INVESTIGATING SOURCES OF PHONOLOGICAL RARITY AND INSTABILITY: A STUDY OF THE PALATAL LATERAL APPROXIMANT IN BRAZILIAN PORTUGUESE

BY

NICOLE W. WONG

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Linguistics
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2017

Urbana, Illinois

Doctoral Committee:

Associate Professor Ryan K. Shosted, Chair and Director of Research
Associate Professor Chilin Shih
Dr. Joseph Roy
Associate Professor Torrey Loucks

ABSTRACT

The palatal lateral is a rare sound in the world's languages; a review of the literature reveals just 23 languages that currently possess the palatal lateral. Similarly, only 15 (or 3.33%) of the languages in the UCLA Phonological Segment Inventory Database (UPSID) (Maddieson and Precoda, 1991) can claim to currently possess the palatal lateral. While UPSID reports that an additional five languages (Basque, Guarani, Iate, Spanish, Turkish) possess the palatal lateral, these languages have either lost the palatal lateral or were included erroneously. Understanding the production and perception of rare speech sounds is important for understanding the distribution of speech sounds cross-linguistically, especially with regards to the establishment of a single phonetic alphabet (i.e. the International Phonetic Alphabet (IPA)) that can be used to describe and transcribe the languages of the world (Ladefoged and Everett, 1996). An investigation of rare speech sounds can also reveal important findings regarding the physical limitations of the vocal tract and human auditory system.

Given that the palatal lateral is a rare speech sound, a complete description of the articulation, acoustics, and perception of this sound does not currently exist. Accounts of the palatal lateral vary with regards to terminology; the palatal lateral has also been referred to as a so-called "phonemically" palatalized lateral (Zilyns'kyj, 1979), a laminal post-alveolar lateral (Ladefoged and Maddieson, 1996), and an alveolopalatal lateral (Recasens, 2013). Furthermore, current literature also does not distinguish between the palatal lateral and a palatalized lateral. The lack of agreement in literature regarding terminology can present problems when attempting to assess whether a palatal lateral in one language is similar to a palatal lateral in another language. This dissertation provides a comprehensive description of the palatal lateral, as a means of initiating cross-linguistic comparisons of the palatal lateral as well as understanding the difference between a palatal and palatalized lateral.

A two-part study of the articulation and acoustics of the palatal lateral in Brazilian Portuguese (BP) was undertaken in this dissertation. Articulatory data was collected using electromagenetic articulography (EMA) from 10 female native speakers of BP from São Paulo state in Brazil, which permitted the simultaneous collection of acoustic information. Study 1 investigated the articulation of the palatal lateral through a battery of measures and compares the palatal lateral against the palatalized lateral approximant, alveolar lateral approximant, palatal approximant, palatal nasal, palatalized nasal, and alveolar nasal. Study 2 analyzes the acoustics of the palatal lateral in comparison to the palatalized lateral approximant, alveolar lateral approximant, and palatal approximant.

A third study was included in the appendix. This study incorporates a phone identification task to understand the role of acoustic saliency in the rareness of the palatal lateral, i.e. compared to other palatal sounds, is the palatal lateral more likely to be misidentified and if so, as which sounds? This task also investigates whether there is a perceived difference between the palatal and palatalized lateral that may not be captured by Study 1 and 2, in addition to whether native speakers of BP are better at distinguishing the two sounds than non-native speakers (here, native speakers of American English). The palatal lateral was compared to the palatalized lateral, palatal approximant, alveolar lateral approximant, palatal nasal, palatalized nasal, alveolar nasal, voiced alveolar stop, and voiced palatalized alveolar stop. 25 (11 male, 14 female) natives speakers of BP and 20 (11 male, 9 female) native speakers of American English with no extensive exposure to BP participated in this study.

Results from Study 1 show that the palatal lateral is articulated laminally with a high front tongue body and concave anterior tongue shape that gradually becomes straighter as the phone progresses. Acoustic results in Study 2 indicate a median F1, F2, and F3 of 367 Hz, 1954 Hz, and 3035 Hz respectively for female speakers of BP. Statistical analysis reveals little or no evidence of significant difference between the palatal lateral and palatalized lateral with regards to the shape of the tongue body, duration of the phone, or formant frequencies.

The perception study included in the appendix finds that while both native and non-native speakers of BP distinguish between the palatal lateral and palatalized lateral at chance level, native speakers of BP perform better than the non-native speakers at correctly identifying the palatal and palatalized nasal. This study also finds that of all the sounds included in this task, the palatal and palatalized lateral are the most likely to be misidentified as the palatal approximant for both participant groups, with the addition of -3 dB of speech-shaped noise greatly increasing the rate of confusion. However, the palatalized lateral is inaccurately identified as a palatal approximant at a confusion rate nearly double or more than the palatal lateral.

This dissertation reveals that the palatal and palatalized lateral are essentially the same sound in BP. Furthermore, there is no evidence that indicates that the palatal or palatalized lateral are composed of two separate phones, i.e. an alveolar lateral approximant followed by a palatal approximant. Findings from the perception study support the proposal that *yeísmo* (i.e. the merger of the palatal lateral in favor of the palatal approximant (Colantoni, 2001; Hualde et al., 2005)) occurs because lateral sounds are less robust against added noise than nasal sounds. I argue here that this contributes directly to the rareness of the palatal lateral.

To all the educators and mentors I have had along the way: thank you for encouraging me on a path that I could never even have dreamed of.

To my family: you have always been my biggest supporters and source of strength. I love you.

ACKNOWLEDGMENTS

I am grateful to my committee members for their support and guidance. Each of them have spent countless hours working with me and teaching me how to be a better scientist. I truly appreciate all the time they have sacrificed to help me grow as a scholar.

I would like to especially thank my adviser Dr. Ryan Shosted, who supported my work not only through his time, but also by funding my electromagnetic articulography research through his own research grant. Without his supervision and expertise, this dissertation would not have been possible.

Special thanks to Dr. Torrey Loucks, for graciously allowing me to collect articulatory and acoustic data in the NeuroSpeech lab using the NDI Wave Speech Research System.

Thanks to Dr. Jont Allen for taking the time to discuss his work on speech perception with me, which inspired the design of the perception study in my dissertation.

Thanks to Marissa Barlaz, Zainab Hermes, Sarah Johnson, and Anna Tendera, for donating hours of their free time to help me collect my articulatory and acoustic data.

Thanks to everyone who helped me recruit participants by spreading the word to all their Brazilian friends, especially Jessica Ramirez and Denise Pozzani.

Many thanks to Zainab Hermes for working with me to create the Matlab scripts to process EMA data.

I was blessed to receive the support and encouragement of so many wonderful friends and colleagues during this process. Jen Chun, thanks for your sage advice and for letting me cause a ruckus at your house every week. You read every page of my dissertation and for that, I could never thank you enough. Marissa Barlaz, I could not have had a better partner in crime. Thanks for commiserating with me. Myeong-Hyeon Kim, thanks for being an endless source of encouragement.

To my family, thanks for putting up with me and for the countless rides to the airport. Without your love and support, I would not be here.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1

INTRODUCTION

1.1 Background and motivation

The exploration of articulatory stability and acoustic similarity to other speech sounds is particularly interesting with regards to the effect on distributions of rare speech sounds cross-linguistically. Investigation of vocal tract limitations and cognitive processing of spoken speech contributes to our understanding of why some sounds are found more or less frequently across the world's languages. Previous work on the distribution of vowels within and across languages incorporates information from both the production and perception to explain why some vowel sounds are more frequent than other, i.e. the theory of adaptive dispersion (Liljencrants and Lindblom, 1972; Lindblom, 1986, 1990), which invokes the principal of maximal contrast (Jakobson, 1941) to propose that speech sounds are positioned in a finite acoustic space such that the greatest amount of perceptual distinctiveness can be achieved. In other words, the frequent appearance of the vowels /i, a, u/ in phone inventories cross-linguistically is a result of the interaction between vocal tract limitations and optimization of cognitive processing: maximal perceptual contrast optimizes ease of cognitive processing, with possible vowel candidates determined by limitations of the oral tract, e.g. how far the jaw can lower or the tongue can retract. Similarly, investigating the production and perception of a rare speech sound contributes to our understanding of the phonetic structure of language systems as a whole.

Ladefoged and Everett (1996) emphasize the importance of the study of rare sounds, particularly with regards to implications for developing a phonetic alphabet that encompasses the majority of speech sounds in the world's languages. A single, coherent, all-encompassing phonetic alphabet such as the International Phonetic Alphabet (IPA) has without a doubt assisted field work in cataloging undocumented languages and the comparison of sound systems cross-linguistically. The study of rare sounds can also inform the discussion of historical sound changes such as assimilation. Ohala (Ohala, 1981, 1989, 1993) argues that sound change is listener-driven, e.g. occurring as a result of the listener misinterpreting perceptual cues in the acoustic signal. Rare sounds may be infrequent due to acoustic characteristics that render the sound less salient to the listener; a better understanding of these characteristics and how they are produced may shed light on their effect on perception as well as rarity.

This dissertation will provide a detailed and comprehensive description of the production, perception, and acoustics of the palatal lateral approximant /ʎ/, a rare sound in the world's languages. Compared to other liquids such as the alveolar lateral approximant /l/, which is found in 174 (or 39.58%) of the languages contained within the UCLA Phonological Inventory Database (UPSID), the palatal lateral approximant is reported in only 20 (or 4.43%) of the languages in UPSID (Maddieson and Precoda, 1991). Unlike research on the alveolar lateral approximant /l/ (O'Connor et al., 1957a; Ainsworth, 1968; Dalston, 1975; Sproat and Fujimura, 1993), there have only been a handful of studies on the articulation (Straka, 1965; Fant, 1960; Martins et al., 2008) or acoustics (Recasens, 1984b; Fant, 1960) of the palatal lateral approximant. In fact, a comprehensive study of the articulation, perception, and acoustics of the palatal lateral approximant does not currently exist.

1.1.1 A preliminary definition of the palatal lateral

For the purpose of clarity, it is important for this dissertation to distinguish between a palatalized lateral approximant and a palatal lateral approximant. A palatalized lateral approximant will be defined here as the product of an alveolar lateral approximant coarticulated with a nearby palatal approximant[1] or high vowel, while a palatal lateral approximant will be defined as a phoneme, i.e. appearing in contrastitive distribution with /l/. I will preliminarily define the palatal lateral approximant as articulated with: (1) a high front

[1]Note that while both the lateral palatal approximant /ʎ/ and the central palatal approximant /j/ are both palatal approximants, the term *palatal approximant* will be used to indicate the central palatal approximant /j/.

tongue body, (2) a laminal articulation, (3) lingual contact along the hard palate, and (4) a lateral airstream channel on one or both sides of the tongue.

The palatal lateral is represented in the IPA by the symbol ʎ. Like other palatal sounds, it is described as being produced with a high front tongue body and extended contact from the alveolar ridge to the back of the hard palate (Keating, 1988; Recasens, 2013). Lingual contact along the hard palate can extend up to 2-3 times as far as the alveolar lateral approximant /l/ (Keating, 1988:87). As with other lateral sounds (Ladefoged and Maddieson, 1996; Ladefoged and Johnson, 2014), the palatal lateral possesses a narrow lateral passage perpendicular to the midline of the tongue, with a lateral airstream channel on one or both sides of the tongue (Recasens, 1984b). Average formant frequencies reported for several languages (see Table 2.6) indicate that the palatal lateral is produced with a relatively low F1 and high F2 and F3 (Fant, 1960; Silva, 1999; Colantoni, 2004), which suggests that the palatal lateral is articulated with a high front tongue body. Like the palatal approximant in English (Stevens, 2000), there is an approximately 1500 Hz difference between the reported average F1 and F2 for the palatal lateral in Brazilian Portuguese (Silva, 1999), Corrientes Spanish (Colantoni, 2004), and Russian (Fant, 1960). The most common terminology for this sound is a 'palatal lateral'; for the sake of simplicity, the term 'palatal lateral' will be used here.

While the above definition of a palatal lateral is quite straightforward, in reality the literature is often inconsistent when identifying palatal laterals cross-linguistically. For example, the palatal lateral in Ukrainian has been referred to as a palatalized lateral (Zilyns'kyj, 1979), a laminal post-alveolar lateral (Ladefoged and Maddieson, 1996) and an alveolopalatal lateral (Recasens, 2013). Several languages such as Russian (Fant, 1960), Irish (Rousselot, 1899), and Bulgarian (Scatton, 1975), are described as possessing a so-called "phonemically" palatalized lateral; these three languages contrast plain and "palatalized" consonants in the same way that English contrasts voiced and voiceless consonants. However, the term *palatalized* implies a phonetic distinction. The so-called "palatalized" phones in all three languages are not the result of coarticulation with a neighboring high vowel or palatal approximant. According to the definition of a palatal lateral as presented above, this suggests that the so-called palatalized lateral in these languages should actually be considered a palatal lateral. For example, the palatalized lateral in Russian (Fairbanks, 1965; Fant, 1960; Kochetov, 2005; Proctor, 2009) can occur in non-palatalizing contexts and phonemically contrasts with the non-palatalized lateral approximant /l/, i.e. in the minimal pair дал /dal/ 'give' and даль /dalʲ/ 'distance'. One might therefore argue that the Russian palatalized lateral is a palatal lateral instead, due to its phonological status[2]. To clarify what constitutes a palatal lateral and a palatalized lateral, the distinction between a palatal articulation resulting from co-articulation with the phonetic context (i.e. next to a palatal approximant /j/ or high front vowels) and a palatal articulation that is inherent to the phone must be made.

No clear distinction is made in the literature between a phonemically and phonetically palatalized lateral with regards to specific articulatory or acoustic differences. A review of the literature does not indicate whether a palatalized lateral (regardless of phonemic status) is the same as a palatal lateral. There is some evidence from the investigation of contextually palatalized and phonemically palatalized velars (Keating and Lahiri, 1993) that suggests there is a significant difference in the place of articulation on the hard palate between contextually palatalized (i.e. sounds that acquire a palatal constriction due to neighboring high vowels or glide) and phonemically palatalized sounds (i.e. as in the previous example from Russian). However, there is no concrete way as of yet to determine whether a sound is a palatalized or palatal lateral with regards to either the articulation or acoustics. As a result, it can be difficult to correctly assess whether a palatal lateral in one language is similar to a palatal lateral in another language, or even whether a palatalized lateral and a palatal lateral in a single language are the same sound (e.g. Portuguese, which has both: *familia* 'family' [fa.ˈmi.lʲə][3] versus *pilha* 'battery' /ˈpi.ʎə/). A precise and detailed description of the articulatory and acoustic differences between a palatalized lateral and a palatal lateral must be completed, in order to provide a clearer understanding of what constitutes a palatal lateral and what differentiates this sound from a palatalized lateral.

[2] One might also contend that the "palatalized" consonant class in Russian is a translation artifact and that modern Russian phonology does indeed treat these sounds as a separate phoneme class. Regardless, the literature does indeed continue to use the term "palatalized" to reference these sounds, which can potentially obfuscate our understanding of what constitutes a palatal and palatalized sound. Here, I am advocating the usage of a single, more theoretically intuitive terminology.

[3] Note that the example of the palatalized is provided in brackets [] to emphasize that this is not considered a phonemic contrast in this dissertation.

1.1.2 Sources of phonological rarity

The investigation of diachronic language change is interesting for the cross-linguistic documentation and categorization of speech sounds, especially trends such as the merger of the palatal lateral towards articulatorily similar sounds such as the palatal approximant /j/ (De los Heros Diez Canseco, 1997; Colantoni, 2004) or lateral approximant /l/ (Benko and Imre, 1972). Such trends suggest that the change could be due to an inherent trait of the phone (e.g. the physical production of the phone or the acoustics). One development in particular may be especially insightful: amongst the handful of languages that are said to possess a palatal lateral, many have already undergone a process similar to *yeísmo*. This term refers to a historical sound change found in Spanish, when the palatal lateral began to be pronounced as a palatal approximant [j] (De los Heros Diez Canseco, 1997; Colantoni, 2004), resulting in a merger between the two sounds. However, this phenomenon is not limited to Spanish and can also be found in other non-Romance languages, such as in the prestige Budapest variety of Hungarian (Benko and Imre, 1972). The sound that merges with the palatal lateral may not always be a palatal sound either; unlike in Budapest, western dialects of Hungarian have merged the palatal lateral with the plain lateral approximant (Benko and Imre, 1972).

Strictly from an impressionistic point of view, the palatal lateral can sound very similar to other palatal sounds such as the palatal approximant /j/ or the voiced palatal fricative /ʝ/. Given the acoustic similarity between the palatal lateral and other palatal sounds, it is interesting that some languages can continue to distinguish such highly similar sounds (e.g. /j/, /ʎ/, and [lʲ] in Portuguese, as in the example provided previously). It is possible that acoustic similarities with neighboring sounds may cause listeners to confuse the palatal lateral with a sound that is both acoustically similar and more salient. A comparison of the formant frequencies for the palatal lateral and the palatal approximant in Table 1.1 supports this possibility; the two sounds demonstrate very similar first, second, and third formant frequencies. Table 1.1 compares the calculated formant frequencies for the Russian palatal lateral (Fant, 1960:167) against the American English palatal approximant (Stevens, 2000:516,526) and lateral approximant (Stevens, 2000:535,546); the contents of this table serve to highlight the acoustic similarities of the palatal lateral to the palatal approximant. This is one potential reason for why the palatal lateral merges with the palatal approximant in languages that have lost this sound (e.g. Spanish and Hungarian). The development and loss of the palatal lateral in these languages, amongst others, will be discussed further in Section 2.1.

	F1	F2	F3
[ʎ]	210 Hz	1700 Hz	2500 Hz
[j]	260 Hz	1770 Hz	2950 Hz
[l]	360 Hz	1100 Hz	2800 Hz

Table 1.1: Calculated formant frequencies of the palatal lateral in Russian and the palatal approximant and lateral approximant in American English.

Additional reasons for the rarity of the palatal lateral may also include articulatory instability or frequency effects; each potential factor will be discussed carefully following extensive articulatory and acoustic analysis of the data. Compared to the alveolar lateral approximant /l/, there is relatively less information on the articulation and acoustics of the palatal lateral. Many of the articulatory experiments (see Section 2.2) were conducted several decades ago and do not include naturally produced speech from multiple speakers. This is problematic, particularly when considering the many technological advances in methodological approaches that have been made since then.

1.1.3 Brazilian Portuguese

This dissertation will provide a detailed description of the palatal lateral specifically for Brazilian Portuguese (BP), which possesses both a palatal lateral as a phoneme and a palatalized lateral that is the result of coarticulation with a following high front vowel /i/. An example of a near minimal pair is the name *Emilio* ‘Emilio’ [ɛ.ˈmi.lʲu] and the two-word phrase *e milho* ‘and corn’ /ɛ.ˈmi.ʎʊ/. Some dialects of BP, unlike European Portuguese, palatalize certain consonants before /i/ (Mateus and d’Andrade, 2000:7; Azevedo, 2005:46), including /l/. Since BP possesses both a palatalized lateral and a palatal lateral, this facilitates the investigation of what constitutes a palatal lateral and whether there are any defining characteristics that distinguish it from a palatalized lateral. Interesting comparisons can also be made to the palatal nasal /ɲ/ and palatalized nasal [n^{j}] in BP. The variety of BP spoken in São Paulo state is one of the dialects that palatalize consonants before /i/, which is not a linguistic feature of all Brazilian Portuguese dialects. For this reason and to minimize unnecessary dialect factors, all the Brazilian participants in this dissertation were from São Paulo state. Further discussion of the distribution and historical development of the palatal lateral in Portuguese will be presented in Section 2.1.2 and Section 2.1.1.

There are several advantages for selecting BP as the language of choice for this study of the palatal lateral. Not only does BP distinguish between /j/, /ʎ/, and [l^{j}], there is also a an interesting parallel with /j/, /ɲ/, and [n^{j}]; this makes BP ideal for studying not just the palatal lateral, but the relationship between the palatal lateral and similar sounds as well. Additionally, 80% of Portuguese speakers reside in Brazil, with just the population of São Paulo city equal to slightly more than all of Portugal (Azevedo, 2005:17). In a census report published in November 2010, the population of Brazil was revealed to be 190,732,694 (IBGE, 2010). Because of Brazil, Portuguese has a substantial representation in the languages spoken worldwide. The overseas Brazilian community is large as well, with a sizable population residing in Urbana-Champaign, Illinois - the presence of local community was especially helpful when collecting EMA data. The implications of linguistic research on Brazilian Portuguese are relevant as well; as Brazil continues to grow as an economy and as a nation, so to will the importance of a Brazil-specific approach to Portuguese, particularly in the area of speech technologies.

1.2 Research goals

This dissertation provides a comprehensive and detailed phonetic description of a rare speech sound, the palatal lateral in BP, through a three-part study of the articulation, acoustics, and perception of this speech sound. The palatal lateral is compared against the following sounds [j, l, l^{j}, n, n^{j}, ɲ], as a means of identifying the articulatory, acoustic, and perceptual characteristics that distinguishes the palatal lateral from similar speech sounds in BP. Because of the presence of the sounds [j, l, l^{j}, n, n^{j}, ɲ] in BP, this language presents a relatively unique opportunity to study the relationship of the palatal lateral with the palatalized lateral, as well as permitting the study of how the parallel relationship between the palatal nasal and palatalized nasal differs. An investigation of potential factors for the rarity of the palatal lateral is included, e.g. the perceptual effects of acoustic similarity to other sounds and the effect of articulatory similarity on phonemic stability. The results are situated in a literature review of the diachronic development of the palatal lateral in other languages, which assists in identifying potential trends in the development or loss of the palatal lateral that may contribute to its lower frequency cross-linguistically.

