffixed with the number 1, and the second consonant was suffixed with the number 2 (e.g., phone labels = "P1", "L2"). In singleton tokens with only one non-anchor consonant (e.g., *lug*), no numbers were added to the consonant label (e.g., phone label = "L"). See Table 5.3 for an illustration of the alphanumeric annotation system. The search window for gestural targets in velocity profiles spanned approximately one segment before and after the acoustic boundary depending on the context. Example annotations are shown in Figures 5.5 and 5.6 for words *lug*, and *plug*. Where the segment was not a peripheral segment, such as /l/ in *plug*, the interval was made as large as possible within the constraints of the surrounding segments, and the interval was instead increased by widening the temporal window to the segment before and/or after when searching for the velocity minima.

TextGrids and their corresponding audio and articulatory files were imported into R, where each speaker was separately processed using the *tardis* and *tadaR* R packages (Kirkham, 2024a,b), as detailed in the methods section (Chapter 3).

Table 5.3: Alphanumeric annotations

Onset Cluster	P	L	U	G
	P1	**L2**		**G0**
Onset Singleton	L	U	G	
	L		**G0**	
Coda Cluster	G	U	L	P
	G0		**L1**	**P2**
Coda Singleton	G	U	L	
	G0		**L**	

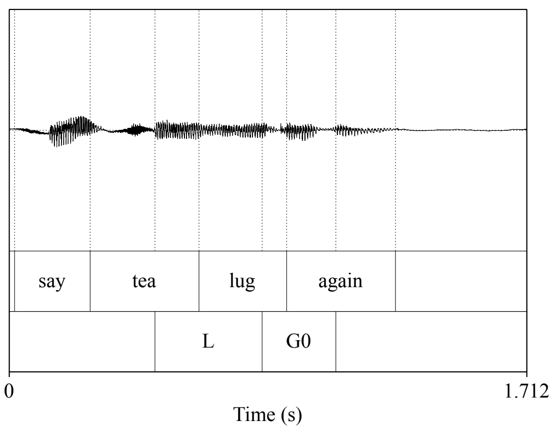

Figure 5.5: Figure showing example Text Grid annotation for *tea lug*

To identify the time of gesture achievement for each segment, the velocities of the relevant sensors, were plotted to ensure trajectories looked sensible and contained the relevant information, i.e., two velocity peaks and a local velocity minimum. The first velocity peak shows the articulator moving away from the previous target and towards the target, before slowing down as it approaches the current target. The velocity minimum is the point at which the articulator is at its slowest, corresponding to the achievement of the target. The second velocity peak then, shows the articulator in the post-target achievement stage, where it moves away from the target, and then begins to slow as it approaches the next target. This is shown in the schematic illustration in Figure 5.7. Once I had established that velocity profiles of all segments met these criteria, the "findpeaks" function from the *pracma* (Borchers, 2022) R package was used to identify the two velocity peaks and minima within each segment. This function identifies the time points of maximum and minimum points

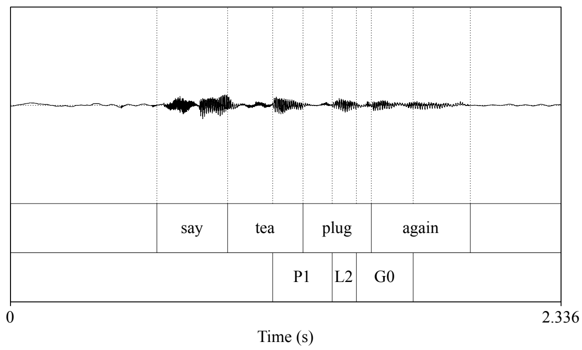

Figure 5.6: Figure showing example Text Grid annotation for *tea plug*

on a dynamic trajectory within defined parameters. To improve the success rate of the "find peaks" function in correctly identifying the peaks, parameters were individually specified for each segment, including the maximum peak height of each peak, and the temporal search window. Results were manually checked for each token.

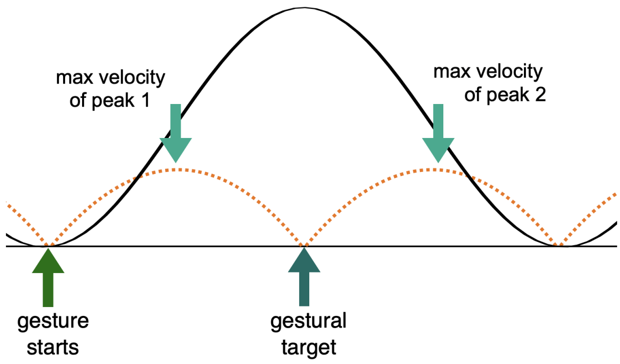

Figure 5.7: Schematic illustration of velocity peaks used for identification of gestural targets. Orange dashed line shows velocity trajectory; solid black line shows displacement trajectory.

Vowel Adjacent and Vowel Remote Lags

Vowel-remote and vowel-adjacent lags were calculated for each cluster/singleton pair, as in Marin and Pouplier (2014). "Vowel adjacent" here refers to the consonant within a cluster which neighbours the vowel, such as /l/ in *plug*, or *lug*, while "vowel remote" refers to the consonant in the cluster which is furthest away from the vowel, such as /p/ in *plug*. Hence, singleton tokens have a vowel-adjacent consonant only. The lag duration captured the distance between the target achievement (i.e, velocity minima) of the vowel adjacent or remote consonant, and the target achievement of a post-vocalic anchor consonant, i.e., /g/ in the example of *plug*. Figures 5.8 and 5.9 show schematic illustrations of vowel adjacent and remote lags for onset and coda contexts. The top half of each figure illustrates vowel-remote and vowel-adjacent lags for a cluster context, while the bottom half shows the vowel-adjacent lag for a singleton context. An example of vowel remote and vowel adjacent lags within the data can be seen in Figure 5.10 which shows a token of *clip* from Onset Lightening speaker S03. Here, velocity and displacement trajectories are shown for the tongue dorsum raising gesture of /k/ in *clip* (the vowel remote consonant), the tongue tip raising gesture of /l/ in *clip* (the vowel adjacent consonant), and the lip closure gesture of /p/ (the anchor consonant). Red dots show velocity minima used to define the gestural target achievement. This example also highlights how identifying the gestural target is sometimes ambiguous, as can be seen in the case of the tongue dorsum raising for /k/ in the top panel of Figure 5.10, where the tongue is held in a raised position.

For this study, I am interested in the comparison between vowel-adjacent lags of /l/ across matched cluster and singleton pairs, such as *plug* and *lug* (note: /l/ is always C2 within a C1 + C2 + V + Anc token), for predictions can be made regarding how the vowel-adjacent /l/ to anchor lags should pattern across such pairs. For example, a C-centre pattern would predict a shorter /l/-to-anchor lag in the cluster context compared to the singleton context. This is because, when moving from a singleton to a cluster, C2 must shift rightwards towards the vowel, and C1 leftwards away from the vowel in order to maintain a constant consonant centre (Marin and Pouplier, 2014). However, this information alone cannot tell us whether or not the C-centre has in fact been maintained, rather it is an 'indirect' measure of C-centre stability (Tilsen and Goldstein, 2012). Answering this question requires a consideration of both C1 and C2 of the cluster. This will be considered in the 'direct' measures of C-centre stability which follow.

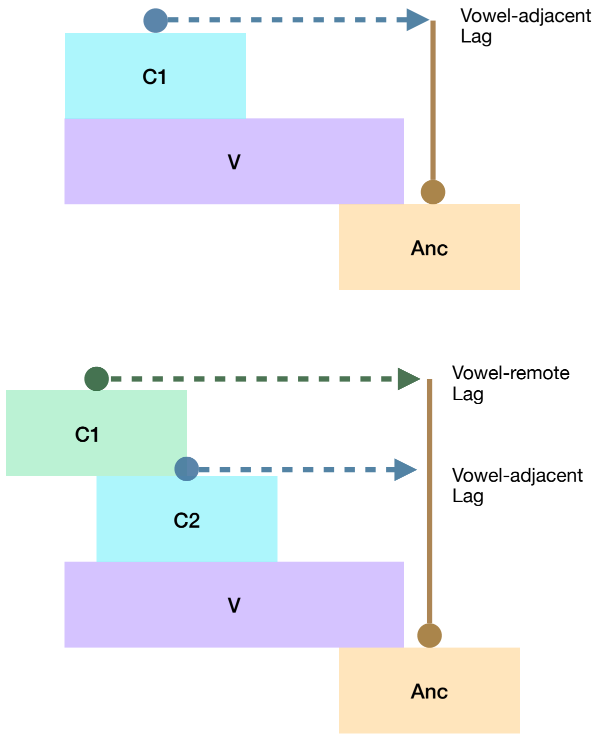

Figure 5.8: Vowel-adjacent and vowel-remote lags for onsets in singleton onsets (top) and cluster onsets (bottom). Arrows show the distance measured for each lag. Circles denote the velocity minima of the relevant gesture.

Stability Measures

"Direct" stability measures are an explicit measure of C-centre to Anchor stability across a singleton / cluster pair to determine whether C-centre stability has been maintained (Tilsen and Goldstein, 2012). A lag from the Consonant Centre to Anchor is taken for both the singleton and cluster tokens within a pair. For singleton tokens, such as *lug*, this lag is the interval between the velocity minima of /l/ and the velocity minima of the anchor consonant, here /g/. For cluster tokens, the lag is the distance between the anchor consonant and the time point equidistant between the velocity minima of C1 and C2. This lag, shown by the grey dashed line in the bottom figure of Figure 5.11, will be referred to as the C-centre lag. The C-centre lags will then be compared within singleton / cluster pairs to determine whether or not a C-centre pattern has occurred; a stable lag across the pair would be indicative of a C-centre pattern. The singleton C-centre lag was also compared to two other cluster lags: the Left-Edge lag and Right-Edge lag. For onset clusters, the Left-

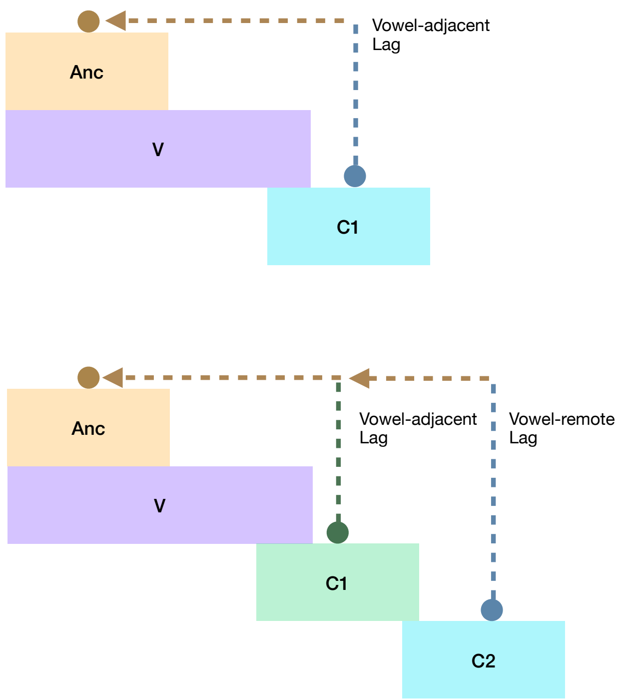

Figure 5.9: Vowel-remote and vowel - adjacent lags for codas, in singleton codas (top) and cluster codas (bottom). Arrows show the distance measured for each lag.. Circles denote the velocity minima of the relevant gesture.

Edge lag was calculated as the distance between the velocity minima of C1 and the anchor (shown by the green dashed line in the bottom figure of Figure 5.11), and the Right-Edge lag was calculated as the distance between the velocity minima of C2 and the anchor (shown by the blue dashed line in the bottom figure of 5.11); the reverse was true for codas, see the bottom figure of Figure 5.12. Comparing the singleton C-centre lag with the cluster C-centre, Left-Edge and Right-Edge lags of the cluster token enabled a test of Centre, Left-Edge or Right-Edge stability across the singleton-cluster pairs.

For each singleton-cluster pairing, the C-centre lag of the singleton token was compared to each of the cluster token lags: Left-Edge, Right-Edge, and consonant C-centre, to determine the most stable interval across the the singleton-cluster pair. Comparisons are made visually and quantified within linear mixed effects models. Specific model details are provided below. Stability measures used here are loosely based on those of Marin and Pouplier (2014); however, the present study differs in the time points between which intervals are drawn. Unlike this study, Marin

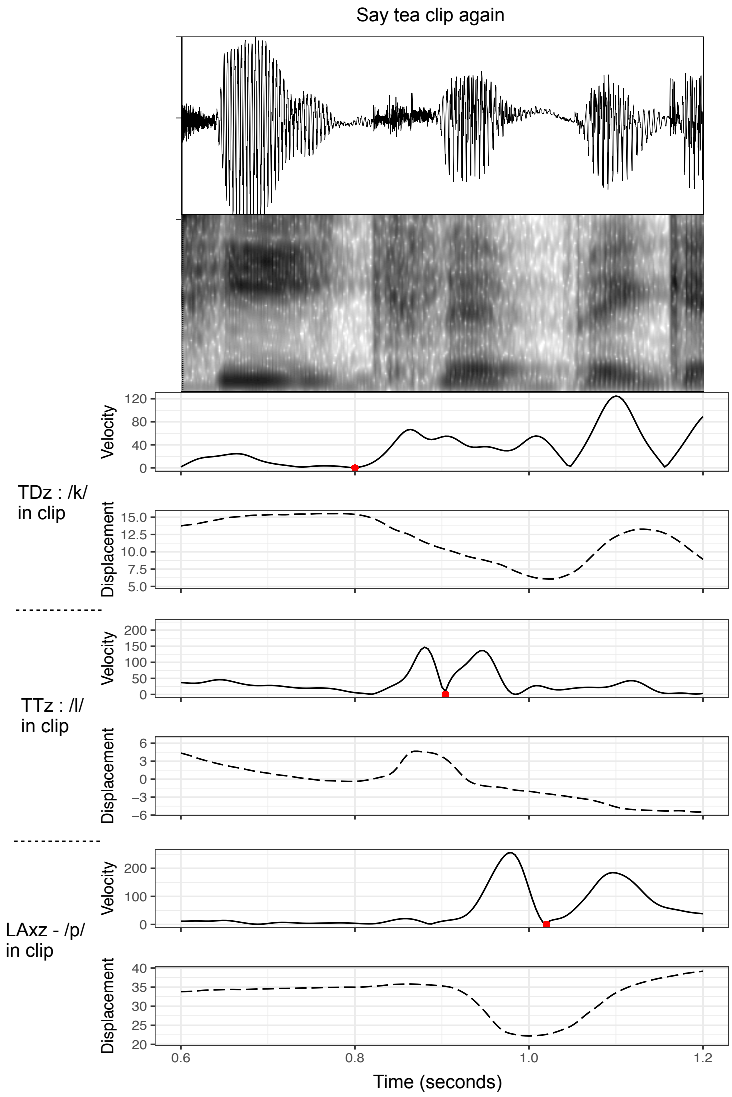

Figure 5.10: Annotated plot showing time aligned acoustic, displacement and velocity trajectories of relevant sensors involved in the consonantal gestures of *clip*. Trajectories are from a single token produced by one speaker of the Onset Lightening dialect: S03. Velocity and displacement trajectories are shown by solid and dashed lines respectively. The top two panels show the velocity and displacement of the tongue dorsum sensor in the vertical dimension, capturing the TD gesture for /k/. The middle two panels show the velocity and displacement of the tongue tip sensor in the vertical dimension, capturing the TT gesture of /l/. The bottom two panels show velocity and displacement of lip aperture in the horizontal-vertical dimension, capturing the lip closure gesture for /p/. Red dots show the velocity minimum, here used to as a proxy for gesture achievement.

and Pouplier (2014) distinguish between the maximum: the velocity minima, the gestural target: 20% of the first velocity peak, and gestural release: 20% of the second velocity peak, and use these landmarks for their stability lag calculations. The present study, on the other hand, looks only at the velocity minima (which I

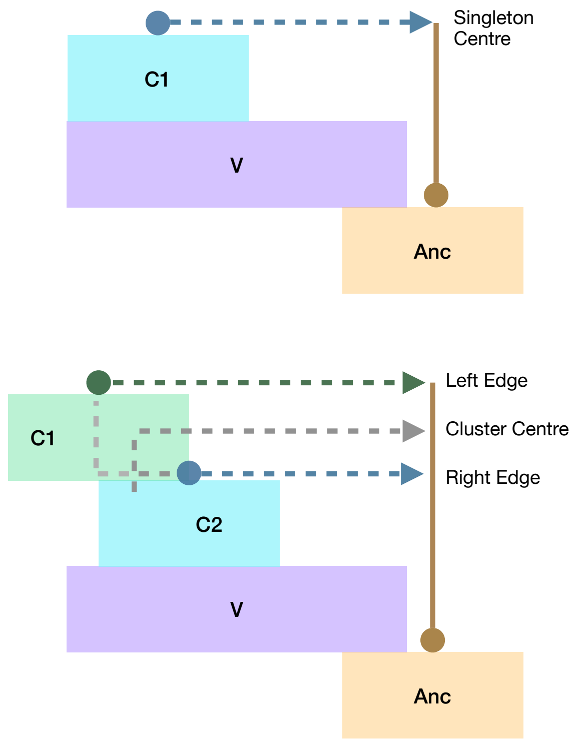

Figure 5.11: Figure showing intervals used within the stability analysis for onsets. The singleton context is shown on the top, and the cluster context is shown on the bottom. Circles indicate the velocity minima of the relevant gesture.

here define as target achievement). This was due to difficulties in establishing 20% of the first and second velocity peaks given the sensitivity of this measure to the size of the velocity peak, which was often variable within the data.

Statistics

A linear mixed effects analysis was conducted using the *lme4* R package (Bates et al., 2015) to model the relationship between lag duration and consonant structure (singleton / cluster) for the Centre, Left Edge and Right Edge measure sets. For each cluster/singleton pair (e.g., *plug / lug*), three models were fitted. The first compared the Centre Set measures, the second compared the Left Edge Set measures, and the third compared the Right Edge Set measures. This structure allowed me to quantify which of the three cluster measures (Centre, Left Edge, or Right Edge measures) were most similar or different to the Singleton Centre measure, and hence, which measure was the most stable across the singleton - cluster pair.

In order to test for the effect of consonant structure and dialect on lag duration,

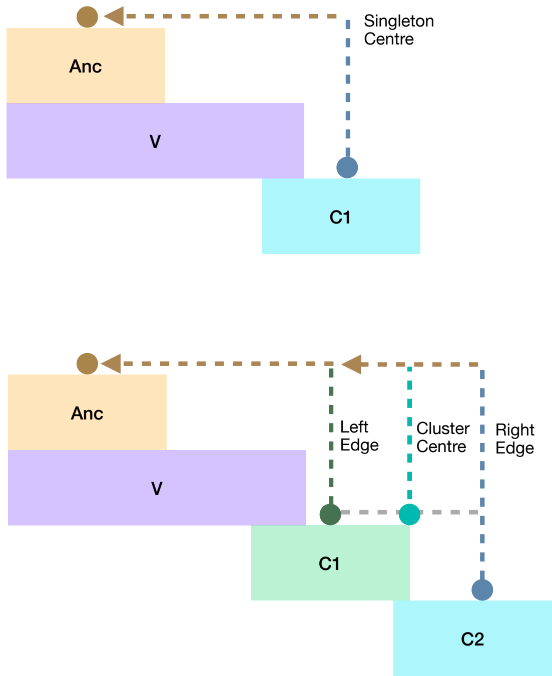

Figure 5.12: Figure showing intervals used within the stability analysis for codas. The singleton context is shown on the top, and the cluster context is shown on the bottom. Circles indicate the velocity minima of the relevant gesture.

full models were compared to nested models, whereby a fixed effect had been excluded using a likelihood ratio test. The structure of the full and nested models are outlined in Table 5.4. The full model, included fixed effects of consonant structure (i.e., singleton versus cluster), dialect, and estimated vocal tract length, a consonant structure by dialect interaction term, a random intercept for speaker, and a random slope for speaker for the effect of consonant structure, see 'Full Model' below. The effect of including a random slope for speaker for the effect of consonant structure means that speakers are allowed to vary in the difference between singletons and clusters, which seems a reasonable assumption to make.

Full Model: *duration: consonant structure + dialect + VT length + dialect * consonant structure + (1 + consonant structure | speaker)*

The model comparison procedure was as follows:

1. First, I tested for an effect of the consonant structure by dialect interaction

Model	Fixed Effects	Interactions
Full model	consonant structure, dialect	timing measure*dialect
Null interaction model	consonant structure , dialect	
Null consonant structure model	dialect	
Null dialect model	consonant structure	

Table 5.4: Table showing components of the full and partial models. Note, also models also included random intercepts for speaker and random slopes of speaker for the effect of consonant structure.

by comparing the full model to a model which was identical to the full model, but with the interaction term removed.

2. If the above was found to be significant, ($p < .05$), no further comparisons were made. If the above was found to be non significant ($p > .05$), steps 3 and 4 were taken.

3. To test for an effect of dialect, the full model was compared to a model which was identical to the full model except for the exclusion of the fixed effect of "dialect" and the "dialect by consonant structure" interaction term. A p-value and model summary for this comparison was then reported.

4. To test for an effect of consonant structure (i.e., singleton or cluster), the full model was compared to a model which was identical to the full model except for the exclusion of the fixed effect of "consonant structure" and the "dialect by consonant structure" interaction term. A p-value and model summary for this comparison was then reported.

5.3 Predictions

5.3.1 Onsets

- **Onset Lightening Dialect:**

  The clear /l/ onsets of Onset Lightening speakers are predicted to show a non

C-centre effect. This follows from the findings of Brunner et al. (2014) and Pouplier (2012) for a non C-centre pattern in the clear /l/ onsets of German speakers.

While the above prediction seemingly conflicts with findings of C-centre stability for /l/ onset clusters in English speakers (e.g., Browman and Goldstein, 1988; Honorof and Browman, 1995; Marin and Pouplier, 2010), the variety of English used for these studies is American English. American English speakers typically produce darker /l/s than that of SSBE speakers used for this study (Turton, 2014). Further, the laterals of Onset Lightening speakers may be considered to occupy an intermediate position between the darker /l/s of American English, and the clearer /l/s of German. As a result, we may also expect an intermediate timing pattern from Onset Lightening speakers. This may look something like a small rightward shift of /l/ within a C1 + /l/ + V structure, which is in the direction of maintaining a C-centre pattern, but is not quite enough to be considered a true C-centre structure.

- **Onset Darkening Dialect:**
  The darker /l/ onsets of Onset Darkening speakers are predicted to show a C-centre effect, as found for American English speakers (Browman and Goldstein, 1988; Marin and Pouplier, 2010).

5.3.2 Codas

- Vowel compression, or a leftward shift of coda /l/ towards the vowel is predicted for both dialects. This prediction is motivated by findings of vowel compression in /l/ coda clusters in American English, where coda /l/ is dark (Marin and Pouplier, 2010). Vowel compression has not been found for languages such as German, where coda /l/ is clear (Pouplier, 2012).

  Greater vowel compression/leftwards lateral shift is predicted for Onset Darkening speakers where /l/ is darker. However, since both Onset Darkening and Onset Lightening dialects have a dark coda /l/, differences are expected to be small.

- In addition, for both onsets and codas, an interaction with vowel and context is predicted due to variation in degrees of coarticulation resistance of surrounding consonants, as determined by the degree of tongue dorsum involvement in the constriction (Brunner et al., 2014). In particular, a difference is predicted to emerge between /p, b/ and /k, g/ clusters, which differ considerably in this regard. The constriction location of the vowel is also expected to interact with the coarticulatory dominance of the first consonant in the cluster, both of which will affect the time it takes for the tongue to reach the constriction location, and hence the timing of the consonant-vowel sequence.

5.4 Onset Results

5.4.1 Lag Measures

Lag measures compare the distance between /l/ and the post-vocalic anchor consonant in singleton and cluster pairs, e.g., *plug* and *lug*. A C-centre pattern would predict a shorter /l/ - to-anchor lag in the cluster context compared to singleton context. This is because a C-centre pattern demands that, when moving from a singleton to a cluster, C2 must shift rightwards towards the vowel, and C1 leftwards away from the vowel, so as to maintain a constant relationship between the consonant centre and the vowel (Marin and Pouplier, 2014).

Figure 5.13 shows lateral to anchor lag durations for onset cluster-singleton pairs. Across the word-pairs, we can see that cluster lags are shorter than singleton lags for both dialects, as would be predicted by a C-centre pattern. The Onset Lightening dialect shows a greater amount of variation in lag duration than the Onset Darkening dialect; however, there are no major dialectal differences. The word pair with the largest difference between the cluster and singleton lags for both dialects is *club / lug*. Perhaps, non coincidentally, this is the only word pair which isn't directly matched segmentally; while the anchor is a bilabial within the cluster context, it is a velar within the singleton context.

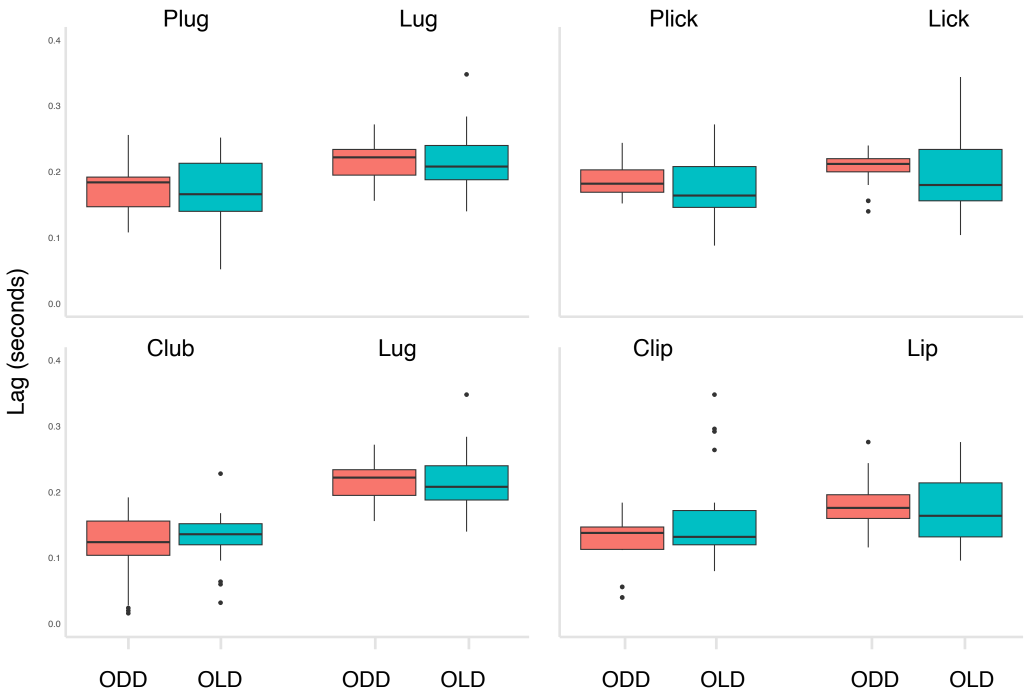

Figure 5.13: Figure showing lateral to anchor lags in seconds for singleton-cluster pairs. A different word pair is shown within each panel, the cluster on the left, and the singleton on the right. Colour indicates dialect.

Measure Set	Singleton Lag	Cluster Lag
Centre Set	singleton centre to anchor	cluster centre to anchor
Left Edge Set	singleton centre to anchor	cluster left edge to anchor
Right Edge Set	singleton centre to anchor	cluster right edge to anchor

Table 5.5: Table showing singleton and cluster lags included within the Centre, Left Edge, and Right Edge measure sets.

Stability Measures

I now focus on the results of the stability analyses. For each singleton/cluster pair, the relevant lag durations included the lateral centre to anchor lag for singleton tokens, such as *lug*, here called the "Singleton Centre", and three lags for cluster tokens: the Left Edge to anchor lag ("Left Edge"), the Right Edge to anchor lag ("Right Edge"), and the Cluster Centre to anchor lag ("Cluster Centre"). These lags are illustrated in Figure 5.11. For each singleton/cluster pair (e.g., *plug / lug*), the Singleton Centre lag was compared to each of the three cluster lags within

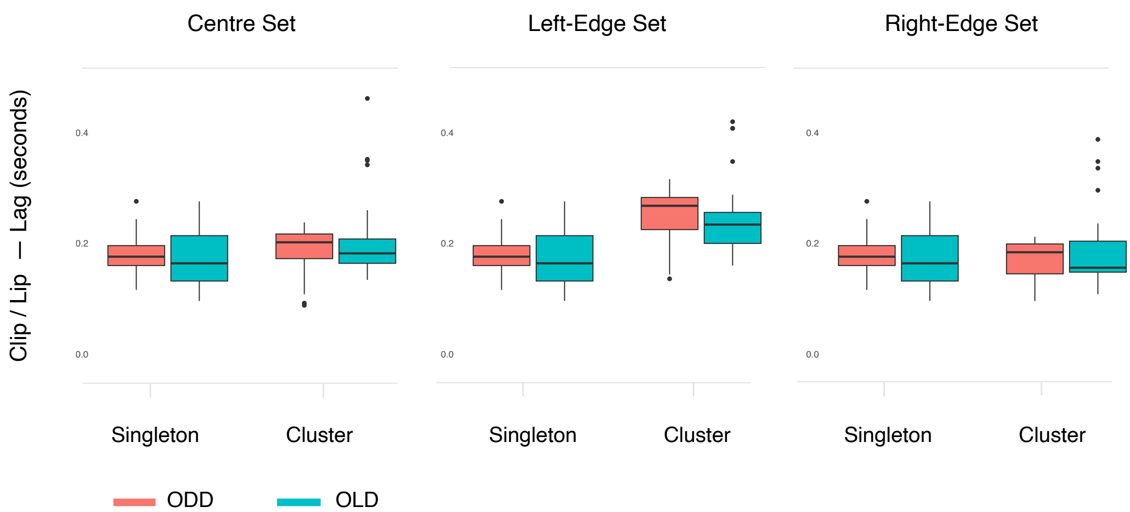

Figure 5.14: Figure showing singleton centre lag of *lip*, alongside three cluster measures of *clip*: Cluster centre to anchor lag (Centre Set), Left edge to anchor lag (Left Edge Set), and Right Edge to anchor lag (Right Edge Set). Dialect is indicated by colour.