A detailed phonetic description of the palatal lateral approximant is obtained through two studies[4]: (1) an electromagnetic articulography (EMA) study of the articulation and (2) a study of the acoustics and the effect of the articulation on the acoustics. A single language (BP) is selected for the purpose of study here, providing a frame of reference against which future comparisons of the palatal lateral approximant in other languages can be made. A discussion of potential reasons for why the palatal lateral approximant is rare is also included

1.3 Outline

Chapter 2 presents a review of the literature on the articulation, acoustics, and perception of the palatal lateral, as well as the cross-linguistic distribution of the palatal lateral and the diachronic development and loss of the palatal lateral in select languages. Chapter 3 and 4 pertain to Study 1 and 2, respectively, with an interim discussion of the results from each study at the end of each respective chapter. In Chapter 3, Study 1 is an EMA study documenting and comparing the articulation of the BP sounds [ʎ, l, lʲ, ɲ, n, nʲ, j]. A comprehensive description of the articulation of the tongue, including a dynamic model of tongue movement, is presented in this chapter. Study 2 is a study of the acoustics simultaneously recorded during the EMA study. A description of the acoustics is provided here, along with an investigation of correlations between tongue position and the resulting

[4]A third study, a study of the perception and confusability of /ʎ/, is also included in the appendix. This study is a perception study of the BP sounds [ʎ, l, lʲ, ɲ, n, nʲ, j, d, dʲ], as perceived by both native speakers of BP and English. Comparisons of the responses between the two groups will be used to understand how native and non-native perception of the palatal lateral differs.

acoustics. Chapter 5 integrates results from Study 1 and 2 in addition to results from a perception study included in the appendix, thereby providing a general discussion of the rarity of the palatal lateral as motivated by articulatory, acoustic, or perceptual factors. The dissertation is concluded in Chapter 6.

Chapter 2

LITERATURE REVIEW

This chapter reviews the current literature on the palatal lateral across different languages. Since the palatal lateral is rather infrequent cross-linguistically, experimental studies of the articulation and acoustics of the palatal lateral are also few. When appropriate, literature on other palatal or contextually palatalized sounds will be included for comparison. Section 2.1 discusses the distribution of the palatal lateral cross-linguistically from a historical linguistics perspective, i.e. the diachronic development and loss of the palatal lateral. Previous research on the articulation (Section 2.2) and acoustics (Section 2.3) of the palatal lateral is reviewed according to research paradigms and technologies. Section 2.4 identifies potential perceptual factors for the rarity of the palatal lateral by reviewing literature on similar sounds such as the palatal approximant and alveolar lateral approximant.

2.1 The history and linguistic distribution of the palatal lateral

	ʎ	l	j	ɲ
# of languages in UPSID	20	174	379	141
% of languages in UPSID	4.43%	38.58%	84.04%	31.26%

Table 2.1: Distribution of sounds in the UPSID language database (Maddieson and Precoda, 1991).

Frequency counts pulled from the UPSID language database (Maddieson and Precoda, 1991) illustrate the imbalanced distribution of palatal and lateral sounds across the world's languages (see Table 2.1). While the nasal counterpart to the palatal lateral is relatively frequent amongst the languages in the database at 31.26%, the palatal lateral is found in just 4.43% of the languages. The palatal lateral is by far the least common of the palatal and lateral sounds listed here. It is possible that the palatal lateral is less frequent due to its high acoustic similarity to the palatal approximant. This may explain why the palatal lateral typically merges with the palatal approximant in languages that have lost this sound. However, despite the acoustic similarity to the palatal lateral[1], the cross-linguistic frequency of the palatal approximant is not similar; languages possessing a palatal approximant comprise 84.04% of the languages in UPSID (4% of which also possess a palatal lateral) (Maddieson and Precoda, 1991).

While one might argue that the distribution of the palatal lateral is simply a result of the Perceptual Magnet Effect (PME), i.e. statistically-driven learning resulting in phone categories defined by cognitive "prototypes" that "attract" variable phonetic input (Kuhl, 2007), this theory does not hold true for cross-linguistic frequency. The PME is an effect that occurs within the mental space of a single speaker; when considering multiple speakers that do not belong to the same linguistic community, one would then expect that the distribution of speech sounds across the world's languages would actually be randomly determined. This however, is not the case. Table 2.1 clearly demonstrates that the frequency of speech sounds is not evenly distributed across the world's languages, indicating that the rarity of the palatal lateral is not simply the result of statistical learning.

Recasens (1984b) finds that greater tongue contact against the palate is positively correlated with less coarticulation with adjacent vowels, with order by degree of palatal contact as follows: palatal approximant, palatal nasal, and palatal lateral. His findings correspond with cross-linguistic frequency rates presented in Table 2.1, which, when ordered from highest to lowest is as follows: palatal approximant, alveolar lateral approximant, palatal nasal, and palatal lateral. One would expect that if a sound demonstrates greater coarticulation, that sound may be more prone to being interpreted incorrectly by the listener, resulting in a higher rate of sound change and therefore causing the sound to be less frequent. The expectation that coarticulation results in sound change follows from Ohala's theory of sound change, where sound change occurs as a result of the listener misinterpreting the acoustic

[1] Please see Table 2.6 for comparisons of the formant frequencies of the palatal lateral and palatal approximant in multiple languages.

signal. The expectation that higher rates of coarticulation would result in lower frequency rates of the sound is reflected in the cross-linguistic frequency of the palatal lateral; Recasens (1984b) finds that of three palatal sounds /ɲ, ʎ, j/, the palatal lateral demonstrates the highest rate of coarticulation, which is also shown to be the least frequent of the three sounds cross-linguistically (Maddieson and Precoda, 1991).

A survey of the literature reveals a total of 33 languages that are said to possess a palatal lateral in their phone inventory (see Table 2.2 and Table 2.3), just 13 languages more than indicated by UPSID. Of the 13 languages, five were simply not included in the database (Catalan, Italian, Portuguese, Slovak, Ukrainian), four languages were reported to posses a palatalized lateral and not a palatal lateral (Bulgarian, Greek, Irish, Russian), two languages had undergone *yeísmo* (French, Hungarian), one language (Arrernte) was not categorized as possessing a palatal lateral (however, no source for the phone inventory was given), and lastly, UPSID follows the conventional belief in the literature (Gim, 1971; Iverson and Sohn, 1994) that the liquid in one language is in fact an alveolar lateral approximant (Korean). Of the 33 languages, 11 have either lost the palatal lateral (i.e. Basque, Hungarian, Spanish, French, Italian), do not conform to the description of a palatal lateral as defined here (Greek, Slovak, Turkish), or there is a discrepancy in the literature (Guarani, Iate, Korean). Of the latter three languages, examination of the literature cited by UPSID for Guarani (Uldall, 1954) and Iate (Lapenda, 2005) revealed that these two languages are not described as possessing a palatal lateral in the original sources, while the existence of the palatal lateral in Korean is disputed (see Section 2.2.4 for more information on the dispute). This results in only 23 languages that can currently claim a palatal lateral as a member of their phone inventory. Note that the majority of these reports are impressionistic accounts, without substantiative articulatory or even acoustic data; articulatory and acoustic studies of the palatal lateral are limited to Spanish, Catalan, Italian, Russian, and Portuguese.

Language	Language family	Comments
Araucanian*	Language isolate	Three-way lateral contrast: dental, alveolar, palatal (Echeverria and Contreras, 1965).
Arrernte	Australian, Pama-Nyungan	Contrasts laminal palato-alveolar lateral with apical dental lateral (Wilkins, 1989; Breen and Henderson, 1990; Tabain, 2009).
Basque*	Language isolate	Restricted to V.CV environment (Hualde and Bilbao, 1992). Reportedly has undergone *yeísmo*[2].
Breton*	Indo-European, Celtic	Contrasts alveolar and palatal laterals (Elmar, 1970).
Bulgarian	Indo-European, Slavic	"Hard" (plain) vs. "soft" (palatalized) contrast as in Russian, palatalization only occurs before back vowels (Benko and Imre, 1972; Scatton, 1975).
Camsa*	Language isolate	Contrasts alveolar lateral approximant and palatal lateral (Howard, 1967).
Catalan	Indo-European, Romance	Contrasts alveolar lateral approximant and palatal lateral (Recasens et al., 1993; Recasens and Espinosa, 2006).
Diyari*	Australian, Pama-Nyungan, Karna	Contrasts palatal, dental, and alveolar laterals (Austin, 1981).
French	Indo-European, Romance	*Ye'ismo* (/ʎ/ > [j]) (Dauzat, 1899).
Greek	Indo-European, Greek, Attic	Occurs before an /i/ that is followed by another vowel (Arvaniti, 1999; Joseph and Tserdanelis, 2003).
Guambiano*	Paezan, Coconuco	Contrasts palatal and alveolar, also refers to palatal lateral as alveo-palatal (Branks and Branks, 1973).
Guarani*	Tupian, Tupi-Guarani	Error in UPSID. Only lateral is alveolar lateral approximant, present in Spanish loanwords (Uldall, 1954).
Hungarian	Uralic	*Yeísmo* (/ʎ/ > [j]), [l]. Only some northern and minority dialects distinguish /ʎ/ and /j/ or /l/ (Benko and Imre, 1972).
Iate*	Language isolate	Error in UPSID. Contrasts dental lateral and velar lateral (Lapenda, 2005).
Irish	Indo-European, Celtic	"Broad" (plain) vs. "slender" (palatalized) contrast, similar to Russian "hard" (plain) vs. "soft" (palatalized) (Rousselot, 1899; O'Nolan, 1934).
Italian	Indo-European, Romance	Has undergone *yeísmo* in several varieties (Rohlfs, 1966; Maiden, 1995b); variable rate of maintenance (Bladon and Carbonaro, 1978).

Table 2.2: Palatal laterals of the world (A-I). Languages that have lost the palatal lateral or if the sound is not a true palatal lateral are indicated in gray. Asterisks indicate languages with a palatal lateral according to UPSID. Language family information retrieved from Ethnologue (Lewis, 2009).

Language	Language family	Comments
Jaqaru*	Aymaran, Tupe	Contrasts with alveolar lateral approximant, similar with other languages within the Jaqi linguistic family (Hardman, 1966, 1978).
Jebero*	Cahuapanan	Contrasts with alveolar lateral approximant (Bendor-Samuel et al., 1961). Contact occurring from dental to palatal (impressionistic) (Valenzuela and Gussenhoven, 2013).
Kalkatungu*	Australian, Pama-Nyungan	Contrasts with dental, alveolar, and retroflex lateral approximants (Blake, 1979).
Khanty*	Uralic	Phonemic palatalization, like other Uralic languages (Abondolo, 1998).
Komi*	Uralic	Phonemic palatalization. Contrasts with dental lateral approximant (Tereshchenko, 1966; Bubrikh, 1949; Lytkin, 1966) .
Korean	Koreanic	Disagreement in literature, impressionistic descriptions (Gim, 1971) report /l/ while ultrasound (Oh and Gick, 2002) suggest /ʎ/.
Koryak*	Chukotko-Kamchatkan	Phonemic palatalization (Zhukova, 1980). Spoken in Eastern Russia (Lewis, 2009).
Mari*	Uralic	Phonemic palatalization (Ristinen, 1960).
Nganasan*	Uralic	Phonemic palatalization (Tereshchenko, 1966).
Ngarinjin*	Australian, Worrorran	Descriptions not consistent, /ʎ/ referred to as palato-dental or palatal, /j/ referred to as "alveolar", /l/ referred to as "alveolar" and "alveo-velar" (Coates and Elkin, 1974).
Portuguese	Indo-European, Romance	Contrasts with alveolar lateral approximant (Mateus and d'Andrade, 2000; Azevedo, 2005).
Quechua*	Quechuan	Contrasts with alveolar lateral approximant (Bills and Troike, 1969).
Russian	Indo-European, Slavic	Phonemic palatalization (Fairbanks, 1965; Cubberley, 2002).
Slovak	Indo-European, Slavic	Commonly produced as /l/ with "secondary palatalization" (Hanulíková and Hamann, 2010).
Spanish*	Indo-European, Romance	Has undergone *yeísmo* (/ʎ/ > [j], [ɟ]) (De los Heros Diez Canseco, 1997; Colantoni, 2004; Hualde et al., 2005).
Turkish*	Turkic	Occurs when vowel environment is [-back] (Clements and Sezer, 1982).
Ukrainian	Indo-European, Slavic	"Soft" and "hard" contrast, large amount of dialectal variation for each aspect (Zilyns'kyj, 1979).

Table 2.3: Palatal laterals of the world (I-Z). Languages that have lost the palatal lateral or if the sound is not a true palatal lateral are indicated in gray. Asterisks indicate languages with a palatal lateral according to UPSID. Language family information retrieved from Ethnologue (Lewis, 2009).

Five languages have lost the palatal lateral as a result of a merger between the palatal lateral and another sound (e.g. *yeísmo*). In the majority of these languages, the merger occurred between the palatal lateral and a palatal approximant. A review of the acoustics of the palatal lateral and palatal approximant reveal that the two sounds are strikingly similar with regards to their average formant frequency (see Section 2.3), with the similarity between the two sounds (as well as the similarity between the palatal nasal and palatal lateral) discussed in depth in a treatment of *yeísmo* in Spanish by Lipski (1989). In conjunction with the frequency results from UPSID, it is entirely possible that the motivating factor for why the palatal lateral typically merges with the palatal approximant (and not the reverse) is the same factor that causes the rarity of the palatal lateral. Of these five languages, only one language (Hungarian) did not merge the palatal lateral with a palatal sound. A detailed explanation of the development and loss of the palatal lateral in these languages will be presented in Section 2.1.1.

The so-called palatal lateral in Greek, Slovak, and Turkish is identified here as a palatalized lateral since it occurs in these languages only before front and/or high vowels (Arvaniti, 1999; Joseph and Tserdanelis, 2003; Hanulíková and Hamann, 2010; Clements and Sezer, 1982). Given some evidence that contextually palatalized sounds are articulated differently from sounds that were originally produced with a palatal articulation (Keating and Lahiri, 1993), these three languages are excluded from the final tabulation of languages possessing a palatal lateral. A number of Uralic and Slavic languages are included in Tables 2.2 - 2.3, many of which are characterized by phonemic palatalization (e.g. Russian, Ukrainian, Komi, Khanty) where the sound inventory of these languages contrast a series of palatal and non-palatal phones. There is evidence that a palatal nasal in these languages is articulated slightly differently than a palatal nasal in languages without this phonemic palatalization contrast (see Section 2.2.3), i.e. in Russian versus Catalan (Recasens and Romero, 1997), but no study has made a direct comparison for the palatal lateral.

For three languages (Guarani, Iate, Korean), there was a discrepancy in the literature regarding the phonological status of the sound. UPSID incorrectly identifies Guarani and Iate as possessing a palatal lateral when in fact, no such sound exists in these languages. In the case of Guarani, reports indicate that the alveolar lateral approximant entered the language through Spanish loanwords (Uldall, 1954). While it is possible that the same may have occurred for the palatal lateral, this is dependent on whether Spanish had already undergone *yeísmo* at the time the words were borrowed, which is not indicated in the literature. With regards to Korean, traditional impressionistic descriptions ascribe an alveolar place of articulation to the liquid when syllable-final (Gim, 1971; Iverson and Sohn, 1994), though imaging studies reveal that the sound is produced with a high front tongue body and laminal articulation (Oh, 2002; Oh and Gick, 2002; Gick et al., 2006), which is characteristic of a palatal sound.

In all of the languages examined above, mergers only occur between the palatal lateral and phonetically similar sounds such as the palatal approximant (i.e. the glide /j/) or the lateral approximant /l/; it is never simply deleted. Similarly, *yeísmo* is beginning to appear in some dialects of BP (Azevedo, 2005), though the majority of BP speakers maintain a distinction between the palatal lateral and palatal approximant. The study of diachronic sound change is especially relevant for the study of rare speech sounds, since mergers such as *yeísmo* can be an indication of inherent traits of the palatal lateral that render it more susceptible to sound change (e.g. acoustic similarity or perceptual ambiguity). Indeed, sound change does not occur randomly and is in fact quite ordered, which allows the development of theories capable of predicting the occurrence of sound change either as a function of the role of the listener (e.g. as follows from Ohala (1981, 1989, 1993)) or the speaker (e.g. resulting from overlapping gestures in Articulatory Phonology, as developed by Browman and Goldstein (1986, 1995)). In the following section, a detailed comparison of the historical development and loss of the palatal lateral cross-linguistically will be presented, in an attempt to identify shared characteristics that might be relevant towards a more accurate definition of the palatal lateral, its historical development, and why it is relatively uncommon among the languages of the world.

2.1.1 The development of the palatal lateral and *yeísmo* in select languages

Amongst the handful of languages that are said to possess a palatal lateral, several have already undergone *yeísmo* or a process similar to *yeísmo* (see Tables 2.2 - 2.3). As mentioned in Section 1.1.2, this term refers to a historical sound change in Spanish that merged the palatal lateral with the palatal approximant [j] (De los Heros Diez Canseco, 1997; Colantoni, 2004). This section will also integrate hypotheses from the literature that propose potential factors for why *yeísmo* occurs. It is entirely possible that what drives the merger of the palatal lateral directly contributes to the rareness of the palatal lateral.

Typical discussions of the palatal lateral and *yeísmo* are focused on the development and loss of the sound in Spanish, including not only Peninsular Spanish (PS) (Hualde et al., 2005; Zampaulo, 2015) but regional dialects as well (Boyd-Bowman, 1952; Mouton and Martos, 2012). The importance of the palatal lateral and *yeísmo* is especially relevant for the study of Spanish; dialectal classifications of Latin American dialects incorporate the presence or absence of the palatal lateral as one of the few phonetic criteria for categorization (Rona, 1964; Honsa, 1975; Resnick, 1975). The literature generally agrees that the palatal lateral in Spanish developed from three separate sources and was eventually lost in all three situations. The first two instances occurred during Proto-Spanish: the first from the sequence /-lj-/ (i.e. an intervocalic cluster of the alveolar lateral approximant and palatal approximant) in Popular Latin, and the second from /-k^jl-, -g^jl-/ (i.e. intervocalic clusters of palatalized velar stops and an alveolar lateral approximant) in Western Romance (Penny, 2000; Pharies, 2007; Zampaulo, 2015). The latter change is also reported as having occurred during "primitive" Castilian from Latin sources (Pharies, 2007). Following this development in Proto-Spanish, orthographic representations suggest that the palatal lateral changed to the voiced palato-alveolar sibilant fricative [ʒ] during Old Spanish (Lapesa and Pidal, 1942; Menéndez Pidal, 1950; Zampaulo, 2015), also referred to as Medieval Castilian (Pharies, 2007).

At the same time that the palatal laterals that developed from Popular Latin and Proto-Spanish sources were lost during Old Spanish, another palatal lateral developed from the Latin sequences /pl-, kl-, fl-/ (i.e. syllable-initial clusters with an alveolar lateral) and /-lː-/ (i.e. a geminate intervocalic alveolar lateral) (Penny, 2000; Pharies, 2007; Zampaulo, 2015). This palatal lateral began to de-lateralize during the mid 17th century and eventually merged with ɟ, which is pronounced as [ʝ] when intervocalic (Lipski, 1989; Penny, 2000; Hualde et al., 2005). The merger was complete within a single generation (Lipski, 1989). However, it was not complete in all Spanish-speaking regions, including rural areas at the center of the Iberian Peninsula which did not begin to experience the merger until the 1930s (Mouton and Martos, 2012). Some of these areas still maintain the distinction, though there have been significant advances in the merger recently (Mouton and Martos, 2012). Accounts of Latin American Spanish also report a large variation in terms of the maintenance or loss (through merger) of the palatal lateral (Lipski, 1994).

This more recent loss of the palatal lateral in Spanish is now known as the phenomenon *yeísmo* (Hualde et al., 2005; Zampaulo, 2015). However, the description of the sound which merges with the palatal lateral can vary; the product of *yeísmo* in Spanish is sometimes described as palatal approximant (De los Heros Diez Canseco, 1997; Colantoni, 2001) or a palatal fricative [ʝ] (Lipski, 1989; Penny, 2000; Hualde et al., 2005). This discrepancy is probably due to allophonic variation, as the realization of this phone changes depending on phonetic context; [ɟ] following a nasal, [l] following a pause, or [j] in all other contexts (Martínez-Celdrán et al., 2003).

Some suggestions for potential driving forces behind the merger of the Spanish palatal lateral have been made. Citing the theory of hypocorrection as put forth by Ohala (1981), Zampaulo (2015) hypothesizes that if listeners are unable to hear the lateral aspect of the palatal lateral, the sound may be reinterpreted as a palatal approximant. Similarly, results from an acoustic study of Corrientes Spanish (Colantoni, 2004) suggest that listeners hear the glide-like element of the palatal lateral present in the transition between the palatal lateral and neighboring vowel and interpret this element as an indication of the presence of a glide /j/ instead of a palatal lateral. This is apparently also true of speakers of Argentinian Spanish (Colantoni, 2004). In defense of this proposal, Colantoni (2004) references work on hidden variation theory by Ohala (1989), who proposes that sound change occurs when the listener misinterprets preexisting acoustic cues in the signal.

However, results from the same study of Corrientes Spanish (Colantoni, 2004) also support the proposal that the palatal lateral undergoes merger due to the expansion of the constriction at the palate (i.e. the constriction loosens as the tongue moves down and away from the roof of the mouth), resulting in a more glide-like articulation; lower F2 values suggest a more open articulation, with the palatal lateral in Corrientes Spanish articulated further back in the vocal tract in comparison too Italian and PS. Likewise, Mouton and Martos (2012) ascribe the merger to a weakened constriction during the articulation of the palatal lateral; this however, does not explain why some products of *yeísmo* result in a sibilant fricative as in Old Spanish (Menéndez Pidal, 1950; Lapesa and Pidal, 1942; Zampaulo, 2015).