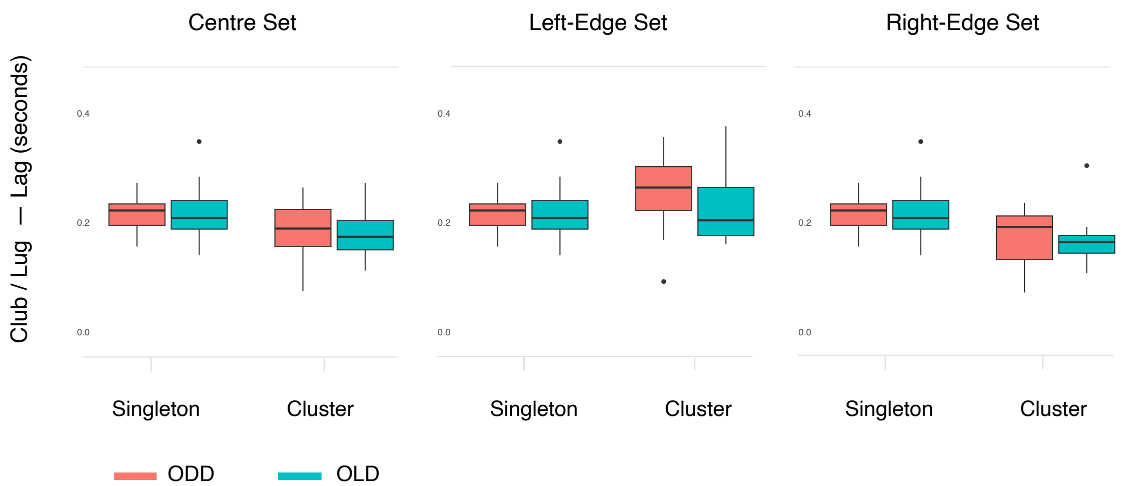

Figure 5.15: Figure showing singleton centre lag of *lug*, alongside three cluster measures of *club*: Cluster centre to anchor lag (Centre Set), Left edge to anchor lag (Left Edge Set), and Right Edge to anchor lag (Right Edge Set). Dialect is indicated by colour.

three separate models. In one model, the Centre Set measures were compared; these included the singleton consonant velocity minima to anchor lag (the Singleton Centre lag), and the consonant cluster midpoint to anchor lag (the Cluster Centre lag). In another model, the Left Edge Set measures were compared; these included the singleton consonant velocity minima to anchor lag (the Singleton Centre lag), and the C1 velocity minima to anchor lag of the cluster (the Left Edge lag). In a further model, the Right Edge Set measures were compared; these included the

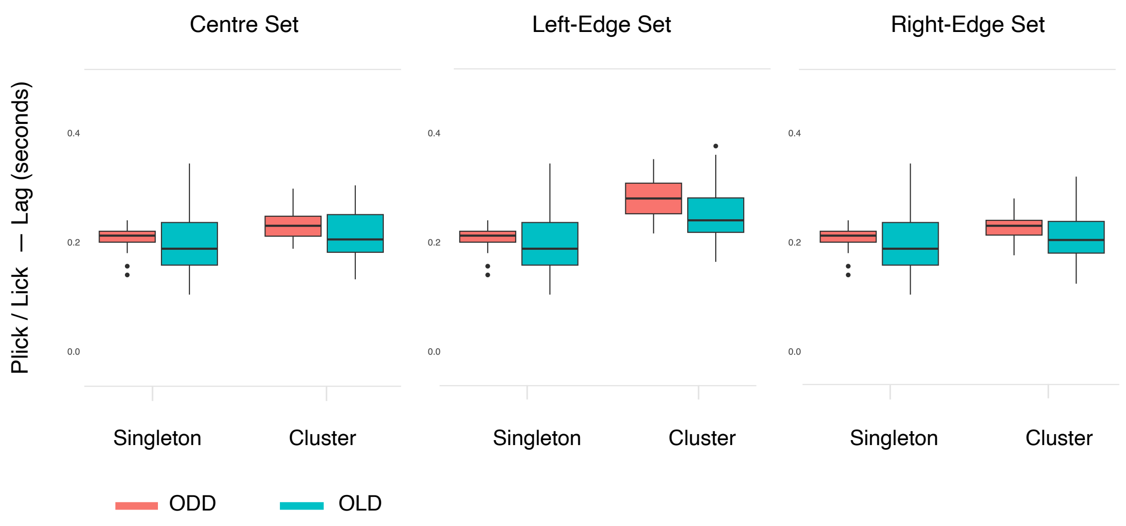

Figure 5.16: Figure showing singleton centre lag of *lick*, alongside three cluster measures of *plick*: Cluster centre to anchor lag (Centre Set), Left edge to anchor lag (Left Edge Set), and Right Edge to anchor lag (Right Edge Set). Dialect is indicated by colour.

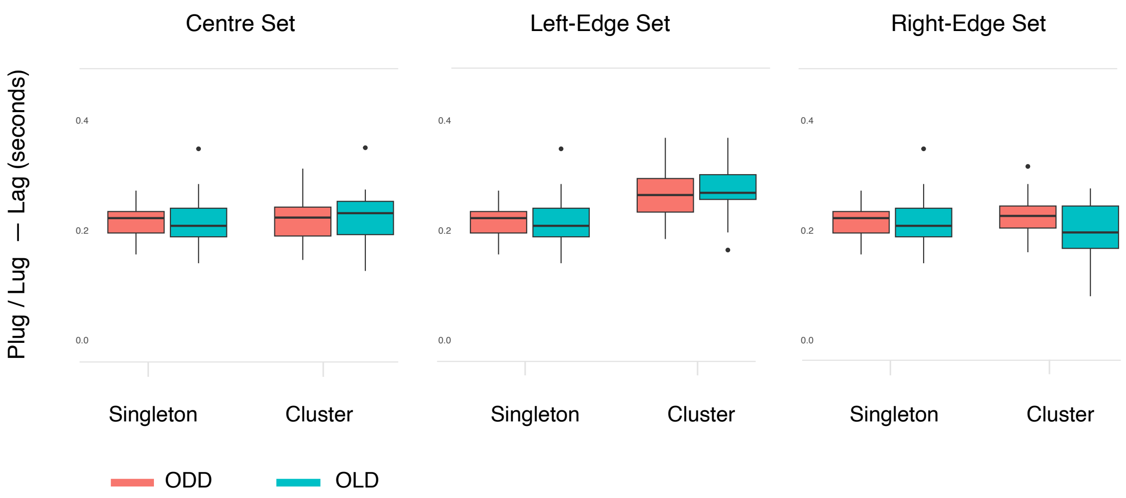

Figure 5.17: Figure showing singleton centre lag of *lug*, alongside three cluster measures for *plug*: Cluster centre to anchor lag (Centre Set), Left edge to anchor lag (Left Edge Set), and Right Edge to anchor lag (Right Edge Set). Dialect is indicated by colour.

singleton consonant velocity minima to anchor lag (the Singleton Centre lag), and the C2 velocity minima to anchor lag of the cluster (the Right Edge lag). The measure sets and corresponding lags are summarised in Table 5.5.

Box pots shown in Figures 5.14 to 5.17 show, for each word pair, the Singleton Centre lag, labelled "Singleton" on three places of the x-axis next to one of three consonant cluster measures. Within the far left panel, titled "Centre Set", the cluster measure shown alongside the singleton measure is the Cluster Centre lag. In the middle panel, titled "Left Edge Set", the cluster measure shown alongside the Singleton Centre lag is the cluster Left Edge lag. Finally, the far right panel, titled, "Right Edge Set", shows the Singleton Centre lag along side the cluster Right Edge lag. Dialect is shown by colour.

In order to quantify the degree of stability between each of three cluster measures (Centre, Left Edge, or Right Edge measures) and the Singleton Centre measure, model comparisons were conducted according to the structure outlined in Section 5.2.2. For each word pair and measure set, the following effects on lag duration were tested for: a fixed effect of consonant structure (singleton vs cluster), a fixed effect of dialect and a consonant structure by dialect interaction term. Full models were compared to partial model where the relevant effect had been removed.

5.4.2 Stability Analysis Results

Stimuli Pair	Measure Set	Effect Tested	Chi Sq	*df*	p-value
Plug - Lug	C-Centre	Interaction	0.051	1	0.822
		Consonant Structure	0.653	2	0.722
		Dialect	0.053	2	0.974
	L-Edge	Interaction	0.512	1	0.474
		Consonant Structure	16.159	2	<**0.001**
		Dialect	0.553	2	0.758
	R-Edge	Interaction	1.291	1	0.256
		Consonant Structure	1.518	2	0.468
		Dialect	2.061	2	0.357
Plick - Lick	C-Centre	Interaction	0.432	1	0.511
		Consonant Structure	9.296	2	**0.009**
		Dialect	2.131	2	0.345
	L-Edge	Interaction	0.981	1	0.322
		Consonant Structure	23.662	2	<**0.001**
		Dialect	2.002	2	0.368
	R-Edge	Interaction	0.130	1	0.718
		Consonant Structure	7.314	2	**0.026**
		Dialect	3.100	2	0.212

Table 5.6: Anova results for /pl/ onset cluster context. Significant p-values ($p < .05$) are shown in bold.

Stimuli Pair	Measure Set	Effect Tested	Chi Sq	*df*	p-value
Club - Lug	C-Centre	Interaction	0.007	1	0.935
		Consonant Structure	2.890	2	0.236
		Dialect	0.012	2	0.994
	L-Edge	Interaction	0.177	1	0.674
		Consonant Structure	1.719	2	0.423
		Dialect	0.335	2	0.846
	R-Edge	Interaction	0.025	1	0.873
		Consonant Structure	8.271	2	**0.016**
		Dialect	0.029	2	0.986
Clip - Lip	C-Centre	Interaction	1.517	1	0.218
		Consonant Structure	3.684	2	0.159
		Dialect	1.581	2	0.454
	L-Edge	Interaction	0.759	1	0.384
		Consonant Structure	11.882	2	**0.003**
		Dialect	0.915	2	0.633
	R-Edge	Interaction	1.289	1	0.256
		Consonant Structure	1.557	2	0.459
		Dialect	1.373	2	0.503

Table 5.7: Anova results for /kl/ onset cluster context. Significant p-values ($p < .05$) are shown in bold

Model comparison results are shown in Tables 5.6 and 5.7 for /pl/ and /kl/ onset clusters respectively. An effect is considered significant if the model comparison which tests for the effect returns a p-value of $< .05$. Before looking at the models, it may be beneficial to briefly revisit the predictions of a C-centre pattern in order to better contextualise the results. A C-centre timing pattern predicts least variability between the singleton and cluster pair in the centre measure set compared to either of the Left Edge or Right Edge set. This follows from the idea that for the C-centre effect to prevail, the temporal relationship between the consonant centre and a fixed anchor point should remain the same, despite changes to the complexity of the onset. A substantial difference in the lag durations between the singleton and cluster pair within the Centre set would be diagnostic of a non C-centre timing pattern. From this hypothesis, a non significant p-value would be expected from the model which tested for the effect of consonant structure in the Centre set. This would imply that there is a non significant difference between the singleton centre to anchor lag and the cluster centre to anchor lag, indicative of a C-centre pattern. The Left Edge and

Right Edge sets can be understood in the same way; a significant effect from the model which tests for an effect of consonant structure within either the Left Edge or Right Edge sets would mean that there were significant differences between the singleton centre to anchor lag and the cluster C2 or C1 to anchor lag respectively.

The only comparison to yield significant results were the models which tested for the effect of consonant structure. Therefore, neither dialect nor the interaction between dialect and consonant structure had a significant effect across the model comparisons. The models where the effect of consonant structure was significant varied for each cluster-singleton pair. For *plug / lug*, consonant structure is significant for the Left Edge set only ($p < .001$), meaning that across the *plug / lug* pair, the C-centre and Right Edge intervals are stable. Likewise, the effect of consonant structure is also significant for the Left Edge set only in the *clip / lip* word pair ($p = .003$). For *plick / lick*, the effect of consonant structure is significant for all sets, meaning that there is a significant difference between the singleton Centre to anchor lag in *lick* and the Left Edge, Right Edge, and Centre to anchor lags in *plick*. For the *club / lug* pair, the effect of consonant structure is significant for the Right-Edge set only, meaning that the C-centre and Left Edge intervals are stable across the pair.

5.4.3 Onset Results Summary

Two sets of measures were taken across singleton and cluster pairs: the duration of the lateral to anchor interval, and a series of stability measures. These were used to determine the most stable interval across the singleton-cluster pair. In addition, stability measures were used to test for the effect of dialect and constant structure on lag duration. Collating the results from the two measures, the following can be observed of the onset singleton-cluster pairs studied here. Contrary to predictions of the coupled oscillator model, across the singleton-cluster pairs, the C-centre interval was not necessarily the most stable interval. Rather the most stable interval across singleton and cluster word pairs varied across word pairs.

A more surprising finding was the lack of a qualitative difference between dialects in onset cluster timing. This finding was unexpected and runs contrary to hypothesis that the difference in lateral darkness between Onset Lightening and Onset Darkening dialects would result in different patterns of lateral onset cluster timing patterns.

5.5 Coda Results

One challenge presented by the coda analysis was the frequent /l/ vocalisation of speakers of the Onset Lightening Dialect in pre consonantal coda contexts (i.e., /l/ clusters *milk, philp, gulp*); interestingly, /l/ in word final singleton contexts were not vocalised (*mill, fill, gull*). Since the timing of the tongue tip raising gesture was used to define the time of lateral target achievement, a reduced or absent tongue tip gesture, characteristic of /l/ vocalisation, meant that the timing of lateral coda clusters could not be measured for Onset Lightening speakers.

As an illustrative example, Figure 5.18 shows z-scored tongue tip raising trajectories for /l/ in *gulp* for each speaker. Trajectories span /l/ plus the preceding and following segment. Onset Lightening speakers are shown in blue, and Onset Darkening speakers are in red. The figure shows 5 out of 8 Onset Lightening speakers to have vocalised realisations of /l/ (S01, S03, S04, S05, S08). For these speakers, the tongue tip vertical displacement trajectory is characterised by either an initial lowering or absence of raising. In contrast S02, S07, S08, and all included Onset Darkening speakers show initial tongue tip raising of varying degrees. One token of L06 shows partial vocalisation with minimal tongue tip raising, and one token from L04 shows full vocalisation - these tokens were removed from the analysis.

Table 5.8 quantifies the instances of vocalisation within the coda /l/ cluster contexts, *gulp, milk*, and *philp* for each dialect. Across the coda cluster contexts, Onset Lightening speakers vocalised 57.83% of tokens, while Onset Darkening speakers, vocalised only 9.52% of tokens. While Onset Lightening speakers vocalised in all coda cluster contexts, there are less vocalised tokens within the *philp* context relative to the *gulp / milk* contexts. Figures of the tongue tip raising trajectories for individual

Dialect	Context	No of vocalised tokens
Onset Lightening	Gulp	18 of 30
Onset Darkening	Gulp	2 of 21
Onset Lightening	Milk	23 of 32
Onset Darkening	Milk	4 of 23
Onset Lightening	Philp	7 of 21
Onset Darkening	Philp	0 of 19
	Total vocalised	OLD: 57.83%; ODD: 9.52%

Table 5.8: Table showing the number of vocalised tokens for each dialect and coda cluster context. No of vocalised tokens shows the number of vocalised tokens out of the total number of times the word was produced by speakers of the dialect group.

speakers for both the coda cluster *gulp* / *milk* / *philp* and singleton *gull* / *mill* / *fill* contexts are included in Appendix 9.4. The few instances of /l/ vocalisation from Onset Darkening speakers were excluded from the analysis. In addition, speaker L07 was removed from the *fill* / *philp* analysis due to a missing lower teeth sensor required to identify the gestural target of [f]. A total of 132 tokens were included in the Onset Darkening dialect coda analysis.

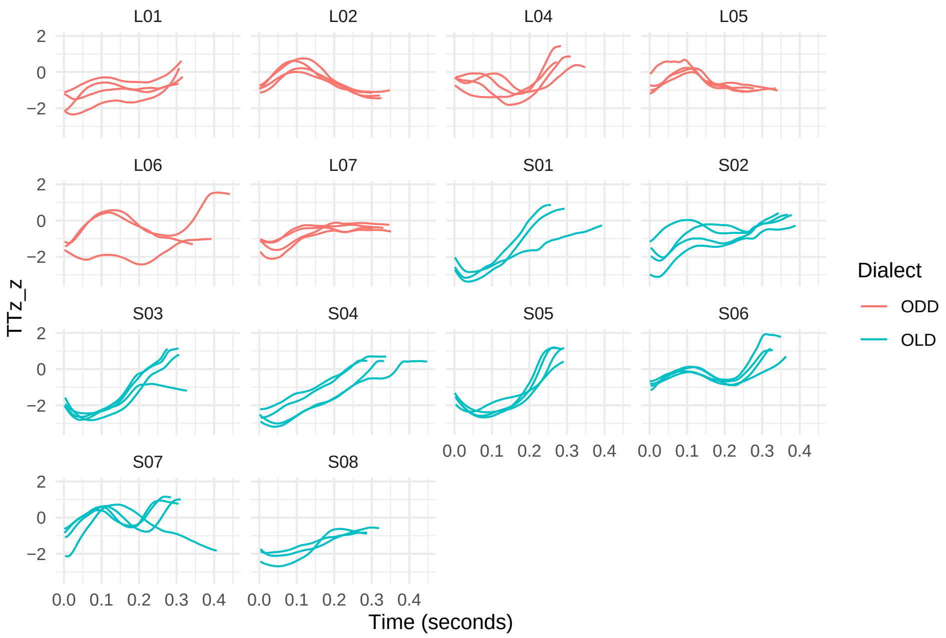

Figure 5.18: Z-scored vertical tongue tip trajectories for /l/ (plus the preceding and following segment) in the coda context of *gulp*.

In light of the large number of vocalised tokens in Onset Lightening speakers, I here present coda timing results exclusively from speakers of the Onset Darkening dialect who only vocalised lateral segments on a few occasions. While a dialect comparison of cluster timing measures is not possible for coda /l/ clusters, looking at coda cluster timing patterns from a single dialect is still useful in that it enables a direct comparison to be made with the findings of cluster timing patterns of dark /l/s in Marin and Pouplier (2010), which have been central to this thesis. Specifically, Marin and Pouplier (2010) found the dark coda /l/s of American English speakers showed a left-ward shift towards the vowel in clusters relative to singletons, contrary to the predictions of a coupled oscillator model. It is therefore predicted that the dark coda /l/s of Onset Darkening speakers will also show a similar left-ward shift in clusters. Lateral to anchor lags and stability measures of singleton-cluster pairs are here presented for coda tokens, *philp / fill*, *milk / mill*, *gulp / gull*.

Lag Measures

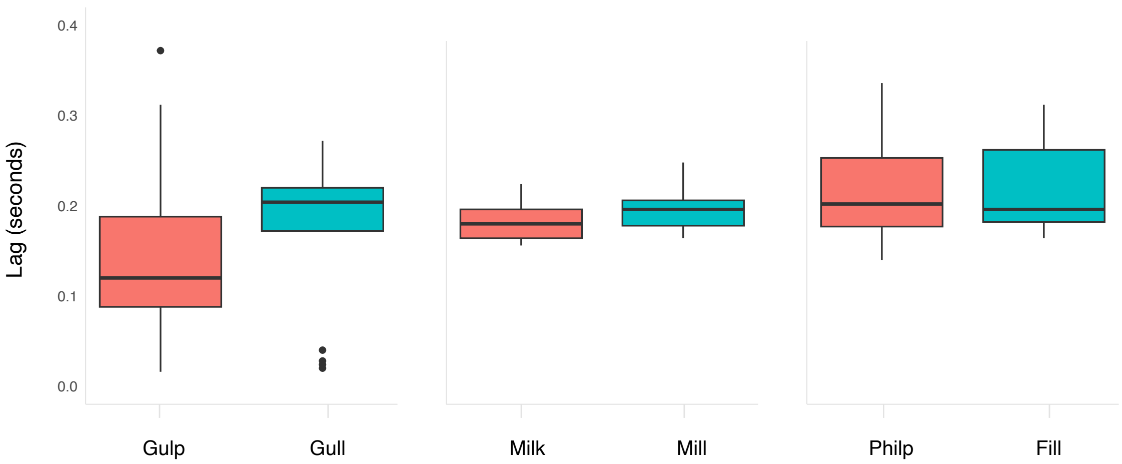

Figure 5.19: Figure showing lateral to anchor lags (seconds) for singleton-cluster coda pairs in the Onset Darkening Dialect. For each word pair, clusters are shown on the left in red, and singletons on the right in blue.

Figure 5.19 shows the lateral-to-anchor lags for each cluster-singleton word pair (from left to right: *gulp / gull ; milk / mill ; philp / fill*) for Onset Darkening speakers. To recap, a sequential timing pattern would predict a similar lateral-to-anchor lag between the singleton and cluster tokens of each word pair. This pattern would imply that the transition from a singleton coda to a cluster coda entails the mere sequential addition of a further coda consonant, and a stable vowel lateral relationship. On the other hand, the leftward shift of laterals found in the dark lateral coda clusters of American English speakers (Marin and Pouplier, 2010) would predict a shorter lateral-to-anchor lag in a cluster context, relative to a coda context, suggesting relatively greater lateral reduction or vowel overlap within cluster context.

Lateral-to anchor lags in *philp / fill* word pair exhibits a lag consistent with a sequential timing patterns, both showing similar lateral-to-anchor lags between singleton and cluster tokens. Lateral-to-anchor lags for *gulp / gull* rather shows a greater lag within the singleton context relative to the cluster context. This is also the case for the *milk / mill* word pair, though differences are small. A larger singleton to anchor lag is consistent with Marin and Pouplier (2010)'s findings for a leftward shift pattern for lateral coda clusters. However, a note should be made on the differences in variability in lateral to anchor lags across word pairs. The *milk / mill* word pair in particular is considerably less variable than *gulp / gull* and *philp / fill* word pairs. This makes it difficult to make conclusions from visual comparison alone. Statistical modelling of stability measures in the subsequent sections will offer a more robust quantification of lag differences across word pairs.

Stability Measures

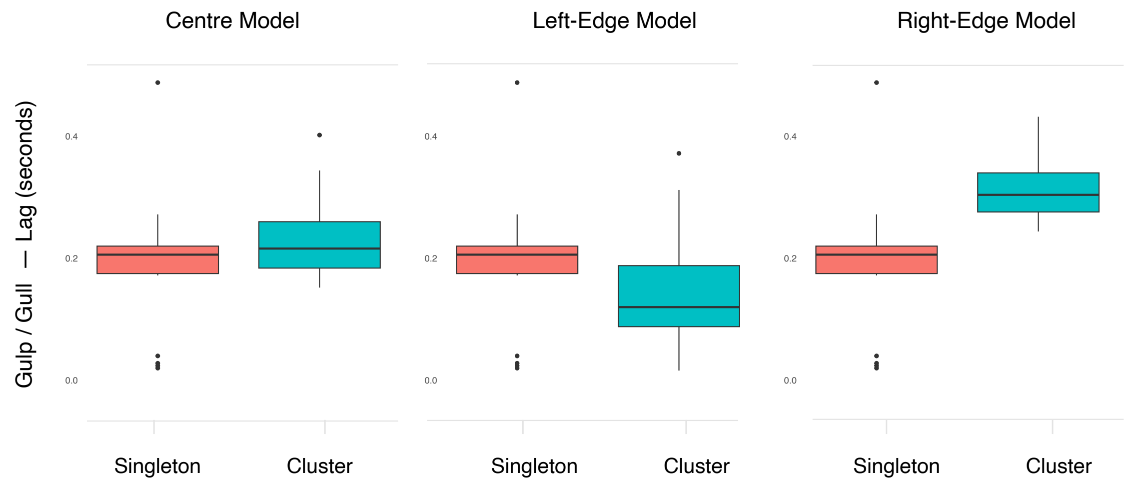

Figure 5.20: Figure showing singleton centre lag of *gull* in red, alongside three cluster measures of *gulp* in blue: the anchor to cluster centre lag (Centre Set), the anchor to left edge lag (Left Edge Set), and the anchor to right edge lag (Right Edge Set).

The stability measures used for the coda analysis mirrored those used within the onset analysis, see Figure 5.12. One measurement was taken for singleton tokens within each singleton-cluster pair, which was the lag between the velocity minima of the anchor consonant to the velocity minima of the singleton coda consonant,

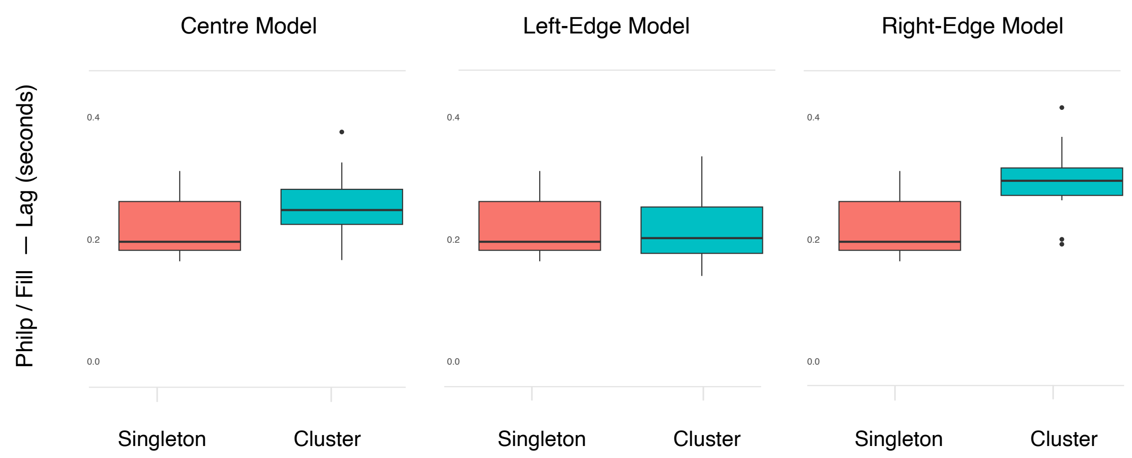

Figure 5.21: Figure showing singleton centre lag of *fill* in red, alongside three cluster measures of *philp* in blue: the anchor to cluster centre lag (Centre Set), the anchor to left edge lag (Left Edge Set), and the anchor to right edge lag (Right Edge Set).

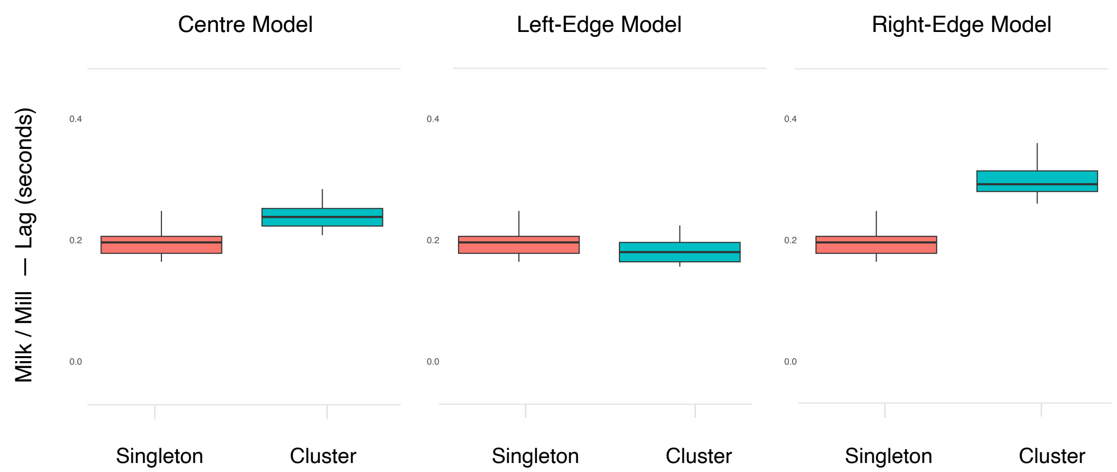

Figure 5.22: Figure showing singleton centre lag of *mill* in red, alongside three cluster measures of *milk* in blue: the anchor to cluster centre lag (Centre Set), the anchor to left edge lag (Left Edge Set), and the anchor to right edge lag (Right Edge Set).

here referred to as the "Singleton Centre". For each word pair, this measure was compared to three cluster measures taken from the cluster token. The first measured the time between the anchor velocity minima and the velocity minima of C1 in the coda cluster, here referred to as the "Left Edge" measure. The second captured the time between the anchor velocity minima and the velocity minima of C2 in the coda cluster, here referred to as the "Right Edge" measure. Finally, the "Cluster Centre" measure captured the time between the anchor velocity minima and the mean time point between the velocity minima of C1 and C2.