Theoretical phonology accounts of *yeísmo* describe the sound process in Spanish as "delateralization" (Lipski, 1989), though Lipski (1989) notes that this does not explain why the alveolar lateral approximant does not delateralize as well. This treatment of the palatal lateral assumes that the sound is a complex segment, composed of two articulators (a coronal and dorsal gesture) simultaneously linked to single timing slots (Lipski, 1989). Like the palatal lateral, the palatal nasal in Spanish is also known to be pronounced as a nasalized glide [j̃] during rapid or otherwise conversational speech (Lipski, 1989; Recasens et al., 1995), though a more recent study finds that the Spanish palatal nasal is fully occluded while the BP palatal nasal is pronounced as the nasalized glide (Shosted et al., 2012b). However, this does not result in a similar sound change (i.e. /nh/ > /j/) affecting the phonemic status of the Spanish palatal nasal.

The development of the palatal lateral in Portuguese is said to resemble that of Spanish (Zampaulo, 2015), though the development is not identical. Like Spanish, the palatal lateral entered Portuguese through more than one source. However, historical sources for the palatal lateral in Portuguese are not the same sources as in Spanish. Latin sequences /kl, pl, fl/ and intervocalic geminate laterals /ll/ which developed into a palatal lateral in Old Spanish (Penny, 2000; Pharies, 2007; Zampaulo, 2015) respectively became a voiceless palatal fricative and a singleton alveolar lateral approximant /l/ in Early Portuguese (Mateus and d'Andrade, 2000). Instead, the Portuguese palatal lateral developed from Latin singleton /l/ or geminate /ll/ laterals when followed by a high front vowel /i/ (Mateus and d'Andrade, 2000) or a palatal approximant /j/ (Azevedo, 2005); any instance of Latin /ll/ resulting in a palatal lateral are strictly loanwords from Spanish (Mateus and d'Andrade, 2000). For example, Latin *filiu* became Portuguese *filho* [fiʎu]'son' (Mateus and d'Andrade, 2000:81,153). The same process is also occurring in modern BP, where casual pronunciations of syllable-initial alveolar lateral approximants followed by an unstressed high front vowel in words like *familia* or *mobilia* result in a palatalized lateral instead (Azevedo, 2005).

Vulgar Latin sequences composed of first a velar stop and then an alveolar lateral approximant (i.e. /kl/ and /gl/) also resulted in a palatal lateral in Portuguese (Mateus and d'Andrade, 2000), e.g. Vulgar Latin *seclam* became Portuguese *selha* [seʎɐ]'pail' (Mateus and d'Andrade, 2000:86). Both developments resemble the development of the palatal lateral in Italian, which was also derived from Vulgar Latin /kl/ and /gl/ (Grandgent, 1927; Maiden, 1995b).

While Portuguese is typically described as a language that still possesses the palatal lateral (Mateus and d'Andrade, 2000; Azevedo, 2005), Azevedo (2005) notes that *yeísmo* does occur in certain varieties of Portuguese. However, its appearance is socially conditioned and considered to be an indication of poor education (Azevedo, 2005). While the development of the palatal lateral in Portuguese may bear a resemblance to the parallel development of the palatal lateral in Spanish in that both languages developed the palatal lateral through multiple Latin sources, Portuguese has not undergone widespread *yeísmo* nor has it lost the palatal lateral at any stage of the language's evolution.

As fellow descendants of Latin, Portuguese and Italian share similar phone inventories and linguistic histories: both developed the palatal lateral due to assimilatory palatalization as a result of the Proto-Romance palatal glide developing from unstressed high front vowels immediately followed by another vowel (Maiden, 1995b; Mateus and d'Andrade, 2000). Unstressed Latin high front vowels (expressed orthographically as *e* or *i*) followed by another vowel of any quality became a palatal approximant in Proto-Romance, resulting in widespread palatalization of immediately preceding consonants. This eventually led to the emergence of the palatal lateral in Italian (Grandgent, 1927; Maiden, 1995b). The origin of the Italian palatal lateral resembles that of the Portuguese palatal lateral in that it is a product of a palatalizing context. However, the palatal lateral in Italian is slightly different from the palatal lateral in other Romance languages, as it is realized as a geminate speech segment (Maiden, 1995b; Recasens et al., 1993); e.g. Proto-Romance *filja* became Italian *figlia* /fiʎʎa/ 'daughter' (Maiden, 1995b). Italian may have possessed a palatal lateral at another point in time as well; while historical descriptions typically describe the alveolar lateral approximant in Italo-Romance as directly changing to a palatal approximant when followed by a tautosyllabic consonant (Maiden, 1995b), Repetti and Tuttle (1987) argue for the presence of a palatal lateral occurring during the transition between the two sounds.

A more modern source of the palatal lateral in Italian may be derived (sporadically) from the Tuscan intervocalic [gl] cluster, which is supposed to be realized as a geminate stop cluster [ggj] but is occasionally produced as a geminate palatal lateral (Maiden, 1995b). One such example is the Tuscan word *strig(i)le(m)* producing Italian *striglia* /striʎʎa/ 'currycomb' (Maiden, 1995b). Castellani (1954) argues that this is a result of hypercorrection, due to the merger of the sequences [ʎʎ] and [ggj] with the geminate voiced palatal affricate [ɟɟ] in non-prestige varieties of Tuscan. As a result, speakers incorrectly reconstruct instances of the geminate voiced palatal affricate as a geminate palatal lateral.

In addition to the aforementioned merger between the palatal lateral and the voiced palatal affricate in certain Tuscan varieties, the intervocalic palatal lateral in Italian also undergoes *yeísmo* in central (to which Tuscan belongs), southern, and northern varieties of Italian (Maiden, 1995b). In these varieties, the palatal lateral becomes a palatal approximant [j] or palatal stop [ɟ] (Maiden, 1995b), with the majority of northern varieties merging the palatal lateral with a palatal approximant (Rohlfs, 1966). If one agrees with the aformentioned assertion by Repetti and Tuttle (1987) that a palatal lateral emerged as a intermediary stage during the historical development of the alveolar lateral approximant to the palatal approximant, then the Italian palatal lateral can be said to have experienced *yeísmo* more than once.

As in Spanish, the palatal lateral in Basque developed from a geminate alveolar lateral approximant (Hualde, 1991). It is possible that this development may have occurred as a result of close proximity between Spanish-speakers and Basque-speakers. The palatal lateral cannot occur word-initially and appears word-finally in some dialects (such as Arbizu or Baztan) only as a result of coarticulation with a preceding high front vowel, e.g. in the case of /mutil/ becoming [mutiʎ] 'boy' (Hualde, 1991). This is most likely a case of palatalization rather than an example of a phonemic palatal lateral. However, more than 20 years has passed since the publication of this analysis of Basque and oral correspondence in 2014 with a speaker of Basque, currently in her twenties, reveals that the palatal lateral in present-day Basque has also already undergone widespread *yeísmo*[3].

One source for the palatal lateral in Hungarian can be found in Proto-Hungarian, which is said to contrast a plain and palatalized lateral (Benko and Imre, 1972). The presence of such a contrast indicates that the palatalized lateral in this instance is a phonemically palatalized lateral. Like the previous languages discussed above, another source for the palatal lateral in Hungarian was the result of palatalization affecting multiple consonants in

[3] As reported in an oral communication with Itxaso Rodriguez, a colleague from the Department of Spanish and Portuguese at the University of Illinois, in June 2014.

the phone inventory (Benko and Imre, 1972; Vago, 1974). Benko and Imre (1972) report that alveolar lateral approximants occurring in a palatalizing context were gradually phonemically replaced with a palatal lateral beginning in the 15th and 16th century, with the change more pervasive in the eastern regions. In contrast, the western regions did not palatalize the alveolar lateral approximant, while simultaneously de-palatalizing the palatal lateral that was present from proto-Hungarian (Benko and Imre, 1972). Both sources for the palatal lateral in Hungarian are reflected in the orthography, as the palatal lateral is represented by the characters *ly* in Hungarian texts (Benko and Imre, 1972; Vago, 1974). More recently, contemporary varieties of Hungarian in the north have begun to demonstrate palatalization of the alveolar lateral approximant (amongst other consonants) before high vowels as recently as the 1970s (Benko and Imre, 1972).

The palatal lateral underwent *yeísmo* in the so-called "standard" Budapest pronunciation, a prestige variety, with only some of the northern and minority dialects maintaining the distinction (Benko and Imre, 1972). The product of *yeísmo* in modern Hungarian is described as articulated with slightly more frication than the English palatal approximant, but is not a true fricative (Benko and Imre, 1972). The palatal lateral has been pronounced as a palatal approximant in standard Hungarian since the mid 19th century, with the merger beginning in the eastern dialects around the 16th century (Benko and Imre, 1972).

However, the country is divided; while eastern dialects (including the prestige dialect spoken in Budapest) merged the palatal lateral with the palatal approximant, western dialects merged the palatal lateral with the alveolar lateral approximant (Benko and Imre, 1972). Despite this appraisal of the state of the palatal lateral, Benko and Imre (1972) observe a slow progression of the palatal lateral merging with the palatal approximant occurring from the east to the west of Hungary, with the potential for the palatal approximant to dominate both eastern and western dialects. As part of the Uralic family (Lewis, 2009), Hungarian is the only language besides Basque that is not of the Indo-European family identified here as undergoing *yeísmo*. It is also the only language reviewed here that has experienced a relatively recent merger of the palatal lateral with a non-palatal sound (i.e. the alveolar lateral approximant).

2.1.2 Brazilian Portuguese: Additional information

While much of the information in this section has been presented previously in the sections above, it will be summarized here in a codified form for ease of reference, particularly with regards to understanding the construction of the word lists in Chapter 3 and Chapter 4 (as well as the perception study in the appendix). As mentioned in Section 1.1.3, BP was selected as the Portuguese variety of choice for this study as it is the variety spoken by the majority of Portuguese speakers (Azevedo, 2005). Additionally, the palatal lateral in the variety of Portuguese spoken in São Paulo state in Brazil has yet to undergo *yeísmo* or a similar form of sound change. Specifically, the speech of residents from the state of São Paulo will be studied here, to eliminate the effects of dialectal variation.

The Portuguese palatal lateral is constrained in terms of where it can appear; Portuguese (including BP) does not permit the palatal lateral to appear in the coda. The appearance of the palatal lateral in word-initial position is also limited, with nearly all instances being low-frequency loanwords except for the word *lhe* (Mateus and d'Andrade, 2000:11). Considering these phonotactic constraints, all target sounds in the following chapters were embedded in an intervocalic and syllable-initial position. Palatalization of the alveolar lateral approximant occurs only when preceding a tautosyllabic unstressed high front vowel /i/ (Mateus and d'Andrade, 2000; Azevedo, 2005), e.g. the word /ˈpa.li.u/ *pálio* 'canopy' is pronounced [ˈpa.lʲu] in BP.

Portuguese also possesses a palatal nasal that contrasts with an alveolar nasal stop (Barbosa and Albano, 2004; Shosted et al., 2012b). Like the palatal lateral, the palatal nasal is palatalized before a tautosyllabic unstressed high front vowel (Mateus and d'Andrade, 2000; Azevedo, 2005). There is some overlap in the historical development of the palatal nasal with the palatal lateral; like the palatal lateral, the palatal nasal also developed as a result of palatalization, which occurred when Latin singleton /n/ and geminate /nn/ was followed by the high front vowel /i/ (Mateus and d'Andrade, 2000). The palatal nasal is also a result of vowel sequences formed from a pretonic vowel and a tonic /i/ in hiatus with an /a/ or /o/ and the sequences [gn] and [ngl] in Latin (Mateus and d'Andrade, 2000).

Again like the palatal lateral, the palatal nasal is only possible word-initially due to a small set of loanwords; these loanwords are higher-frequency in BP than EP (Mateus and d'Andrade, 2000). The contrast between the palatal nasal, alveolar nasal, and palatalized nasal permits a parallel comparison of the same contrasts in the BP laterals. As a result, this dissertation will be able to investigate not only what distinguishes a palatal lateral from a palatalized lateral approximant and an alveolar lateral approximant, but also whether it is different from what distinguishes a palatal nasal from a palatalized nasal stop and an alveolar nasal stop.

2.2 Articulation of the palatal lateral

2.2.1 X-Ray

X-ray images of the historical Spanish palatal lateral (no longer present in modern Spanish (Hualde et al., 2005)) indicate that the sound was produced with a laminal articulation (Straka, 1965). However, an x-ray study of the Russian palatal lateral (traditionally referred to as a palatalized lateral) does not corroborate Straka (1965)'s findings; x-ray tracings of the Russian palatal lateral reveal a flat or neutral tongue blade and tongue tip contact with the teeth, indicating presence of an apical articulation (Fant, 1960:163). If the two sounds are both considered a palatal lateral under the definition provided in Section 1.1.1, one would expect that both the historical Spanish palatal lateral and the Russian palatal lateral should demonstrate corresponding midsagittal tongue profiles. Given that both studies were conducted on a single speaker of the language, it is possible that the difference in tongue blade contact is due to speaker-specific articulatory strategies as opposed to language-specific differences. A survey of multiple speakers from a single language would be beneficial for understanding which articulatory strategy (laminal or apical) can be generalized to the production of the palatal lateral in that language.

Fant (1960) finds that the Russian alveolar lateral approximant and palatal lateral demonstrate clear articulatory differences: the alveolar lateral approximant has an apical constriction at the alveolar ridge with tongue dorsum retraction, while the palatal lateral has extended contact of the tongue tip and tongue blade reaching from the alveolar ridge to the anterior region of the hard palate (Fant, 1960:163). For both laterals, Fant (1960:162) observes a "lateral outlet" along the lateral edge of the tongue, which extends from near the posterior molars to the front of the tongue. Observations of cine x-rays (Koneczna and Zawadowski, 1956) lead Ladefoged and Maddieson (1996:187) to conclude that the Russian alveolar lateral approximant is distinguished by a retracted tongue dorsum and the palatal lateral by a raised and fronted tongue blade. They also observe an apical versus laminal articulation respectively for the Russian alveolar lateral approximant and palatal lateral, which contrasts directly with observations made here of x-ray tracings made by Fant (1960).

2.2.2 Palatography and electropalatography (EPG)

The earliest study of the palatal lateral linguopalatal relationship utilized palatography to obtain information regarding linguopalatal contact. Palatograms for the historical Spanish palatal lateral show extended palatal contact (Navarro Tomás, 1968). This finding is corroborated by more recent electropalatography (EPG) studies on the Catalan palatal lateral (Recasens, 1984b; Recasens et al., 1993; Recasens and Pallarès, 2001; Recasens and Espinosa, 2006). Descriptions of the palatograms (Navarro Tomás, 1968) by Lipski (1989) claim that the palatal lateral, palatal nasal, and palatal approximant are nearly identical in terms of linguopalatal contact, with the palatal lateral encompassing the areas of contact occupied by both the alveolar lateral approximant and the voiced palatal fricative /ʝ/ .

An EPG study comparing the Catalan palatal sounds (i.e. the palatal approximant, palatal nasal, and palatal lateral) reveals that the more tongue dorsum contact registered against the palate, the less coarticulation that occurs between the consonant and the adjacent vowel (Recasens, 1984b). Findings indicate that tongue dorsum contact decreases in the following order: palatal approximant, palatal nasal, palatal lateral, and alveolar nasal. These results suggest that the palatal lateral is more susceptible to the influence of neighboring vowels and may be articulated with greater variability than other palatal sounds.

A similar study (Recasens et al., 1993) finds that Italian palatal consonants are articulated at a more anterior place of articulation than their Catalan counterparts, with the Italian geminate palatal consonants demonstrating increased linguopalatal contact and de-

Vowel context	Catalan ɲ	Italian ɲ
i	92.6	164
a	85.2	142.7
u	85.6	187.3

Table 2.4: Reproduced mean closure durations for intervocalic syllable-initial Catalan and Italian palatal nasals in ms (Recasens et al., 1993).

creased coarticulation at the front palatal zone. Their findings support the hypothesis that increased closure duration results in increased linguopalatal contact, which in turn results in greater resistance to coarticulation. Mean closure duration is reproduced in Table 2.4, with closure defined as the achievement of the maximum activated electrodes. EPG results reveal that the closure location for both the Italian and Catalan palatal lateral and palatal nasal is alveolopalatal (i.e. encompassing the postalveolo-prepalatal regions) and not truly palatal. This study (Recasens et al., 1993) defines the postalveolo region as the third and fourth row on the artificial palate (Reading EPG system), prepalatal as the fifth and sixth row, and palatal as the sixth through eighth row. Both sounds are articulated with a single articulator, (lamino)predorsal - i.e. articulated with the back of the tongue blade and/or predorsum, and are not considered complex sounds. Linguopalatal contact is smaller and more anterior for the palatal lateral than the palatal nasal; the authors attribute this to manner requirements, suggesting that the tongue body must be more anterior to allow airflow to escape over the sides of the tongue.

Another EPG study by Recasens and Espinosa (2006:312) finds that the palatal lateral in Majorcan Catalan demonstrates greater coarticulation with the adjacent vowel when utterance-initial, suggesting that the palatal lateral is less able to resist coarticulatory influences when syllable-initial as compared to when it is in the coda. Their results corroborate previous findings (Recasens et al., 1993; Recasens and Pallarès, 2001) regarding the place of articulation of palatal sounds in Catalan: the palatal lateral in Catalan is articulated further forward than the other Catalan palatal consonants [c] and [ɲ], which demonstrated more contact along the hard palate - i.e. up to row 8 (the most posterior row) on the artificial palate (Recasens and Espinosa, 2006:305-306).

Recasens and Espinosa (2006) also report mean durations of closure (defined as the

Syllable position	Vowel context	ʎ	ɲ
Initial	i	190.29	177.14
Initial	a	160.88	146.36
Initial	u	158.57	179.12
Intervocalic	i	63.14	127.94
Intervocalic	a	66.86	60.00
Intervocalic	u	65.00	64.86
Final	i	167.94	157.94
Final	a	186.00	151.71
Final	u	160.88	147.65

Table 2.5: Reproduced mean closure durations for Majorcan Catalan /ʎ/ and /ɲ/ in ms (Recasens and Espinosa, 2006:309).

duration for which the palate indicates rows showing more than 80% electrode activation), reproduced here in Table 2.5. An Analysis of Variation (ANOVA) reveals significant effect of syllable position, vowel context, and phone identity on duration. Compared to earlier findings by Recasens et al. (1993), the more recent findings from Recasens and Espinosa (2006) of the Catalan palatal nasal bear a stronger resemblance to the Italian palatal nasal. This discrepancy in average durations is probably due to how closure was calculated in the two studies; the duration for which maximum electrode activation (see Recasens et al. (1993)) is maintained is likely shorter than the duration for which rows maintain more than 80% electrode activation.

A study comparing the palatal nasal in BP and PS (Shosted et al., 2012b) finds that the two sounds are articulated differently. The PS palatal nasal is consistently articulated with an occlusion at the alveolar and front palatal area similar to observations of the Catalan palatal nasal by Recasens et al. (1993) and Recasens and Espinosa (2006), while the BP palatal nasal is typically articulated with an open channel (i.e. lack of complete occlusion) similar to that of an approximant. Whether this is also the case for the BP palatal lateral is unknown, due to the lack of EPG studies on the BP palatal lateral.

While EPG is effective at capturing linguopalatal contact, it is unable to determine tongue position when there is no palatal contact, as in situations like the BP palatal nasal (Shosted et al., 2012b). Methods such as electromagnetic articulography (EMA) will be useful for investigating whether the palatal lateral in BP is similar to that of the palatal nasal, manifesting incomplete occlusion against the hard palate.

2.2.3 Electromagnetic Articulography (EMA)

An electromagnetic articulography (EMA) study comparing Russian and Catalan finds a much longer lag between when the tongue tip and tongue dorsum articulation is achieved for the Russian palatalized nasal (approximately 35 ms) than the Catalan palatal nasal (approximately 15 ms) (Recasens and Romero, 1997). Results indicate that the Russian palatalized nasal is similar to a [nj] sequence, except with a shorter lag between when the second most anterior sensor (Tongue Lamina (TL)) and most anterior sensor (Tongue Tip (TT) achieve their maxima. Recasens and Romero (1997) conclude that there is a difference in articulatory complexity between the Russian palatalized nasal and palatal nasal in Catalan. This suggests that there may be a difference between palatalized and palatal nasals, even when the palatalization occurs phonemically - according to the definition of the palatal lateral in this dissertation, the Russian palatalized nasal should also be considered a palatal nasal. Regardless, there is no indication yet that this same distinction between palatalized and palatal nasals also exists for palatalized and palatal laterals, or whether this distinction is simply an effect of language. However, this study was conducted on only two subjects in total, with only a single speaker of Catalan and a single speaker of Russian. A larger scale study may reveal greater insights into how palatalized and palatal sounds are articulated differently.

An EMA study on Russian alveolar and palatalized liquids (Kochetov, 2005) finds that the Russian so-called palatalized lateral has a higher and more retracted TT than the alveolar lateral approximant, with an average vertical difference of 5 mm and horizontal difference of 6 mm. The more posterior tongue tip articulation is suggested to be a result of the tongue blade raising to the palate to create the palatalized variant. The alveolar lateral approximant is also reported to have a longer TT constriction duration. Results for the so-called palatalized lateral reveal that the third most anterior tongue sensor (TB) achieves its target and is released nearly at the same time as the most anterior tongue sensor (TT), with an average achievement lag of about -5 ms (i.e. the time between when the maxima of each sensor is achieved) and release lag of about 9 ms (i.e. the time between when the constriction is released at each sensor). The length of the achievement lag reported for this study is noticeably shorter than the lag reported for the Russian palatalized nasals (Recasens and Romero, 1997), indicating that the nasal may not be an exact parallel to the lateral in Russian. Finally, Kochetov (2005) reports that greater articulatory differences and the resulting increase in acoustic differences between the alveolar lateral approximant and its palatalized counterpart produce a contrast that is more "phonologically stable" in comparison to the rhotic and its palatalized counterpart, which demonstrate smaller articulatory differences in the study.