Stability intervals are shown for the *gulp / gull* pair in Figure, 5.20, for *philp / fill*, in Figure 5.21 and for *milk / mill*, in Figure 5.22. For each word pair, Singleton Centre lag durations are shown in red alongside three cluster lags in blue: the Cluster Centre (left), the Left Edge cluster measure (middle), and the Right Edge cluster measure (right). For a sequential timing pattern to exist, the Singleton Centre lag would be most similar to Cluster Left Edge lag, hence showing Left-Edge stability. From visual inspection a sequential timing pattern appears to prevail for word pairs *philp / fill*, (Figure 5.21), and *milk / mill* (Figure 5.22). That is, for these word pairs, the singleton centre shown in red, is most similar to the cluster lag in the Left Edge set than it is in the Centre or Right Edge sets. For the *gulp / gull* pair, (Figure 5.20), on the other hand, the Singleton Centre lag is most similar to the Cluster Centre lag, suggesting a rightward shift of /l/ away from the vowel in the cluster context relative to the singleton context.

To quantify the variability between the singleton centre to anchor lag and each of the three cluster lags model comparisons were performed for the Centre, Left Edge, and Right Edge sets of each word pair. Model comparisons followed the same structure as those in the onset analysis. Unlike the onset analysis, only the effect of consonant structure was tested for given that all speakers in the coda analysis belonged to the Onset Darkening dialect. A full model was compared to a partial model where the effect of consonant structure had been removed. The full model included fixed effects of consonant structure and estimated vocal tract length and a random intercept for speaker. Results were considered significant when p was $<$ .05. Results of the model comparisons are shown in Table 5.9. The prediction of the C-centre model is that the Singleton Centre lag and the Cluster Left Edge lag (Left Edge model) should not be significantly different, while the Centre and Right Edge sets should be significantly different. Results of the model comparison show that for the *philp / fill* and *gulp / gull* word pairs, the Left Edge sets are not significantly different, while the Centre and Right Edge sets are. For the *milk / mill* word pair, all measure sets are shown to be significantly different. This suggests that *philp / fill* and *gulp / gull* word pairs show the predicted Left Edge stability indicative of a sequential timing pattern, while the *milk / mill* word pair is not stable along any of the three measures. These results are somewhat surprising given that visually, there is a larger difference between singleton and cluster lags in the Left Edge set for the *gulp / gull* word pair than there is for the *milk / mill* word pair. This may, however, be a result of relatively smaller amount of variation in lag measures within the *milk / mill* pair.

Stimuli Pair	Measure Set	Chi Sq	*df*	p-value
Milk - mill	C-Centre	38.641	1	<**0.001**
	Left Edge	7.121	1	**0.0076**
	Right Edge	76.305	1	<**0.001**
Philp - fill	C-Centre	6.335	1	**0.0118**
	Left Edge	0.524	1	0.469
	Right Edge	22.055	1	<**0.001**
Gulp - gull	C-Centre	4.135	1	**0.0402**
	Left Edge	3.333	1	0.0679
	Right Edge	27.034	1	<**0.001**

Table 5.9: Anova results for coda cluster contexts. Model comparison results are shown for the C-centre, Left Edge and Right Edge sets of each word pair. Significant results ($p < .05$) are shown in bold.

5.5.1 Coda Results Summary

The coda analysis, unlike the onset analysis, looked at Onset Darkening speakers only. This was due to the frequent vocalisation of coda /l/s in the Onset Lightening Dialect, which meant that the tongue tip gesture, here used to define lateral timing, was not measurable. Two groups of measures were considered; lateral to anchor lags and stability measures, which were used to determine the most stable interval across the singleton and cluster pairs.

Results were mixed across measures and word pairs. First, lateral to anchor lags were visually presented across singleton and cluster contexts of each word pair. The word pair *philp / fill* appeared to showed stable lateral-to-anchor lags across singleton and cluster contexts, suggestive of Left-Edge stability, i.e., a sequential timing pattern. The *gulp / gull* and *milk / mill* word pairs showed shorter lateral to anchor lags in the cluster context, suggesting a leftward shift of /l/ in the cluster context compared to the singleton context. This pattern was inline with findings of Marin and Pouplier (2010) for the dark /l/s of American English speakers. The stability analysis then compared the Singleton Centre lag to Cluster Centre, Left Edge and Right Edge lags in order to determine the most stable interval across each word pair. Statistical comparisons showed no significant difference between the Single Centre lag and the Left Edge cluster lag for *philp / fill* and *gulp / gull* word pairs, suggesting a pattern of Left Edge stability for these word pairs. The *milk / mill* word pair showed a different pattern; for this word pair, the Singleton Centre lag was not significantly different from any of the cluster lags, hence suggesting that neither the C-centre, Left Edge nor Right Edge intervals are stable across the *milk / mill* word pair. These results did not straightforwardly map onto the visual differences seen between cluster and singleton lags for each word pair. I suggested this to be a result of differences in variability across word pairs which were better accounted for within the statistical models.

5.6 Discussion

This analysis set out to answer the following question: How do spatial differences in /l/ darkening between dialects found in Chapter 4, affect patterns of /l/ cluster timing? It was predicted that the differences in /l/ darkening between dialects, found to manifest in greater tongue body lowering for Onset Darkening speakers, would result in subsequent differences in /l/ cluster timing. For onset /l/ clusters, a C-centre timing pattern was predicted for Onset Darkening speakers, while a non C-centre timing pattern was predicted for Onset Lightening speakers, in line with findings for non C-centre timing patterns for clear onset /l/ in German (Pouplier, 2012). For /l/ coda clusters, both dialects were predicted to show a left-ward shift towards the vowel, following findings for this pattern in dark coda /l/s in American English Marin and Pouplier (2010). Results found non C-centre timing patterns for lateral onset cluster across both Onset Lightening and Onset Darkening dialects. Contrary to expectation, no significant differences in onset cluster timing were found between dialects. Coda cluster results were presented for Onset Darkening speakers only, given that /l/ was often vocalised in lateral coda clusters for Onset Lightening speakers. Results of the /l/ coda cluster analysis for Onset Darkening speakers revealed a pattern of Left Edge stability for two word pairs *gulp / gull* and *fill /*

philp, and no stable timing patterns for *milk / mill*. While these results did not reveal the predicted timing patterns found in Marin and Pouplier (2010), findings for Left Edge stability are consistent with the local timing pattern in coda clusters predicted by the coupled oscillator model. Results of the onset and coda analyses are below discussed in turn.

A non C-centre timing pattern was found for onset lateral clusters in both dialects. Lateral-to-anchor lags showed temporal reorganisation from singletons to clusters in the direction of maintaining a stable Consonant Centre, i.e., reduced lateral-to-anchors lags in clusters relative to singletons. However, the subsequent stability analysis showed that such reorganisation was not enough to result in a clear C-centre effect. This however, was not entirely surprising given previous observations of non C-centre patterns within the literature as discussed at the start of this chapter (e.g, Goldstein et al., 2009; Mücke, Hermes, and Tilsen, 2020). Further, that the onset results here showed evidence of reorganisation in the direction of maintaining a stable consonant centre certainly positions them within the realm of findings from other cluster timing studies.

How can the findings for a non C-centre timing pattern in lateral onsets be explained within an articulatory framework of speech timing? In agreement with Mücke, Hermes, and Tilsen (2020) and Goldstein et al. (2009), I suggest that these finding are not incompatible with a coupled oscillator model of speech timing. For the *plug / lug* and *clip / lip* word pairs, non-significant differences were observed between the Singleton Centre lag and the both the Cluster Right Edge and Cluster Centre lags. This suggests that laterals within these onset clusters, relative to their singleton counterpart, exhibit somewhere in between a rightward shift towards the vowel, and no shift at all. If we accept that asymmetrical couplings between onset constants and the vowel exist and hypothesise that the multi-gestural composition of /l/ results in multiple couplings with the vowel (Goldstein et al., 2009), then we would expect a greater overall coupling between the lateral and the vowel. The reduced movement of the lateral towards the vowel thus makes sense; the stronger coupling between the lateral and the vowel somewhat reduces the flexibility of /l/s movement (Goldstein et al., 2009). One way to concretely test such predictions would be through the application of computational modelling (e.g., Turk, Elie, and Šimko, 2023, Goldstein et al., 2009). While beyond the scope of the present study, this is a promising avenue which I hope to explore within future work. However, this does not account for the different stability patterns observed for the *plick / lick* and *club / lug* pairs. The *club / lug* pair was the only word pair to show stability of the C-centre and Left-Edge intervals. One plausible reason for this could be that the anchor consonant differs across the singleton and cluster token of this pair, and that the timing measures are confounded by this difference (e.g., Iskarous and Pouplier, 2022). For the *plick / lick* word pair, no stable intervals were observed. While this has no immediately obvious explanation, it is at least possible that the relative unfamiliarity of the word *plick* interacts with the timing patterns here.

The most surprising finding to come from this study was the absence of a significant difference in lateral onset cluster timing patterns between dialects. Such a difference between dialects was anticipated to follow from differences in lateral darkening between the dialects, as was illustrated within the previous chapter. Lateral darkening was predicted to have an effect on the timing of lateral onset clusters given previous findings for cross-linguistic differences in lateral coda cluster timing for liquids which differed in darkness (Marin and Pouplier, 2014). That no timing difference was found between dialects which differed in onset lateral darkness may be considered analogous to the idea of motor equivalence; here rather, it is not the stability of the acoustic signal which is being preserved in the face of differences in articulation, but rather the timing pattern. Before I engage with possible ways in which a stable timing pattern may be maintained despite articulatory differences /l/ darkening, we may reasonably consider the possibility that laterals behave differently when occurring in clusters. While the previous chapter showed an acoustic and articulatory difference in lateral darkening between dialects for laterals in high vowel, non word-final contexts, the stimuli of the previous chapter did not include lateral clusters.

Below, I demonstrate that there is indeed an articulatory difference between dialects in laterals in the onset cluster contexts. Figure 5.23 compares the tongue body vertical displacement trajectories for the C + /l/ + V temporal window of onset cluster tokens. The decision to look at tongue body displacement was motivated by the previous chapter's finding that the vertical tongue body displacement is the most successful in capturing lateral darkening. In Figure 5.23, tongue body displacement trajectories are presented for each cluster token, and each dialect. The TBz values of each token, were centred by subtracting the mean TBz position from each value raw TBz position value. Centring the data allowed me to resolve the issue of comparability presented by the considerable variation of raw TBz values between speakers, while preserving the magnitude of displacement. From Figure 5.23, differences can be observed between dialects in terms of the overall magnitude of tongue body lowering and the utilisation of vertical space, with Onset Darkening speakers often beginning with a higher TBz value than Onset Lightening speakers, and lowering to a lower TBz value than Onset Lightening speakers. To visualise the difference in TBz displacement magnitude more clearly, Figure 5.24 shows measure of TBz displacement magnitude. Magnitude was here calculated as the maximum centred TBz position minus the minimum centred TBz position. Results are here shown by dialect for *plug*, *plick*, *club*, and *clip* tokens. For both dialects, there is a higher magnitude of tongue body lowering for the back vowel contexts *plug* - *club* relative to front vowel contexts *plick* - *clip*. Across all contexts, Onset Darkening speakers show a higher magnitude of tongue body displacement than Onset Lightening speakers.

Linear model comparisons were performed to further quantify the effect of dialect on the measure of TBz magnitude presented here. A full model, containing fixed effects of word, dialect, and estimated vocal tract length, and a random intercept of speaker, was compared to a partial model, which was identical except for the exclusion of the fixed effect of dialect. Random sloped were not included due to convergence issues. The model comparison is considered significant if $p = <.05$. Table 5.10 presents a summary of the model comparison outcome. The effect of dialect on the measure of TBz magnitude to be significant ($p = 0.0121$), suggesting there to be a significant dialectal difference in the magnitude of tongue body lowering for /l/ in onset clusters.

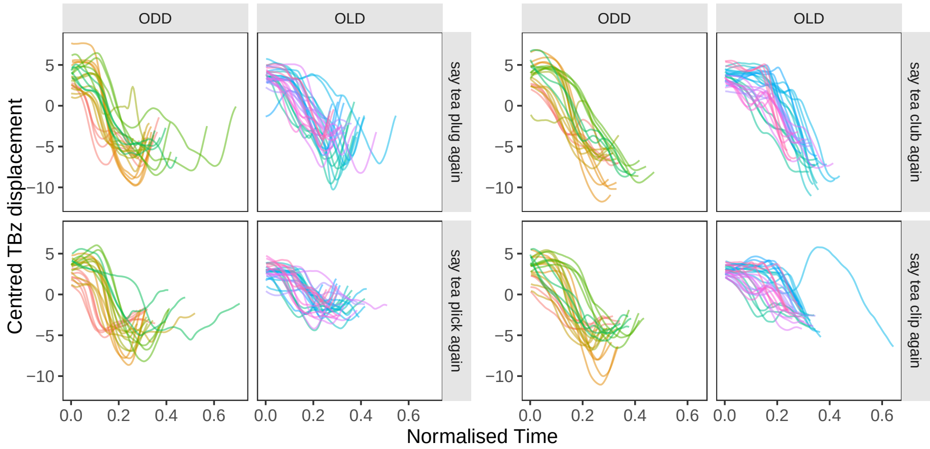

Figure 5.23: Figure showing centred vertical tongue body displacement trajectories over the C+/l/+V temporal window of onset cluster tokens: *plug*, *plick*, *club*, *clip*.

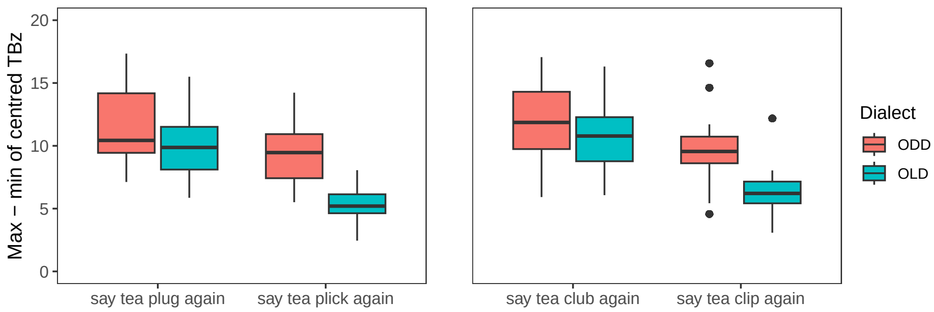

Figure 5.24: Figure showing maximum minus minimum centred TBz values over the C+/l/+V temporal window of onset cluster tokens: *plug*, *plick*, *club*, *clip*.

Effect Tested	χ^2	df	p-value
dialect	6.3	1	**0.0121**

Table 5.10: Model comparisons testing for effect of dialect on TBz magnitude of /l/ within onset clusters.

The above has shown a dialectal difference in lateral darkening for laterals in onset clusters. We are now presented with a very interesting question: how can a stable timing pattern of lateral onset cluster timing be preserved across dialects which differ in lateral darkening? This presents a somewhat of a paradox given the well reported effects of lateral darkening on timing (e.g., Sproat and Fujimura, 1993). One possibility which could explain both a difference in lateral darkening and a stable onset cluster timing pattern between the dialects, is that dialects differ in the dynamics of vertical tongue body movement, which may affect onset cluster timing patterns. However, differences in timing may compensated for by changes to the duration of the vowel, such that an overall stable pattern of timing is maintained. While this is a reasonable hypothesis, it is not one that will here be tested given that it is difficult to segment the tongue body gesture of the lateral from that of the adjacent vowel gesture with the precision required for this analysis. A second possibility is that dialects may again differ in the dynamics of the vertical tongue body movement, which may also affect onset cluster timing patterns. However, such differences in timing may be compensated for by differences in the velocity profiles of the vertical tongue body movement, thus enabling a stable timing pattern to be maintained. It is plausible that Onset Darkening speakers could produce a tongue body lowering gesture which is larger in magnitude, as shown in Figure 5.23, but has a higher velocity. This would allow the target to be achieved within the same amount of time as a smaller but slower tongue body gesture. This hypothesis will be further investigated within the following chapter.

To take an alternative perspective, that a difference in lateral cluster timing was not found between dialects is perhaps not all that surprising. The systematic spatial variability in laterals observed between the dialects here occurred in the vertical tongue body displacement gesture. The articulatory landmark used to measure lateral timing, however, was the velocity minimum of the tongue tip raising gesture. It is quite possible then for differences in lateral cluster timing between dialects to exist that the tongue tip measure is not sensitive to. In such a case, the tongue body gesture may be variably timed to the tongue tip gesture, while the relationship between the tongue tip gesture and the anchor consonant may be stable between dialects. However, if this is the case, how then do we explain the variable patterns in lateral cluster timing between clear and dark laterals across languages (Marin and Pouplier, 2010), where lateral timing was defined in terms of the tongue tip gesture? This remains an open question. The final and subsequent study of this thesis will aim to provide some answers to this question, while working within the practical constraints of what is measurable.

Turning to the coda analyses, coda cluster timing results were presented from the Onset Darkening Dialect only. This was due to the large proportion of vocalised tokens within the coda clusters of Onset Lightening speakers, which rendered the tongue tip gesture unmeasurable. For the coda tokens examined here, statistic analyses found two of the three word pairs, namely *gulp / gull* and *philp / fill* to show a pattern of Left Edge stability, i.e., a sequential timing pattern. This timing patterns is the typical timing pattern predicted for coda clusters (e.g., Browman and Goldstein, 1988). However, it was not the timing pattern which was predicted here. Rather, the prediction made for the dark coda /l/s in this study was that they would show the same patterns as found in Marin and Pouplier (2010) for dark lateral coda clusters of American English. Marin and Pouplier (2010) observed lateral codas of American English speakers to show a leftward shift towards the preceding vowel when in a cluster context, relative to their temporal position within a singleton context. Given the findings for a sequential timing pattern within lateral coda clusters in German (e.g., Pouplier, 2012), and Romanian (e.g., Marin and Pouplier, 2014), where /l/ is comparatively clearer, it was hypothesised that the leftward shift reported for American English clusters was the result of a difference in lateral darkness between languages (Marin and Pouplier, 2014). The findings for a sequential timing pattern in coda clusters *gulp* and *philp* in the Onset Darkening Dialect here, runs contrary to this hypothesis.

Stability analysis results for the *milk / mill* word pair on the other hand, revealed no interval to be stable across the singleton - cluster word pair. This raises the question of whether the landmarks used in this analysis were sufficient to capture the intervals of stability between singleton and cluster contexts. For example, it is possible that different results may be found if intervals were defined by gestural onsets, rather than gestural targets. While incompatible with the predictions of the C-centre model, the finding for no stable interval in the *milk / mill* pair is not incompatible with the findings of Marin and Pouplier (2010) for a leftward shift of /l/ in the cluster context. Taking together the slightly shorter lateral to anchor lag *milk* compared to *mill* from Figure 5.19, and the finding for no stable intervals across the word pair, it seems that the /l/ in *milk* exhibits a slight left-ward shift relative to the singleton context, however, not enough to result in C-centre stability. Further, we may reasonably ask why singleton-cluster timing in the *milk / mill* word pair behaves differently to the *gulp / gull* and *philp / fill* word pairs. Since the /ɪ/vowel context was not unique the *milk / mill* word pair, also occurring in *philp / fill*, it is likely that differences here rather owe to the different post-lateral consonants (/k/ in *milk*, vs /p/ in *philp* and *gulp*). One possibility is that timing differences between word pairs result from the different articulators used to produce /k/ and /p/, (i.e, the tongue dorsum vs the lips). While the lip closure required to produce /p/ is independent from the lingual gestures of /l/, the tongue dorsum raising gesture of /k/ is not.

5.7 Summary

This chapter has presented findings of an articulatory study of lateral cluster timing patterns in two dialects of British English that differ in lateral darkening: An Onset Lightening dialect, and and Onset Darkening dialect. The central aim of the study was to test the effects of dialect mediated spatial differences in lateral darkening on patterns of cluster timing. Due to the large number of vocalised lateral codas in the Onset Lightening dialect, I was only able to test the effects of lateral darkening on patterns of onset cluster timing. The key finding of this study, and perhaps the most unexpected, was the finding that dialects do not differ in onset lateral cluster timing, despite dialectal differences in lateral darkening for these tokens, which also carry over to cluster contexts. This finding raises further important questions about what facilitates the existence of the stable timing pattern in lateral onset clusters, despite differences in lateral darkening between dialects. To arrive at an answer to this question, we must consider several possibilities including the limitations of the measure used. The final study of this thesis hopes to shed light on these issues.

Chapter 6

Dialect Variation in the Timing of the Tongue Body Gesture of /l/

6.1 Introduction

This chapter takes an explicit look at the tongue body gesture in /l/, focusing on how knowledge of the timing of the tongue body gesture can enrich our understanding of the relationship between lateral timing and lateral darkening, as well as its role in cluster timing. The inability to consistently measure the timing of the tongue body lowering gesture of /l/ in Chapters 4 and 5 left open questions. Within the first study, Chapter 4, the inability to reliably measure the tongue body gesture of /l/ meant the relative timing relationships between the tongue tip and tongue body gestures of /l/ could not be examined, thus limiting our understanding on the ways in which dialects differed temporally in /l/ darkening. Chapter 5 investigated how differences in lateral darkening affected patterns of lateral cluster timing. However, since the timing of the tongue body gesture of /l/ could not be measured across all tokens, lateral timing was defined by the timing of the tongue tip gesture of /l/ only. The main finding of the second study was that lateral cluster timing did not differ between dialects, in spite of a dialectal difference in /l/ darkening (which, primarily manifested in differences in the spatial magnitude of tongue body displacement). This finding raised an interesting question: how could dialects differ in the darkness of /l/ spatially, but not differ temporally in patterns of lateral cluster timing? Various solutions were proposed, one of which concerned the timing of the tongue body gesture. In order to confidently speak of a relationship between lateral darkening and lateral timing, we must look at the timing of both the tongue tip *and* the tongue body gestures of /l/, especially given the key role of the tongue body gesture in differentiating dialects in their degree of darkening.

This chapter aims to address both of these issues by exploring how the timing of the tongue body gesture of /l/ patterns at an inter-gestural and inter-segmental level. For this to be possible, this chapter focuses on a substantially narrower pool of tokens. Specifically, I focus on a high front vowel context, in which the tongue tip and the tongue body gestures of /l/ can be measured. While we cannot assume that the results from this subset of tokens generalise to all cases in Chapter 5, it provides a more comprehensive account of how lateral darkness interacts with cluster timing.

The structure of the chapter is as follows. I first provide a review of timing measures which utilise information on the timing of the tongue body / dorsal gesture of /l/, namely measures of inter-gestural timing, or tip delay. A review of measures of lateral cluster timing is not here provided; for this, the reader is referred back to the literature review provided prior to the previous study, Chapter 5. The empirical part of this chapter calculates temporal measures using both the tongue tip and tongue body gestures of /l/. These include measures of inter-gestural timing, as well as the cluster timing measures which were used within the previous chapter. I end the chapter by discussing the role of the tongue body gesture in exposing variation between dialects and cluster / singleton contexts.

6.2 Inter-gestural timing in /l/

Since Sproat and Fujimura (1993), measures of inter-gestural timing have been used to capture variation in /l/ darkening; however, such measures have also posed considerable methodological challenges. In the context of /l/, a measure of inter-gestural timing is the temporal distance between the dorsal and coronal gestures of /l/, hence the measure is often referred to as a measure of "Tip Delay" (Sproat and Fujimura, 1993, p. 298). Below I provide examples of how Tip Delay has been implemented using three different articulatory techniques: point tracking, Electropalatography, and Ultrasound Tongue imaging, before discussing potential methodological challenges.

6.2.1 Point Tracking

One of the most widely cited studies to implement a measure of inter-gestural timing is Sproat and Fujimura (1993). In their study, Sproat and Fujimura (1993) used X-ray microbeam, a point tracking technique which tracks the position of pellets attached to flesh points on the tongue, lips and gums to calculate a measure of inter gestural timing, or Tip Delay in /l/ within a high front vowel context. Pellets were attached to the tongue tip, body and dorsum of four American English speakers and one British English speaker. Tip Delay was calculated as the time between the extremum of the tongue tip raising gesture and the tongue body lowering gesture. For clear word-initial /l/s, they found the tongue tip gesture to precede the tongue body gesture, while for word-final dark /l/s before a intonation boundary they found that the tongue body gesture preceded the tongue tip gesture. They explained their findings in terms of differences in gestural attraction between vocalic and consonantal gestures; while vocalic gestures (i.e., the dorsal gesture of /l/) are attracted to the syllable nucleus, consonantal gestures (i.e., the coronal gesture of /l/) are attracted the syllable margin (Sproat and Fujimura, 1993, p. 306).

6.2.2 Electropalatography

Electropalatography (EPG) is an articulatory technique which measures lingual-palatal contact during speech, and it has been used to measure Tip Delay in Scottish Standard English and Southern Standard British English speakers (Scobbie and Pouplier, 2010). Laterals were recorded in a high front vowel context and a range of morphosyntactic positions. Tip Delay was calculated as the temporal distance between the time at which the tongue tip made contact with the alveolar palate region, and the time at which the palatal contact made by the tongue dorsum showed retraction from the fronted position of the high front vowel /i/. Their findings revealed a large distinction in gestural timing between onset and coda contexts. For Southern Standard British English speakers, the tongue tip gesture preceded the tongue dorsum gesture in onset position, while the tongue dorsum gesture preceded the tongue tip gesture in coda position. For Scottish Standard English speakers, the tongue dorsum gesture preceded the tongue tip gesture for all contexts, however, the degree to which the tongue dorsum preceded the tongue tip gesture was far greater within coda relative to onset position.

6.2.3 Ultrasound Tongue Imaging

Ultrasound imaging, an imaging technique which tracks the position of the tongue surface during speech, has been used by Strycharczuk and Scobbie (2015) to measured Tip Delay in SSBE speakers across a range of morphosyntactic contexts. Within their study, laterals were recorded at a frame rate of 121.5 fps. Tip Delay was calculated as the lag between 20% of the peak velocity of the dorsal and apical regions of the tongue which showed the greatest displacement during the lateral. Their results showed that, in the context of /u/, Tip Delay increased, and hence /l/ was darker at a stronger morphosyntactic boundaries.

6.2.4 Challenges in measuring tip delay

As we can see, Tip Delay in /l/ has successfully been calculated using a range of articulatory techniques. However, this measure also presents certain methodological challenges. Firstly, to measure the relative timing between gesture requires the articulatory device to record at a high temporal precision. However, this is not always available; for example, Lee-Kim, Davidson, and Hwang (2013) raise this as a concern for early ultrasound machines which had a limited frame rate in comparison to point tracking techniques such as EMA, which can record articulatory movement at 250Hz. This was not an issue, however, for the above described Ultrasound study of Strycharczuk and Scobbie (2015) who recorded their data at 121.5 fps.

Beyond temporal precision of the measurement device, other challenges arise, such as whether the most relevant parts of the gestures are being captured. For example, when using EPG to measure Tip Delay, Scobbie and Pouplier (2010) discuss how being limited to measuring only tongue to palate contact meant that the apical and dorsal gestures of /l/ were captured at different stages of the the gesture. Given the pre-existing contact of the tongue dorsum with the palate during the preceding vowel, the dorsal gesture of /l/ could be measured from the onset of movement towards the target, while the tongue tip only made contact with the alveolar region of the palate upon achievement of the apical target. Such issues are clearly relevant to capturing precise timing relationships between gestures.

A further methodological challenge for measures of inter-gestural timing, and one that has been encountered within studies of this thesis, is the difficulty in locating the tongue body gesture of /l/. The difficulty in capturing the tongue body gesture typically owes to the segmental environment of /l/. The vowel-like lowering gesture undergoes considerable coarticulation with surrounding vowels, hence becomes difficult to isolate, particularly in the context of back vowels, where a similar tongue movement is recruited (e.g., Strycharczuk, Derrick, and Shaw, 2020). It is important to acknowledge this issue as a problem of conceptualisation; from the perspective of Articulatory Phonology, gestures should not be treated as discrete units, but as continuously varying forces acting on articulatory parameters. Hence, if gestures are not discrete units, discrete timing measures cannot always be taken. For example, if a TB gesture for /l/ and a tongue body gesture for the following vowel are overlapped and have similar spatial targets, then it stands to reason that there will be few identifying signatures that allow us to separate the two gestures. This does of course prove inconvenient for the purpose of extracting precise measures of articulatory timing.

While it may be challenging to obtain measures of gestural timing, they are critical to our understanding of variation in /l/. The suggested relationship between gestural timing and spatial magnitude in /l/ is one example of this. For example, Strycharczuk, Derrick, and Shaw (2020) discuss this relationship with reference to the apical gesture of /l/: assuming a fixed temporal window, a tongue tip raising gesture which occurs relatively later in that window will be smaller in magnitude (i.e., undershot) than a gesture which occurs earlier within the window. Hence, increasing Tip Delay is considered to be a precursor to vocalisation.

6.2.5 Role in cluster timing

Knowledge of gestural timing can also be used to increase the precision of measures of cluster timing. As described in the previous chapter, onset cluster timing refers to how an onset consonant is timed relative to a following vowel or anchor point when the consonant is the only consonant in the onset (a singleton), compared to when the consonant is part of an onset consonant cluster. Testing whether a C-centre organisation has been maintained involves measuring the distance between the temporal midpoint of the gestures in the consonant onset and a fixed anchor point, and comparing this between singleton and cluster contexts. If the interval is the same across the singleton and cluster context, then a C-centre pattern has been maintained (Browman and Goldstein, 1988). Within /l/ cluster timing studies, it is common for only a single gesture of /l/ to be considered in such calculations, typically the apical raising gesture (Marin and Pouplier, 2010, 2014). However, if we use information gained from measures of inter-gestural timing then more accurate cluster timing measures can be made, especially if the laterals show great variability in the timing of the anterior and posterior gestures. For example, the temporal midpoint of a singleton /l/ in onset position can be calculated as the midpoint between the achievement of the apical and dorsal gestures of /l/, rather than the achievement of a single one of the gestures. Further, using knowledge of the relative temporal order of the apical and dorsal gestures of /l/, the temporal midpoint of a C/l/ cluster onset can be calculated as the midpoint between the leftmost and rightmost gesture. Only through measuring the relative timing of the gestures of /l/ can the rightmost gesture of the consonant cluster be known. The rationale for examining the timing of the individual gestures of /l/ presented here is in keeping with the arguments pursued by Browman and Goldstein, 1986, that there are no segments, only constellations of gestures. From this perspective, English /l/s, comprising a dorsal and an apical gesture, are, in themselves clusters. To gain a complete understanding of how laterals are timed in cluster and singleton contexts then, it is necessary to consider the timing patterns of individual gestures which each contribute to the overall temporal alignment with the vowel.

6.3 Research Questions

The overarching goal of this chapter then is to determine whether analysing the timing of the tongue body gesture in /l/ can help to explain how laterals vary across singleton and cluster onsets within the two dialects examined here. The specific researched questions of this chapter are as follows:

- RQ: *How can an analysis of both the tongue tip and tongue body gestures of /l/ enhance our understanding of the relationship between /l/ darkening and /l/ cluster timing?*
- Sub RQ: *How do patterns of inter-gestural timing in /l/ differ between dialects and singleton vs cluster contexts?*

6.4 Methods

Electromagnetic articulography data was examined from 6 Onset Darkening speakers and 8 Onset Lightening speakers. Two Onset Darkening speakers were excluded from this analysis due to missing tongue body sensor. Data was collected and processed according to the procedures outlined within the general methods section (Chapter 3).

6.4.1 Stimuli

For this analysis, a single word pair was examined; namely, the singleton - cluster pair: *plick / lick*. There were four repetitions of each token per speaker, but for some speakers, this number was less due to mis pronuciations or processing errors. A total of 107 tokens were included in the analysis; 60 tokens were produced by Onset Lightening speakers, and 47 tokens were produced by Onset Darkening speakers. The *plick / lick* pair was chosen because, unlike other word pairs used in previous studies of this thesis, it offers a context where the tongue body lowering gesture of /l/ is clearly identifiable. In the cluster token *plick*, the non-lingual segment preceding /l/ and the following high front vowel allows the lowering gesture of the tongue body to be clearly and unambiguously observed. This clear lowering of the tongue body can be seen in the velocity and displacement trajectories shown in Figure 6.1, which is discussed later in this section. Note, for this figure, the tongue dorsum (TD) sensor is used to capture tongue body lowering; this is discussed in greater detail in the next section.

As with the previous chapter, the anchor point for the *plick / lick* pair is the velocity minimum of the tongue dorsum raising gesture of /k/. The anchor point is the point relative to which various time points within the /l/ or /pl/ onset will measured.

6.4.2 Identifying the Tongue Body lowering gesture

Before gestural and segmental timing measures could be calculated, the time of gesture achievement had to be identified for both the tongue tip and tongue body gestures of /l/. The previous chapter identified the time of the tongue tip raising gesture of /l/ and the time of gesture achievement of the anchor consonant /k/ for all vowel contexts. I here supplement these time points with the time of the tongue body lowering gesture of /l/ for the *plick / lick* pair. With this information, it is possible to look at the relative timing between the tongue tip and tongue body gestures of /l/, and compare how the different gestures of /l/ pattern in measures of lateral cluster timing. The primary motivation for looking at the tongue body lowering gesture is the findings of Chapter 4, which found differences in /l/ darkening between dialects to manifest in differences in variation in the vertical displacement of the tongue body. In addition, others have also found the lowering of the tongue body to be a robust correlate of /l/ darkening (Lee-Kim, Davidson, and Hwang, 2013; Proctor et al., 2019; Sproat and Fujimura, 1993; Strycharczuk, Derrick, and Shaw, 2020).

As with the previous chapter, the time of gesture achievement is here defined as the velocity minimum of the relevant sensor in the relevant dimension. The time of achievement of the tongue body gesture of /l/ was defined as the velocity minimum of the tongue body or tongue dorsum sensor in the vertical dimension. The sensor used (tongue body or tongue dorsum) varied between speakers depending on which sensor showed the clearest extremum. In practice, both the tongue body and tongue dorsum sensors are positioned on the anatomical tongue body region about 1cm apart, thus the two are assumed to be highly correlated. Details of which sensors were used for each speaker are provided in Appendix 3 – Section 9.2. From visual inspection, consistent displacement trajectories were observed for both the tongue body and tongue dorsum sensors. Further, regardless of which sensor was used here, this gesture will hereafter be referred to as the tongue body lowering gesture of /l/.

To identify the relevant time point of the tongue body (or tongue dorsum) velocity profile, the velocity data was manually inspected according to the point in time when there was a clear lowering in the displacement data, as illustrated in Figure 6.1. It is clear to see from Figure 6.1 that the point of maximal lowering unambiguously corresponds to a velocity minimum between the largest two velocity peaks.

After visually assessing the velocity and displacement profiles, a temporal window containing the relevant velocity minimum was manually specified. The *findpeaks* function of the pracma R package (Borchers, 2022) was then used to automatically identify the velocity minima within the specified temporal window. This process was performed separately on each speaker for each prompt (i.e., *plick/lick*) and, where necessary, temporal windows were individually specified for each token to ensure the relevant time point was correctly extracted. All outputs were manually checked for accuracy.

6.4.3 Calculating temporal measures

Once the time of the tongue body lowering gesture had been identified for all tokens, temporal measures were calculated and compared across contexts (singleton and cluster) and dialects (ODD and OLD). Three temporal measures were calculated for both *plick* and *lick* tokens: (i) The time between the tongue tip and tongue body gestures of /l/; (ii) The time between the achievement of the tongue tip gesture of /l/ and the anchor /k/; (iii) The time between the achievement of the tongue body gesture of /l/ and the anchor /k/. The first measure will be referred to as a measure

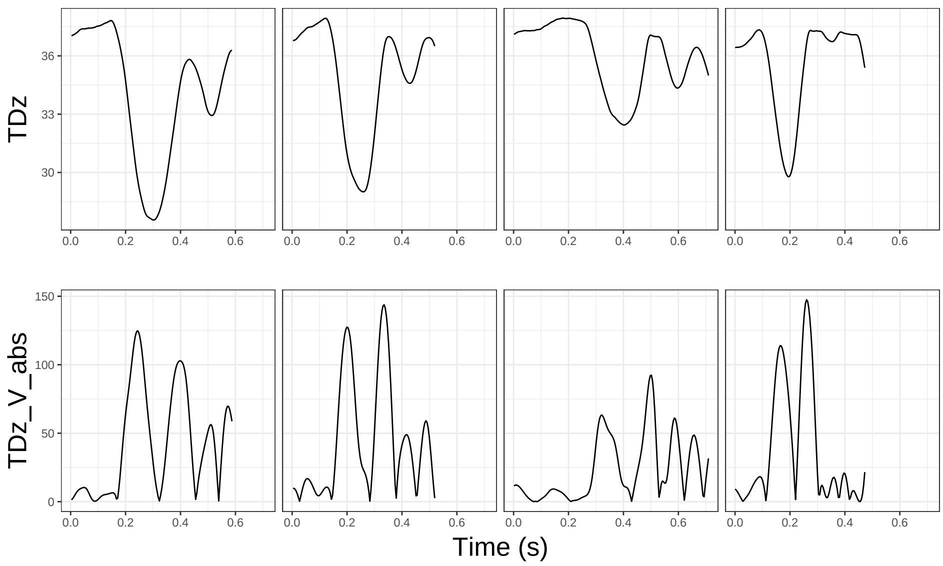

Figure 6.1: Tongue body displacement and velocity (captured here by the TD sensor) during /l/ plus the preceding and following segment for Onset Lightening speaker S07.

of inter-gestural timing, and the last two will be referred to as lateral gesture to anchor lags.

Inter-gestural timing between the TT and TB gestures of /l/ was calculated as the time of the achievement of the tongue tip raising gesture of /l/ (velocity minimum of TTz) minus the time of the achievement of the tongue body lowering gesture of /l/ (velocity minimum of TBz or TDz):

$$\text{Time of TT /l/} - \text{Time of TB /l/}$$

A negative inter-gestural timing value means that the TT gesture of /l/ temporally precedes the TB gesture of /l/; a value of 0 means that both gestures occur at the same time; and a positive value means that the TB gesture precedes the TT gesture.

The tongue tip lateral gesture to anchor lag was calculated as the time of achievement of the anchor consonant /k/ (defined as the velocity minimum of the TD raising gesture of /k/) minus the time of the achievement of the tongue tip raising gesture

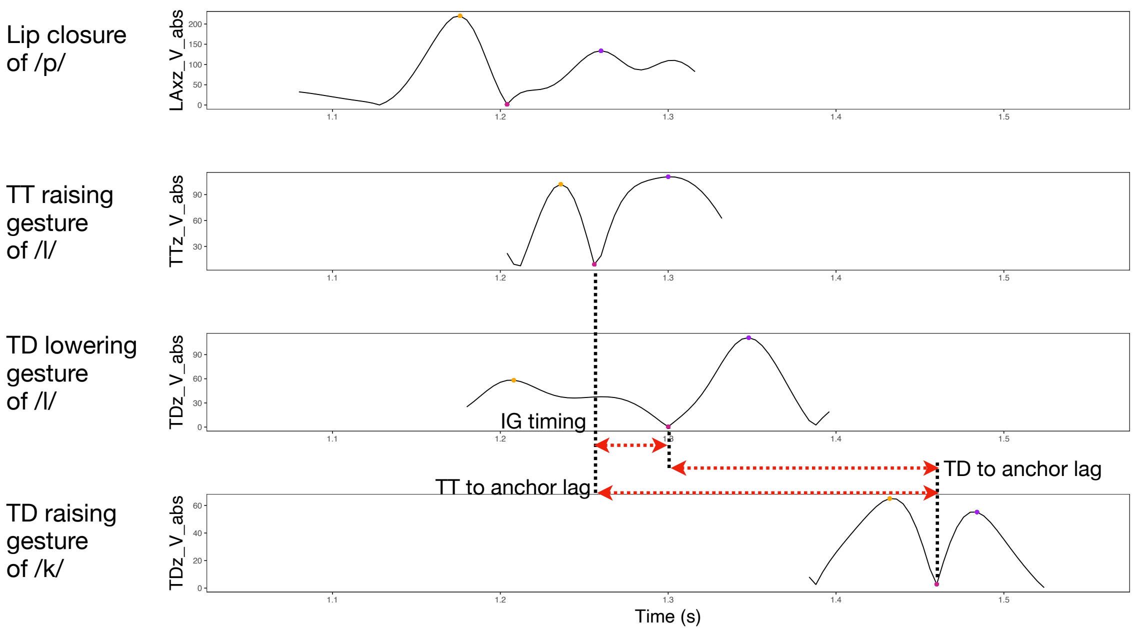

Figure 6.2: Velocity profiles for *plick* showing measurement intervals. Trajectories are shown for a single token from one Onset Lightening speaker, S01. Temporal measures include: (i) inter-gestural timing, between the velocity minima of the TT raising and TD lowering of /l/; (ii) TT to anchor lag, between the velocity minima of the TT raising of /l/ and TD raising of /k/; (iii) TD to anchor lag, between the velocity minima of the TD lowering of /l/ and TD raising of /k/.

(velocity minimum of TTz). The tongue body lateral gesture to anchor lag was calculated as the time of achievement of the anchor consonant /k/ (defined as the velocity minimum of the TD raising gesture of /k/) minus the time of the achievement of the tongue body lowering gesture (velocity minimum of TBz or TDz):

$$\text{Time of TD /k/} - \text{Time of TT /l/}$$
$$\text{Time of TD /k/} - \text{Time of TB /l/}$$

Figure 6.2 shows an illustration of the above described temporal measures for an Onset Lightening speaker's production of *plick*. While *plick* is used for purposes of illustration, the temporal measures shown here are also identical for *lick*, except for the obvious absence of a lip closing gesture for /p/. Each panel shows absolute velocity profiles corresponding to the achievement of each gesture of *plick*. For each gesture, the velocity minimum is shown between the preceding and following velocity peak. The top panel shows absolute velocity of a derived measure of lip aperture in the x/z dimension during the lip closure for /p/ in *plick*. The second panel shows the absolute velocity of the tongue tip in the vertical dimension during the tongue tip raising gesture of /l/. The third panel shows the absolute velocity of the tongue dorsum in the vertical dimension during the tongue body/dorsum lowering of /l/. The bottom panel shows the absolute velocity of the tongue dorsum in the vertical dimension during the tongue dorsum raising gesture of /k/. Red arrows indicate the temporal distances measured. Temporal measures are compared across singleton and cluster contexts, and dialects.

6.4.4 Stability measures

Stability measures were calculated to determine the most stable interval across the *plick / lick* pair. With information on the timing of both the TT and TB gestures of /l/, a more nuanced stability analysis is performed compared to that of the previous chapter. Rather than simply comparing the singleton centre to anchor lag to a range of cluster to anchor lags, C-centre, Left-Edge, and Right-Edge lags could be compared across both singleton and cluster tokens. These are described below:

C-centre to anchor lag

The C-centre of the singleton token *lick* was defined as the temporal midpoint between the velocity minimum of the tongue tip raising gesture and tongue body lowering gesture of /l/ subtracted from the time of the velocity minimum of the tongue dorsum raising gesture of the anchor /k/. This is illustrated in the top diagram of Figure 6.3.

The C-centre of the cluster *plick*, was defined as the temporal midpoint between the achievement of the leftmost gesture in the consonant cluster and the achievement of the rightmost gesture in the consonant cluster subtracted from the time of the velocity minimum of the tongue dorsum raising gesture of the anchor /k/. Again time of gesture achievement is defined as the velocity minimum of the relevant gestural dimension. The leftmost gesture in the consonant cluster of *plick* was always the lip closure gesture of /p/; however, the rightmost gesture of the consonant cluster varied between speakers; for most it was the tongue body lowering gesture of /l/,

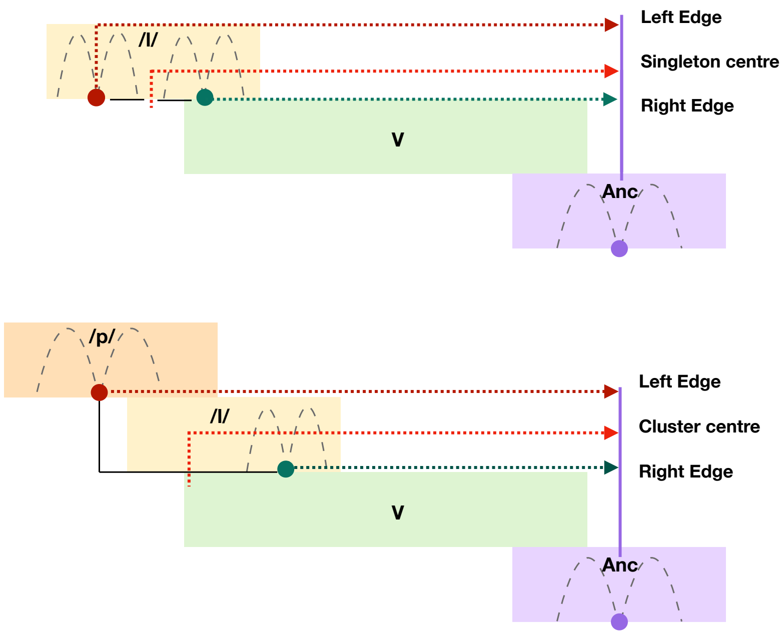

Figure 6.3: Schematic diagram of stability intervals for *“lick* (top), and *plick*(bottom). Circles denote velocity minima. Arrows show the duration measured for each lag.

but for some it was the tongue tip raising gesture of /l/.

Left Edge to anchor

The Left Edge lag of *lick* was the time of achievement of the leftmost gesture of /l/, which was the TT raising velocity minimum for all tokens, subtracted from the time of the velocity minimum of the tongue dorsum raising gesture of /k/.

The Left Edge lag of *plick* was the time of achievement of the leftmost gesture in the cluster subtracted from the anchor.

Right Edge to anchor

The Right Edge lag of the singleton was the time of achievement of the rightmost gesture of /l/ (which was the TB gesture of /l/ for all tokens) subtracted from the time of the velocity minimum of the tongue dorsum raising gesture of /k/.

The Right Edge lag of the cluster was the time of achievement of the rightmost gesture in the cluster subtracted from the anchor. The rightmost gesture in the cluster varied between the tongue tip and tongue body gesture of /l/ for different speakers.

6.4.5 Statistical measures

Model comparisons were used to quantify the effect of consonant structure (singleton or cluster), and the effect of dialect (ODD or OLD) on: (i) lateral inter-gestural timing; (ii) lateral gesture to anchor lags; (iii) stability measures. The procedure for each is outlined below:

Lateral inter-gestural timing

A full linear mixed-effects model was created using the *lme4* R package (Bates et al., 2015). The full model included the measure of inter-gestural timing as a function of fixed effects of dialect, consonant structure (singleton or cluster) and estimated vocal tract length, a dialect by consonant structure interaction, a random intercept for speaker, and a random slope for speaker for the effect of consonant structure.

To test for the effect of (i) a dialect by consonant structure interaction, (ii) consonant structure, and (iii) dialect, the full model was compared to partial models where the relevant effect had been excluded. First an effect of the dialect by consonant structure interaction was tested; if this was found to be non-significant ($p > .05$), then further model comparisons were performed to test for fixed effects of consonant structure and dialect. If the interaction was found to be significant ($p < .05$), no further model comparisons were conducted and the effect of the interaction is reported as the major finding.

Lateral gesture to anchor lags

For each of the lateral gesture to anchor lags (TT to anchor and TB to anchor), model comparisons were conducted in the same way as above to test for effects of consonant structure and dialect. The effect of a dialect by consonant structure interaction was not tested, and hence a dialect by consonant structure term was not included in the full model. This is because the interaction was found to be non significant for the inter-gestural timing measure, which is considered to be highly correlated with lateral gesture to anchor lags.

Stability measures

Three sets of model comparisons were performed for each of three stability lags: (i) C-centre to anchor, (ii) Left Edge to anchor, and (iii) Right Edge to anchor. For each, a full model was created including fixed effect terms of consonant structure and dialect, a dialect by consonant structure interaction, a random intercept for speaker, and a random slope for speaker for the effect of context. Full models were then compared to partial models where the relevant effect had been excluded. As with the lateral inter-gestural timing model, the effect of the dialect by consonant structure interaction was first tested; if the interaction was found to be non-significant ($p > .05$), then further model comparisons were performed to test for fixed effects of context and dialect. If the interaction was found to be significant, ($p < .05$), no further model comparisons were conducted.

6.5 Results

6.5.1 Inter-gestural timing results

Figure 6.4 shows inter-gestural timing patterns between the tongue tip raising and tongue body lowering gestures of /l/ for each context (singleton - *lick*, and cluster - *plick*), and each dialect (ODD - Onset Darkening Dialect, and OLD - Onset Lightening Dialect). Inter-gestural timing is here measured as the time of the tongue tip raising gesture (TTz velocity minimum) minus the time of the tongue body lowering gesture (TBz velocity minimum). A negative value means that the tongue tip gesture precedes the tongue gesture of /l/. A value of 0 means that both gestures occur at the same time, and a positive value means that the tongue body gesture occurs before the tongue gesture.

For both dialects, the tongue tip and tongue body gestures of /l/ are temporally closer (i.e., closer to 0) in the cluster context compared to the singleton context. In the singleton *lick* context, negative values show that the tongue tip gesture pre-

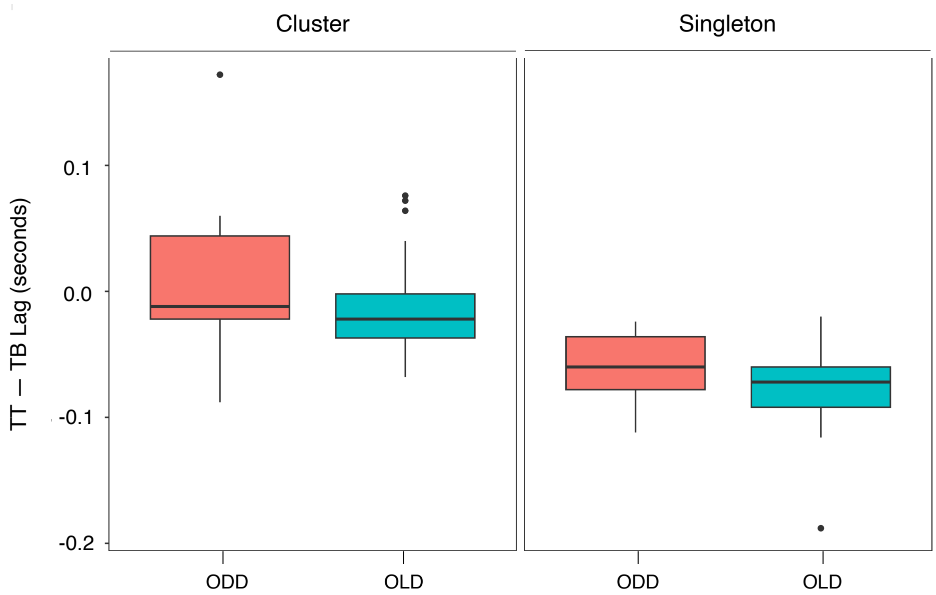

Figure 6.4: Inter-gestural timing (time of TT gesture - time of TB gesture) for cluster tokens (*plick*) on the left, and singleton tokens (*lick*) on the right.

cedes the tongue body gesture for both dialects. In the cluster context of *plick*, gestures appear to be produced near simultaneously for both dialects. A small dialectal differences can be observed in the singleton context whereby Onset Lightening speakers show a larger negative value than Onset Darkening speakers; this suggests that the tongue tip gesture occurs earlier relative to the tongue body gesture for Onset Lightening speakers compared to Onset Darkening speakers in this context. However, dialectal differences are not present in the cluster condition, with both dialects showing values close to 0, though a larger amount of variation can be observed for Onset Darkening speakers. Results of model comparisons (Table 6.1) confirm a significant effect of consonant structure (singleton versus cluster) on inter-gestural timing, and a non significant effect of dialect and a non-significant interaction between dialect and consonant structure.

Measure Set	Effect Tested	Chi Sq	*df*	p-value
IG timing TT−TB	Interaction	0.0545	1	0.816
	Consonant Structure	14.336	2	$<$**0.001**
	Dialect	1.472	2	0.479

Table 6.1: Model comparison results of inter-gestural timing measures

To get a sense of the degree of between speaker variation, a further by-speaker analysis of inter-gestural timing was performed. Another way to look at inter-gestural timing is to compare the lag durations of the tongue tip gesture of /l/ to anchor, and the tongue body gesture of /l/ to anchor. If the tongue tip to anchor lag is longer than the tongue body to anchor lag, this means that the tongue tip gesture is further away from the anchor consonant, and hence occurs before the tongue body gesture. Conversely, if the tongue body to anchor lag is longer than the tongue tip to anchor lag, then the tongue body gesture precedes the tongue tip gesture. If lag durations are the same, this suggests the tongue tip and tongue body gestures to occur at the same time. Lag durations for individual speakers are shown in Figure 6.5 for the Onset Darkening dialect, and in Figure 6.6 for the Onset Lightening dialect.

Figures 6.5 and 6.5 show that all speakers of both dialects produce the tongue tip raising gesture before the tongue body lowering gesture within the singleton *lick* context, as Figure 6.4 suggests. For the cluster *plick*, context, variation can be observed. Within the Onset Darkening dialect, 4 speakers, L02, L04, L05, and L06, produce the tongue tip gesture of /l/ before the tongue body gesture, while two speakers, L01 and L07, produced the tongue body gesture before the tongue tip gesture. Within the Onset Lightening dialect, 6 speakers, S01, S03, S04, S05, S07, and S08, produced the tongue tip gesture of /l/ before the tongue body gesture, while 1 speaker, S02, produced the tongue body gesture before the tongue tip gesture, and 1 speaker, S06, produces the tongue tip and tongue body gestures near synchronously. These speaker-level patterns of /l/ gesture timing show that, for most speakers, the tongue tip gesture precedes the tongue body gesture of /l/ in both singleton and cluster contexts. For some speakers, however, the cluster context prompts a reversal in gestural order, such that the tongue body gesture is shifted to precede the tongue tip gesture. These patterns do not appear to be dialect specific.

6.5.2 Lateral gesture to anchor lag results

Figure 6.7 shows lateral gesture to anchor lags for both the tongue body lowering gesture of /l/, and the tongue tip raising gesture of /l/. This is a measure of the time between the achievement of either the TT or TB gesture of /l/ and the time of achievement of the TD gesture of the anchor consonant /k/. The left panel shows

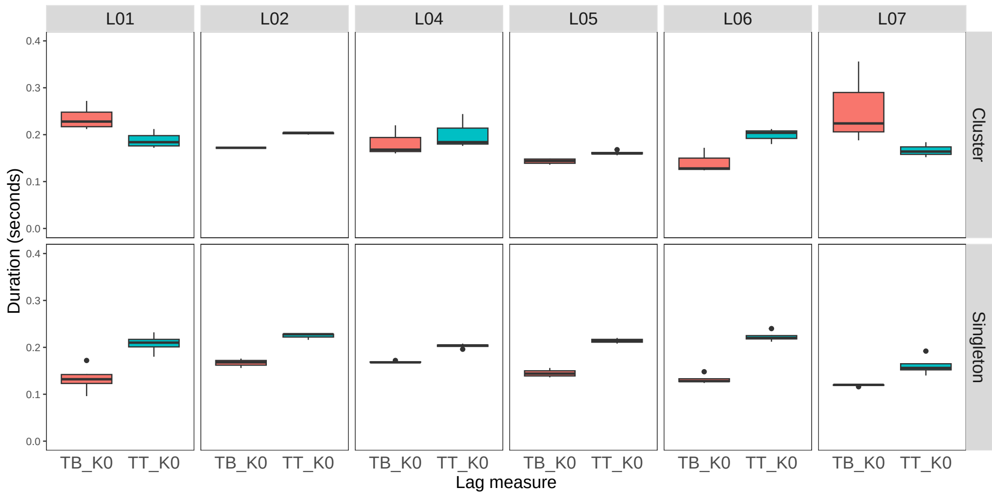

Figure 6.5: Lateral gesture to anchor lags for Onset Darkening speakers. TB K0 shows distance between the tongue body gesture of /l/ and the anchor consonant /k/. TT K0 shows distance between the tongue tip gesture of /l/ and the anchor consonant. Lags are shown for the cluster context *plick* on the top panels, and the singleton context *lick* on the bottom panels.