2.2.4 Ultrasound

An ultrasound study by Gick et al. (2006:63) categorizes the Korean liquid as a palatal lateral with two gestures: tongue tip closure and raising/fronting of the tongue body. While traditional impressionistic accounts of Korean describe the liquid as an alveolar lateral approximant that appears syllable-final or as an alveolar tap appearing as its allophone in the onset (Gim, 1971; Iverson and Sohn, 1994), ultrasound images of the Korean liquid by Oh (2002) and Oh and Gick (2002) bear an undeniable resemblance to magnetic resonance (MR) images of the European Portuguese palatal lateral (Martins et al., 2010). These images of the Korean liquid (Oh and Gick, 2002; Oh, 2002; Gick et al., 2006) reveal that when syllable-final, the sound is articulated laminally with a high front tongue body.

Another ultrasound study of the so-called palatalized lateral in Russian (Proctor, 2009) indicates a similar articulatory configuration; a high tongue body and downward-pointing tongue tip is present in the sound for all vowel contexts collected (i.e. /e, u, a, i/). This study suggests that the tongue body is more highly constrained (i.e. less variation in possible tongue shapes observed across vowel contexts) during the production of palatalized liquids than palatalized alveolar stops. Proctor (2009) argues that this is evidence of complexity in palatalized liquids; a more constrained tongue body is necessary to produce two different tongue gestures - palatal approximation as a result of palatalization and an anterior dorsal gesture preserved from the non-palatalized liquid counterpart.

Recasens and Rodríguez (2015) fit a Smoothing Spline ANOVA (SS-ANOVA) to the tongue contour and divided the lingual spline into four zones (alveolar, palatal, velar, and pharyngeal) in order to calculate coarticulatory resistance in Catalan consonants. Each zone was defined by a polygon, incorporating the confidence intervals produced by the SS-ANOVA; the area of each polygon was then used to compare the amount of coarticulation (which would presumably manifest as greater variability) between the speech sounds to adjacent vowels. Results indicate that palatal consonants (including the palatal lateral) are the most resistant to coarticulation in comparison to alveolar sounds and trills, demonstrating very little variability at the palatal, velar, and pharyngeal zones. Like Proctor (2009), Recasens and Rodríguez (2015) suggest that the tongue body is highly constrained during the production of palatal sounds.

2.2.5 Magnetic Resonance Imaging (MRI)

While there is a distinct dearth of literature on the articulation of the palatal lateral in the previous articulatory methodologies, more recently there has been a surge in the study of the European Portuguese palatal lateral using magnetic resonance imaging (MRI) (Martins et al., 2008, 2010; Teixeira et al., 2012). An early study from Martins et al. (2008) on a single speaker of European Portuguese indicates a more palatoavleolar place of articulation for the Portuguese palatal lateral, which is articulated at a more anterior place than the European Portuguese palatal nasal. This corresponds to previous findings from EPG studies of Catalan and Italian (Recasens et al., 1993; Recasens and Pallarès, 2001; Recasens and Espinosa, 2006).

In a larger scale MRI study involving seven speakers of European Portuguese, Martins et al. (2010) find a high degree of inter-speaker variability in the production of the alveolar lateral approximant and the palatal lateral, with more variability observed in the alveolar lateral approximant than the palatal lateral. This study finds that one out of the seven speakers articulated the palatal lateral exclusively at the palatal region. Corresponding to previous EPG studies (Recasens, 1984b; Proctor, 2009), results reveal that vowel context only slightly influences the articulation of the palatal lateral. Any variation that occurred in the tongue shape always occurred behind the occlusion formed by the tongue against the palate or around the tongue root; both locations which are not specified by the articulatory requirements for the palatal lateral.

Another study of European Portuguese laterals (Teixeira et al., 2012) reports that both the palatal lateral and alveolar lateral approximant are articulated with a lateral compression of the tongue. Of the seven speakers in this study, a single lateral channel was observed in the coronal view for only one speaker, while the remaining subjects manifested two moderately asymmetrical lateral channels. Additionally, midsagittal views of the posterior tongue body reveal a convex tongue body shape for both the palatal lateral and alveolar lateral approximant. Midsagittal MR images (Teixeira et al., 2012) reveal that the palatal lateral in European Portuguese is produced with a complete occlusion at the alveolopalatal region using the tongue blade and/or pre-dorsum. The length of lingual contact is approximately 0.8 to 2.4 centimeters across all vowel contexts (i.e. /i, a, u/) and speakers. This is consistent with previous findings (Oliveira et al., 2011) that report greater contact for the European Portuguese palatal lateral than the alveolar lateral approximant. The tongue anterior is always high when articulating the palatal lateral, with the tongue tip tucked behind the lower incisors. Some images indicate that tongue root retraction and pharyngealization is occasionally present, though it is unclear whether this is speaker-dependent or resulting from coarticulation with the adjacent low back vowel.

2.3 Acoustics of the palatal lateral

Due in part to the infrequent appearance of the palatal lateral cross-linguistically, there are relatively few accounts of its acoustic characteristics. Table 2.6 compares reported average formant frequencies for a number of languages. The contents of this table serve to highlight acoustic similarities of the palatal lateral to the palatal approximant.

According to the acoustic model of speech production (Stevens, 2000:513,526-527), a sound articulated with a palatal constriction is expected to manifest a high second formant frequency with a Helmholtz resonance for the first formant frequency. Because of the narrowing at the oral cavity anterior, the third formant frequency increases and draws closer to the fourth formant frequency. As an approximant, turbulence is unexpected as the constriction is too wide to result in a significant drop in average pressure.

An acoustic model of lateral production is more complex (Stevens, 2000:543-547), as lateral sounds can be produced with lateral contact alone one side of the tongue, or along the midline of the tongue with no contact along the lateral edges of the tongue. Considering a lateral sound as produced with lateral contact along only one side of the tongue is modeled by interconnecting resonators, the lateral airstream that flows over the non-occluded side of the tongue is modeled as a side branch (Fant, 1960; Stevens, 2000). The length of this side branch varies; Stevens (2000) estimates a side branch of approximately 2.5 cm in length and a volume of 2 to 3 cm for the English alveolar lateral approximant, while estimations based on the same sound in Russian by Fant (1960) are slightly larger and longer. This side branch adds a zero to the vocal tract transfer function, thereby modeling the antiresonance that arises from a lateral articulation (Stevens, 2000). Modeling the effects of the side branch is important as the presence of an antiresonance affects the acoustic signal by attenuating overlapping formants (Johnson, 2011). According to calculations by Stevens (2000), the frequency for this zero lies within a range of 3500 and 4000 Hz, with the lowest possible antiresonance in the range of 2200 to 4400 Hz. However, the frequency of this zero can vary according to vowel environment and speaker-intrinsic differences in size or shape of the oral tract. Since small changes in the zero results in large changes in spectral peak amplitudes, substantial variation in the spectral shape can be expected in the frequency range covered by the zero.

	Brazilian Portuguese ʎ	Corrientes Spanish ʎ	Catalan ʎ	Russian ʎ / lʲ	Italian ʎ	Catalan j	English j	English l
F1	300	377.6	–	210	280	–	260	360
F2	1870	1816.6	1600-2000	1700	–	1925-2150	1770	1100
F3	2900	2549.1	–	2500	–	–	2950	2800

Table 2.6: Average formant frequencies as reported in Hz for BP (Silva, 1999), Corrientes Spanish (Colantoni, 2004), Catalan (Recasens, 1984b), Italian (Vagges et al., 1978), Russian (Fant, 1960:167), and English (Stevens, 2000:516, 526, 533, 546).

Recasens (1984b) finds a large range of F2 for palatal consonants in Catalan (as produced by a single male speaker), with minimum values around 1600 Hz and maximum values around 2400 Hz. This is consistent with the second formant frequencies of the palatal sounds reported in Table 2.6; reported F2 values for the palatal laterals and palatal approximant fall within this range, while the alveolar lateral approximant does not. Ladefoged and Maddieson (1996) cite previous research on Breton (Bothorel, 1982) indicating that the Breton "laminal post-alveolar lateral" demonstrates a lower range of F1 than the apical alveolar lateral, which deviates from the F1 value reported for the palatal lateral in Corrientes Spanish (Colantoni, 2004) but holds true for the Russian lateral (Fant, 1960) and Italian palatal lateral (Vagges et al., 1978). This is consistent with observations of Bulgarian (Tilkov, 1979) referenced by Ladefoged and Maddieson (1996), for which the so-called palatalized lateral is reported to have an F1 that is 100 Hz to 150 Hz lower than the plain apically-articulated lateral. An observation made by Ladefoged and Maddieson (1996:197) regarding the proximity of F1 and F2 for laminal laterals does not apply to palatal laterals, as there is an approximately 1500 Hz difference between the F1 and F2 for the Russian lateral, Corrientes Spanish palatal lateral, and BP palatal lateral. While the authors suggest that the first and second formant frequency are closer for laminal laterals than apically-articulated laterals, they do not quantify the amount of proximity.

Ladefoged and Maddieson (1996:193) suggest that laminal sounds, including the palatal lateral as defined here, can sometimes have slower transitions than apical sounds. Slower transitions suggest that laminal articulations take longer to achieve than apical articulations. In a study of Corrientes Spanish, Colantoni (2004) reports longer transitions for the Corrientes Spanish palatal lateral than the PS (Quilis et al., 1979) and BP (Silva, 1999) palatal laterals, while being 10 ms shorter than the palatal approximant in PS (Aguilar, 1997). Colantoni (2004) concludes that longer transitions are a potential factor for why *yeísmo* occurs, suggesting that the long transitions are misinterpreted as perceptual cues for an approximant.

Similarly, Colantoni (2004) reports a consensus in the literature that palatal laterals are long segments; including other liquids, the palatal lateral is identified as the longest consonant cross-linguistically. The intervocalic palatal lateral is reported to be 20% longer than the intervocalic alveolar lateral approximant in Spanish, with the transition from the palatal lateral to the following vowel nearly double the length of the alveolar lateral approximant (Lipski, 1989). Comparatively, the Spanish palatal nasal is 10-15% longer than the alveolar nasal, with formant transitions from the palatal nasal to the following vowel approximately 50% longer than the alveolar nasal (Lipski, 1989). Colantoni (2004) finds that the palatal lateral in Corrientes Spanish has a mean duration of 70 ms with a standard deviation of 20 ms, which is slightly longer than in other Spanish dialects. Silva (1999:57) reports that the duration of the palatal lateral in BP is 101 ms when word-initial and 78 ms when intervocalic.

Statistical analysis of Corrientes Spanish palatal lateral formant frequencies in stressed and unstressed syllables reveals that the first formant frequency is significantly higher in unstressed syllables, though the second and third formant frequencies are unaffected (Colantoni, 2004). This suggests that stress has a significant affect on tongue height during the production of the palatal lateral. Colantoni (2004) finds a positive correlation between higher F2 values and higher rates of maintenance (i.e. resistance against using the palatal approximant allophone of the palatal lateral), suggesting that speakers who are more likely to produce the palatal lateral as a palatal approximant are articulating the sound with a more retracted tongue body.

A study of Jebero (also referred to as Shiwilu) phonology touches briefly upon the acoustics of the palatal lateral in this endangered language spoken in Peruvian Amazonia (Valenzuela and Gussenhoven, 2013). The study reports an average duration of 170 ms for the Jebero palatal lateral, which is similar to the values reported for the word-initial and word-final Catalan palatal lateral (Recasens and Espinosa, 2006). The approximants /l, w, j/ have a mean duration of 112 ms in Jebero, corresponding to Colantoni (2004)'s observation that the palatal lateral is longer than other liquid sounds. Presumably, the duration of the Jebero palatal lateral was measured based upon the acoustics (in contrast with Recasens and Espinosa (2006), who report closure in terms of articulation), however no details are provided as to how phone duration was determined.

Valenzuela and Gussenhoven (2013) report that the palatal lateral is articulated with the tongue tip to anterior tongue blade raised towards the palate and the tongue tip placed behind the upper incisors. It is occasionally articulated with some lateral frication. However, this description is based off of the authors' impressionistic observations and not articulatory data. A figure comparing the formant frequencies of Jebero approximants reveals that the palatal lateral has a low F1 around 300 Hz, high F2 around 2500 Hz, and a relatively high F3 that moves from around 3000 Hz to 4000 Hz. The observed low F1 and high F2 would result in a difference of approximately 2200 Hz, which is 700 Hz larger than the approximate F2-F1 difference observed in the palatal laterals reported in Table 2.6.

2.4 Perception of the palatal lateral

Since descriptions of how the palatal lateral is perceived are infrequent in the literature, discussions of similar sounds (e.g. other palatal or lateral sounds, as well as palatalized sounds) will be included here as a means for comparison. There is also a sizable body of literature regarding the perception of the alveolar lateral approximant in English which will be briefly discussed.

Recasens et al. (1995) cites previous EPG studies of the palatal nasal in Catalan (Recasens, 1984a) and palatal nasal, palatalized nasal, and alveolar nasal in Irish (Recasens et al., 1991) to provide a gestural account for the difference in the perception of palatal and palatalized sounds. Recasens et al. (1995) utilize this articulatory information to propose that palatal sounds are simple segments while palatalized consonants are complex sounds composed of two gestures, one of which is the palatal gesture. The authors suggest that the overlap between the two gestures in palatalized sounds is not complete, resulting in the clearer presence of a palatal approximant during the production of palatalized sounds. However, this study does not provide clear evidence of whether a longer lag between the two gestures truly correlates with an increase in the perception of a palatal lateral.

Guion (1998) investigates the role of perception in the palatalization of the voiceless velar stop before front vowels, a common cross-linguistic sound process that produces the palatoalveolar affricate. Results from English-speaking participants support the hypothesis that velar palatalization is a byproduct of what occurs when listeners assess fast speech; Guion (1998) finds that velars preceding front vowels resemble a palatoalveolar both in terms of acoustics and perception. There is a greater overlap in peak spectral frequencies between [k] and [tʃ] before front vowels, but not before back vowels, an effect which is exacerbated when participants produced the same stimuli at a faster speaking rate. A forced-choice perception task presented the voiceless velar stop and voiceless palatoalveolar affricate before /i, a, u/ in two conditions (the first being the entire consonant with the vowel portion digitally removed and the second containing only the first 30 ms of the consonant); results from this task indicate that participants accurately identify the consonant involved at above 90% accuracy except for when the velar stop preceded the high front vowel. In this context and exclusively in this context, participants performed at approximately chance level, identifying the velar stop incorrectly as the palatoalveolar affricate 47% of the time. These findings suggest that velar palatalization is perceptually motivated.

A study of the perception of Russian palatalized and non-palatalized contrasts in labials (/p/ and /p^j/) and coronals (/t/ and /t^j/) by both Russian-speaking listeners and Japanese-speaking listeners (Kochetov, 2004) identifies certain characteristics that appear to be global and not language specific. There is an effect of syllable-position on perception, with both Russian and Japanese listeners performing better when the contrasts were placed in the onset position as opposed to the coda. Results also revealed longer reaction times when the contrasts were presented in the coda position. This effect extends to both the palatalized and non-palatalized sounds.

In particular, Kochetov (2004) finds that both Russian and Japanese listeners were more accurate in identifying the segments /p/ and /t^j/ than /p^j/ and /t/. This asymmetry in phone identification is interesting, as one would expect the division to be between either palatalized sounds (i.e. /p^j/ and /t^j/) or place of articulation (/p/ and /t/). While the difference in performance was moderate, the Russian listeners were significantly more accurate than the Japanese listeners; furthermore, the Japanese listeners performed above chance level on stimuli that were phonotactically impossible in their native language. Kochetov (2004) takes this as evidence for language-independent syllable position effects on the perception of the plain-palatalization contrast, resulting in asymmetries in the accurate identification of consonants within this contrast.

Multiple studies have investigated the perception of the alveolar lateral approximant in English across several different varieties and by native and non-native listeners. Some acoustic features which have been identified as perceptual cues for the correct identification of the alveolar lateral approximant include transition lengths (O'Connor et al., 1957b), overall duration (Fant, 1960:167), presence of a transient at onset (Fant, 1960:167), as well as the first (Fant, 1960) and the third formant frequency (Yamada and Tohkura, 1992).

The importance of the first and third formant frequency in the identification of the alveolar lateral approximant may differ according to the English variety spoken by the listener. American English-speaking listeners attend to the third formant frequency as the strongest cue for the presence of an alveolar lateral approximant (Yamada and Tohkura, 1992; Lotto et al., 2004), while listeners who speak Standard Southern British English (SSBE) have been shown to switch between the first and second formant frequency as the most important cue (Knight et al., 2008). This directly contradicts with American English speakers, for whom the second formant frequency is an unreliable cue (Iverson et al., 2005:3267). It is unknown whether sociolinguistic factors influence which formant frequency listeners attend to when identifying the palatal lateral.

2.5 Summary

Accounts[4] of the diachronic development of the palatal lateral strongly suggest that cross-linguistically, the palatal lateral typically results from an alveolar lateral approximant occurring in a palatalizing context. At some point, the palatalizing context disappears and the palatalized lateral enters into the phone inventory. Sound changes that occur to the palatal lateral are similar across the languages reviewed here as well; the palatal lateral typically loses its lateral manner of articulation and merges to another palatal sound, which may be articulated with or without an additional constriction of the airflow resulting in frication, e.g. as in the case of Hungarian or Spanish. Within the last 100 years, only one language was found to merge the palatal lateral with a non-palatal sound (Hungarian). Attempts to explain why this phenomenon occurs have been unsatisfactory, largely relying on speculation.

[4]Note that the historical development of the palatal lateral is better documented in some languages than others. In particular, the majority of languages with a documented development and/or loss of the palatal lateral are primarily Latin in origin.

The articulatory studies reviewed here provide a foundation from which to begin a detailed articulatory analysis of the palatal lateral in BP. However, with the exception of the recent MRI studies, none of the articulatory studies were focused on the Portuguese palatal lateral and only a few studies (namely, Silva (1999) and Colantoni (2004)) discuss the acoustics of this sound. While an EPG study (Shosted et al., 2012b) of the palatal nasal in PS and BP find that the palatal nasal is fully occluded in the former and approximated in the latter, it is unknown whether the BP palatal lateral is also produced without complete occlusion as well. None of the studies on the Portuguese palatal lateral were able to collect high-quality acoustic information paired with articulatory data. This dissertation seeks to add to the body of literature on the palatal lateral by providing a detailed analysis of the articulation and acoustics of the palatal lateral in BP by utilizing EMA to collect a large corpus of high-quality acoustic and articulatory data from 10 native speakers of the language.

Contrary to stereotypical IPA descriptions of place of articulation (Ladefoged, 1988), the palatal place of articulation is much less precise in reality. In a cross-linguistic discussion of laterals, Ladefoged and Maddieson (1996) distinguish between pre-palatal and laminal post-alveolar (also referred to as palato-alveolar), while Recasens (2013) argues for distinguishing between an alveolopalatal and palatal place of articulation in the IPA. However, a language that contrasts palatal and alveolopalatal sounds has yet to be found, so including this distinction in the IPA may be unnecessary. This dissertation will include sounds articulated both at the palatal and alveolopalatal regions, which may reveal whether these sounds behave differently with regards to articulation or acoustic features (differences in perception will be investigated in the appendix).

The literature on the articulation and acoustics of the palatal lateral is sparse, especially in comparison to the literature on the alveolar lateral approximant. Furthermore, there are only a handful of articulatory or acoustic studies specifically focused on the palatal lateral in Portuguese (Martins et al., 2008; Teixeira et al., 2012), and only one on the palatal lateral in BP (Silva, 1999). To the extent of the author's knowledge, the literature on the perception of the palatal lateral is nonexistent. From the current state of the literature, it is difficult to establish a cross-linguistic description of the articulation and acoustics of the palatal lateral that is both comprehensive and, regarding the difference between a palatalized and palatal lateral, unambiguous. This dissertation will provide a detailed account of the articulation and acoustics of the palatalized and palatal lateral in BP, which can be used as a starting point for future cross-linguistic comparisons.

Chapter 3

STUDY 1: ELECTROMAGNETIC ARTICULOGRAPHY (EMA)

3.1 Introduction

This study is designed to provide a detailed description of the articulation of the Brazilian Portuguese (BP) palatal lateral, in addition to identifying articulatory characteristics that distinguish the palatal lateral from similar sounds in BP. Articulatory data was collected using electromagnetic articulography (EMA).

3.2 Methodology

3.2.1 Participants

Ten female speakers of BP from São Paulo state were recruited for the purpose of this study. Nine were currently residing in Urbana-Champaign, Illinois, while one was visiting from São Paulo. Their ages ranged from 20 to 55 years old, with an average age of 33.9 years. Participants included exchange students, instructors, and homemakers. They were compensated $25 for their participation. Participants will be referred to as BP 1-10.

3.2.2 Materials

Participants were asked to produce a set of words in the following carrier phrase: *Diga* ___ *para nós* [ʤi.gɐ ___ pa.rɐ nɔʃ] 'Say ___ four times'. Since BP phonotactics do not permit the palatal lateral syllable-finally and since the palatal lateral rarely appears word-initially in Portuguese (Mateus and d'Andrade, 2000:11-12) (see Section 2.1.2 for more information

Palatal	**Plain**	**Palatal-ized**
[ˈba.ʎa] *bálha*	[ˈpa.la] *pála*	[ˈpa.lʲa] *pália*
[ˈpa.ɲa] *pánha*	[ˈpa.na] *pána*	[ˈpa.nʲa] *pánia*
[ˈpa.ja] *páia*		

Table 3.1: List of nonsense words for production task. Phonetic representation in brackets [], orthographic representation in italics.

on the distribution of the palatal lateral and other target sounds in this study), the word list (see Table 3.1) was composed of nonsense words that contain the target consonant in a syllable-initial and word-medial context; i.e. [ˈpa.Ca], where 'C' represents the target consonant. Each nonsense word began with the voiceless bilabial plosive [p], which provided a clear boundary between the carrier phrase and the nonsense word (a useful visual aid when annotating). However, voiced [b] was used instead of voiceless [p] for the nonsense word *bálha*, because the word *palha* [ˈpa.ʎa] 'straw' is a real word. The first syllable of each nonsense word was always stressed.