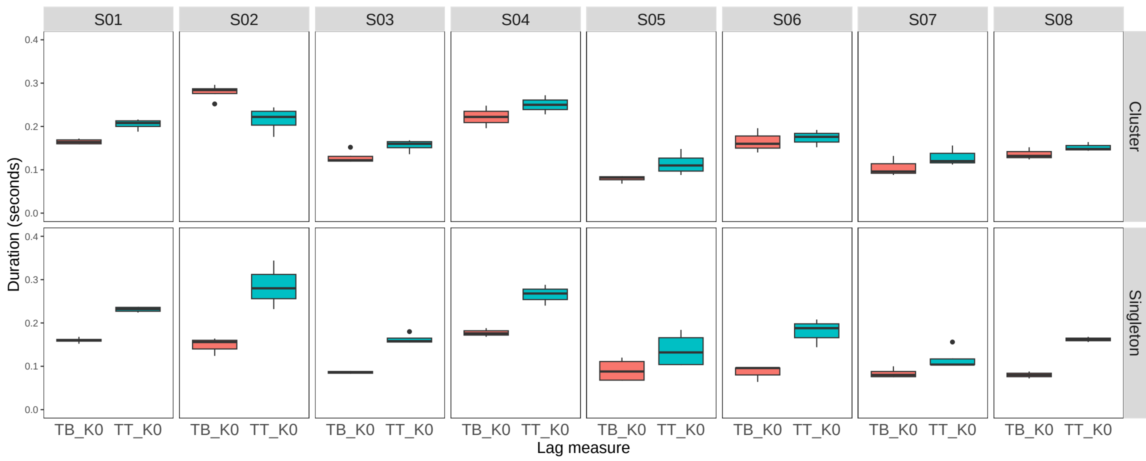

Figure 6.6: Lateral gesture to anchor lags for Onset Lightening speakers. TB K0 shows distance between the tongue body gesture of /l/ and the anchor consonant /k/. TT K0 shows distance between the tongue tip gesture of /l/ and the anchor consonant. Lags are shown for the cluster context *plick* on the top panels, and the singleton context *lick* on the bottom panels.

lags for the cluster context *plick*, and the right panel shows lags for the singleton context *lick*. Dialect is shown by colour. For both dialects, the tongue body to anchor lag is longer in the cluster contexts compared to the singleton context, while the tongue tip to anchor lag slightly shorter in the cluster context. This means that the tongue body gesture shifts leftwards away from the vowel in the cluster context, compared to the singleton context. There appears to be no clear differences between dialects.

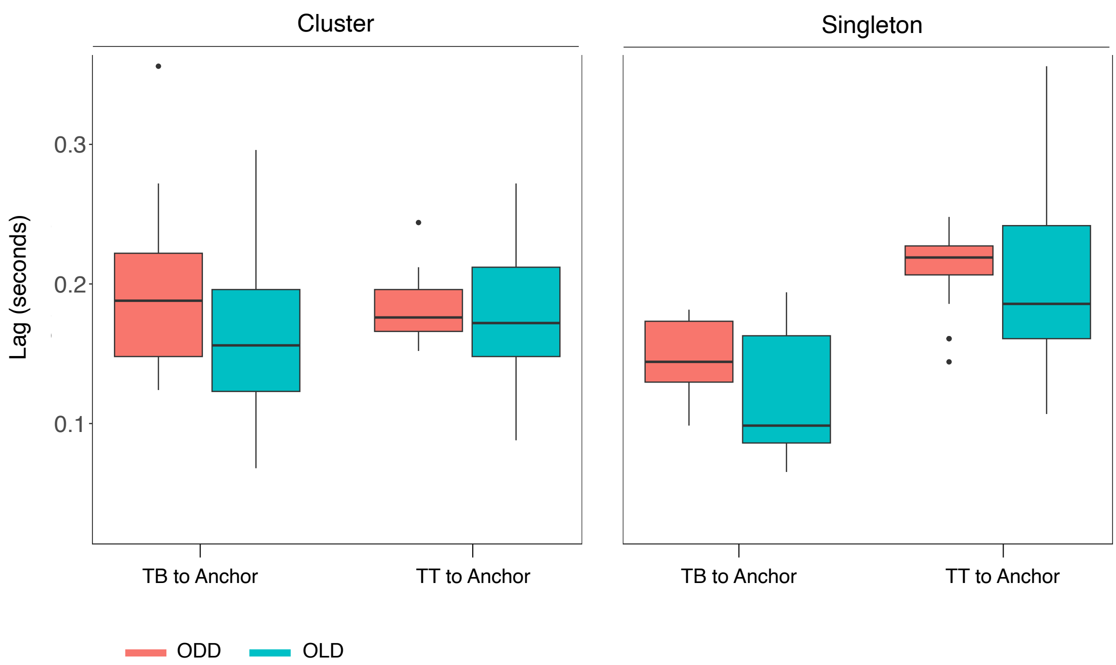

Figure 6.7: TB gesture of /l/ to anchor lags and TT gesture of /l/ to anchor lags for *plick*(left), and *lick*(right).

To quantify differences in lateral gesture to anchor lags between dialects and consonant structure, model comparisons were performed. To test for effects of consonant structure (singleton versus cluster) and dialect (Onset Darkening Dialect versus Onset Lightening Dialect), a full model was compared to partial models where the relevant effect had been excluded. An effect was considered significant if it returned a p-value of $< .05$. Note, for this model, the random slope for speaker for the effect of consonant structure was excluded due to convergence issues. Model outcomes are reported in Table 6.2. Results showed the effect of consonant structure to be significant for both the tongue tip to anchor lag, and the tongue body to anchor lag. However, the effect of dialect, and a dialect by consonant structure interaction was not significant for either lags.

6.5.3 Stability analysis results

To determine the most stable interval across the *plick/lick* pair, three intervals were compared: the C-centre to anchor lag, the Left Edge to anchor lag, and the Right Edge to anchor lag, as shown in Figure 6.8, (repeated here for reference), and Table 6.3.

Measure Set	Effect Tested	Chi Sq	*df*	p-value
TB gest of /l/ to anchor	Interaction	0.0967	1	0.756
	Consonant Structure	35.864	2	<**0.001**
	Dialect	2.629	2	0.269
TT gest of /l/ to anchor	Interaction	0.0133	1	0.908
	Consonant Structure	15.636	2	< **0.001**
	Dialect	0.648	2	0.723

Table 6.2: Model comparisons of lateral to anchor lags. Significant effects ($p < .05$) as shown in bold.

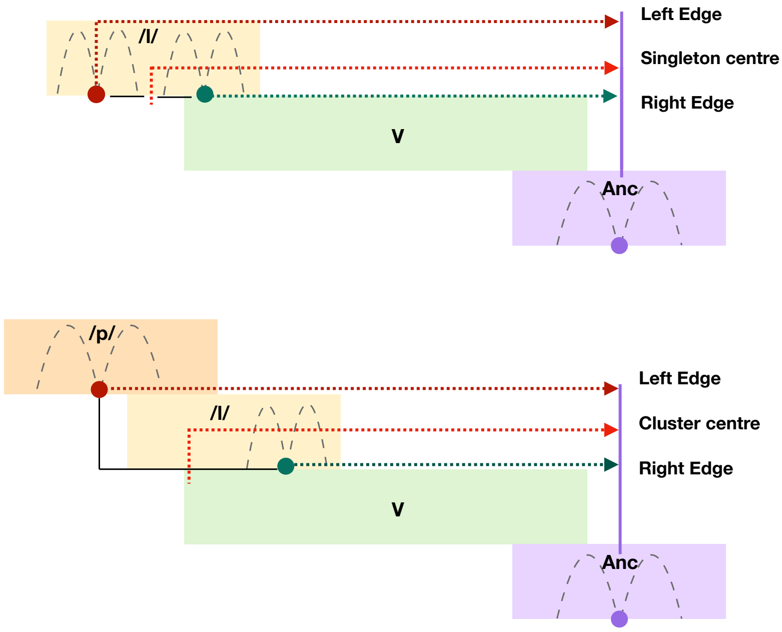

Figure 6.8: Diagram illustrating stability intervals for singleton *lick* (top), and cluster *plick* (bottom). Circles denote velocity minima. Arrows show duration measured for each lag.

For context, a C-centre timing pattern would predict a non significant effect of consonant structure (i.e., singleton or cluster) in the comparison of the C-centre intervals, but a significant effect of consonant structure in the Left Edge and Right Edge comparisons. In addition to the effect of consonant structure, I also tested for the effect of dialect.

Stability measure results are first presented in Figure 6.9. Within the leftmost panel, C-centre intervals are compared between the *plick* / *lick* word pair. Left Edge intervals are compared in the middle panel, and Right Edge intervals are compared

Comparison	Cluster *plick*	Singleton *lick*
C-centre	C-centre to anchor	C-centre to anchor
Left Edge	Left Edge to anchor	Left Edge to anchor
Right Edge	Right Edge to anchor	Right Edge to anchor

Table 6.3: Stability measure comparison structure

in the rightmost panel. Dialect is shown by colour. Visually, intervals are the most similar across the *plick / lick* word pair for the Right Edge intervals; hence the Right Edge to anchor appears to be the most stable interval across the singleton-cluster pair. Both dialects pattern very similarly; however, greater variation can generally be observed for Onset Lightening speakers compared to Onset Darkening speakers.

Table 6.4 presents the outcomes of models comparisons quantifying the effect of consonant structure (singleton or cluster), dialect (Onset Darkening dialect and Onset Lightening dialect), and the interaction between the two for each interval. For all intervals, the C-centre, Left Edge, and Right Edge, the effect of the consonant structure by dialect interaction was non-significant, as was the effect of dialect. Consonant structure had a significant effect on the interval ($p < .05$), and did so on all intervals. This means that there was a significant difference between: (i) the C-centre lag in *lick* and the C-centre lag in *plick*, (ii) the Left Edge lag in *lick* and the Left Edge lag in *plick*, and (iii) the the Right Edge lag in *lick* and the Right Edge lag in *plick*. This suggests that the singleton and cluster context, *plick* and *lick*, are fundamentally different in temporal structure, for both dialects.

6.5.4 Results interim summary

Considerable differences in patterns of inter-gestural timing in /l/ were observed between singleton and cluster contexts. In the singleton context, the tongue tip raising and tongue body lowering gestures of /l/ were found to be temporally further apart. Relative to the anchor point (the time of the velocity minimum of the tongue dorsum raising gesture of /k/), the tongue tip gesture of /l/ is achieved earlier, and the tongue body gesture is achieved later in the singleton context compared to the

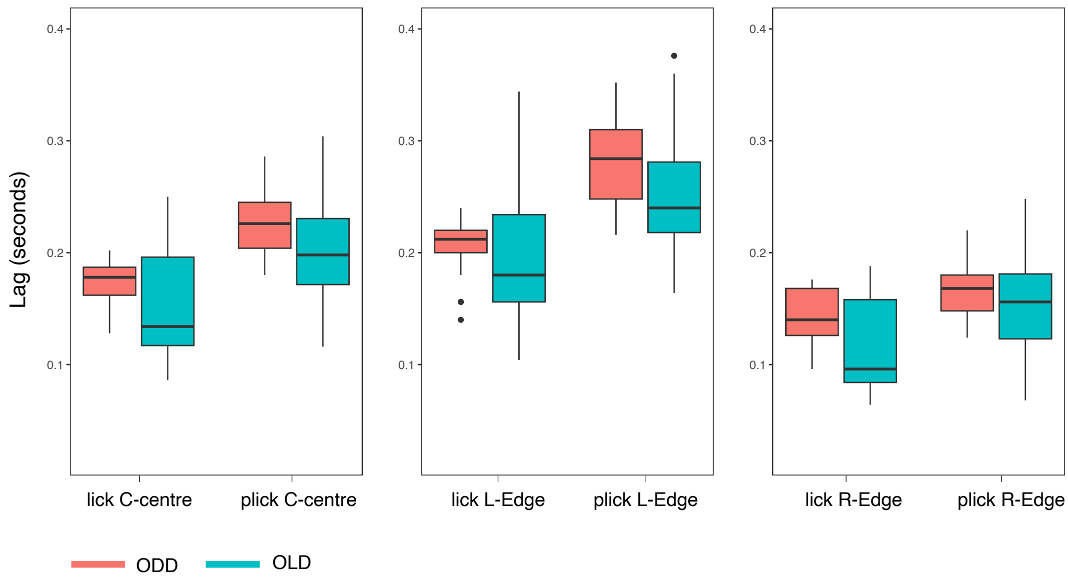

Figure 6.9: Comparison of C-centre, left edge and right edge intervals across the singleton - cluster pair *plick / lick*.

Measure Set	Effect Tested	Chi Sq	*df*	p-value
C-Centre	Interaction	0.002	1	0.961
	Consonant Structure	21.228	2	**<0.001**
	Dialect	1.451	2	0.484
Right Edge	Interaction	1.113	1	0.292
	Consonant Structure	12.419	2	**0.002**
	Dialect	3.682	2	0.159
Left Edge	Interaction	0.981	1	0.322
	Consonant Structure	23.662	2	**<0.001**
	Dialect	2.002	2	0.368

Table 6.4: Model comparisons for *plick / lick* pair. Significant effects ($p < .05$) are shown in bold.

cluster context. This suggests the cluster context triggers a compression effect, such that the gestures of /l/ in an onset cluster are achieved within a shorter temporal window than when in a singleton onset. Interestingly, this pattern occurred across both dialects; the effect of dialect was non-significant.

Measures of cluster timing found none of the intervals measured (C-centre, Left Edge and Right Edge) to be stable across singleton and cluster contexts. Least variability across contexts was found for the Right Edge lag; however, the effect of consonant structure was still significant for this lag. Again, dialects patterned similarly in cluster timing measures, and the effect of dialect was non-significant.

6.5.5 Potential explanations for lack of a dialect effect

Findings for a lack of effect of dialect on measures of cluster timing across the *plick / lick* pair here are in agreement with findings of the previous chapter. The previous chapter found no effects of dialect on /l/ cluster timing across a broader range of segmental contexts, using only the tongue tip raising gesture to temporally define /l/. Further, the issues raised within the discussion of the previous chapter remain relevant here. Specifically, we may continue to ask: how can dialects differ spatially in the tongue body gesture of /l/, but show no temporal differences in patterns of gestural or segmental timing in /l/? After factoring in the role of the tongue body gesture, there are two further potential solutions to this problem which we can explore. The first possible solution is that dialects differ in the velocity of the tongue body gesture of /l/, such that speakers of the Onset Darkening Dialect can achieve greater tongue body displacement than speakers of the Onset Lightening Dialect, but within the same temporal window. The second possibility is that there may be compensatory differences in vowel length between dialects. The measures of /l/ cluster timing employed here measure the temporal lag between a point within the consonant onset and a point within a post vocalic anchor consonant. Therefore, it is possible that systematic differences in vowel length occur between dialects, which affect the duration of this interval. However, vowel duration is not methodologically straightforward to test given the difficulty in precisely measuring vowel duration when /l/ is the neighbouring consonant.

For now, I return to the possibility that differences in velocity may be responsible for the seemingly incompatible results between spatial magnitude and timing in /l/. To explore this as a potential solution, absolute velocity profiles of tongue body lowering are compared between dialects.

Absolute velocity is considered during the time window between the first and second velocity peak of the tongue body lowering gesture of /l/ in *plick* and *lick*, as shown in Figure 6.10. This interval was selected because it captures the velocity

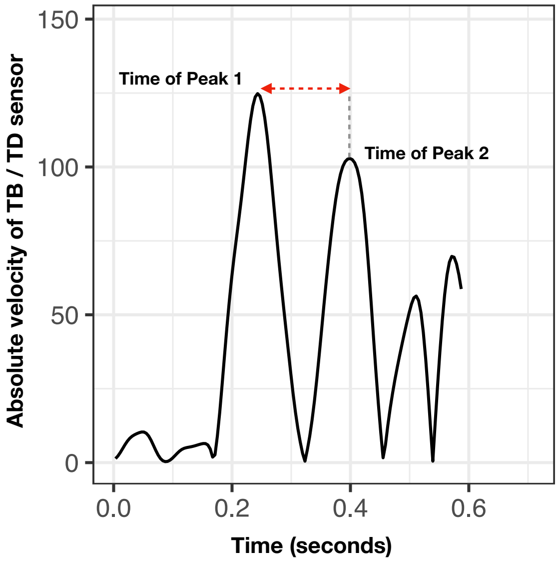

Figure 6.10: Temporal window of TBz/TDz absolute velocity measured. Red arrow shows the interval measured between maximum peak 1 and peak 2.)

trajectory on approach to the target (between peak 1 and the velocity minimum) and the trajectory away from the target (between the velocity minimum and peak 2). Figures 6.11 and 6.12 shows individual absolute velocity trajectories during the tongue body lowering gesture of /l/ for the above specified temporal window for each context. Figure 6.11 shows absolute velocity trajectories of Onset Darkening speakers, and Figure 6.12 shows absolute velocity trajectories for Onset Lightening speakers. For each figure, the cluster context is shown on the top panels, and the singleton context is shown on the bottom panels. Since velocity is closely linked with vocal tract size, individual VT length estimates (as presented in Chapter 3) are reported along side speaker codes. From visual inspection, velocity profiles of Onset Darkening speakers show a much higher magnitude of velocity from peak one the velocity minumum compared to Onset Lightening speakers. Velocity magnitude does not appear to correlate with estimated VT length across dialects, for example, Onset Darkening speaker, L02 has relatively low vocal tract lenghth estimate of 14.1 cm, but has a high TBz velocity, while Onset Lightening speaker S01 has a larger vocal tract length estimate of 16.2 cm, but has a low TBz velocity.

Figure 6.13 visualises the differences in the velocity of peak 1 between dialects more clearly. Here, the velocity of peak 1 is shown for each dialect and context (singleton and cluster). A clear dialectal difference can be observed in Figure 6.13,

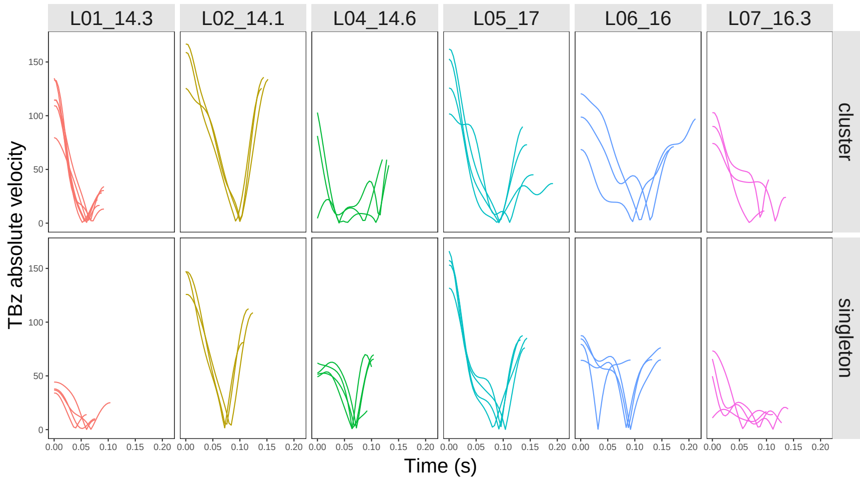

Figure 6.11: Absolute velocity of TB lowering gesture in *plick* (top) and *lick* (bottom) for Onset Darkening speakers.

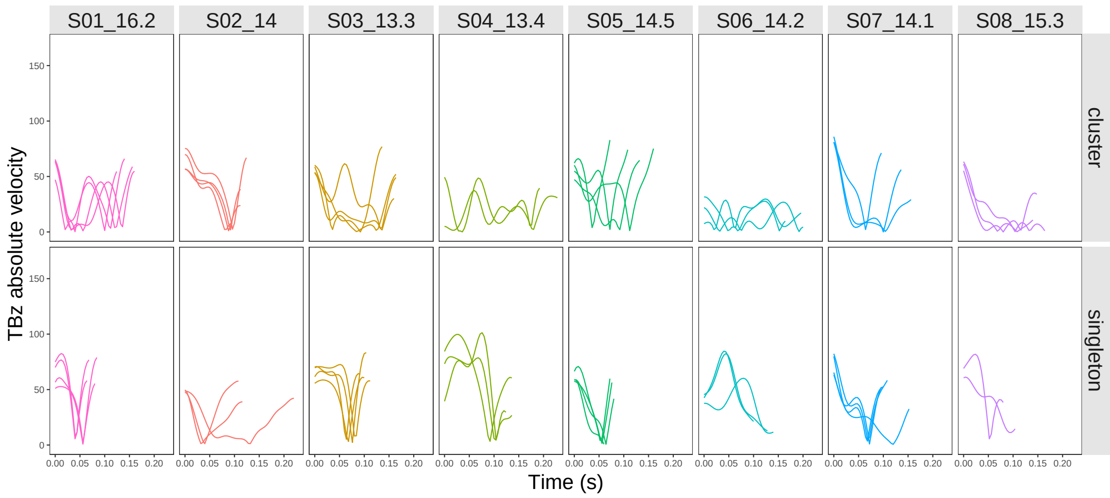

Figure 6.12: Absolute velocity of TB lowering gesture in *plick* (top) and *lick* (bottom) for Onset Lightening speakers.

with speakers of the Onset Darkening Dialect showing a higher velocity value of peak one than speakers of the Onset Lightening Dialect. To statistically verify an effect of dialect on peak velocity, a model comparison was performed whereby a full model was compared to a partial model where the fixed effect of dialect had been removed. The full model contained fixed effects of consonant structure (*plick* vs *lick*), dialect and estimated vocal tract length, and a random intercept of speaker, and a random slope of speaker for speaker for the effect of consonant structure. This was compared to a partial model which was identical, except for the exclusion of the dialect term. Results of the comparison was significant at $p < 0.01$ (see Table 6.5), suggesting a significant effect of dialect on the peak absolute velocity of the tongue body lowering gesture in *plick* and *lick*.

Given that estimated vocal tract lengths are unbalanced across dialects, with the Onset Darkening dialect containing more speakers with longer estimated vocal tract lengths, a further test for collinearity between dialect and vocal tract length was performed. The variance inflation factors function of the *car* R package was used to test for collinearity between dialect and estimated vocal tract length on the full model. Results did not find high collinearity between dialect and vocal tract length (VT length - 1.238; Dialect - 1.238; Consonant structure - 1.0).

These findings suggest that Onset Darkening speakers, are indeed producing a tongue body lowering gesture of /l/ which is greater in displacement than that of Onset Lightening speakers speakers, but at a higher velocity than Onset Lightening speakers, therefore explaining how a stable timing pattern can be maintained across dialects.

Effect Tested	χ^2	df	p-value
dialect	9.868	1	$<$**0.0017**

Table 6.5: Model comparison for the effect of dialect on the absolute velocity of Peak 1 of the TB lowering gesture of /l/.

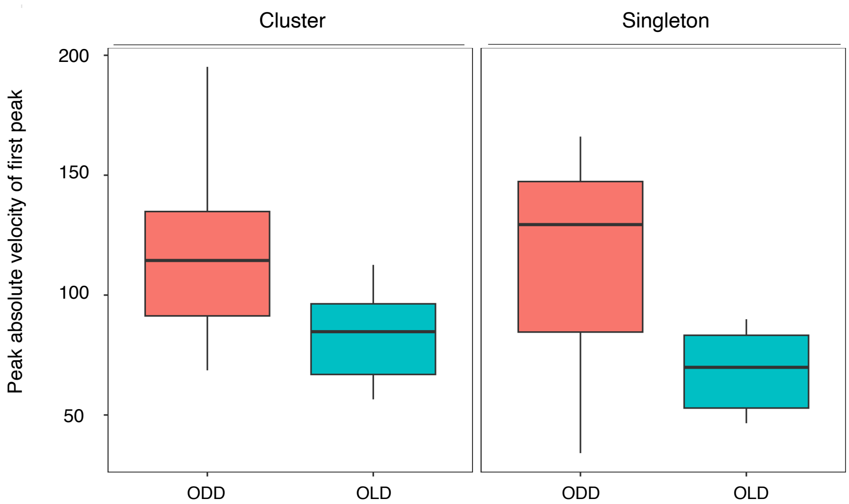

Figure 6.13: Absolute velocity of Peak 1 of the TB lowering gesture of /l/ within the cluster *plick* context (left), and the singleton *lick* context (right).

6.5.6 Speaker comparison of velocity and timing

The previous sections have failed to find an effect of dialect on patterns of lateral timing, but have found a significant difference in the velocity of the tongue body lowering gesture between dialects. In order to check that differences in velocity do not yield different timing patterns at the speaker level, I here compare the lateral timing patterns of a high velocity Onset Darkening speaker (L02), and a low velocity Onset Lightening speaker (S01). Speakers L02 and S01 were selected for comparison due the clear differences in the velocity of the tongue body lowering gesture between these speakers, as can be seen in Figure 6.14.

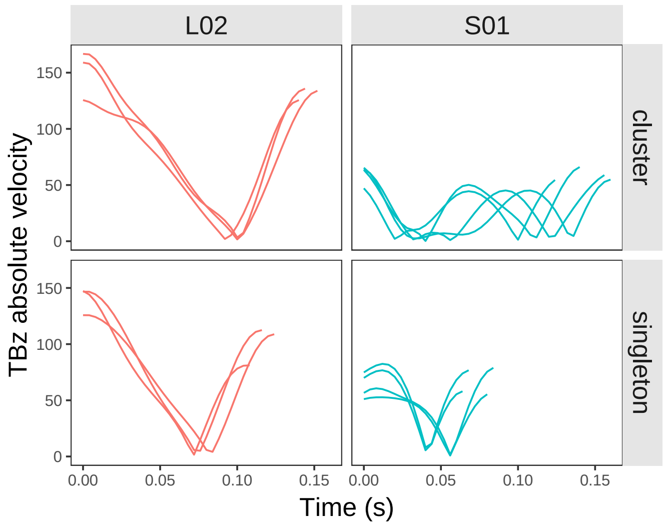

Figure 6.14: Absolute velocity the TB lowering gesture of /l/ within the cluster *plick* context (top), and the singleton *lick* context (bottom) for speaker for L02 and S01.

Figure 6.15 shows lateral gesture to anchor lags for the tongue tip and tongue body gestures in the singleton *lick* and cluster *plick* context for L01 and S02. Both speakers pattern similarly, showing greater tongue tip to anchor lag durations in both singleton and cluster contexts. Further, the tongue tip gesture temporally precedes the tongue body gesture for both speakers in both contexts. Figure 6.16 shows the duration of C-centre, Left Edge and Right Edge intervals for the *plick / lick* pair for speakers L02 (red) and S01 (blue). Again, there are no clear differences to be observed between speakers. Both speakers show a larger C-centre and Left Edge lag in the cluster context relative to the singleton context, and a relatively stable Right Edge lag across cluster and singleton contexts.

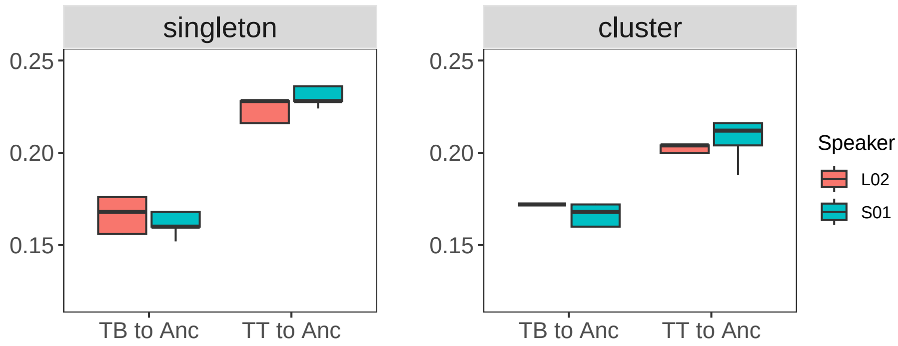

Figure 6.15: TB gesture of /l/ to anchor lags and TT gesture of /l/ to anchor lags for *plick*(left), and *lick*(right) for speakers S01 and L02.

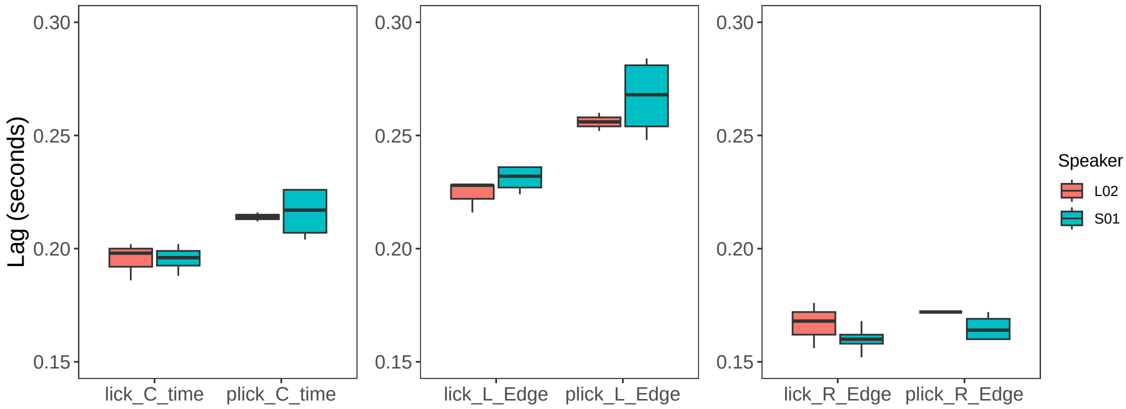

Figure 6.16: Comparison of stability intervals across the singleton - cluster pair *plick / lick* for S01 and L02. The left panel shows C-centre to anchor durations for *lick* and *plick*, the middle panel shows Left Edge to anchor durations for *lick* and *plick*, and the right panel shows Right Edge to anchor durations for *lick* and *plick*.

The individual speakers considered here, a high velocity Onset Darkening speaker L02, and a low velocity Onset Lightening speaker S01, show no clear differences in patterns of inter gestural timing in /l/, or in patterns of /l/ cluster timing across the *plick / lick* pair. These findings echo the findings of previous sections for a lack of dialect difference in /l/ timing across dialects. As such, this small scale speaker comparison lends further support for the argument that dialects, which differ in the magnitude of tongue body lowering in /l/, mediate velocity in order to maintain a stable lateral timing pattern.

6.6 Discussion

6.6.1 Main findings

This chapter sought to resolve methodological challenges presented by the previous chapters, which left open questions regarding the role of the tongue body gesture in patterns of temporal coordination of /l/. The main findings of this analysis were firstly, that laterals are temporally different in singleton versus cluster onsets. This was observed through differences in timing patterns at the inter-gestural level, and consonant cluster level. The second important finding of this analysis was that there was no effect of dialect on patterns of inter-gestural or lateral cluster timing, despite differences in the magnitude of the tongue body lowering gesture of /l/ between dialects. Subsequent analysis revealed a significant effect of dialect on the absolute velocity of the tongue body lowering gesture of /l/, with Onset Darkening speakers exhibiting a greater TBz velocity than Onset Lightening speakers. These findings suggest that temporal stability in /l/ is maintained across dialects through mediation of the velocity of the tongue body gesture. The implications of the results of this chapter are discussed below.

The lack on an effect of dialect at the inter-gestural level was surprising, given the literature reporting a relationship between gestural timing and /l/ darkening (Scobbie and Pouplier, 2010; Sproat and Fujimura, 1993). It is also interesting to consider that there was a greater effect of context on inter-gestural timing than there was of dialect. Considerably different patterns of gestural timing were observed for /l/ in the singleton *lick* context, compared to the cluster *plick* context. Further, stability analyses showed that no interval (C-centre, Left Edge or Right Edge) was stable across singleton and cluster contexts. Along with challenging the predictions of a coupled oscillator model for a stable C-centre (Browman and Goldstein, 1988), these results serve to reaffirm the conclusion that, for the speakers within this study, onset /l/s are temporally different in singleton versus cluster contexts. Since this temporal difference occurs across both dialects, it appears that temporal differences across singleton and cluster contexts in /l/ occur regardless of the darkness of /l/.

In discussing these results, a caveat should be noted. The timing analyses presented in this chapter and Chapter 4 relied upon identifying the time of the velocity minimum of relevant gestures within the singleton/cluster onsets as well as the anchor consonant. Lag durations were then calculated as the distance between these time points. However, in some cases, variation in velocity trajectories made it challenging to identify the relevant velocity minimum, particularly in cases where the velocity magnitude was small but variable. Examples of variation in velocity profiles can be seen in Figures 6.11 and 6.12 which show the velocity trajectories of the tongue body lowering gesture of /l/ in *plick / lick*. These figures show that while the velocity minimum is clearly identifiable in some cases, it is ambiguous in others. It is at least possible then that the challenges of reliably identifying the velocity minimum may be the reason why no stable timing pattern was found across the singleton - cluster context here.

6.6.2 The relation between time and space in stop-lateral clusters

This analysis has shown that temporal stability is maintained in stop-lateral clusters, despite spatial variability in the magnitude of the tongue body lowering gesture of /l/ between dialects. Such temporal stability is achieved through mediation of velocity; in this way, spatial differences in the degree of tongue body lowering between dialects is compensated for by differences in velocity. This suggests that, across dialects, speakers are privileging a particular timing relationship, which is invariant to spatial differences. This finding has implications for a coupled-oscillator model. Consistent with a feed-forward model, these findings suggest no obvious interaction between space and time (e.g., Iskarous and Pouplier, 2022), and suggest the coupling relations which mediate timing to be invariant to spatial differences. However, the finding of temporal stability was somewhat unexpected in light of previous studies which have reported a relationship between spatial and temporal variation, for example Shaw and Chen (2019)'s finding for an effect of articulator position on gestural timing.

6.7 Conclusion

This chapter has considered the role of the tongue body lowering gesture of /l/ in developing our understanding of the temporal behaviour of /l/ across speakers of an Onset Darkening Dialect, and an Onset Lightening Dialect. Results have shown that, through examination of the timing of both the apical and dorsal gestures of /l/, new insights can be gained into the variable nature of gestural relationships, particularly between onset singleton and cluster contexts. While the timing of the tongue body gesture did not reveal the expected differences in lateral cluster timing between Onset Lightening and Onset Darkening dialects, I have shown that systematic differences in the velocity of the tongue body lowering gesture between dialects can explain this pattern. This finding has theoretical implications for a coupled-oscillator model, largely supporting a feed-forward model.

Chapter 7

Discussion

7.1 Aims and contributions

The overarching goal of this thesis was to explore the relationship between spatial and temporal variation in consonant cluster production, using the English lateral as a test case. English laterals were selected as the focus of investigation as they show systematic dialectal variation in darkening, which has previously been described in terms of both spatial and temporal measures (e.g., Turton, 2014). In this way, the relationship between spatial and temporal variation in lateral darkening could be explored within a language. This novel experimental design enabled spatio-temporal variation in /l/ to be investigated systematically in a way that avoids cross linguistic confounds and taps into a more nuanced level of variation. This is particularly relevant for Articulatory Phonology, which does not typically investigate between-speaker and between-group variation within a language. By examining dialect-level variation, this thesis develops a perspective in which articulatory parameters are shared by speakers of a dialect, thus enriching understandings of how articulatory models may be applied to group-level variation.

The specific goal of this thesis was to determine the effect of /l/ darkening on patterns of /l/ cluster timing. The motivation for this derived from the hypothesised relationship between the darkness of /l/ and coda cluster timing patterns across languages (Marin and Pouplier, 2014). I here expanded this to investigate how differences in onset darkening in /l/ between two dialects of the same language affected patterns of lateral onset cluster timing.

7.2 Summary of findings

This thesis comprised three studies, each playing a role in investigating the relationship between spatial and temporal variation in /l/ darkening. A summary of each study is provided below.

7.2.1 Dialect variation in /l/ darkening

The first study sought to capture the ways in which Onset Lightening and Onset Darkening dialects differ articulatorily in /l/ darkening, and the contexts that condition variation. Establishing a dialectal difference in /l/ darkening had a broader relevance to the subsequent aims of the thesis, which investigated the ways in which such differences in /l/ darkening affected patterns of cluster timing.

The study intentionally elicited a wide range of variation in /l/ through looking at /l/ within three vowel contexts (FLEECE, KIT, THOUGHT), and four morphosyntactic contexts (word initial, mono-morphemic intervocalic, pre-boundary intervocalic, word final pre-vocalic, and word final pre-consonantal). This yielded insights into the ways in which dialectal variation in /l/ darkening interacted with vocalic context and morphosyntactic context. Clear effects of dialect were found in minimum F2−F1 for front vowel contexts in all but the word final pre consonantal morphosyntactic context. Articulatorily, /l/ darkening was captured through vertical tongue body displacement using both static (TBz minimum) and dynamic (fPCA of vowel + /l/ displacement) measures. Tongue body displacement revealed dialectal differences in both front vowel contexts, and back vowel contexts, which varied with morphosyntactic position. This chapter thus established articulatory differences in /l/ darkening between dialects. In so doing, the chapter revealed interactions between dialect and vocalic context, and dialect and morphosyntactic context. Such interactions highlight the sensitivity of /l/ to contextual variation, and the importance of studying /l/ across a wide range of contexts.

7.2.2 Variation in /l/ cluster timing

The second study built upon the established dialectal difference in /l/ darkening. Motivated by a hypothesised relationship between /l/ darkening and cluster timing (Marin and Pouplier, 2014), this study investigated how differences in /l/ darkening between dialects affected patterns of /l/ cluster timing. Consonant cluster timing measures were taken for /l/ in singleton and cluster onsets. Given the difficulty in identifying the time of the tongue body gesture (found in the previous study), the time of /l/ was defined as the time of achievement of the tongue tip raising gesture of /l/. The analysis focussed primarily on onset clusters due to a substantial number of coda tokens within the Onset Lightening Dialect being vocalised, hence the tongue tip raising gesture could not be identified for these tokens.

Stability measures determined the most stable interval across matched cluster/singleton word pairs and results were compared across dialects. Indirect measures were suggestive of a C-centre pattern (the /l/ to anchor lag was shorter in the cluster word compared to the singleton word.) However, the C-centre was not the most stable interval when compared with other measures of cluster timing. Interestingly, there was no effect of dialect – and therefore no effect of /l/ darkening – on /l/ cluster timing.

7.2.3 Dialect variation in the TB gesture of /l/

Bringing together aspects of the prior two studies, the final study of this thesis investigated the relationship between the timing and displacement of the tongue body gesture of /l/ across the two dialects. A restricted analysis was performed on the singleton/cluster word pair *lick / plick*, which provided a context whereby the tongue body gesture of /l/ could be identified. Timing measures were performed and included measures of inter-gestural timing (between the tongue tip and tongue body gesture of /l/), and stability measures of various intervals (Left Edge Right Edge and C-centre) over the singleton/cluster pair. Unlike the previous study, and the vast majority of /l/ cluster timing studies, timing measures here considered both the tongue tip and tongue body gestures of /l/. Timing differences were observed between singleton and cluster contexts, with gestures of /l/ showing temporal compression in the cluster context compared to the singleton context. However, once again, no dialect differences were observed despite spatial differences in the tongue body lowering gesture between dialects. Subsequent analysis revealed that both temporal stability and spatial variability between dialects could be achieved through differences in gestural velocity between the two dialects. Specifically, Onset Darkening speakers produced a larger but faster tongue body lowering gesture, while Onset Lightening Dialect speakers produced a smaller but slower tongue body gesture. This suggests that Onset Darkening speakers compensate for the larger spatial magnitude by executing the gesture faster, thus maintaining the same temporal relationship in clusters as seen in Onset Lightening speakers.

7.3 Implications for Articulatory Phonology

7.3.1 On the absence of C-centre stability in clusters with laterals

A coupled oscillator model of syllable timing predicts a C-centre timing pattern for onset clusters with a complex syllable organisation, such as English. The pattern is said to arise due to competing phase relationships, whereby consonants in a (C)CCV cluster are coupled in-phase with the vowel, but anti-phase with one another. The result of this is that consonants are equally displaced around the vowel, such that the temporal midpoint of the consonants (relative to a fixed anchor point) remains stable as onset complexity increases (Browman and Goldstein, 2000).

Across matched singleton/cluster word pairs, where /l/ was the singleton onset consonant (e.g., *lug*), or C2 in an onset cluster (e.g., *plug*), conclusive evidence for C-centre stability was not found. Timing analyses measured the stability of three intervals across each word pair: the Left Edge, Right Edge, and C-centre (see Section 5.2.2 for details of the specific measures used). Stable intervals were defined as those that were not significantly different across singleton and cluster contexts, and the findings show that these varied across word pairs. The C-centre and Right Edge intervals were stable for *plug / lug* and *clip / lip* word pairs. For the *club / lug* word pair, the C-centre and Left Edge intervals were stable, and for the *plick / lick* word pair, no intervals were stable. The lack of conclusive evidence for a C-centre organisation found in these data is consistent with findings for non C-centre stability in /l/ onset clusters in English (Goldstein et al., 2009), and German (Brunner et al., 2014; Mücke, Hermes, and Tilsen, 2020). On the other hand, these findings contrast with the observed C-centre stability of lateral onsets in American English (Browman and Goldstein, 1988; Honorof and Browman, 1995; Marin and Pouplier, 2010), hence contributing to a mixed picture of /l/ cluster timing.

One limitation of the word list which should not be overlooked is the unmatched segmental context within the *club / lug* pair, where the anchor constant is /b/ in *club* and /g/ in *lug*. This may explain why the Left Edge interval was stable for this word pair and not for others. In addition, it is also possible that the relative unfamiliarity of *plick* contributed to the results for this word pair, where no interval was stable.

Inter-gestural timing patterns between the tongue tip and tongue body gestures of /l/ were examined within in a restricted analysis of the *plick / lick* word pair. This word pair was specifically selected for this analysis as it offered a segmental context whereby the tongue body gesture of /l/ could be unambiguously identified. Results revealed compression effects, where by the tongue tip and tongue body gestures of /l/ were temporally closer in the cluster context *(plick)*, than the singleton context *(lick)*.

Considering differences in relative gestural timing from an AP perspective, relative differences between the timing of lateral gestures in singleton and cluster contexts may be modelled by adjusting the coupling strength between gestures. For example, Goldstein et al. (2009) modelled findings for non C-centre timing patterns found in /sp/ and /pl/ clusters by adjusting the coupling strength between the consonants and the following vowel. For the /pl/ cluster, they show that adjusting the coupling strength parameter was not sufficient to model patterns in the data, and rather required coupling both gestures of /l/ in-phase with the vowel. Extending this to my data, the relative differences in inter-gestural timing in lateral gestures in the singleton context *(lick)* compared to the cluster context, *(plick)*, could be modelled by varying the coupling strength between lateral gestures in each of the contexts, such that there is a stronger anti-phase coupling between the tongue tip and tongue body gestures of /l/ in the singleton context, compared to the cluster context. A second possibility could be that both gestures of /l/ are coupled in-phase with the following vowel in the cluster context, but only one gesture of /l/ is coupled to the vowel in the singleton context, with the multiple in-phase couplings of /l/ resulting in a gestural compression within the cluster context. However, this latter possibility seems less plausible, for it requires a more fundamental restructuring of coupling relations which are presumed phonologically stable within the AP model.

7.3.2 Problematising the C-centre metric

While C-centre stability has been successfully found in many cluster timing studies (see Section 2.2.3), it is not entirely surprising that a robust pattern of C-centre stability was not found here. C-centre stability has been widely problematised as a metric for identifying syllable structure (Mücke, Hermes, and Tilsen, 2020; Shaw et al., 2011; Sotiropoulou and Gafos, 2022). Shaw et al. (2011) refer to the use of C-centre stability and Right Edge stability to diagnose syllable structure as the "static invariance" view. Using simulations based on their Moroccan Arabic data, Shaw et al. (2011) show how factors such as consonant compression can affect the metric. For example, if a cluster context has a compression effect, and the duration of the C2 in an onset cluster is reduced relative to when the consonant occurs in a singleton context, this may result in Right Edge stability, and a misdiagnosis of a simplex syllable structure. To resolve this issue, and thus prevent such misdiagnoses, Shaw et al. (2011) propose an alternative approach, the "dynamic invariance" approach (Shaw et al., 2011). The dynamic invariance approach to syllable structure instead focuses on relations between parameter values across a range, meaning that invariance resides in the relationship between parameters rather than static structures. A similar sentiment is echoed by Sotiropoulou and Gafos (2022) who list "duration of the pre-vocalic consonant [...], the voicing of the initial stop, vowel initiation in relation to the cluster, and compensatory effects between IPI [- temporal interval between consonants -] and the duration of the pre-vocalic consonant" (Sotiropoulou and Gafos, 2022, p. 39) as phonetic parameters which should be considered within such an approach.

While it was not the primary goal of this thesis to test for syllable structure, findings here largely suggest the static invariance view to be problematic. In particular, findings for the stability of intervals to vary with segmental context lend support to the view that syllable structure cannot be reduced to a single metric. Findings for gestural compression in /l/ in *plick* compared to *lick* further expose how the C-centre metric can be influenced by conflating phonetic parameters.

However, we should also here consider the possibility that a stable C-centre timing pattern was not found within this study because gesture targets (velocity minima), rather than onsets were measured. Measuring the gestural target rather than onset may also have implications the relative timing between the tongue tip and tongue body gestures of /l/. The question of which aspects of the gesture to measure in a temporal analysis comes down to a question of which aspect of a gesture is under coordinative control. While the coupled oscillator model considers the gesture onset to be under coordinative control (Browman and Goldstein, 1988), others has since suggested that that this may not be the case (e.g., Shaw and Chen, 2019; Turk and Shattuck-Hufnagel, 2020). Within this study, the decision to derive measures of consonant cluster timing and inter-gestural timing from gestural targets (i.e., the velocity minimum of the relevant gestural dimension) was ultimately made on the methodological grounds. Given the variability observed in gestural velocity across the data, it was considered more methodologically robust to identify the velocity minimum than it was to identify the movement onset, which is typically measured as 20% of the relevant velocity peak (e.g., Marin and Pouplier, 2014; Pouplier et al., 2022), hence highly sensitive to variability in velocity magnitude. However, the point remains valid, that different timing patterns may have emerged had a different time point within the gesture been measured.

A final note on the use of the C-centre metric as a measure of syllable structure is the possibility for an imperfect translation between planning and production.

The coupled oscillator model is a model of speech planning (Nam, Goldstein, and Saltzman, 2009), thus it is possible that the phasing relations between planning oscillators are not captured by temporal measures of speech production. One issue, discussed in detail in Simko and Cummins (2010), is that all articulators within the AP Task Dynamic model are assigned the same abstract mass; however, mass is important for rate of movement. In other words, the same timing relations may look different between two syllables if the involved articulators differ in their masses and/or stiffness, despite the fact that the identical timing relations could have been specified in planning. Evaluating this possibility requires computational modelling, but even then it is difficult to derive accurate estimates of mass differences between articulators (Simko and Cummins, 2010).

7.3.3 The role of space and time in AP

The central finding of this thesis is that spatial differences in the magnitude of the tongue body gesture of /l/ do not result in temporal differences in /l/ cluster timing. Further analysis of the *plick / lick* pair, where the tongue body gesture of /l/ could be easily identified, revealed such temporal stability to result from compensatory differences in velocity between dialects. A tongue body lowering gesture of /l/ with a greater magnitude was produced by speakers of the Onset Darkening dialect, but at a higher velocity. That is, the Onset Darkening dialect produced a larger, faster tongue body lowering gesture, while the Onset Lightening dialect produced a slower tongue body lowering gesture of a smaller magnitude. In this way, global timing was found to be invariant to spatial differences.

These findings suggest that laterals within the two dialects have identical coupling relations between gestures, but have different constriction degree and stiffness values for the tongue body lowering gesture. Stiffness mediates the speed at which a gestural target is achieved; for example, vowels and consonants differ in stiffness; vowels are longer and slower than consonants and so have a lower stiffness value (Saltzman and Munhall, 1989, p. 66). Speakers of the Onset Darkening dialect who exhibit a larger tongue body lowering gesture of /l/ at a higher velocity would hence be assigned a relatively larger stiffness value than speakers of the Onset Lightening dialect. Previous work has explored the possibility of language-specific gestural parameters (Iskarous, McDonough, and Whalen, 2012), so it is not a stretch to extend this possibility to dialect-specific gestural parameters.

The finding that /l/ cluster timing is stable, despite differences in the magnitude of the tongue body lowering gesture of /l/, has a number of implications for how spatio-temporal variation is handled within an AP model. I here draw upon Iskarous and Pouplier (2022) who provide a detailed review of this relationship within an AP framework. The coupled oscillator model predicts no interaction between intrinsic properties of a gesture and the coupling relations between associated planning oscillators (Iskarous and Pouplier, 2022). This is compatible with the findings here; no differences in timing were found across spatially variable lateral gestures. The implication for these data is that there is no interaction between the coupling relations that mediate timing and the intrinsic spatial properties of the tongue body gesture of /l/.

However, this contrasts with findings which *do* suggest a relationship between gestural timing and intrinsic gestural properties. For example, examining onset timing patterns in Polish, Pastätter and Pouplier (2017) found a correlation between the degree of coarticulatory resistance of a vowel-adjacent-consonant and the amount of onset-vowel overlap in a CCV cluster. A consonant is considered more coarticulation resistant to the effects of adjacent segments if it recruits greater tongue body movement (Recasens and Espinosa, 2009). They found that the more coarticulation resistant the vowel-adjacent-consonant, the less overlap was observed between the vowel-adjacent-consonant and the vowel. Pastätter and Pouplier (2017) suggested that their results could be modelled by adjusting coupling strength in line with the coarticulation resistance of the consonant. Given that tongue body displacement is such a crucial determinant of coarticulation resistance, it is plausible that onset laterals within this study differ in coarticulation resistance between dialects. Against the backdrop of Pastätter and Pouplier (2017)'s results, the findings of stable patterns of lateral cluster timing between dialects is thus surprising. One possibility for this is that dialects examined here are not different enough in darkening, an idea I discuss in more detail in Section 8.1.

Shaw and Chen (2019) also reported an interaction between spatial properties and timing; they found that the position of the tongue at the beginning of the vowel in a CV sequence mediated the initiation time of the vowel gesture in Mandarin. Specifically, they found the vowel gesture to be initiated earlier when the tongue position at the start of the vowel was further away from the vowel target, and the vowel gesture to be initiated later when the tongue was closer to the vowel target. Shaw and Chen (2019)'s findings conflict with the predictions of a coupled oscillator model which predict no interaction between spatial properties of gestures and the coupling relations mediating the time of gesture initiation. However, if we focus, not on the time of gesture initiation, but on global temporal stability, a parallel may be drawn between Shaw and Chen (2019)'s findings, and those of findings for temporal stability here. All other things being equal, it takes longer to achieve a spatial target which is further away than one which is closer. We may then consider the relationship reported by Shaw and Chen (2019) between the distance of the tongue position from the vowel target, and the time of vowel initiation, to be a compensatory mechanism to preserve global stability. When there is greater spatial distance between the tongue position at the start of the vowel and the vowel target, the vowel gesture is initiated earlier to allow adequate time for the gesture to be achieved – a possible mechanism for maintaining temporal stability on a global level. To test this hypothesis would of course, require global temporal stability to be explicitly measures in these data.

Considered together, studies which have investigated the relationship between spatial properties of gestures and their patterns of timing have reported mixed results. One the one hand, studies have found a relationship between gestural timing and intrinsic spatial properties of gestures (Pastätter and Pouplier, 2017; Shaw and Chen, 2019), while on the other hand, findings of this study report no such relationship. The question remains of how to make sense of these collectively conflicting results. One possibility is that these low-level patterns of temporal variability are part of a synergetic relationship, the collective result of which is global temporal stability. One reason why global temporal stability may be so important to preserve, could be due to the fact that all studies cited here look exclusively at onset position. Consistent with the general finding for codas to be temporally more variable than onsets (Honorof and Browman, 1995; Marin and Pouplier, 2010), onsets are considered highly sensitive to temporal variation. Demands of perceptual recoverability require that onset gestures must be coordinated within precise temporal constraints. Further, it is possible that the spatio-temporal relationships of non-onset gestures pattern very differently to those here, and that the patterns observed here are restricted to onset contexts only. To explore this hypothesis empirically would require comparison of the spatio-temporal relations between a controlled set of gestures in onset and non-onset position.

7.4 Models of lateral darkening

7.4.1 Variation in /l/ darkening

The model of lateral darkening assumed throughout this thesis is based on Sproat and Fujimura (1993), who propose that /l/ comprises a consonantal tongue tip raising gesture and a vowel like dorsal raising and retraction gesture. For dark coda /l/, the tongue dorsum gesture (captured by tongue body lowering) occurs before, or simultaneously with the tongue tip gesture. In clear onset /l/, the tongue tip gesture precedes the tongue dorsum gesture. Patterns of inter-gestural timing are explained through rules of attraction: the vocalic gesture is attracted to the syllable nucleus and the consonantal gesture of /l/ is attracted to the syllable boundaries (as described in Section 4.1.3). In addition, differences in /l/ darkening are also said to manifest spatially, with dark /l/s exhibiting greater tongue body lowering and greater tongue dorsum raising and retraction than clear /l/s. Unlike other models of /l/ darkening, Sproat and Fujimura (1993) do not consider there to be an allophonic distinction between clear and dark /l/. Darkness of /l/ is instead considered to be a gradient phonetic property mediated by context (i.e. prosodic boundary strength), and duration.

In the present study, /l/ darkening was found to vary with interactions between dialect and vowel and dialect and morphosyntactic context. Such interactions have also been reported in previous studies (Strycharczuk, Derrick, and Shaw, 2020; Strycharczuk and Scobbie, 2015). Greater variation in /l/ darkening with morphosyntactic context was found within the high front vowel context relative to the THOUGHT context, and greater dialectal variation in /l/ darkening was found in non word final, front vowel contexts.

The results here support the idea that /l/ can be affected by processes of both lightening and darkening (Mackenzie et al., 2018; Strycharczuk, Derrick, and Shaw, 2020). When /l/ is preceded by FLEECE in a non word-final context, coarticulatory effects of the vowel result in lightening. Lightening has the effect of increasing the upper limit on the light to dark scale, and provides greater scope for mid range variation to be detected (Strycharczuk, Derrick, and Shaw, 2020). Darkening effects can be observed in the word final pre consonantal context, and when /l/ is preceded by THOUGHT; in such contexts, /l/ is dark, with relatively little variation. Such interactions are relevant the question of whether /l/ darkening is gradient or categorical. Greater morphosyntactic variation is more likely to be observed in the FLEECE context than a back vowel context.

Vocalisation was observed throughout the studies of this thesis. In Chapter 4, instances of partial and full vocalisation were observed in both Onset Lightening and Onset Darkening speakers to a similar extent; however, greater variation in the contexts of partial vocalisation was observed for Onset Darkening speakers, suggesting a degree of instability, indicative of an ongoing change for these speakers. In Chapter 5, /l/ vocalisation was observed in over 50% of tokens within the lateral coda cluster contexts of *gulp, milk* and *philp* for Onset Lightening speakers, while in the same contexts, Onset Darkening speakers vocalised in only 10% of cases. While vocalised tokens were variably observed in *gulp, milk* and *philp* contexts for Onset Lightening speakers, there were a higher proportion of vocalised tokens within *gulp, milk* contexts relative to the *philp* context, suggesting an effect of the combination of the preceding vowel and post-lateral consonant (Hardcastle and Barry, 1989).

Taken together, dialect differences in the contexts of vocalisation were observed; Onset Darkening speakers rarely vocalised in lateral coda clusters, while Onset Lightening speakers did so in over half of cases. Both dialects variably vocalised in word final pre consonantal contexts, however, Onset Darkening speakers showed greater variability in the contexts of partial vocalisation, perhaps indicative of an ongoing change. The high degree of intra speaker variation in vocalisation observed across speakers of both dialects, which was greater that expected (e.g., Hardcastle and Barry, 1989) also suggests a high degree of instability. While much is known about vocalisation in Southern British English - the Onset Lightening dialect here - comparatively little is known about vocalisation within Lancashire / Manchester English - the Onset Lightening dialect. These findings thus contribute to our understanding of the interaction between /l/ vocalisation and morphosyntactic / vocalic environment, across varieties where /l/ is articulatorily distinct.

7.4.2 Methodological challenges

In measuring /l/ darkening, several methodological challenges were encountered. Much of these owed to the ambitiously varied segmental contexts of /l/ within the stimuli list, which was designed to elicit a wide range of variation. The first challenge challenge was the difficulty in measuring the tongue body gesture of /l/ independently of the vowel given that the tongue body is recruited for both the vowel and lateral gesture. This was particularly difficult for low/back vowels. A further obstacle was the varying degrees of vocalisation observed for both dialects, in word final /l/ tokens. In such cases, the tongue tip raising gesture of /l/ was either reduced or altogether absent, hence could not be measured.

Capturing variation in /l/ thus came at considerable methodological cost. A decision had to be made on whether to (i) only measure what can be measured across the entire data set, or (ii) subset the data and perform different analyses on different tokens. In other words, a lot could be said about a narrow subset of data, such as a single segmental context, or *something* could be said about all of the data. It was important, at least to the initial analysis of the thesis, that the measure of /l/ darkening used could be applied to /l/ across all contexts. This decision constrained analyses to measures which did not rely upon the time of achievement of the tongue body gesture of /l/, or the timing or magnitude of the tongue tip gesture. Working within these constraints, fPCA was performed on dynamic tongue body lowering across the vowel plus lateral window. This measure proved useful in that it would be used on vocalised and non-vocalised tokens, and did not require single time point measures to be made, which were found to be problematic, with the success of the measure resting on whether or not the 'correct' time point had been privileged.

7.5 Limitations

Limitations of this analysis include, firstly, the possibility that that the lack of temporal differences in /l/ cluster timing between dialects is due to systematic dialectal differences in vowel duration. Measures of /l/ cluster timing in this study measured the duration of the interval beginning at a point in the onset to the post-vocalic anchor target, therefore the post lateral vowel was always included with the interval. It is possible that dialects differ in vowel duration in such a way that compensates for between-dialect timing differences in the lateral. Given the difficulty in precisely measuring vowel or lateral duration when vowels and laterals neighbour one another, this limitation is difficult to test in the current data set. Secondly, it possible that use of gestural targets, compared to, for example, gestural onsets, within measures of cluster stability in this analysis were insufficient to capture dialectal differences in lateral cluster timing.