The word list was randomized within block. Each block was presented 40 times, resulting in a total of at least 40 repetitions per word. If malfunctions were apparent during the recording (e.g. sensors flickering from green to red), the speaker was asked to repeat the problematic word before moving on.

3.2.3 The NDI Wave Speech Research System

Articulatory data was collected using the NDI Wave Speech Research System (NDI; Waterloo, Ontario, Canada) located in the NeuroSpeech laboratory in the Speech and Hearing Science building at the University of Illinois at Urbana-Champaign. The NDI Wave system tracks the position of a set of disposable sensors in an electromagnetic field produced by a field generator, at a sampling rate of 100 Hz. As many as 16 sensors can be connected through a set of wires to a system control unit (SCU), which records the sensor data and transmits it to the computer. Sensors with five degrees of freedom are used, which permits the capture of 3D position information (x, y, and z coordinates) in addition to rotation information (pitch and roll). Audio was collected simultaneously by a Countryman Isomax E6 head-mount microphone, so that acoustic and articulatory events may be correlated.

By tracking the position of sensors over time, EMA is able to provide information regarding the timing, velocity, direction, and magnitude of movement. It can also capture, to some degree of accuracy, the amount of contact between the tongue and the hard palate. An approximation of the front portion of the hard palate can be obtained by having the subject trace the roof of her mouth with the sensor close to the tip of her tongue. The contour can then be used to estimate how much of the tongue blade touches the hard palate during speech.

Placement of sensors

While the tongue is extended, three sensors are placed on the tongue midline: one sensor is placed 1 cm from the tongue tip (TT), with subsequent sensors following 1 cm apart (TM and TB respectively). Two sensors are placed on the vermilion border; one sensor on the upper border and one on the lower. Four reference sensors are placed: one sensor on the nose bridge, one on the chin, and two more adjacent to one another on top of the right zygomatic process. A total of nine sensors are used. Sensors are temporarily adhered to the surface of the tongue using Histoacryl Topical Skin Adhesive (TissueSeal), which is a medical grade topical adhesive, and to the surface of the face using medical tape. At the end of each recording session, sensors exposed to saliva were discarded for hygienic reasons.

3.2.4 Data processing

Prior to data analysis, a suite of MATLAB (2014) scripts[1] (Wong and Hermes, 2015) was used to rotate the raw data to the occlusal plane, correct for head movement, and apply a

[1]Included in appendix.

2nd order low-pass Butterworth filter of 10 Hz with a normalized cut-off frequency of 0.2 Hz.

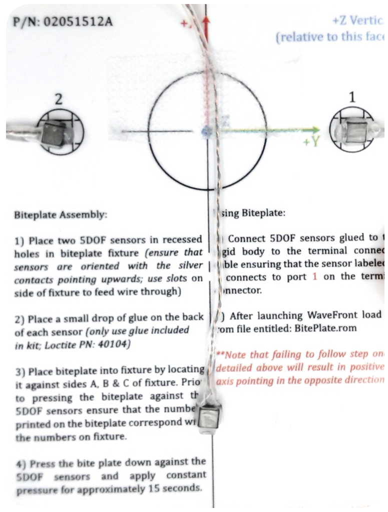

Figure 3.1: Image of bite plate with three sensors glued to surface.

The occlusal plane (or bite plane) is defined by the front incisors and the back molars (Mosby, 2013). Rotating the head to the occlusal plane allows data from different speakers to be compared, as it normalizes the movement of the tongue by orienting the speaker's oral tract in the same space as another speaker. During data collection, information regarding the occlusal plane is acquired by gluing three sensors on a bite plate provided by NDI and inserting the bite plate into the participant's mouth. The position of the three sensors is used to calculate the position and orientation of the notional plane of the bite plate. The bite plate is inserted until the front incisors are between the two anterior sensors (see Figure 3.1) and the participant is able to bite down on the bite plate with the back molars. When completely inserted, the front incisors are between sensors 1 and 2 and the posterior sensor (unnumbered) is between the back molars. The sensors are placed in the same position every time, guided by the markings on the bite plate. A new bite plate is used for each participant. After the bite plate sweep(s) acquisition, the bite plate is removed and the participant may speak normally.

The data from the bite plate sweep is used to calculate the position and orientation of the (notional) occlusal plane, which is used to reorient the head, essentially by centering and rotating the occlusal plane so that it is flat with respect to the three-dimensional space projected by the field generator. To do so, we assume that the unnumbered posterior sensor (see Figure 3.1) to be the center of the space occupied by the speaker's head and translate the raw data such that the posterior bite plate sensor is now at the origin of the coordinate system. In doing so, we perform a translation of the data, including data from all other sensors. Then, the angle between the position of the occlusal plane and the x-y plane is calculated to find the rotation matrix. The rotation matrix is then multiplied against all of the data from all of the sensors across every sweep, which effectively rotates the data such that the occlusal plane is lying flat on the x-y plane projected by the field generator.

Next, the data is corrected for head movement. It is important to differentiate this kind of correction from the reorientation relating to the occlusal plane. Head movement correction is performed to ensure that any movement observed in the tongue sensors' positions will be indicative of tongue/jaw movement and not from, say, movement of the entire head. For example, a sensor's forward movement may be ascribed to forward tongue movement, forward head movement, or both, in the absence of this type of correction. To correct for head movement, we subtract the "cheek bone" sensor position (i.e. the sensor glued at the zygomatic process region) at the point which the rotation matrix was calculated from the cheek bone sensor position at every sample point in a sweep. This creates a vector containing the change in position of the "cheek bone" sensor. This vector is subtracted from every other sensor in the sweep, which effectively removes the movement of the head by performing a translation of the data. This is repeated for each sweep; a unique delta vector must be calculated for each sweep.

Lastly, a low-pass 2nd order Butterworth filter with a cutoff frequency of 20 Hz is applied to remove high-frequency components from the signal. These components are presumably unrelated to speech. This type of filter was chosen as it provides minimal distortion of the information within the filter window of 0 Hz - 20 Hz. Before applying the Butterworth filter, the Fast Fourier Transform (FFT) was plotted for several randomly selected sweeps from each speaker and a cutoff frequency of 20 Hz was chosen. The cutoff frequency was normalized by multiplying the cutoff frequency by two and then dividing by the sampling rate. The Butterworth filter requires the signal to be forward and reverse filtered so that it is zero-phase filtered, which is done by using the *filtfilt*() function in MATLAB (2014). The first application of the filter creates a phase shift to the right (which causes the articulatory data to be out of sync with the acoustics), so a second application (reverse filtering) is required to reverse the shift. The double application of the filter creates a lower cutoff and sharper roll off, thus impacting more of the data at the edges of the filter cutoff. Therefore, a slightly higher cutoff was chosen, as compared to what would have been chosen for a 1st order Butterworth filter.

3.2.5 Identifying errors

Unlike the Carstens system, errors during recording using the NDI WAVE system do not manifest as large leaps in position. Therefore, measures such as the Root Means Square (RMS) measure (Carignan, 2013:46) are not appropriate when identifying errors in the data collected here. Using pilot data collected from four speakers, new measures were developed to identify errors produced by the NDI WAVE system. Presented below are the different types of errors observed and the measures that have been developed to identify each error.

Three types of errors have been observed: 1) No data, 2) Position, and 3) Velocity. The rate of errors for all three types appear to related to speaker- or acquisition- dependent factors, such as the presence and depth of a tongue groove or how well the sensors and wires are secured. It is possible that contact between sensors or disturbance of the wires (resulting in movement in the housing) may cause these momentary errors.

Error type I: No data

The simplest error observed were sample points for which no data is recorded, i.e. NaN (Not a Number). Empty data points can appear sporadically across both sensors and sweeps, but typically do not occur for extended periods of time, i.e. when sensors become worn out or are not securely installed. This type of error can be observed on the computer monitor when sensor icons flicker from green (fully operational) to red (out of field or loose wires).

Fortunately, the average duration of sequentially occurring empty data points is fairly low with an average duration of one to three data points (i.e. approximately 10 to 30 ms), depending on speaker. A script to replace each empty data point with the preceding value (Wong and Hermes, 2015) was implemented into the data processing suite.

Error type II: Position

In general, the NDI Wave system is highly accurate with regards to tracking the position of the sensor. Rigorous testing of the system's accuracy during speech reports that when within 200 mm^3 of the field generator at a 300 mm^3 volume setting, 95% of the data had an error margin of less than 0.5 mm (Berry, 2011). Occasionally however, sensors on the tongue appear in physiologically impossible positions, e.g. when the most posterior sensor is more anterior than the most anterior sensor. These errors are identified by determining whether the x-position of sequential sensors supersedes the position of the anterior sensor, i.e. if the TB sensor moves in front of the TM sensor in the x-dimension. This method assumes that the most anterior tongue sensor (TT) does not manifest position-type errors. This assumption is reasonable, as position errors were never observed for TT in the pilot data.

Errors in position were identified with regards to whether they occurred during the target speech sound. Four files with TM position errors were identified; these files were excluded completely from analysis. 278 files with TB position errors were identified; the majority of the errors were found in the data collected from BP 10 (252 files) and BP 3 (22 files). These files were included only for analyses that required only the first two sensors, e.g. the tongue blade angle. Files containing both TM and TB position errors occurring during the target speech sound were not found.

Error type III: Velocity

Like position errors, velocity errors are only apparent in tongue sensors. This type of error manifests as rapid shifts in velocity that are out of sync with other tongue sensors. Since velocity errors result in abrupt changes in position, all velocity errors are also position errors. As a result, they are identified using the same method as position errors and the percentage of position errors that result from velocity errors was not calculated. Presence of a velocity error can be confirmed by creating video animations of the problematic sweep in MATLAB (2014).

3.2.6 Resampling

Some of the following analyses require the tongue trajectories to be composed of the same number of sample points. For example, Smoothing Spline ANOVA (SS-ANOVA) requires contours to contain the same number of samples; unlike general additive models, SS-ANOVA cannot fit unbalanced data. In order to produce the following analysis of horizontal movement for each tongue sensor, the trajectories of each of the three tongue sensors for all 10 speakers were re-sampled using the *resample*() function in MATLAB (2014) so that all phones were represented by the same number of samples (See Figure 3.2).

This resampling does not apply an interpolation of the curvature formed by the three tongue sensors, as the interpolation is applied separately to the trajectory of each individual tongue sensor for the entire duration of the annotated phone by speaker. The default interpolation method 'linear' was used; *resample*() also applies an antialiasing FIR low-pass filter.

The function *resample*() assumes that the previous and following sample points of a given trajectory equal zero, which can result in warped trajectory edges. To reduce or remove the effect of the perturbations, the edges of each trajectory were padded with additional samples to avoid edge oscillations. To do so, two times the length of the original sample length for each individual trajectory was included to either side. After resampling, there was a total of 100 samples per trajectory. Proportionally, the number of padded resample points to be removed from either side was 40 samples. This resulted in a total of 21 samples per each newly resampled trajectory of any given tongue sensor throughout the duration of a single phone.

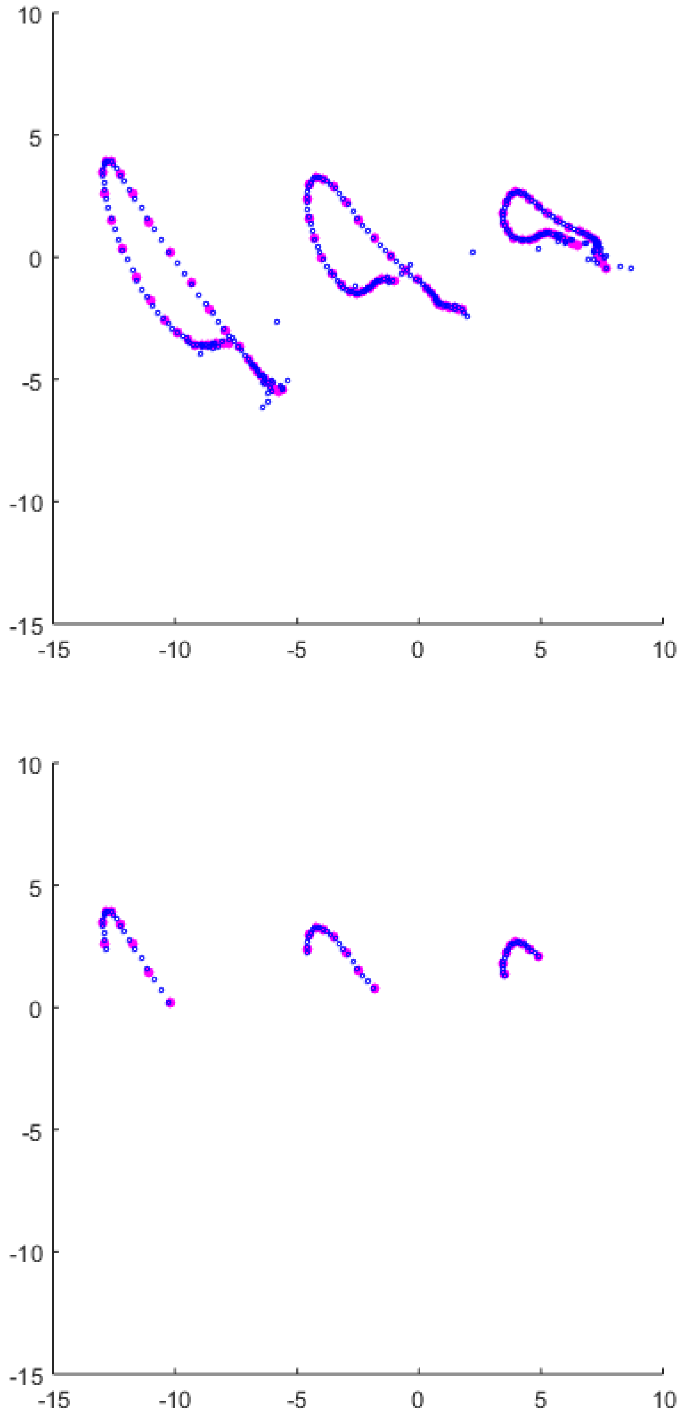

Figure 3.2: Tongue sensor trajectory during [l] produced by Speaker 1 (tongue anterior left, posterior right). Original sample points in magenta, resampled points in blue. Phone shown with padded samples (top) and without padded samples (bottom).

Checking a small sample set of the data confirmed that removing the padded samples also removed edge oscillations introduced by *resample*() (see Figure 3.2). Edge oscillations are clearly visible in Figure 3.2, left, with blue resampled points positioned far away from the magenta original sample points. However, once the padded samples are removed from the resampled trajectory, the introduced edge oscillations are also removed as well - resulting in a reasonably accurate fit of each sensor's trajectory during a single phone's duration.

3.3 Results

This section will begin with a catalog of general articulatory characteristics (e.g. tongue position, tongue blade angle), culminating in models of the shape and trajectory of the tongue anterior. Note that some figures (see Figure 3.10 for example) may refer to sounds /ʎ/, [l^j], /ɲ/, and [n^j] respectively by orthographic representations *lh*, *lj*, *nh*, and *nj* for the convenience of plotting.

3.3.1 Data annotation

The materials were designed specifically for ease of annotation. The target consonant was contained within a nonsense word that began with a bilabial stop and was followed by the word *para*, which also began with a bilabial stop. This facilitated quick identification of the VCV sequence containing the target consonant.

Prior to data analysis, the beginning and end of each phone was hand-annotated in Praat (Boersma and Weenink, 2014) by viewing the spectrogram and identifying abrupt drops in amplitude (see Figure 3.3 and Figure 3.4), with the assumption that constrictions between the tongue blade and hard palate or alveolar ridge (depending on phone identity) would result in reduced amplitude. This serves to identify when the associated articulations begin and end for each phone in the following analysis. Occasionally, the decrease in amplitude marking the phone boundaries was not obvious; in such cases, the boundaries of the formant frequency steady state were annotated (i.e. the portion of the formant frequency that is relatively stable and without change).

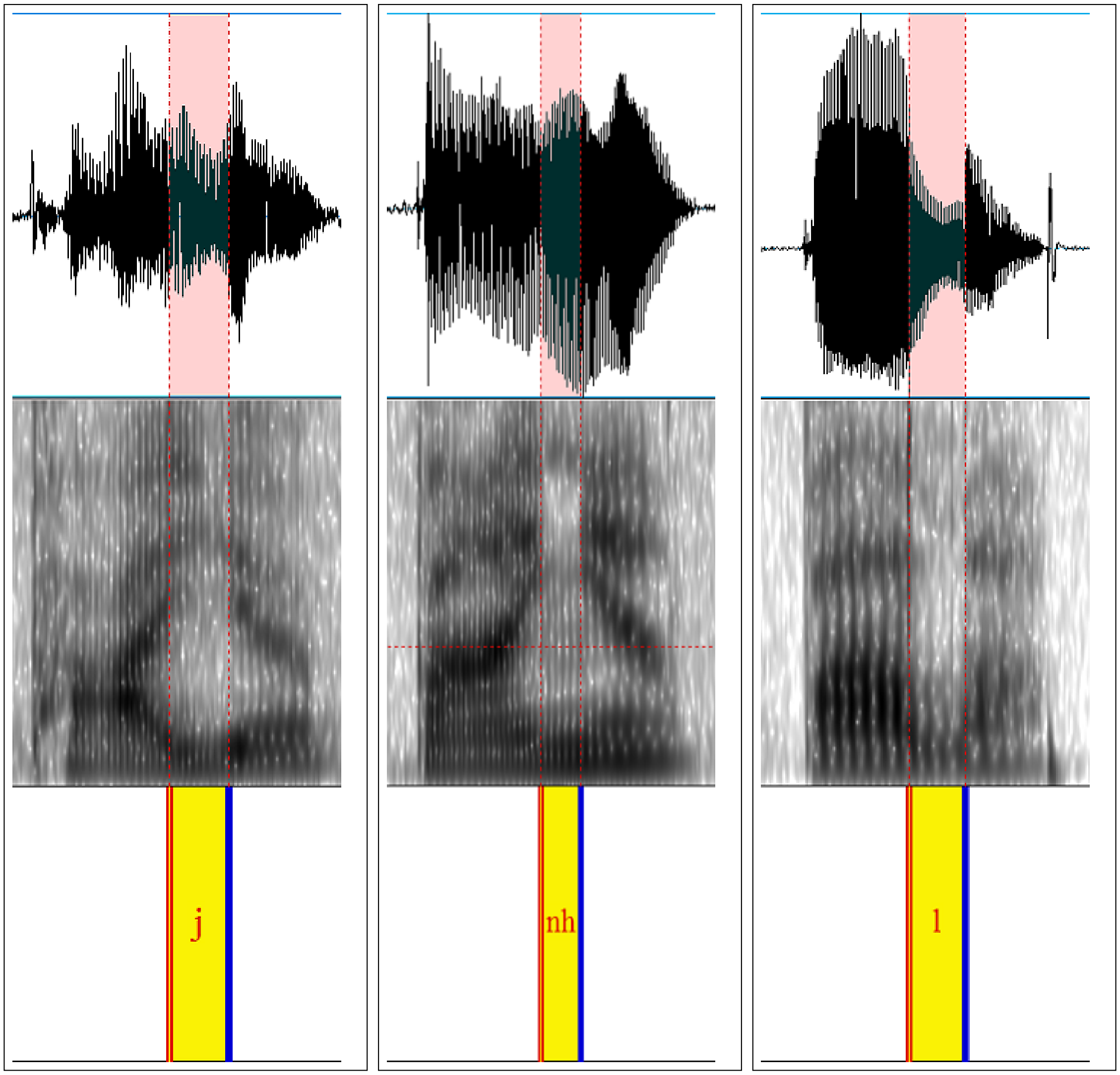

Figure 3.3: Annotated phone boundaries for a palatal approximant (left), palatal nasal (middle), and alveolar lateral approximant (right), as produced by BP 8 (left), BP 9 (middle), BP 2 (right).

Sample annotations of a palatal approximant, palatal nasal, and alveolar lateral approximant are presented in Figure 3.3; these sounds were produced by respectively by Speakers 8, 9, and 2. This figure illustrates productions where there is complete or nearly complete closure during the production of these sounds. Obviously, there is some variation in the degree of closure; sounds produced for this study ranged from complete closure to no closure (approximant).