Chapter 8

Conclusions

This thesis sought to investigate the nature of the relationship between spatial and temporal variation in speech planning and production, through the test case of the English lateral. Empirical findings showed differences in lateral darkening between Onset Lightening and Onset Darkening dialects to manifest spatially, in the magnitude of the tongue body lowering gesture. Further, despite being spatially variability, laterals showed global temporal stability across dialects. These findings have theoretical importance for the relationship between spatial and temporal variability in AP, largely supporting supporting a feed-forward model, where coupling relations are invariant to intrinsic gestural properties. Findings also offered valuable insights into patterns of micro-variation in /l/ darkening across two closely related varieties of British English.

8.1 Future Research

In the final study of the thesis, inter-gestural timing relations were examined for /l/ within *plick* – *lick* tokens. Effects of context (singleton / cluster) were observed, but effects of dialect were not. Given the known effects of increasing boundary strength on /l/ darkening (e.g., Sproat and Fujimura, 1993; Turton, 2014), it would be interesting to explore the relationship between morpho-syntactic context and dialectal patterns of inter-gestural timing to further enrich understandings of the categorical or gradient nature of /l/ darkening. Given the methodological challenges encountered in this study, an analysis of inter-gestural timing of /l/ in multiple contexts would, of course, have to work within the constraints of what was measurable. For example, to measure the timing of tongue body gesture, a high front vowel context is required, while measures of the tongue tip gesture must be restricted to non vocalised contexts.

Another avenue to explore is the possibility that spatial differences in the tongue body gesture of /l/ did not not affect patterns of /l/ timing in the data here because dialects were not spatially different enough for the differences to have temporal effects. It is quite possible that the dialectal comparison performed here captures a fuzzy middle ground. Perhaps a more extreme Onset Lightening Dialect, such as British Asian English (Kirkham, 2017) compared to Manchester or Lancashire English would expose different results. A further way in which the scope of the study could be broadened could be to include non-lateral consonant clusters (e.g., Marin and Pouplier, 2010; Pouplier et al., 2022). This would allow an explicit comparison to be made between lateral and non-lateral clusters within and between dialects. Through this, we could better untangle the lateral-specific temporal patterns from those which are consistent across segmental contexts.

Finally, a natural next step for this study is to explore how Task Dynamic models of the gestural coordination patterns of /l/ may offer insights into gestural coordination patterns of otherwise difficult to obtain lateral data. This has the potential to make predictions for how both gestures of /l/ are temporally coordinated for /l/ within all vowel and morpho-syntactic contexts. In addition, Task Dynamic models may also help us to understand how variation in parameters may model variation in these data. For instance, we may ask whether stiffness can capture the differences in velocities between dialects, and whether variation in coupling strengths can model the differences in the inter-gestural timing relations of /l/ between cluster and singleton contexts. Further, phonetic applications of Dynamic Field Theory, which have been used to model *distributions* of gestural targets (e.g., Kirkham and Strycharczuk, 2024; Tilsen, 2019), offer a promising means through which variability in lateral gestures may be modelled at the dialect and speaker level.

Chapter 9

Appendices

9.1 Appendix 1

Below shows a section of the participant consent form provided to participants prior to the experiment, as referred to in Chapter 3.

Participant Consent Form

Title of Research Project: *Acoustics and articulation of liquids in British English dialects*

Name of Researcher: Emily Gorman

Participant Identification Number for this project: **Please initial box**

1. I confirm that I have read and understand the information sheet explaining the above research project and I have had the opportunity to ask questions about the project. ☐

2. In order to confirm that you do not have a latex allergy, please respond 'yes' or 'no' to the following questions:

a) Do you have a known allergy to latex? ☐

b) Have you ever worn latex gloves? ☐

c) Have you ever been in contact with other latex items? (e.g. latex balloon, condom) ☐

d) Did you have any unusual or allergic reactions within 30 minutes of touching latex? ☐

e) Do you have any food allergies? If 'yes' would you consider your allergy to be severe? ☐

3. I understand that my participation is voluntary and that I am free to withdraw at any time without giving any reason and without there being any negative consequences. In addition, should I not wish to answer any particular question or questions, I am free to decline. ☐

4. I confirm that I do not have a cardiac pacemaker ☐

Figure 9.1: Extract of the participant consent form used to that ensure participants did not have a latex allergy.

9.2 Appendix 3

Below provides details of which sensors were used for each speaker to identify to tongue body lowering gesture of /l/ in *plick / lick* – from Section 6.4.2.

To identify the tongue body lowering gesture for /l/ in *plick*, the tongue body sensor was used for 5 Onset Darkening speakers (L01 L02, L05, L06, L07), and 1 Onset Lightening speaker (S02), while the tongue dorsum sensor was used for 1 Onset Darkening speaker (L04) and 7 Onset Lightening speakers (S08, S07, S06, S05, S04, S03, S01). To identify the lowering gesture for /l/ in *lick*, the tongue body sensor was used for 5 Onset Darkening speakers (L01, L02, L04, L05, L06), and 1 Onset Lightening speaker (S02), while the tongue dorsum sensor was used for 2 Onset Darkening speakers (L04, L07) and 7 Onset Lightening speakers (S01, S08, S07, S06, S05, S04, S03).

9.3 Appendix 4

Below displays the tongue tip vertical displacement trajectories for each speaker across FLEECE, KIT and THOUGHT vowel contexts, and word initial, mono-morphemic intervocalic, pre-boundary intervocalic, word final pre-vocalic, and word final pre consonantal contexts.

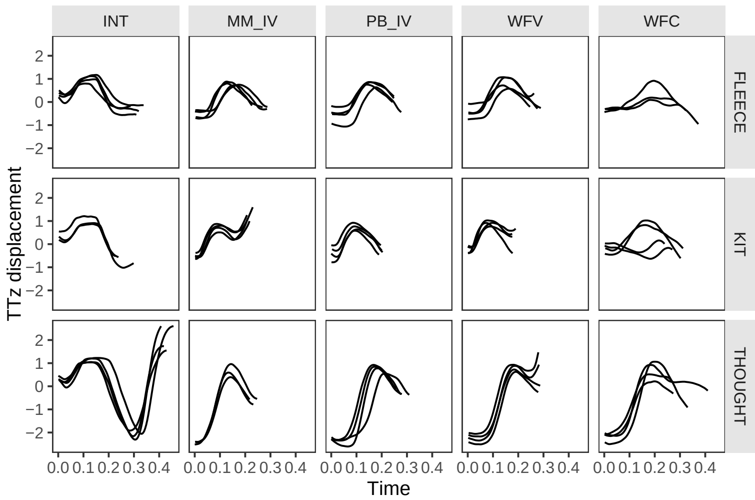

Figure 9.2: Z scored vertical tongue tip displacement for speaker L01. INT = initial; MM IV = mono-morphemic intervocalic; PB IV = pre-boundary intervocalic; WFV = word final pre vocalic; WFC = word final pre consonantal.

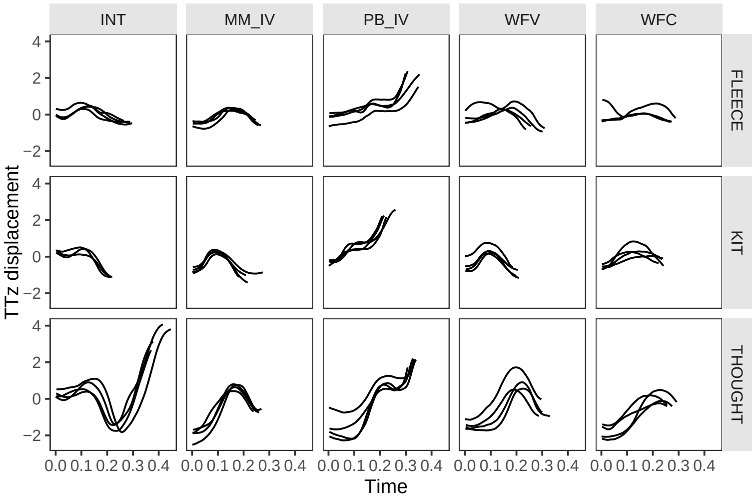

Figure 9.3: Z scored vertical tongue tip displacement for speaker L02. INT = initial; MM IV = mono-morphemic intervocalic; PB IV = pre-boundary intervocalic; WFV = word final pre vocalic; WFC = word final pre consonantal.

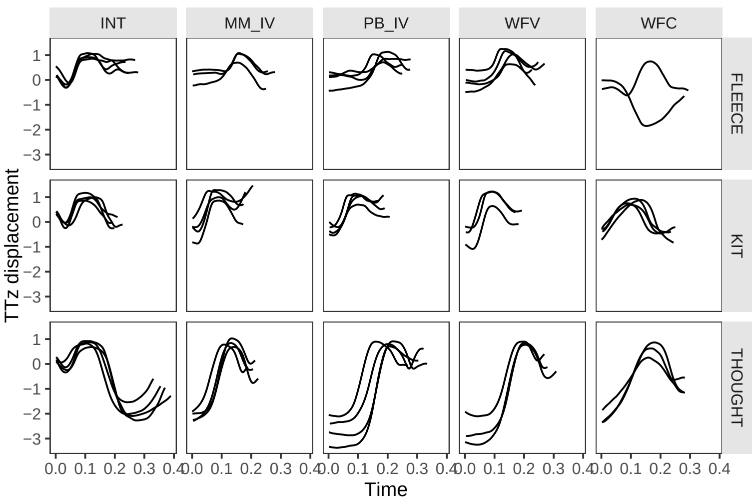

Figure 9.4: Z scored vertical tongue tip displacement for speaker L03. INT = initial; MM IV = mono-morphemic intervocalic; PB IV = pre-boundary intervocalic; WFV = word final pre vocalic; WFC = word final pre consonantal.

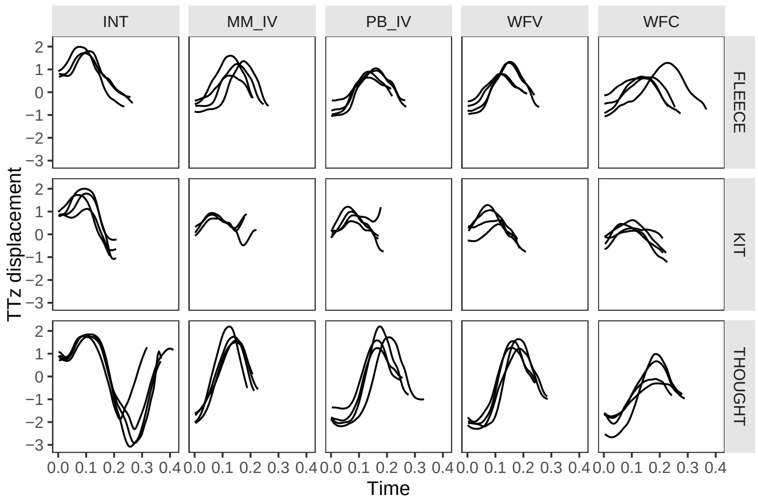

Figure 9.5: Z scored vertical tongue tip displacement for speaker L04. INT = initial; MM IV = mono-morphemic intervocalic; PB IV = pre-boundary intervocalic; WFV = word final pre vocalic; WFC = word final pre consonantal.

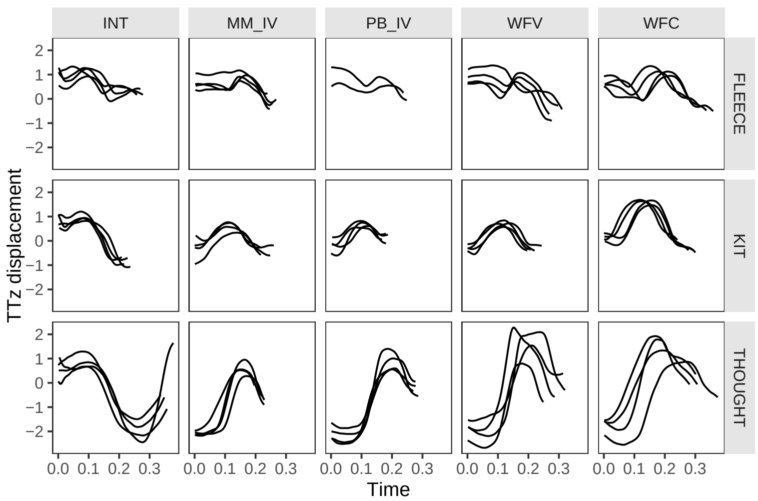

Figure 9.6: Z scored vertical tongue tip displacement for speaker L05. INT = initial; MM IV = mono-morphemic intervocalic; PB IV = pre-boundary intervocalic; WFV = word final pre vocalic; WFC = word final pre consonantal.

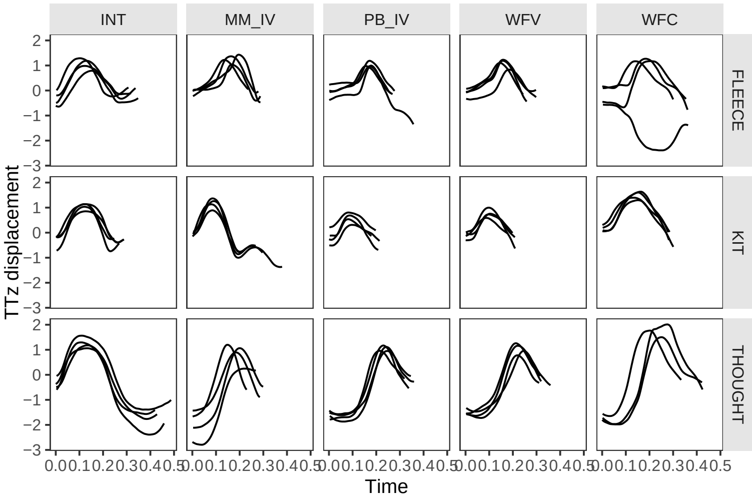

Figure 9.7: Z scored vertical tongue tip displacement for speaker L06. INT = initial; MM IV = mono-morphemic intervocalic; PB IV = pre-boundary intervocalic; WFV = word final pre vocalic; WFC = word final pre consonantal.

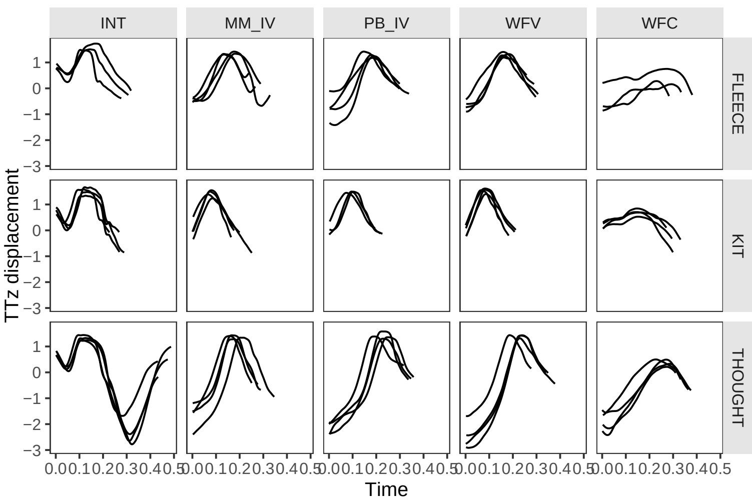

Figure 9.8: Z scored vertical tongue tip displacement for speaker L07. INT = initial; MM IV = mono-morphemic intervocalic; PB IV = pre-boundary intervocalic; WFV = word final pre vocalic; WFC = word final pre consonantal.

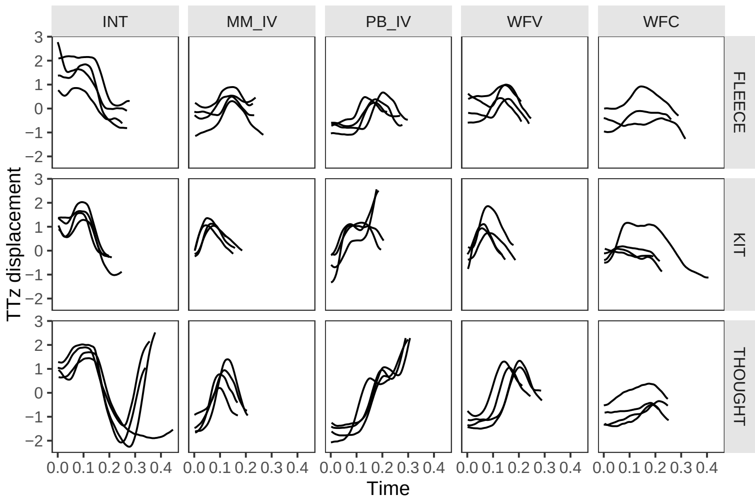

Figure 9.9: Z scored vertical tongue tip displacement for speaker L08. INT = initial; MM IV = mono-morphemic intervocalic; PB IV = pre-boundary intervocalic; WFV = word final pre vocalic; WFC = word final pre consonantal.

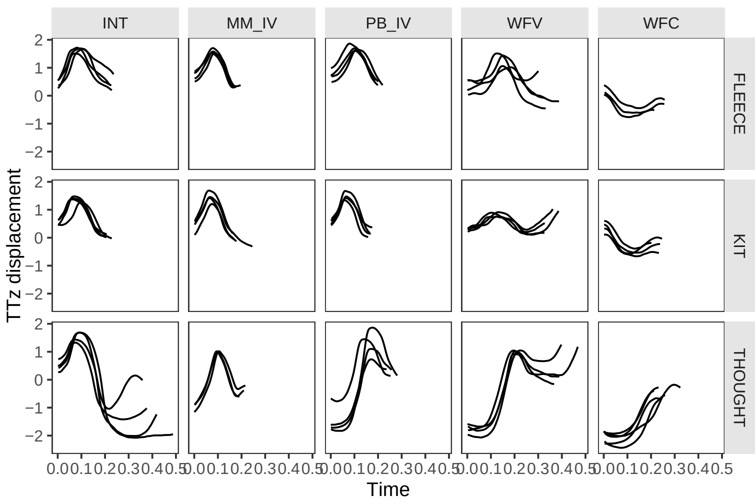

Figure 9.10: Z scored vertical tongue tip displacement for speaker S01. INT = initial; MM IV = mono-morphemic intervocalic; PB IV = pre-boundary intervocalic; WFV = word final pre vocalic; WFC = word final pre consonantal.

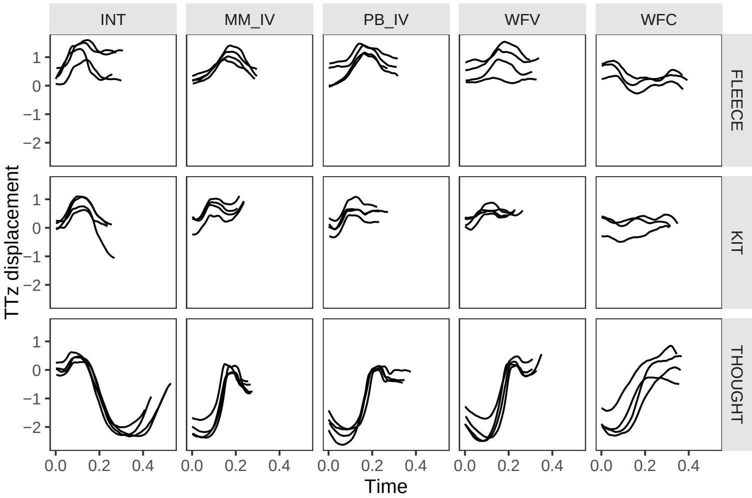

Figure 9.11: Z scored vertical tongue tip displacement for speaker S02. INT = initial; MM IV = mono-morphemic intervocalic; PB IV = pre-boundary intervocalic; WFV = word final pre vocalic; WFC = word final pre consonantal.

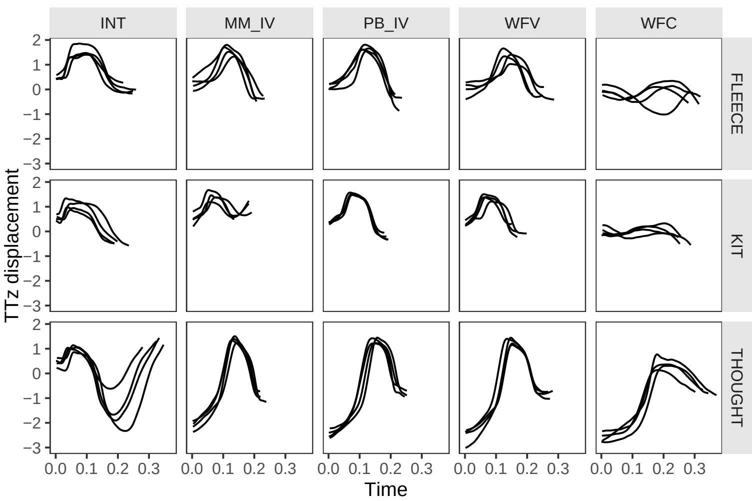

Figure 9.12: Z scored vertical tongue tip displacement for speaker S03. INT = initial; MM IV = mono-morphemic intervocalic; PB IV = pre-boundary intervocalic; WFV = word final pre vocalic; WFC = word final pre consonantal.

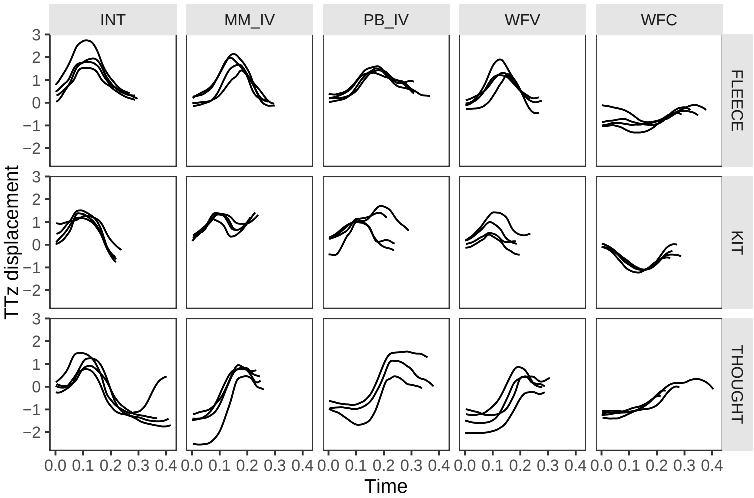

Figure 9.13: Z scored vertical tongue tip displacement for speaker S04. INT = initial; MM IV = mono-morphemic intervocalic; PB IV = pre-boundary intervocalic; WFV = word final pre vocalic; WFC = word final pre consonantal.

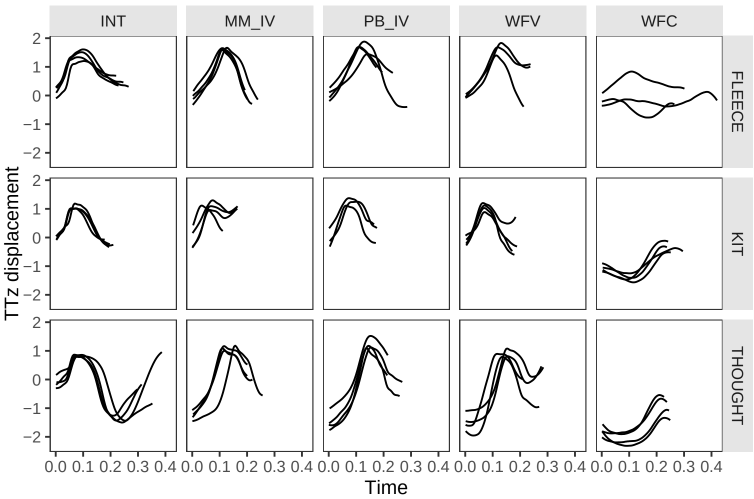

Figure 9.14: Z scored vertical tongue tip displacement for speaker S05. INT = initial; MM IV = mono-morphemic intervocalic; PB IV = pre-boundary intervocalic; WFV = word final pre vocalic; WFC = word final pre consonantal.

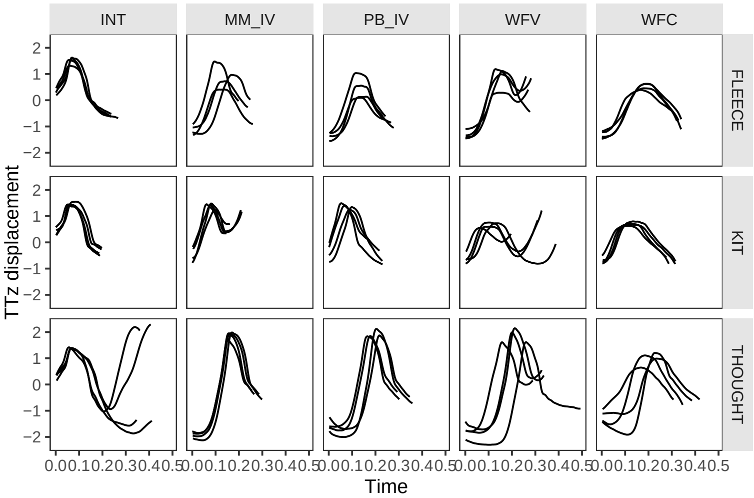

Figure 9.15: Z scored vertical tongue tip displacement for speaker S06. INT = initial; MM IV = mono-morphemic intervocalic; PB IV = pre-boundary intervocalic; WFV = word final pre vocalic; WFC = word final pre consonantal.

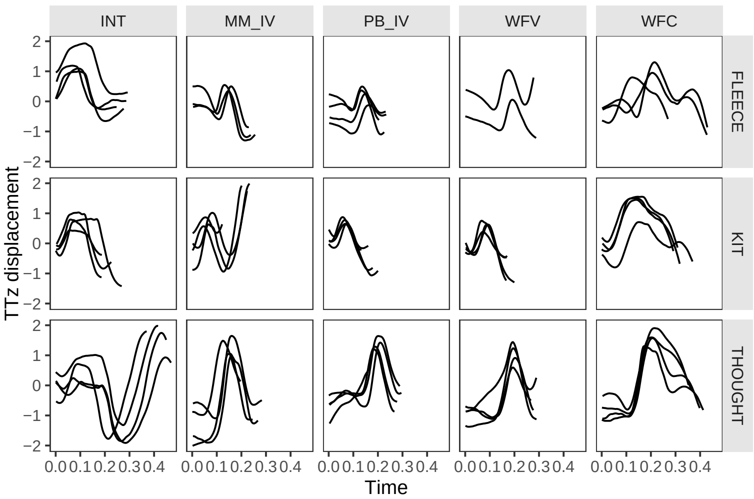

Figure 9.16: Z scored vertical tongue tip displacement for speaker S07. INT = initial; MM IV = mono-morphemic intervocalic; PB IV = pre-boundary intervocalic; WFV = word final pre vocalic; WFC = word final pre consonantal.

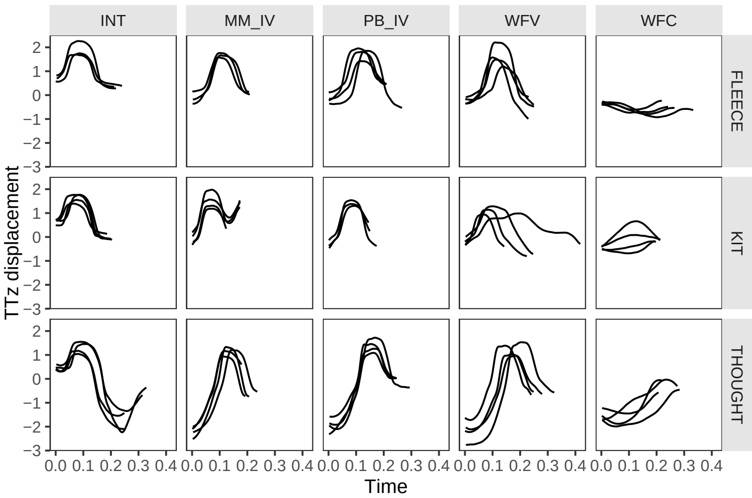

Figure 9.17: Z scored vertical tongue tip displacement for speaker S08. INT = initial; MM IV = mono-morphemic intervocalic; PB IV = pre-boundary intervocalic; WFV = word final pre vocalic; WFC = word final pre consonantal.

9.4 Appendix 5

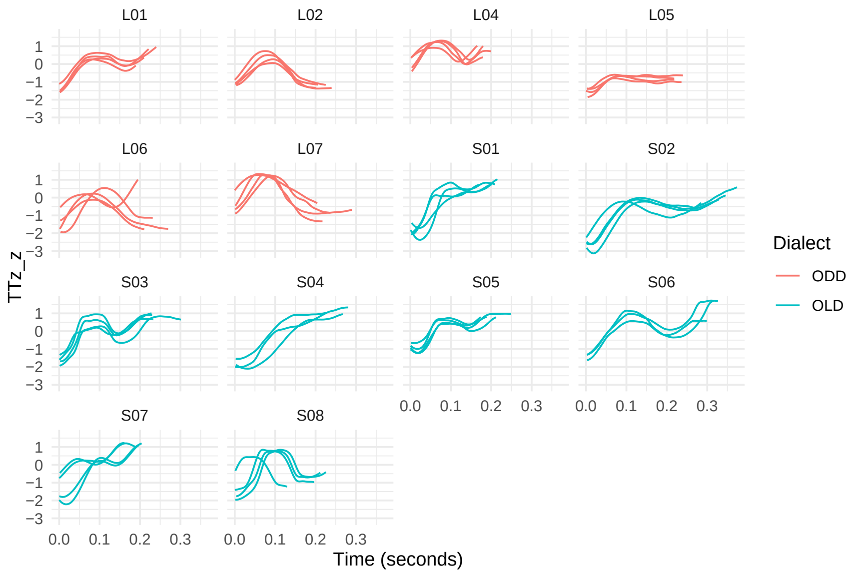

Figure 9.18: Z scored vertical tongue tip displacement for /l/ plus the preceding and following segment in *gull*

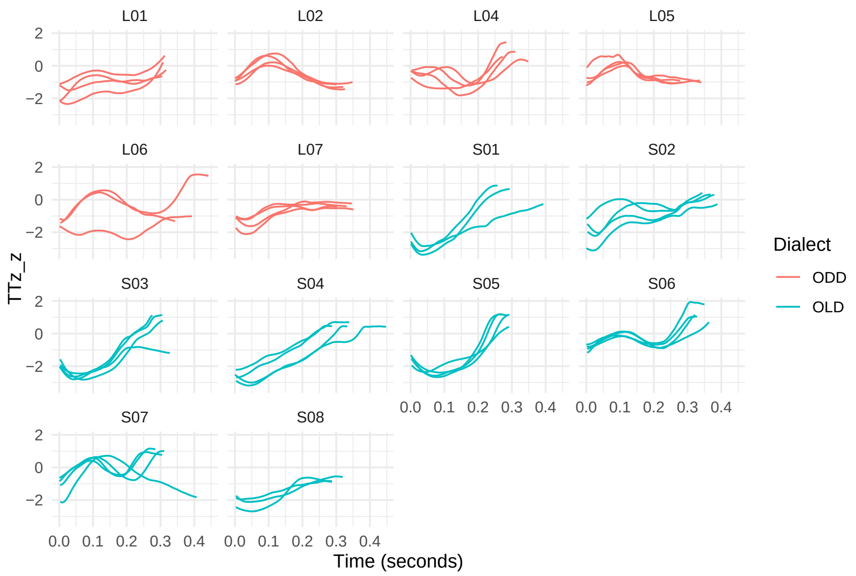

Figure 9.19: Z scored vertical tongue tip displacement for /l/ plus the preceding and following segment in *gulp*.

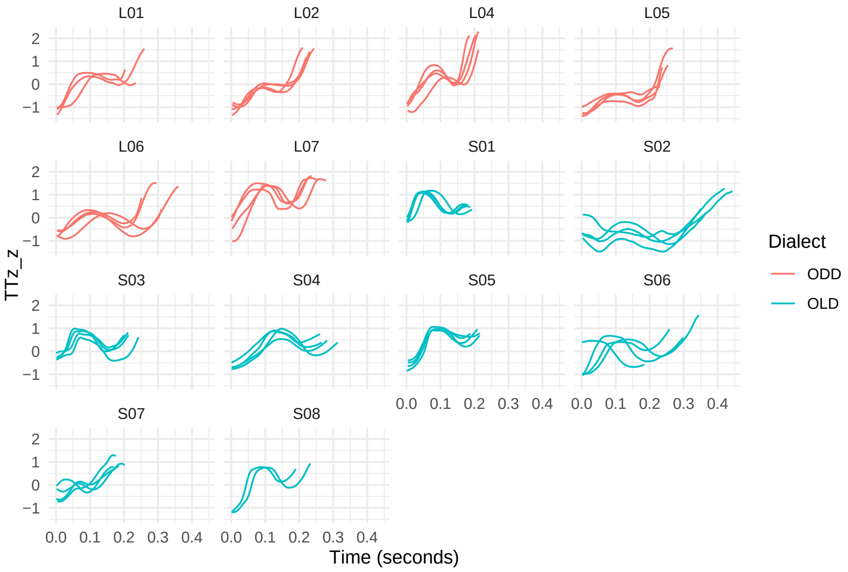

Figure 9.20: Z scored vertical tongue tip displacement for /l/ plus the preceding and following segment in *mill*

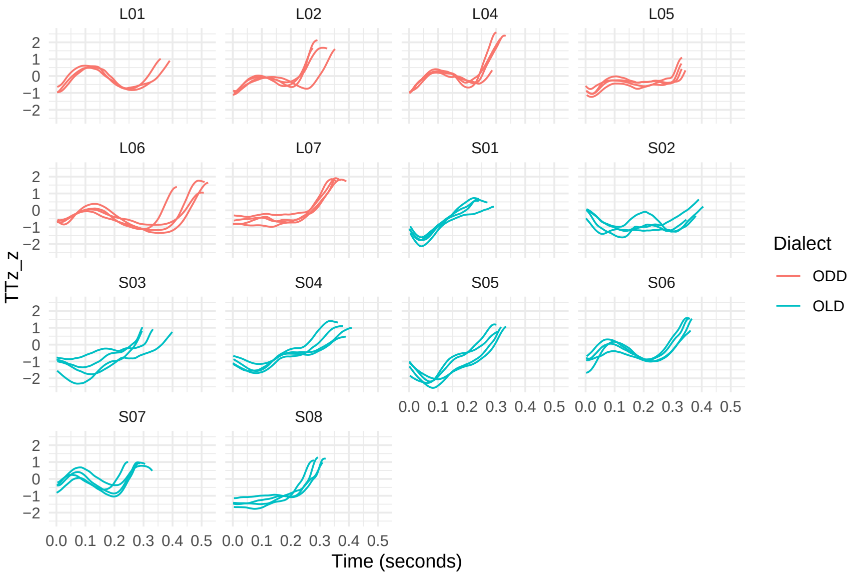

Figure 9.21: Z scored vertical tongue tip displacement for /l/ plus the preceding and following segment in *milk*. Speaker L07 was excluded from the coda analysis due to vocalisation for this word pair.

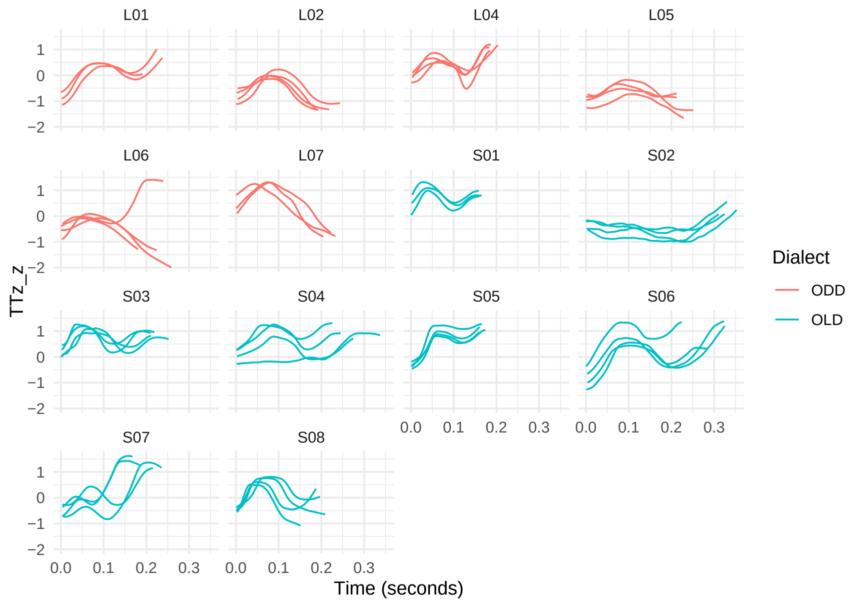

Figure 9.22: Z scored vertical tongue tip displacement for /l/ plus the preceding and following segment in *fill*

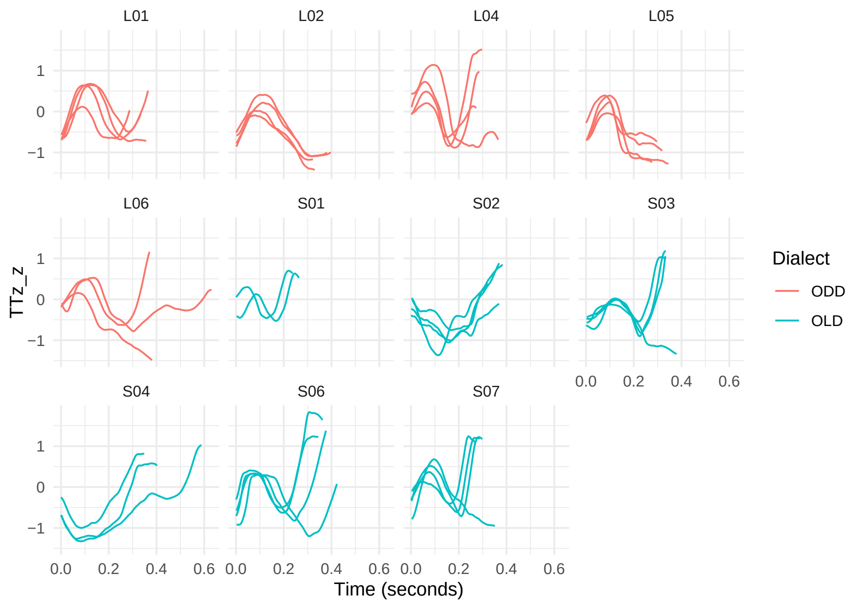

Figure 9.23: Z scored vertical tongue tip displacement for /l/ plus the preceding and following segment in *philp*.

Bibliography

Barbier, Gillaume et al. (2015). "Human vocal tract growth: A longitudinal study of the development of various anatomical structures". In: *Interspeech, 2015-16th Annual Conference of the International Speech Communication Association.*

Barreda, Santiago (2021). "Fast Track: Fast (nearly) automatic formant-tracking using Praat". In: *Linguistics Vanguard* 7.1, p. 20200051.

Bates, Douglas et al. (2015). "Fitting linear mixed-effects models using lme4". In: *Journal of Statistical Software* 67.1, pp. 1–48.

Beal, Joan (2008). "English dialects in the North of England: Phonology". In: *Varieties of English* 1, pp. 122–144.

Bladon, R Anthony W and Ameen Al-Bamerni (1976). "Coarticulation resistance in English/l". In: *Journal of Phonetics* 4.2, pp. 137–150.

Boersma, Paul and David Weenink (2011). *Praat: doing phonetics by computer [Computer program]. Version 5.2.29.* http://www.praat.org/.

Borchers, Hans W. (2022). *pracma: Practical Numerical Math Functions.* R package version 2.4.2. https://CRAN.R-project.org/package=pracma.

Browman, Catherine P and Louis Goldstein (1988). "Some notes on syllable structure in articulatory phonology". In: *Phonetica* 45.2-4, pp. 140–155.

Browman, Catherine P. and Louis Goldstein (1992). "Articulatory phonology: an overview". In: *Phonetica* 49.3-4, pp. 155–180.

Browman, Catherine P and Louis Goldstein (1995). "Dynamics and articulatory phonology". In: *Mind as motion: Explorations in the dynamics of cognition* 175, p. 194.

Browman, Catherine P. and Louis Goldstein (2000). “Competing constraints on intergestural coordination and selforganization of phonological structures.” In: *Bulletin de la Communication Parlée* 5, pp. 25–34.

Browman, Catherine P and Louis M Goldstein (1986). “Towards an articulatory phonology”. In: *Phonology* 3, pp. 219–252.

Brunner, Jana et al. (2014). “Timing of German onset and word boundary clusters”. In: *Laboratory Phonology* 5.4, pp. 403–454.

Burgdorf, Dan Cameron (2022). *The Consonant/Vowel Distinction: A Cognitive Difference Evidenced by Glides*. Cornell University.

Carter, Paul and John Local (2007). “F2 variation in Newcastle and Leeds English liquid systems”. In: *Journal of the International Phonetic Association* 37.2, pp. 183–199.

Cole, Amanda and Patrycja Strycharczuk (2024). “The search for linguistically coherent accents: Unsupervised clustering of diphthong variation in Southeast England”. In: *English World-Wide* 45.1, pp. 1–29.

Fant, Gunnar (1971). *Acoustic theory of speech production: with calculations based on X-ray studies of Russian articulations*. 2. Walter de Gruyter.

Gick, Bryan et al. (2006). “Toward universals in the gestural organization of syllables: a cross-linguistic study of liquids”. In: *Journal of Phonetics* 34.1, pp. 49–72.

Giles, Stephen B and Kenneth L Moll (1975). “Cinefluorographic study of selected allophones of English /l/”. In: *Phonetica* 31.3-4, pp. 206–227.

Goldstein, Louis, Dani Byrd, and Elliot Saltzman (2006). “The role of vocal tract gestural action units in understanding the evolution of phonology”. In: *Action to language via the mirror neuron system*, pp. 215–249.

Goldstein, Louis et al. (2007). “Dynamic action units slip in speech production errors”. In: *Cognition* 103.3, pp. 386–412.

Goldstein, Louis et al. (2009). “Coupled oscillator planning model of speech timing and syllable structure”. In: *Proceedings of the 8th Phonetic Conference of China and the International Symposium on Phonetic Frontiers in Phonetics and Speech Science*.

Gubian, Michele, Francisco Torreira, and Lou Boves (2015). "Using functional data analysis for investigating multidimensional dynamic phonetic contrasts". In: *Journal of Phonetics* 49, pp. 16–40.

Haggard, Mark (1973). "Abbreviation of consonants in English pre-and post-vocalic clusters". In: *Journal of Phonetics* 1.1, pp. 9–24.

Haken, Hermann, JA Scott Kelso, and Heinz Bunz (1985). "A theoretical model of phase transitions in human hand movements". In: *Biological cybernetics* 51.5, pp. 347–356.

Hardcastle, William and William Barry (1989). "Articulatory and perceptual factors in /l/ vocalisations in English". In: *Journal of the International Phonetic Association* 15.2, pp. 3–17.

Honda, Masaaki, Akinori Fujino, and Tokihiko Kaburagi (2002). "Compensatory responses of articulators to unexpected perturbation of the palate shape". In: *Journal of Phonetics* 30.3, pp. 281–302.

Honorof, Douglas N and Catherine P Browman (1995). "The center or edge: How are consonant clusters organized with respect to the vowel". In: *Proceedings of the XIIIth international congress of phonetic sciences*. Vol. 3. Congress Organisers at KTH and Stockholm University Stockholm, Sweden, pp. 552–555.

Hughes, Arthur, Peter Trudgill, and Dominic Watt, eds. (2012). *English Accents and Dialects: An Introduction to Social and Regional Varieties of English in the British Isles*. Fifth. London: Hodder.

Iskarous, Khalil (2010). "Vowel constrictions are recoverable from formants". In: *Journal of Phonetics* 38.3, pp. 375–387.

Iskarous, Khalil, Joyce McDonough, and Douglas H Whalen (2012). "A gestural account of the velar fricative in Navajo". In: *Laboratory Phonology* 3.1, pp. 195–210.

Iskarous, Khalil and Marianne Pouplier (2022). "Advancements of phonetics in the 21st century: A critical appraisal of time and space in Articulatory Phonology". In: *Journal of Phonetics* 95, p. 101195.

Johnson, Keith (2012). *Auditory and Acoustic Phonetics*. Chichester: Wiley-Blackwell.

— (2020). "The ΔF method of vocal tract normalisation for vowels". In: *Laboratory Phonology* 11.1.

Johnson, Wyn and David Britain (2007). “L-vocalisation as a natural phenomenon: Explorations in sociophonology”. In: *Language Sciences* 29.2-3, pp. 294–315.

Katz, Jonah (2012). “Compression effects in English”. In: *Journal of Phonetics* 40.3, pp. 390–402.

Kelso, JA (1984). “Phase transitions and critical behavior in human bimanual coordination”. In: *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology* 246.6, R1000–R1004.

Kelso, James AS (2009). *Coordination dynamics.*

Kelso, John AS, Elliot L Saltzman, and Betty Tuller (1986). “The dynamical perspective on speech production: Data and theory”. In: *Journal of Phonetics* 14.1, pp. 29–59.

Kirkham, Sam (2017). “Ethnicity and phonetic variation in Sheffield English liquids”. In: *Journal of the International Phonetic Association* 47.1, pp. 17–35.

— (2024a). *tadaR: R interfaace to Task Dynamic Application (v.1.0.0).* URL: `https://doi.org/10.5281/zenodo.13329512`.

— (2024b). *tardis: AG01 data processing tools (v.1.0.0).* URL: `https://doi.org/10.5281/zenodo.13752979`.

Kirkham, Sam and Patrycja Strycharczuk (2024). “A dynamic neural field model of vowel diphthongisation”. In: *Proceedings of the 13th International Seminar on Speech Production*.

Kirkham, Sam et al. (2019). “Dialect variation in formant dynamics: The acoustics of lateral and vowel sequences in Manchester and Liverpool English”. In: *The Journal of the Acoustical Society of America* 145.2, pp. 784–794.

Kochetov, Alexei (2020a). “Research methods in articulatory phonetics I: Introduction and studying oral gestures”. In: *Language and Linguistics Compass* 14.4, pp. 1–1.

— (2020b). “Research methods in articulatory phonetics II: Studying other gestures and recent trends”. In: *Language and Linguistics Compass* 14.6, e12371.

Krakow, Rena A (1999). “Physiological organization of syllables: a review”. In: *Journal of Phonetics* 27.1, pp. 23–54.

Krakow, Rena Arens (1989). *The articulatory organization of syllables: A kinematic analysis of labial and velar gestures*. Yale University.

Lee-Kim, Sang-Im, Lisa Davidson, and Sangjin Hwang (2013). "Morphological effects on the darkness of English intervocalic /l/". In: *Laboratory Phonology* 4.2, pp. 475–511.

Léger, A. et al. (2024). "A battle of articulatory methods for the study of English /l/: UTI vs MRI". In: *Ultrafest XI: Extended Abstracts.* DOI: 10.5281/zenodo.12578650: University of Aizu, pp. 38–42.

Lehiste, Ilse (1970). *Suprasegmentals.* MIT Press.

— (1980). "Phonetic manifestation of syntactic structure in English". In: *Annual Bulletin of the Research Institute of Logopaedics and Phoniatrics* 14, pp. 1–27.

Lenth, Russell V. (2025). *emmeans: Estimated Marginal Means, aka Least-Squares Means.* R package version 1.11.2-8. URL: https://CRAN.R-project.org/package=emmeans.

Liu, Zirui, Yi Xu, and Feng-fan Hsieh (2022). "Coarticulation as synchronised CV co onset – Parallel evidence from articulation and acoustics". In: *Journal of Phonetics*, p. 101116.

Löfqvist, Anders and Vincent L Gracco (1999). "Interarticulator programming in VCV sequences: Lip and tongue movements". In: *The Journal of the Acoustical Society of America* 105.3, pp. 1864–1876.

Mackenzie, Sara et al. (2018). "North American /l/ both darkens and lightens depending on morphological constituency and segmental context". In: *Laboratory Phonology* 9.1.

Marin, Stefania and Marianne Pouplier (2010). "Temporal organization of complex onsets and codas in American English: Testing the predictions of a gestural coupling odel". In: *Motor Control* 14.3, pp. 380–407.

— (2014). "Articulatory synergies in the temporal organization of liquid clusters in Romanian". In: *Journal of Phonetics* 42, pp. 24–36.

McAuliffe, Michael et al. (2017). "Montreal Forced Aligner: Trainable Text-Speech Alignment Using Kaldi". In: *Proc. Interspeech 2017*, pp. 498–502. DOI: 10.21437/Interspeech.2017-1386.

Medizinelektronik, Carstens (2014). *AG501 Manual.* https://www. ag500.de/. (Visited on 2024).

Meenakshi, Nisha et al. (2014). “Comparison of speech quality with and without sensors in electromagnetic articulograph AG 501 recording”. In: *Fifteenth Annual Conference of the International Speech Communication Association*. Citeseer.

Mooshammer, Christine et al. (2012). “Bridging planning and execution: Temporal planning of syllables”. In: *Journal of phonetics* 40.3, pp. 374–389.

Mücke, Doris, Anne Hermes, and Sam Tilsen (2020). “Incongruencies between phonological theory and phonetic measurement”. In: *Phonology* 37.1, pp. 133–170.

Munhall, Kevin G, Anders Löfqvist, and JA Scott Kelso (1994). “Lip–larynx coordination in speech: Effects of mechanical perturbations to the lower lip”. In: *The Journal of the Acoustical Society of America* 95.6, pp. 3605–3616.

Munhall, K.G. et al. (2009). “Talkers alter vowel production in response to real-time formant perturbation even when instructed not to compensate”. In: *Journal of the Acoustical Society of America* 125.1, pp. 384–390.

Nam, Hosung (2007). *A gestural coupling model of syllable structure*. Yale University.

Nam, Hosung, Louis Goldstein, and Elliot Saltzman (2009). “Self-organization of syllable structure: A coupled oscillator model”. In: *Approaches to phonological complexity* 16, pp. 299–328.

Nam, Hosung and Elliot Saltzman (2003). “A competitive, coupled oscillator model of syllable structure”. In: *Proceedings of the 15th international congress of phonetic sciences*. Vol. 1. 2253-2256.

Nance, Claire (2014). “Phonetic variation in Scottish Gaelic laterals”. In: *Journal of Phonetics* 47.1, pp. 1–17.

Pastätter, Manfred and Marianne Pouplier (2017). “Articulatory mechanisms underlying onset-vowel organization”. In: *Journal of Phonetics* 65, pp. 1–14.

Pouplier, Marianne (2003). “Units of phonological encoding: Empirical evidence”. PhD thesis. Yale University.

— (2012). “The gestural approach to syllable structure: Universal, language- and cluster-specific aspects.” In: *Speech Planning and Dynamics.* Ed. by Susanne Fuchs et al. Frankfurt am Main: Peter Lang, pp. 63–96.

Pouplier, Marianne et al. (2022). “Language and cluster-specific effects in the timing of onset consonant sequences in seven languages”. In: *Journal of Phonetics* 93, p. 101153.

Proctor, Michael and Rachel Walker (2012). "Articulatory bases of sonority in English liquids". In: *The sonority controversy* 18, p. 289.

Proctor, Michael et al. (2019). "Articulatory characterization of English liquid-final rimes". In: *Journal of Phonetics* 77, p. 100921.

Rebernik, Teja et al. (2021). "A review of data collection practices using electromagnetic articulography". In: *Laboratory Phonology* 12.1, p. 6.

Recasens, Daniel (2004). "Darkness in [l] as a scalar phonetic property: implications for phonology and articulatory control". In: *Clinical Linguistics & Phonetics* 18.6-8, pp. 593–603.

— (2012). "A cross-language acoustic study of initial and final allophones of /l/". In: *Speech Communication* 54.3, pp. 368–383.

Recasens, Daniel and Aina Espinosa (2005). "Articulatory, positional and coarticulatory characteristics for clear /l/ and dark /l/: evidence from two Catalan dialects". In: *Journal of the International Phonetic Association* 35.1, pp. 1–25.

— (2009). "An articulatory investigation of lingual coarticulatory resistance and aggressiveness for consonants and vowels in Catalan". In: *The Journal of the acoustical society of America* 125.4, pp. 2288–2298.

Recasens, Daniel and Clara Rodríguez (2016). "A study on coarticulatory resistance and aggressiveness for front lingual consonants and vowels using ultrasound". In: *Journal of Phonetics* 59, pp. 58–75.

— (2017). "Lingual articulation and coarticulation for Catalan consonants and vowels: An ultrasound study". In: *Phonetica* 74.3, pp. 125–156.

Saltzman, Elliot L and Kevin G Munhall (1989). "A dynamical approach to gestural patterning in speech production". In: *Ecological psychology* 1.4, pp. 333–382.

Scobbie, James M and Marianne Pouplier (2010). "The role of syllable structure in external sandhi: An EPG study of vocalisation and retraction in word-final English /l/". In: *Journal of Phonetics* 38.2, pp. 240–259.

Shaw, Jason A and Wei-rong Chen (2019). "Spatially conditioned speech timing: Evidence and implications". In: *Frontiers in Psychology* 10, p. 468853.

Shaw, Jason A et al. (2011). "Dynamic invariance in the phonetic expression of syllable structure: a case study of Moroccan Arabic consonant clusters". In: *Phonology* 28.3, pp. 455–490.

Simko, Juraj and Fred Cummins (2010). “Embodied task dynamics.” In: *Psychological review* 117.4, p. 1229.

Simpson, Adrian P. (2001). “Does articulatory reduction miss more patterns than it accounts for?” In: *Journal of the International Phonetic Association* 31.1, pp. 29–39.

Sotiropoulou, Stavroula and Adamantios Gafos (2022). “Phonetic indices of syllabic organization in German stop-lateral clusters”. In: *Laboratory Phonology* 13.1.

Sproat, Richard and Osamu Fujimura (1993). “Allophonic variation in English /l/ and its implications for phonetic implementation”. In: *Journal of Phonetics* 21.2, pp. 291–311.

Stevens, Kenneth N. (1998). *Acoustic Phonetics*. Cambridge, MA: MIT Press.

Strycharczuk, Patrycja, Donald Derrick, and Jason Shaw (2020). “Locating delateralization in the pathway of sound changes affecting coda /l/”. In: *Laboratory Phonology* 11.1.

Strycharczuk, Patrycja and James M. Scobbie (2015). “Velocity measures in ultrasound data. Gestural timing of post-vocalic /l/ in English”. In: *Proceedings of the XVIII International Congress of Phonetic Sciences*, pp. 1–5.

Strycharczuk, Patrycja et al. (2020). “Gestural delay and gestural reduction. Articulatory variation in /l/-vocalisation in Southern British English”. In: *The corpus phonology of English: Multifocal analyses of variation*, pp. 9–29.

Szalay, Tünde et al. (2022). “Reconsidering lateral vocalisation: Evidence from perception and production of Australian English /ll”. In: *The Journal of the Acoustical Society of America* 152.4, pp. 2106–2116.

Thompson, Austin and Yunjung Kim (2019). “Relation of second formant trajectories to tongue kinematics”. In: *The Journal of the Acoustical Society of America* 145.4, pp. 323–328.

Tilsen, Sam (2018). “Three mechanisms for modeling articulation: selection, coordination, and intention”. In: *Ithaca, NY: Cornell University.*

— (2019). “Motoric mechanisms for the emergence of non-local phonological patterns”. In: *Frontiers in Psychology* 10, p. 2143.

— (2020a). “A different view of gestural activation: learning gestural parameters and activation with an RNN”. In: *Manuscript, Cornell University.*

Tilsen, Sam (2022). “An informal logic of feedback-based temporal control”. In: *Frontiers in Human Neuroscience* 16.

Tilsen, Sam and Louis Goldstein (2012). “Articulatory gestures are individually selected in production”. In: *Journal of phonetics* 40.6, pp. 764–779.

Tilsen, Samuel (2020b). “Motoric mechanisms for the emergence of non-local phonological patterns in selection co-ordination theory”. In: ICU Linguistics Colloquim.

Tolfree, Laura (1999). “South East London English: discrete versus continuous modelling of consonantal reduction”. In: *Urban Voices: Accent Studies in the British Isles*. Ed. by Paul Foulkes and Gerard J. Docherty. London: Arnold, pp. 163–184.

Trudgill, Peter (1986). *Dialects in Contact*. Oxford: Blackwell.

Turk, Alice, Benjamin Elie, and Juraj Šimko (2023). “PlanArt: A modular platform for computational modeling of articulatory planning”. In: *Proceedings of the 20th International Congress of Phonetic Sciences*.

Turk, Alice and Stefanie Shattuck-Hufnagel (2020). *Speech timing: Implications for theories of phonology, phonetics, and speech motor control*. Vol. 5. Oxford University Press, USA.

Turton, Danielle (2014). “Some /l/s are darker than others: accounting for variation in English /l/ with ultrasound tongue imaging”. In: *University of Pennsylvania Working Papers in Linguistics* 20.2, pp. 189–198.

— (2015). “Determining categoricity in English /l/-darkening: A principal component analysis of ultrasound spline data”. In: *Proceedings of the 18th International Congress of Phonetic Sciences (ICPhS 2015)*. Newcastle University.

Turvey, Michael T (1990). “Coordination.” In: *American psychologist* 45.8, p. 938.

Vollmann, Ralf et al. (2017). “The spread of/l/-vocalization in Styria”. In: *Phonetik in und über Osterreich, S. Moosmller, M. Sellner, C. Schmid, Eds. Wien: Verlag der Osterreichischen Akademie der Wissenschaften*, pp. 123–137.

Wang, Jane-Ling, Jeng-Min Chiou, and Hans-Georg Müller (2016). “Functional data analysis”. In: *Annual Review of Statistics and its application* 3, pp. 257–295.

Wells, John C. (1982). *Accents of English: Volumes 1–3*. Cambridge: Cambridge University Press.

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*.

Ying, Jia et al. (2021). “Evidence for active control of tongue lateralization in Australian English /l/”. In: *Journal of Phonetics* 86, p. 101039.

Yuan, Jiahong and Mark Liberman (2009). “Investigating /l/ variation in English through forced alignment”. In: *Tenth Annual Conference of the International Speech Communication Association*.