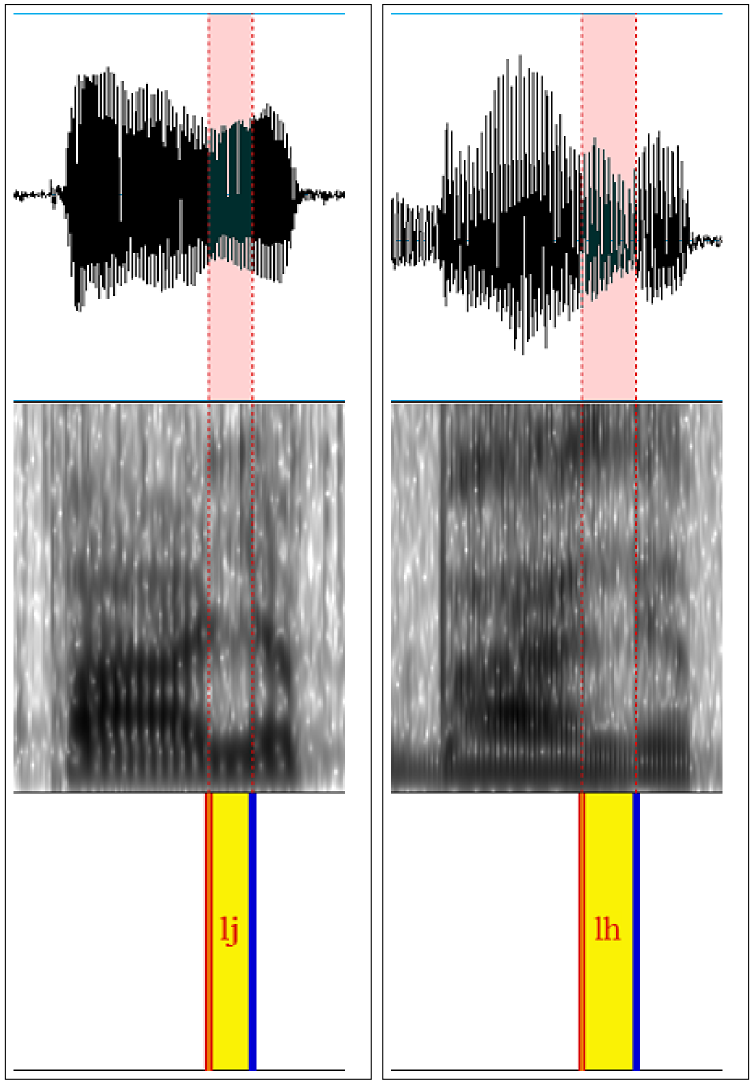

Figure 3.4: Annotated phone boundaries a palatalized lateral (left) and palatal lateral (right), as produced by BP 5 (left) and BP 3(right).

It was observed during the process of annotation that certain sounds were more likely to be articulated with an approximated closure. Note that the palatal approximant in Figure 3.3 is articulated with a complete or nearly complete closure. While the palatal approximant was more likely to be articulated with a more approximated closure, this sound was still articulated with a complete closure not infrequently.

Sample annotations of a palatalized lateral and palatal approximant are illustrated in Figure 3.4. These productions were obtained from speakers 5 and 3, respectively. Like the sounds presented in Figure 3.3, the sounds in Figure 3.4 demonstrate a complete closure.

However, none of the speakers produced these sounds consistently with a complete closure. These productions were selected to create spectrograms for the purpose of illustration, as the phone boundaries are clearly visible and easy to annotate.

3.3.2 Visualizing the distribution of possible tongue positions

After using the *scale*() function in R (R Core Team, 2015) to normalize the position data for each speaker, the position of each instance of every sound was plotted using the R package *ggplot2* (Wickham, 2009) to create two-dimensional contour maps. The code used to create the contour maps was based on R code used to visualize vowel space (DiCanio, 2013)[2]. The highest density of sounds is represented by the innermost ring, which typically represents the highest point when used to depict geographical contours. A contour map was created for three points in time: the beginning, middle and end of the phone. The x-axis represents the mid-sagittal line, with the most anterior part of the tongue on the left and most posterior part of the tongue on the right. The y-axis represents the height of the sensor. Units for both axis are scaled.

Specifically, *geom_density2d*() from the R package *ggplot2* (Wickham, 2009) was used to create the contours. This plotting function uses the function *kde2d* from the R package *MASS* (Venables and Ripley, 2002), which provides a two-dimensional kernel density estimation, i.e. a non-parametric method of density estimation. A simpler form of density estimation is the histogram, but unlike the histogram, kernel density estimation is smooth as a result of interpolation. The default settings were used when creating the density estimations below; the default number of grid points in either direction is 25 (when creating histograms, this would be referred to as bucket or bin size).

Note that the diagrams in this section are purely for visualization only, not statistical analysis. The initial observations of the data distribution made here will be confirmed in the following sections, where rigorous statistical analyses will be performed. This method of visualizing data distribution has many benefits over more traditional methods such as

[2]Included in appendix and available at http://christiandicanio.blogspot.com/2013/10/visualizing-vowel-spaces-in-r-from.html, accessed 12/15/2015

scatter plots and histograms; kernel density estimation avoids the issue of over-plotting - an issue that scatter plots cannot overcome, while still functioning as an effective method of visualizing asymmetric data distributions.

When examining the contour maps produced by the kernel density estimation, note that there are occasionally smaller contours that are separate from the dominant contour plot. These contours indicate clusters of data points which have a higher density relative to the surrounding area; however, the contours themselves do not illustrate any specific density information. This prevents comparisons between the density of the smaller isolated contours against the larger contours; to make such comparisons, information regarding density level can be included using color.

The irregularity of the rings are an indication of asymmetry present in the data, providing a visual understanding of articulatory tendencies - i.e. overshoot or undershoot of the expected articulatory targets. Whether the irregularity of the rings are significant will be investigated in depth in the following sections. Again, the contour maps are intended to provide only a visualization of the data; interpretation of the size and shape of the rings should be reserved for the statistical analyses to follow.

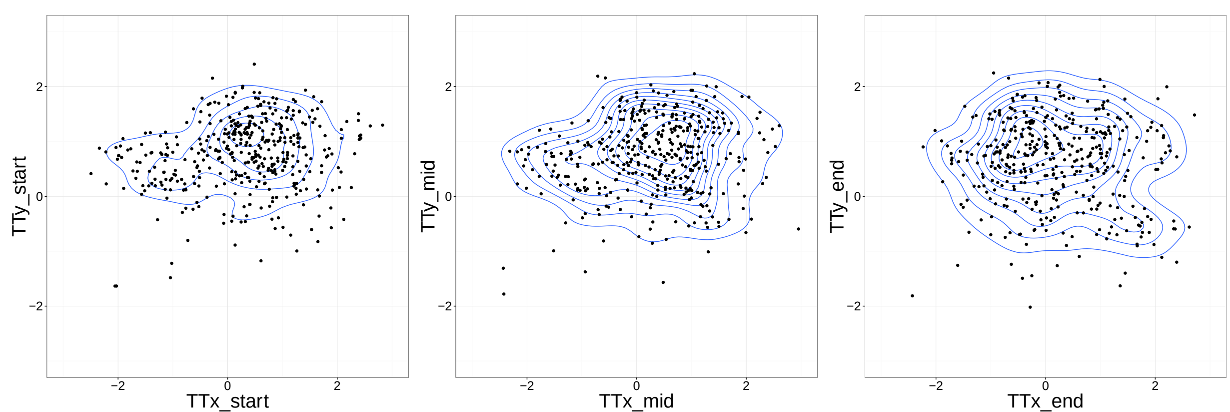

Figure 3.5: Contour maps superimposed over distribution of TT sensor position at start, middle, and end (from left to right, respectively) of /ʎ/.

A single sensor TT for the sound /ʎ/ was isolated for the purpose of facilitating the reader's understanding of the following contour maps. In Figure 3.5, contour maps are superimposed over the data distribution, illustrating how each contour captures the varying level of density in the data. From this figure, one can observe that the innermost ring, representing the highest density, is clearly more forward at the end of the phone (see Figure 3.5, right). Forward movement of the innermost ring indicates that the TT sensor is more anterior at the end of the palatal lateral than at the start. One particular benefit of utilizing contour maps to represent data distribution is that as opposed to plotting averages or ellipses (as is done in many graphical representation of vowel distributions), contour maps are capable of illustrating asymmetry in the data. This is particularly evident in the TT sensor at the start of the palatal lateral (see Figure 3.5, left), where the presence of some data points pull the edges of the two outer most rings further to the left of the plot; this set of low-density data points indicate that while the majority of the productions of the palatal lateral begin with the TT sensor further back in the oral cavity, there are some tokens that are articulated with a more forward TT position. Note too that at the middle and end of the phone, the number of rings in the plot has increased; each additional ring indicates a change in level of density, illustrating how change in data density decreases much more rapidly at the middle and end of the phone than at the start.

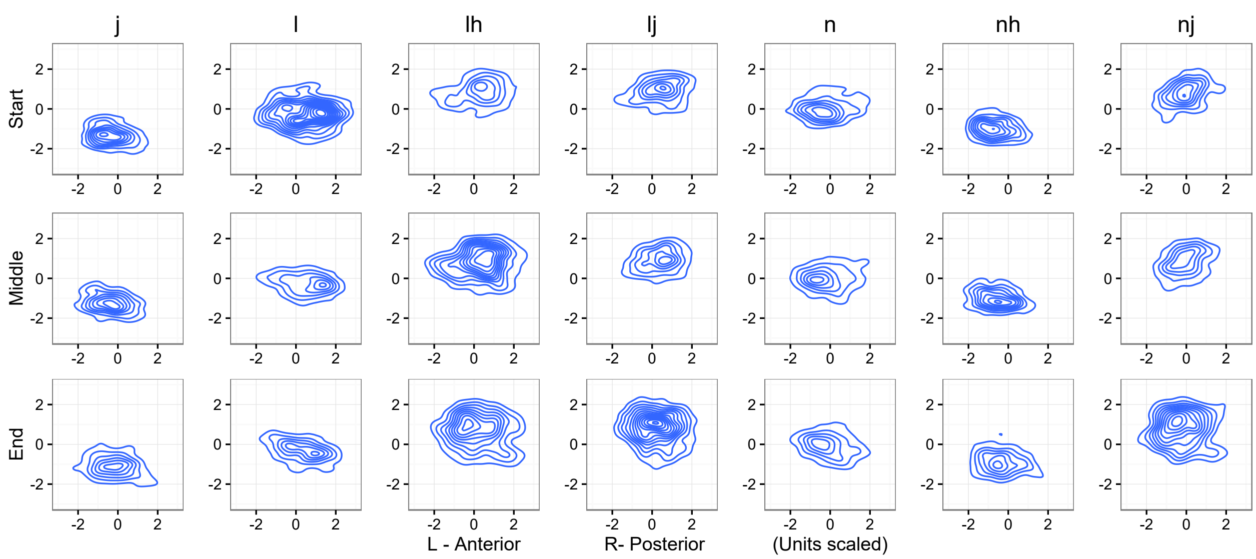

Figure 3.6: Position of TT sensor at three points in time: start, middle, and end of the phone, from top to bottom.

While it may not seem as if there is movement in the TT sensor from start to finish, closer observation of the point at which the highest density occurs reveals small shifts in position (see Figure 3.6). That point shifts backwards for the palatal approximant, backwards and slightly down for the plain lateral, forward for the palatal lateral, slightly upward and forward for the palatalized lateral, slightly upward for the plain nasal, and slightly upward and forward for the palatalized nasal. The TT sensor does not demonstrate a great deal of movement for the palatal nasal.

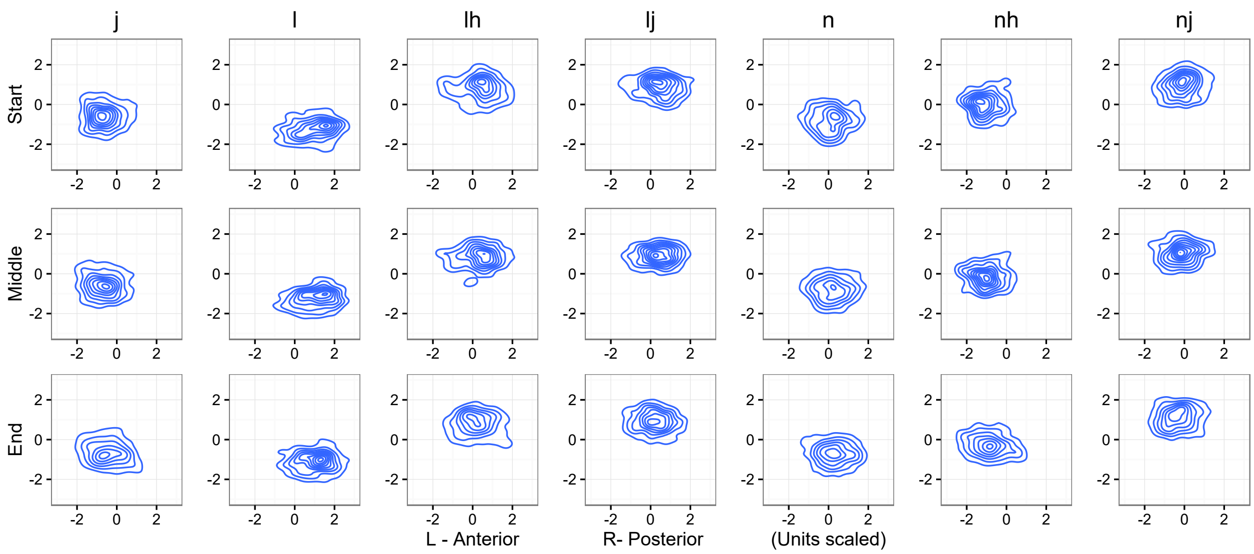

Figure 3.7: Position of TM sensor at three points in time: start, middle, and end of the phone, from top to bottom.

The TT sensor for the palatal nasal and the palatal approximant are both consistently lower than the other palatal and palatalized sounds, which have the highest TT position overall. The plain lateral and the plain nasal are relatively similar, occupying a more neutral height position in comparison. The palatal nasal and the palatal approximant also resemble one another in that they both demonstrate the tightest distributions. In contrast, the other palatal and palatalized sounds demonstrate increasingly more spread out distributions as the phone progresses. In general, the lateral sounds are articulated with a more retracted TT position, while the nasal sounds are articulated with a more anterior TT position.

The TM sensor is more consistent or constrained than the TT sensor (see Figure 3.7), with a much tighter distribution. Again, the palatal approximant and the palatal nasal pattern together, with the lower sensor height contrasting with the higher sensor height observed in the group formed by the other palatal and palatalized sounds. While the overall height of the palatal approximant and the palatal nasal is higher for the TM sensor as opposed to the TT sensor, the overall height of the plain lateral and nasal decrease. There is very little movement during the phone duration for nearly all the sounds; the more notable movement can be observed for the palatal nasal, which lowers slightly and retracts, while the palatalized nasal raises slightly. The tongue body also appears to retract slightly at the middle of the phone for the palatal lateral.

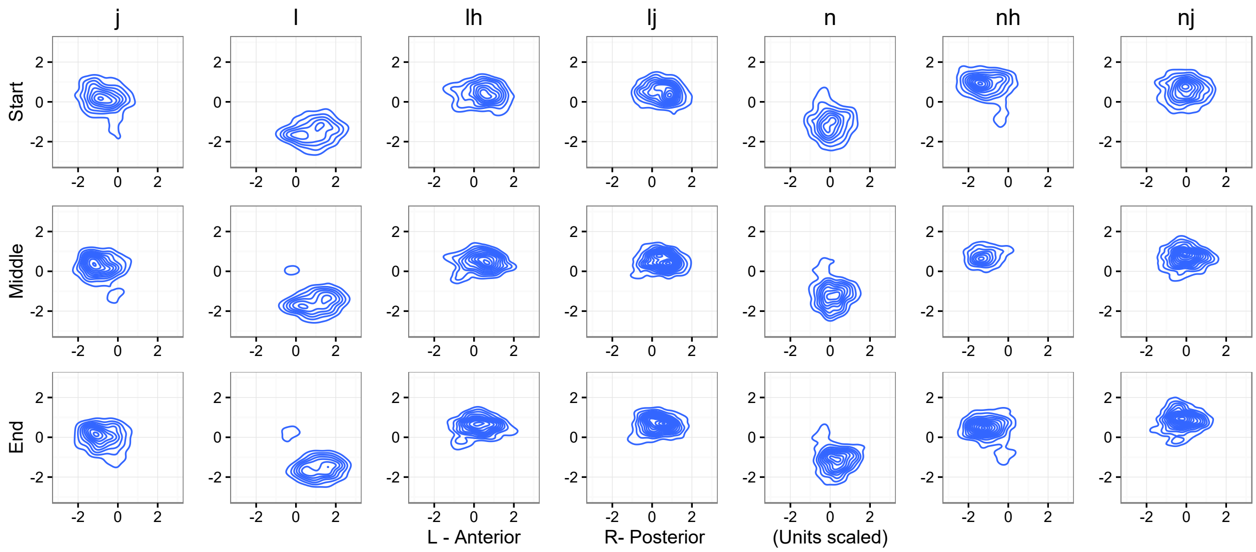

Figure 3.8: Position of TB sensor at three points in time: start, middle, and end of the phone, from top to bottom.

The TB sensor is more indicative of a palatal gesture, with all of the palatal and palatalized sounds having a nearly equally high tongue body position (see Figure 3.8). When comparing the TB sensor to the TM sensor, an increase in height is observed for the palatal nasal and palatal approximant, while a drop in tongue height is observed for the plain sounds. However, it is very difficult to discern movement with regards to changes in position throughout the duration of each phone. The TB sensor is the most variable of the three sensors utilized, with some clusters of points (or contours) that are entirely separate from the main body, as seen in the palatal approximant and the plain lateral. Contours that are distinct from the main body, as in the case for the alveolar lateral approximant at the middle and end of the phone, indicate the presence of a separate region of high density of instances where the sound is articulated at that location. This suggests that there is a larger range of articulatory configurations permitted at this point of the tongue when articulating these sounds.

3.3.3 Tongue blade angle

Some differences between palatal and palatalized sounds will be explored here. If palatalization is simply the addition of a secondary gesture, as suggested by classic phonological theory, then one would expect that palatalized alveolar sounds be articulated with an alveolar (apical) gesture and tongue body fronting. In contrast, palatal sounds should be articulated with only a palatal (laminal) gesture. In this dissertation, apical and laminal articulations are operationalized by measuring the angle formed between the occlusal plane and the line connecting the first and second tongue sensor in MATLAB (2014)[3]. An example is provided in Figure 3.9. While this method is loosely based on the method used in Simonsen et al. (2008) for capturing degrees of retroflexion in Norwegian retroflex stops, their method requires the inclusion of the third tongue sensor, which we find to be more prone to errors.

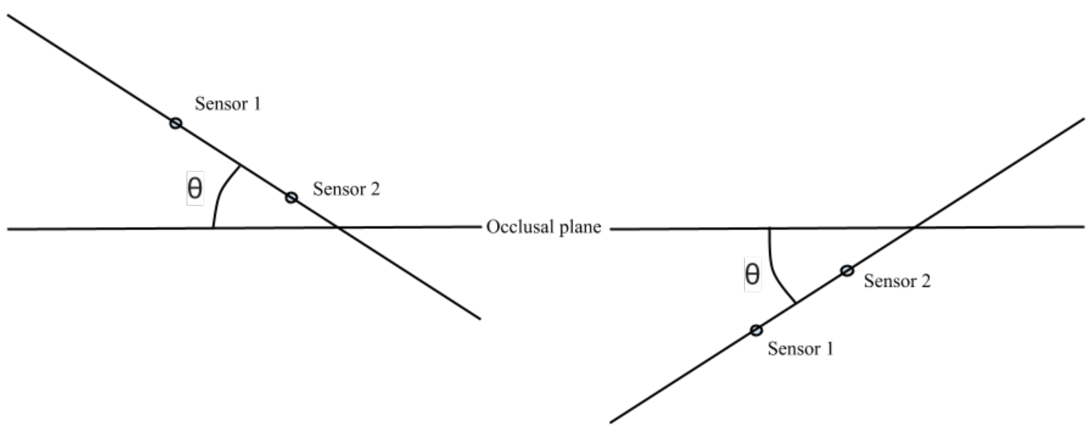

Figure 3.9: Sample representation of an apical articulation (left) and a laminal articulation (right), with θ indicating the measured angle. Head is facing left.

Apical articulations are produced with the tongue tip, which presupposes that the tongue tip and tongue blade are angled upwards for a constriction at the alveolar ridge or palatal area. Laminal articulations are produced using the tongue blade with the tongue tip pointed

[3]Script included in appendix.

downwards, which indicates that the tongue blade is angled downwards for a constriction at the alveolar ridge or palatal area. Therefore, an apical articulation should result in a positive angle θ while a laminal articulation should result in a negative θ (see Figure 3.9). θ was calculated for the beginning, middle, and end of each speech segment.

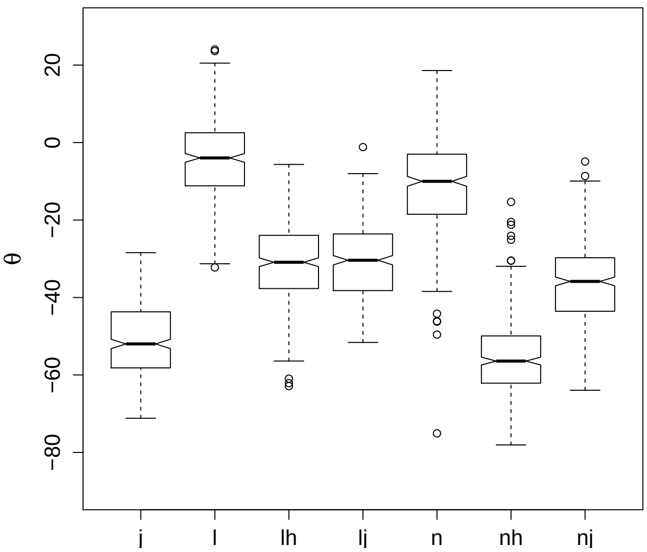

Figure 3.10: Angle in degrees between occlusal plane and line through TT and TM, beginning of sound.

Figure 3.10 reveals that none of the speech sounds studied here demonstrate a median angle above zero degrees, though /l/ and /n/ have a median angle that is close to zero (-4 and -10 degrees respectively). As a result, a neutral tongue angle is more characteristic of apically articulated sounds here, rather than strictly positive angles. The palatal and palatalized sounds fall into two groups, with the palatal approximant and palatal nasal in one group and the palatal and palatalized lateral and palatalized nasal in the second group. The first group has a more negative median angle of -52 and -56 degrees respectively, while the second group has a median angle of -31, -30, and -36 degrees respectively. The notches (representing the confidence intervals) overlap for only the palatal and palatalized lateral, indicating that there is strong evidence that they share similar medians.

The same groupings appear when tongue angles were measured at the middle of the speech segment (see Figure 3.11). Again, the palatal and palatalized laterals are the only two sounds with overlapping notches. Overall, the median degree drops about two to four

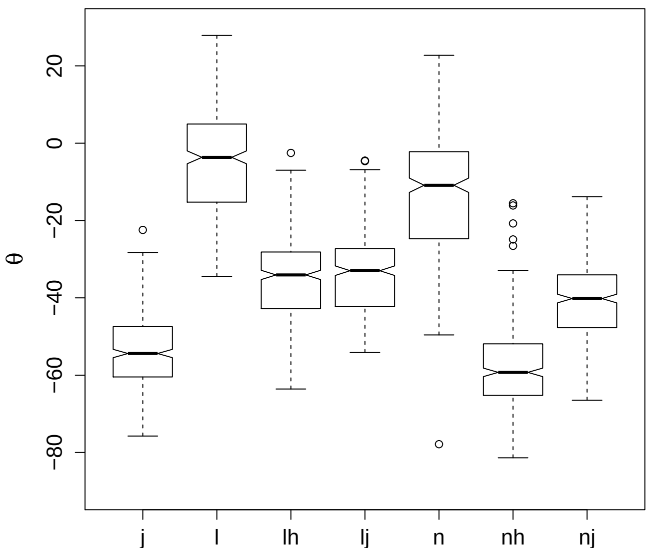

Figure 3.11: Angle in degrees between occlusal plane and line through TT and TM, middle of sound.

degrees for all the palatal and palatalized sounds. A smaller change of 0.9 and -0.3 degrees is observed for /n/ and /l/ respectively.

Figure 3.12 illustrates another drop in median tongue angles for two groups: /n/ and /l/ experience a large drop of 11 and 12 degrees respectively, while median tongue angles for /ʎ/, [l^j], and [n^j] drop by about five, six, and five degrees respectively. The median tongue angle for palatal approximant /j/ and palatal nasal /ɲ/ rises by two and one degrees, respectively. Overall, median tongue angles for the palatal approximant and the palatal lateral drop less than two degrees from beginning to end of the speech segment.

Combined, statistical analysis does not find significant differences between the palatal and palatalized laterals with regards to the angle of the anterior tongue body. In contrast, the palatal nasal is more similar to the palatal approximant, while the palatalized nasal is more similar to the palatal and palatalized laterals. The median angle of the anterior tongue body for all of the palatal and palatalized sounds was consistently lower than that of the two alveolar sounds /n/ and /l/, which both demonstrated similar tongue angles.

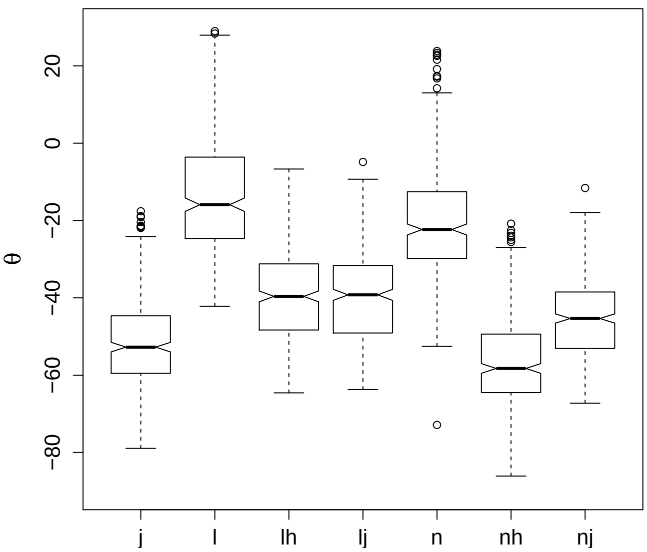

Figure 3.12: Angle in degrees between occlusal plane and line through TT and TM, end of sound.

3.3.4 Smoothing Spline ANalysis of VAriance (SS-ANOVA)

SS-ANOVA is a linear model that fits smoothing splines to curves. It is used for data sets where there are balanced repeated measures across the independent variable and a non-linear relationship between the dependent and independent variable. It was first used in phonetic science by Davidson (2006) for the comparison of tongue contours obtained by ultrasound. This method utilizes 95% Bayesian confidence intervals to identify sections of the curve that are significantly different. Overlapping confidence intervals indicate that for the duration of the overlap, curves do not show evidence of difference. The data was pooled for all speakers and scaled using the *scale*() function in R (R Core Team, 2015). The SS-ANOVA plots were generated using an R script by Mielke (2013)[4], which utilizes the package *gss* (Gu, 2014) for R (R Core Team, 2015) to fit the SS-ANOVAs.

SS-ANOVA is used here to compare the height of the three tongue sensors at the beginning, middle, and end of each phone segment. In doing so, movement in the x-dimension is not considered in this part of the analysis. While Figures 3.13 - 3.15 may look as if the curves are composed of several measurements, the smoothing spline is interpolating the data between three points (i.e. the measurements taken from the beginning, middle, and end of

[4]Included in appendix and available at *http://phon.chass.ncsu.edu/manual/tongue_ssanova.r*, accessed 08/30/2015.

each phone). Movement in the x-direction will be compared separately across phone for each tongue sensor. Tongue height at the beginning, middle, and end of each phone segment for each of the three tongue sensors was identified using MATLAB (2014). An SS-ANOVA of change in tongue blade angle over time by phone will also be presented.

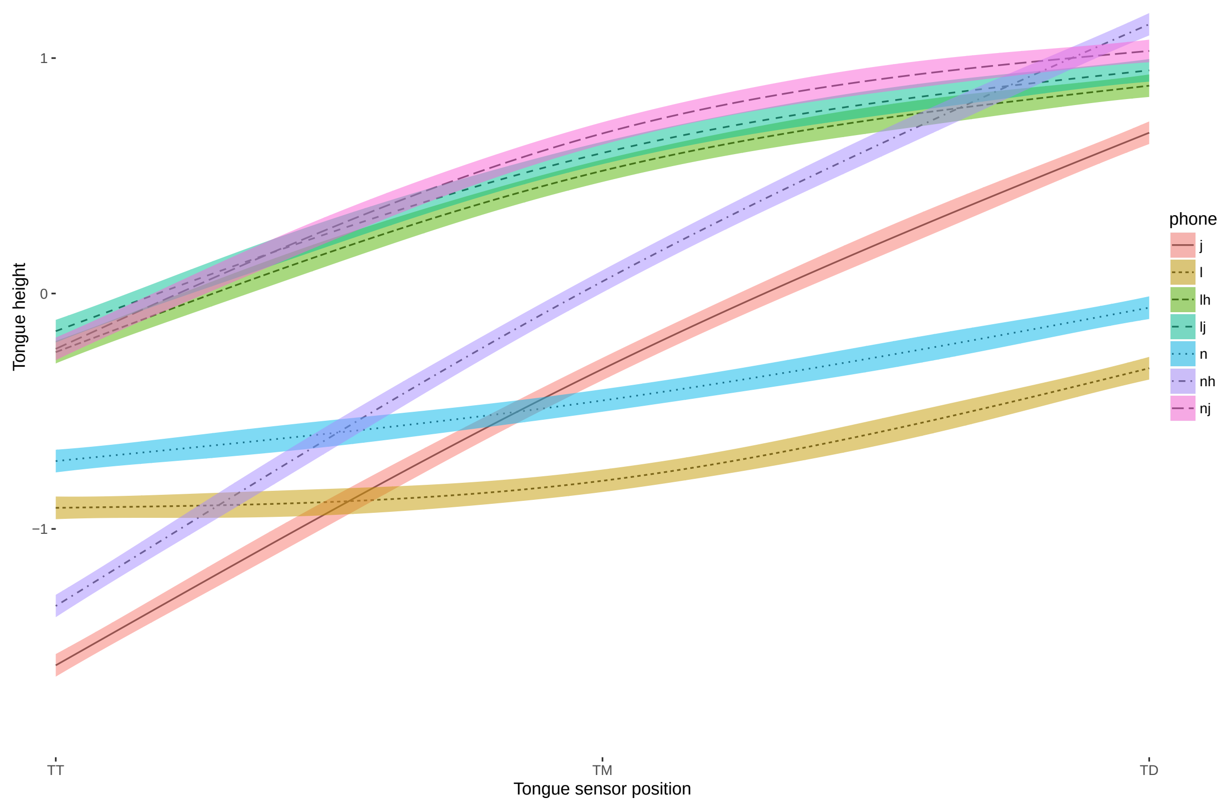

Figure 3.13: Smoothing splines for TT, TM, and TB height at beginning of phone segment for all 10 speakers.

At the beginning of the phone (see Figure 3.13), overlapping confidence intervals for the palatal lateral and the palatalized lateral indicate that there is no evidence of significant differences between the height at each of the three sensors for these two sounds. Confidence intervals for the palatalized nasal overlap with confidence intervals for the palatal and palatalized lateral at the TT and TB sensor, while there is a very marginal amount of separation between the palatalized nasal and the palatal and palatalized lateral at the TM sensor. All three sounds demonstrate a laminal articulation.

The palatal nasal and the palatal approximant appear to have similar relative sensor configurations, albeit with significantly different overall height. Both sounds demonstrate a greater negative angle between TT and TB than the palatalized nasal and the palatal and palatalized lateral, though the palatal nasal has a higher overall tongue position than the palatal approximant. In contrast, the plain nasal and plain lateral have a much more neutral change in height. However, the plain nasal has a significantly higher tongue position.

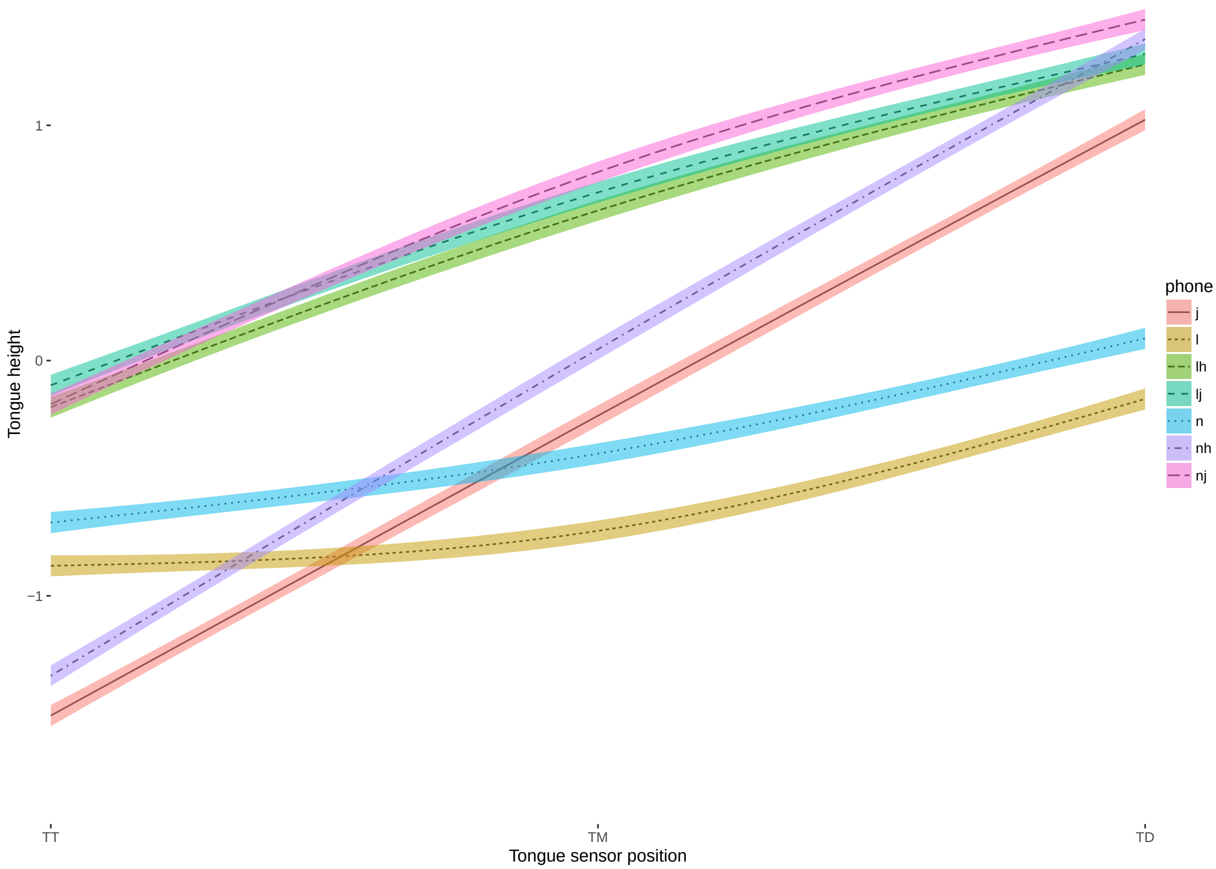

Figure 3.14: Smoothing splines for TT, TM, and TB height at middle of phone segment for all 10 speakers.

The same grouping of phones can be found at the middle of the phone segment (see Figure 3.14); we observe similar changes in tongue height across sensors for the palatal and palatalized lateral with the palatalized nasal, the palatal approximant with the palatal nasal, and the plain lateral with the plain nasal. Other than a small increase in tongue height at the TB sensor for the palatal approximant and the palatal nasal, there are no large observable differences in the height of the tongue sensors for the remaining phones.

Separations between confidence intervals are present at this point of the phone duration that were not present at the beginning of the phone. Confidence intervals for the palatalized lateral and palatal lateral demonstrate a slight amount of separation between the TT and TM sensor, while confidence intervals for the palatalized nasal are completely separated from the palatal and palatalized lateral at the TM and TB sensor. It is clearer here that the palatal and palatalized laterals have a lower tongue position than the palatalized nasal at the TM and TB sensor, with the palatalized lateral demonstrating just a slightly bit more tongue height than the palatal lateral at the TT and TM sensor. Additionally, all three phones (palatalized nasal, palatal and palatalized lateral) demonstrate a higher tongue position at the TB sensor in comparison to the palatal nasal, which was not evident at the beginning of the phone.

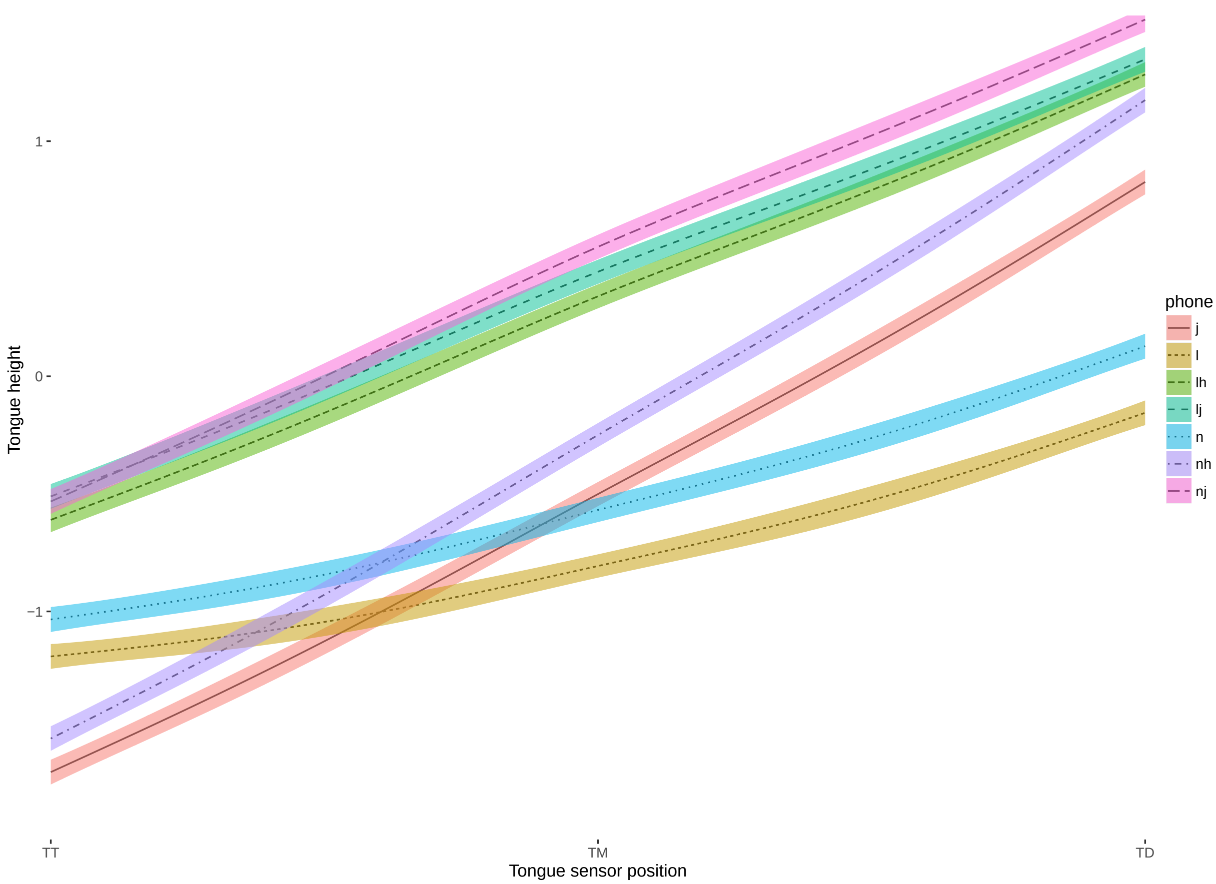

Figure 3.15: Smoothing splines for TT, TM, and TB height at end of phone segment for all 10 speakers.

By the end of the phone (see Figure 3.15), greater separation between the palatalized and the palatal and palatalized lateral can be observed. At each point during the phone (i.e.

from beginning, middle, to end), the palatalized lateral maintains a higher tongue position than the palatal lateral. Overall, the height of the palatal approximant and palatal nasal drops slightly at each tongue sensor in relation to the palatalized nasal and palatal and palatalized lateral. Tongue height also drops at TT for the plain lateral and plain nasal.

While there is concavity in the curves representing the sensor height observed for palatal and palatalized sounds at the beginning and middle of the speech sound, this concavity disappears by the end of the speech sound. Similarly, the convex shape observed in the curves representing the sensor height observed for plain nasal and lateral is gone by the end of the speech sound. Overall, there does not appear to be a drastic difference in tongue height at the beginning, middle, or end of the sound segment for any of the phones compared here. As a whole, this indicates that there is very little movement in the y-dimension (presumably representing tongue contact with the hard palate) required to create a palatal articulation. Additionally, the palatal gesture reaches its climax towards the beginning of the phone.

Note that in Figures 3.16 through 3.18, the direction of the y-axis is reversed. During the data processing (see Section 3.2.4 for more), the head was rotated such that it was facing left. Therefore, forward movement of the tongue results in increasingly larger negative x-values. The reversed y-axis is used here as a more intuitive indication of forward movement of the tongue; as the sensor moves forward in the mouth, the plotted trajectory rises.

Figure 3.16 illustrates the horizontal movement observed in the TT sensor over the percent of phone duration. In general, the three lateral sounds (palatal, palatalized, and alveolar) are articulated with the most retracted TT position. It is possible that the more posterior TT position of the lateral sounds are due to the retracted tongue body that would be characteristic of the so-called "dark" (Sproat and Fujimura, 1993) or velarized /l/ that BP is described as possessing (Mateus and d'Andrade, 2000). Of the three lateral sounds, TT position is significantly more retracted for the alveolar lateral throughout phone production in comparison - though onset TT position for the palatalized lateral briefly overlaps with the alveolar lateral.

Confidence intervals for the trajectory of the TT sensor during the last half of the palatal and palatalized lateral are overlapping throughout the duration of both phones, indicating that there is no evidence of significant difference during the latter half of the two phones.

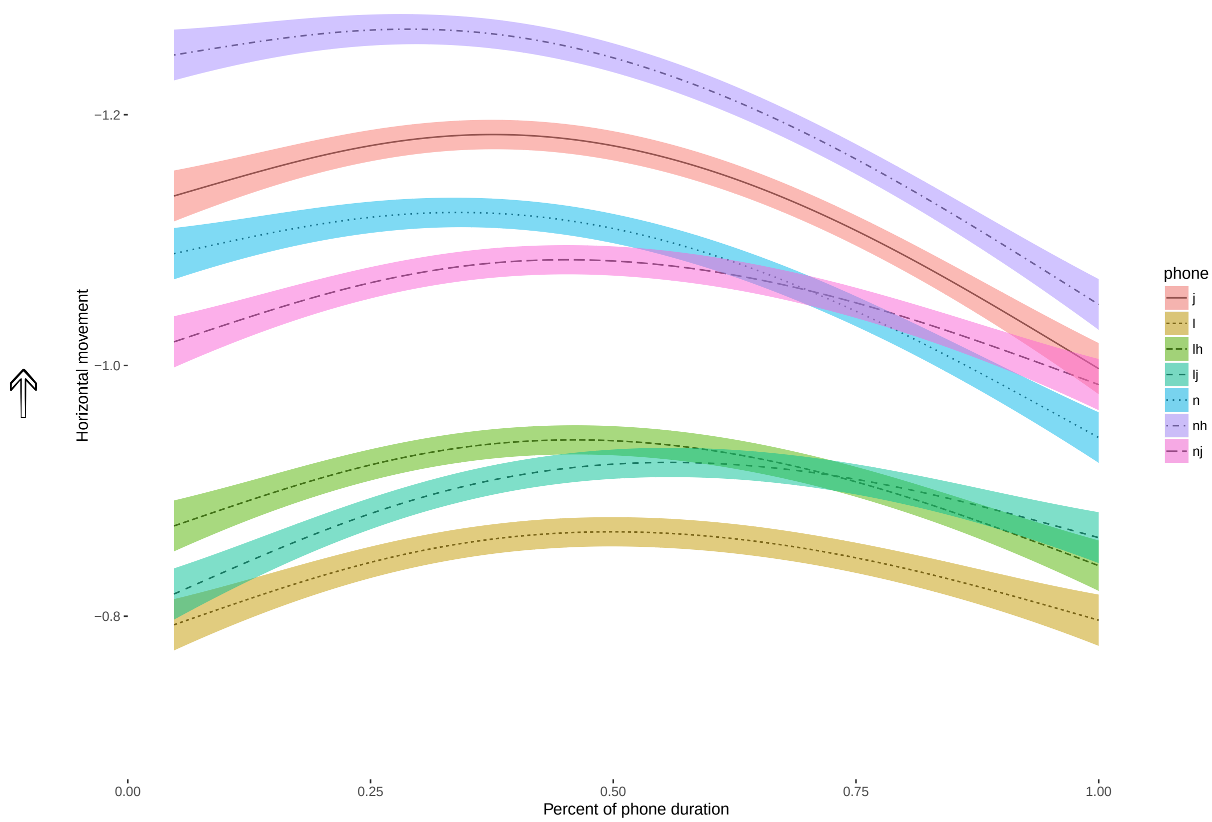

Figure 3.16: SS-ANOVA of horizontal movement in TT sensor, with a reversed y-axis where up and down indicate forward movement and retraction respectively.

However, TT position is significantly different between the palatal and palatal lateral during the previous half of the phones, with the palatalized lateral initially articulated with a more posterior TT position compared to the palatal lateral. As the palatalized lateral is articulated, the TT sensor moves to a more anterior position, eventually overlapping with the TT trajectory plotted for the palatal lateral. In contrast, during the production of the palatalized nasal, onset TT sensor is significantly more retracted in comparison to the alveolar nasal. The trajectories of the two sounds cross at the last quarter of the phone duration, with alveolar nasal TT position retracting to a significantly more posterior position than the palatalized nasal.

Consistently throughout the phone, the most anterior predicted TT position is recorded for the palatal nasal. The trajectory of TT during the production of the palatal nasal is similar to that of the palatal approximant, through the two do not overlap at any point.

Both sounds have a more retracted final TT position in comparison to the TT position at the phone onset. The alveolar nasal also follows this same pattern.

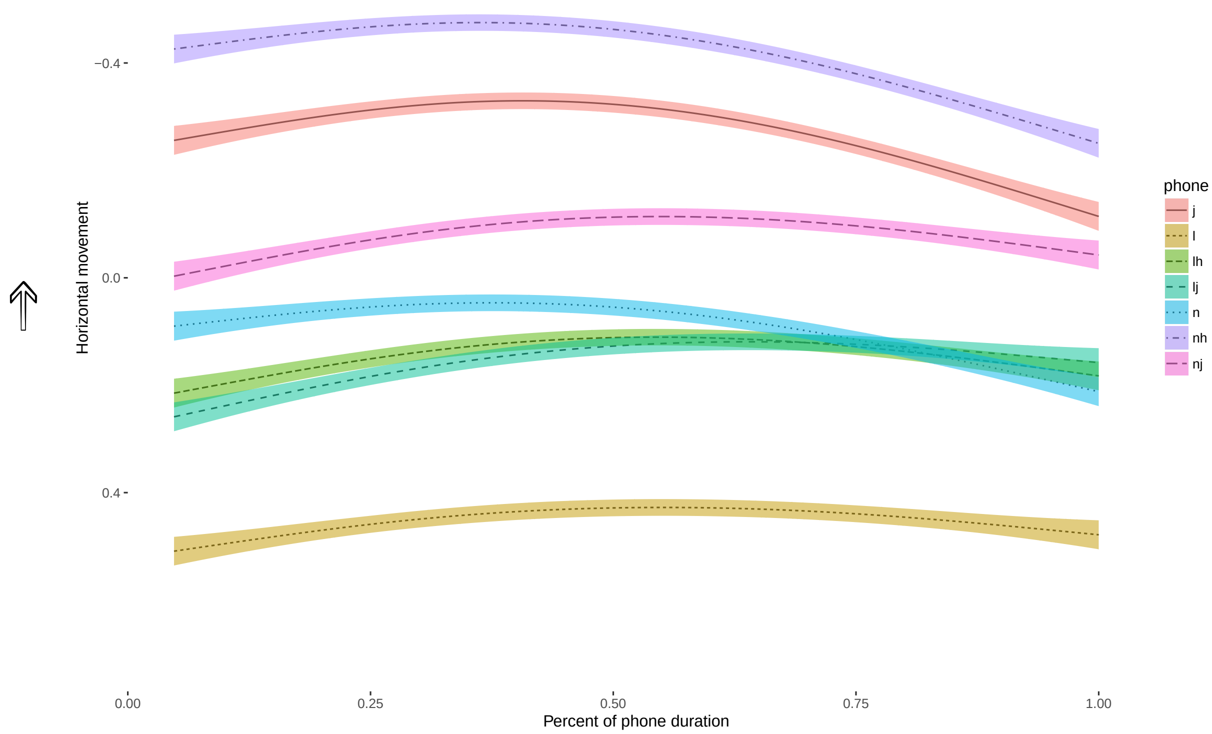

Figure 3.17: SS-ANOVA of horizontal movement in TM sensor, with a reversed y-axis where up and down indicate forward movement and retraction respectively

Horizontal tongue position is more spread out for TM (Figure 3.17) compared to TT, with less overlapping of trajectories, though the relative order of position for each phone remains relatively similar. Again, TM position is the most posterior during the production of the alveolar lateral and the most anterior during the paltal nasal. There is less overall change in x-position over time, which is particularly noticeable in the alveolar lateral.

Confidence intervals for the palatal and palatalized lateral overlap throughout the entire duration of the phones, with a significant separation from the alveolar lateral. TM sensor trajectories are largely similar for the three laterals (and the palatalized nasal, as well), beginning and ending with a more retracted tongue body and stabilizing for the greater duration of the central portion of the phone duration. It is likely that the movement backwards observed at the edges of the phone duration is a result of co-articulation with the surrounding back vowel /a/.

In contrast with the TT sensor, TM remains more anterior during the production of the palatalized nasal in comparison to the alveolar nasal. The modeled TM trajectory for the two sounds do not overlap at any point. Additionally, while onset alveolar nasal TT position was more anterior in comparison to the palatalized nasal, onset alveolar nasal TM position is significantly more posterior in comparison to the palatalized nasal. However, both the TT and TM sensor move to an increasingly more posterior position as the alveolar nasal progresses. Similarly, both the TT and TM sensor move to an increasingly more anterior position as the palatalized nasal progresses.

Note that while the palatalized nasal converges towards the palatal approximant at the end of the phone, the palatalized and palatal lateral do not. If the palatalized lateral and nasal are sequences composed of two separate speech sounds, i.e. an alveolar lateral or nasal followed by a palatal approximant, one would expect the earlier portion of these sounds to resemble the alveolar lateral or nasal and the latter portion to resemble the palatal approximant. While there is some evidence of this occurring with regards to the palatalized nasal, this is certainly not the case for the palatalized lateral or palatal lateral.

Only the TM trajectories for three sounds overlap; the palatalized and palatal lateral overlap throughout the entire phone and the alveolar nasal overlaps with the palatalized and palatal lateral for just the last quarter of the phone. While initial tongue position is significantly more anterior at TM for the alveolar nasal than the palatal and palatalized laterals, the tongue continually retracts beginning at about 25% of the phone duration to overlap with the palatal and palatalized laterals for the last quarter of the phone duration.

TM position for the palatalized nasal is the third most anterior and is significantly different from all other sounds; it is similar to the laterals as it maintains a mostly stable position for the majority of the phone and some apparent co-articulation with /a/ at the boundaries. The palatal nasal and palatal approximant are articulated with the first and second most anterior TM position, and like TT, resemble the alveolar nasal in terms of horizontal displacement.

The horizontal movement of the TB sensor (Figure 3.18) is very similar to the TM

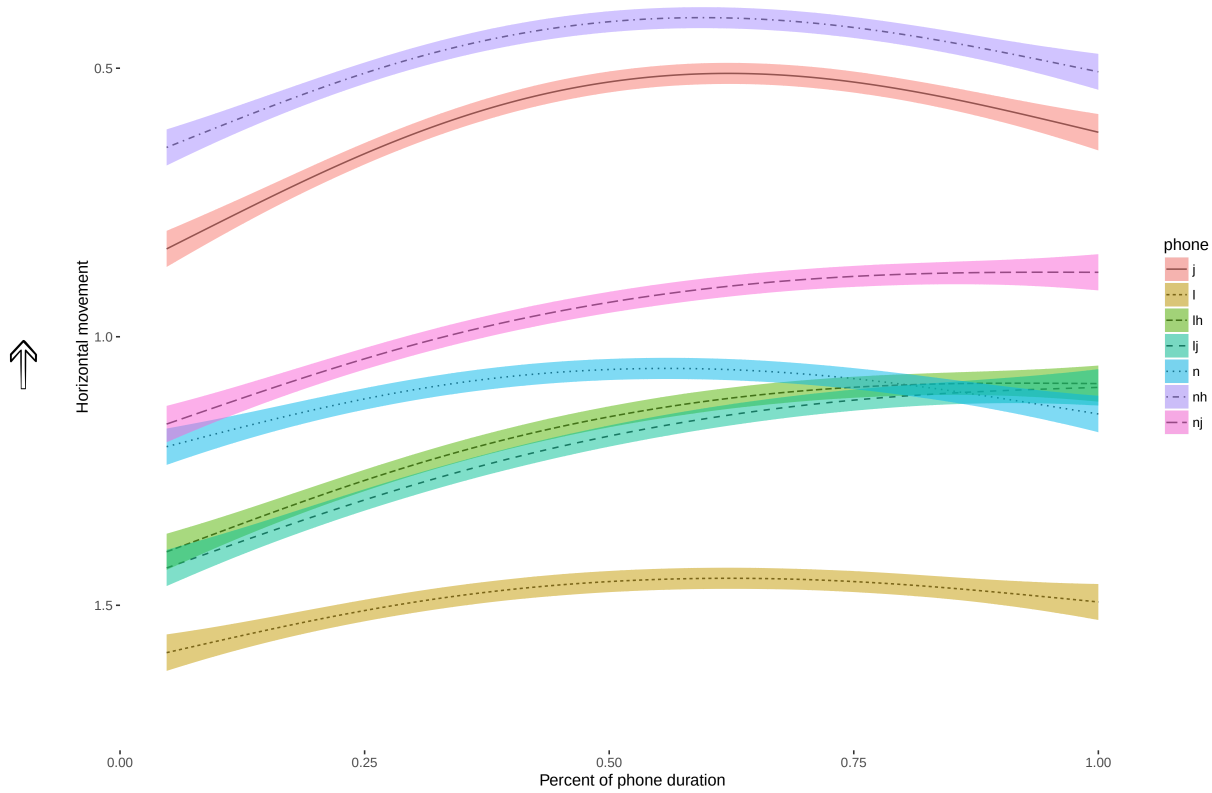

Figure 3.18: SS-ANOVA of horizontal movement in TB sensor, with a reversed y-axis where up and down indicate forward movement and retraction respectively

sensor (Figure 3.17) across phones, though there are more overlapping confidence intervals. While the overall trajectories appear to be the same, there are some differences regarding the spacing between certain sounds. Here, the palatal approximant moves up closer to the palatal nasal, resulting in a consistently larger division between these two sounds and the other sounds. The palatal and palatalized laterals and alveolar and palatalized nasals move closer to the alveolar lateral, with TM overlapping at the onset of the palatalized nasal and alveolar nasal.

The confidence intervals for the palatal and palatalized lateral are completely overlapping throughout the entirety of the phone, indicating that there is no evidence that the TB position is significantly different during the production of the palatal and palatalized lateral. In contrast to the TM sensor, TB movement during the production of the palatal and palatalized lateral becomes increasingly more anterior as the phones progress, indicating that palatalizing gestures are created by the TB portion of the tongue (or about 3 cm from the tongue tip).

Again, TB horizontal movement during the production of the palatalized nasal resembles the movement during that of the palatalized and palatal lateral, though not with regards to overall position. Likewise, the trajectory of the TB sensor during the palatal nasal and approximant resemble one another, while the alveolar lateral and nasal are similar.

While the palatalized and palatal nasal are significantly different throughout the phone and across tongue sensors, it is interesting to observe that the same pattern is not present for the palatalized and palatal lateral. Except for TT during the first half of the phone, the horizontal position of the three sensors for the palatalized and palatal lateral do not demonstrate evidence of significant difference.

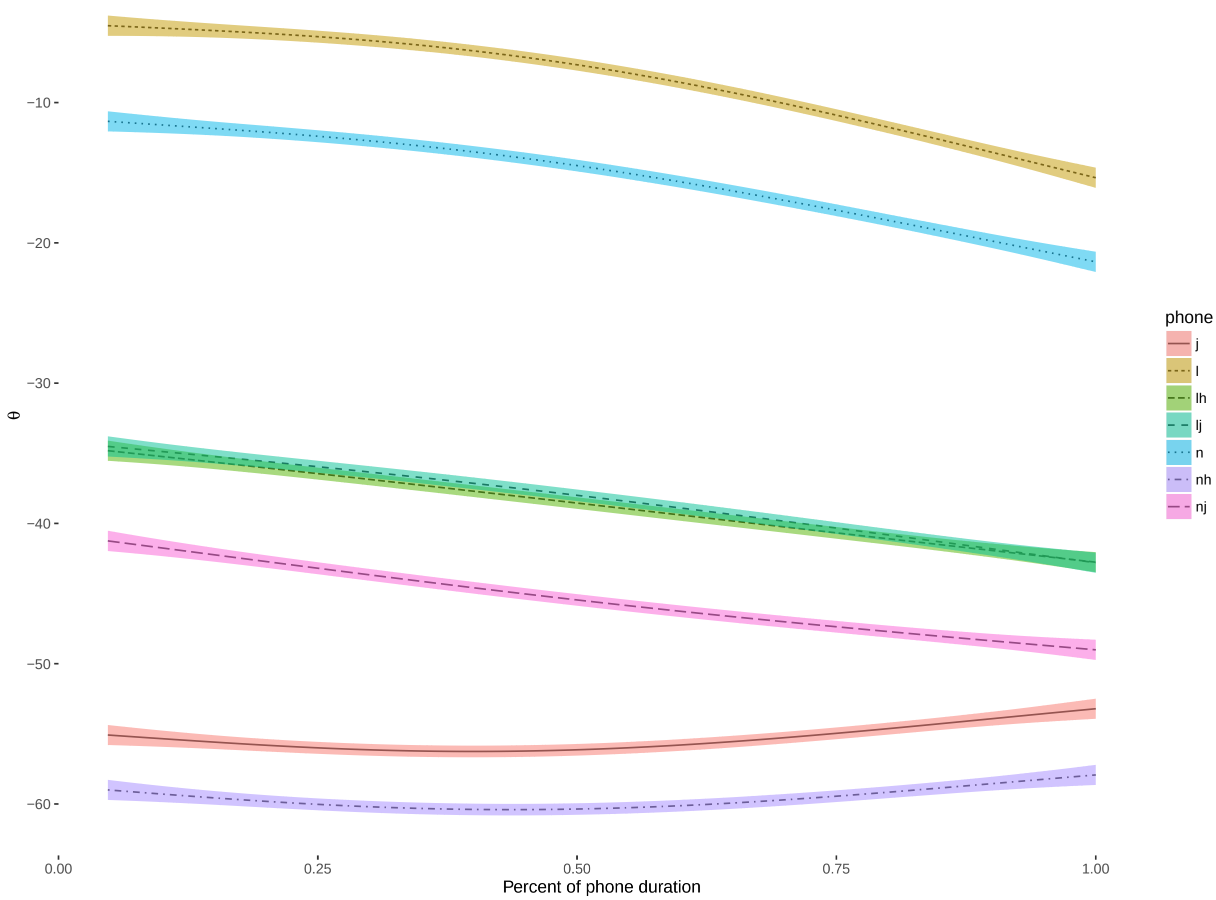

Figure 3.19: SS-ANOVA of change in tongue blade angle throughout phone duration.

Incorporating the same interpolated and scaled data that was used to create Figures 3.16 - 3.18, an SS-ANOVA of the change in tongue blade angle over time (see Figure 3.19) reveals that except for the palatal nasal and palatal approximant, all the sounds tend to be articulated with a steeper tongue blade angle as the phone progresses. There is relatively little change in tongue blade angle for the palatal nasal and palatal approximant until the last quarter of the phone duration, at which point the tongue blade begins to flatten out somewhat. Of all the sounds, the palatal nasal is clearly articulated with the steepest tongue blade angle, followed by the palatal approximant.

The difference in tongue blade angle between the alveolar nasal and alveolar lateral approximant is mirrored in the palatalized nasal and palatal/palatalized lateral; there is a consistent difference of approximately six to seven degrees between the alveolar nasal and alveolar lateral approximant, as well as the palatalized nasal and palatal/palatalized lateral, demonstrating that lateral sounds are articulated with a relatively more moderate tongue blade angle than their nasal counterparts.

In comparison to the boxplots from Section 3.3.3, the difference in tongue blade angles by phone is much more evident in Figure 3.19. While the previously observed groupings are still present (i.e. the alveolar nasal with the alveolar lateral approximant, the palatal/palatalized lateral with the palatalized nasal, and the palatal nasal with the palatal approximant), the only two sounds that do not demonstrate evidence of significant differences are the palatal lateral and palatalized lateral.

3.3.5 Dynamic modeling of anterior tongue shape

While oftentimes phonetic studies confine the scope of their analysis to discrete moments during the phone (e.g. the beginning, middle, or end of the phone), these analyses lend to a myopic understanding of speech. Real speech is of course dynamic, with the tongue profile constantly changing throughout. In order to truly understand the articulatory characteristics of speech, especially speech sounds produced with an approximate place of articulation, the vocal tract configuration must be studied throughout the duration of the phone and not at a single static point in time.

Creating a dynamic model is particularly useful for such an analysis, as it facilitates the prediction of the anterior tongue shape as the phone progresses. The entire phone duration may then be considered for analysis, as opposed to simply specific points in time such as the previous analyses which focused upon the beginning, middle, and end of the phone duration. As described in Section 3.3.4, the trajectories for each tongue sensor were re-sampled in MATLAB (2014) using the *resample*() function[5]. A model of the contour created by the three tongue sensors could then be fitted for each phoneme for each re-sampled point in time, creating a model of how the anterior tongue shape changes over time. Prior to resampling, tongue data was normalized using the *scale*() function in R (R Core Team, 2015).

A 2nd degree natural spline was used to model the anterior tongue shape. This model was chosen over a linear model as tongue movement is not typically linear, it is cyclic. Interaction between the method of interpolation and the model used here is unlikely, as a linear interpolation method was used to resample the data. Additionally, the two methods of interpolation were applied over different aspects of the data set; *resample*() was used to interpolate the trajectory of each individual tongue sensor over a period of time, while the 2nd degree natural spline was used to interpolate the shape formed by the tree tongue sensors.

The model predictions for the articulation of the palatal approximant neatly fits the literature's description of a palatal sound as being produced with a laminal articulation, with Figure 3.20 demonstrating a high tongue position towards the TB sensor and a low tongue position at the TT sensor. Approximate sensor placement along each contour can be estimated by visually dividing each contour into thirds. The highest tongue position is achieved at about a third of the way into the phone, such that the tongue approaches the articulatory target rather early on in the articulation of the palatal approximant. The overall tongue shape remains approximately the same throughout the duration of the phone, with very little movement. Since the tongue tip is relatively stationary, it appears as if there is a pivot point near the tongue tip and the tongue is rotating around that point. Compared to the other sounds modeled here, the palatal approximant is articulated with the lowest tongue tip position, with the palatal nasal articulated with a similar though not quite as

[5]Script used to re-sample the tongue sensor trajectories included in appendix.

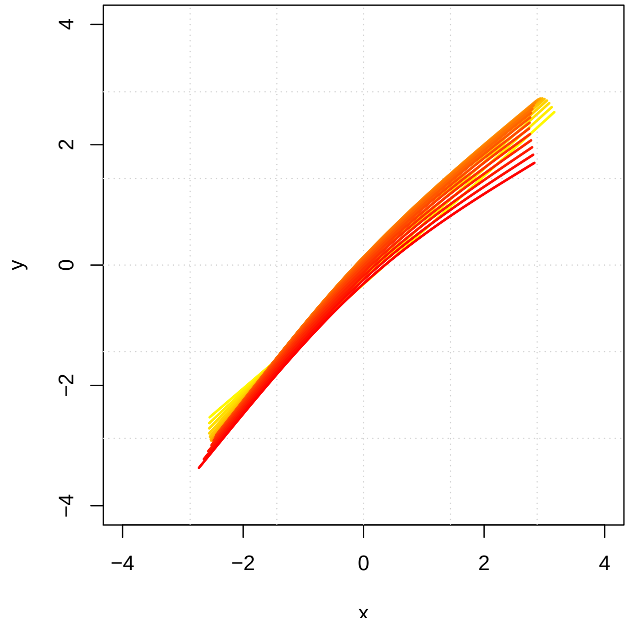

Figure 3.20: Dynamic model of /j/, yellow to red represents first to last sample. Tongue oriented anterior (left) to posterior (right). Axes are scaled.

extreme tongue tip position.

A comparison of the models created for the palatal lateral /ʎ/ (Figure 3.21) and palatalized lateral [lʲ] (Figure 3.22) illustrates that the two sounds are very similar. Both sounds are articulated with a relatively lower TB and higher TT sensor position at the onset of the phone, following which the TB gradually rises and TT gradually lowers. Unlike the palatal approximant, the position of the tongue anterior changes quite a bit as the phone progresses. Despite subtle differences between the two sounds (the palatalized lateral has a more anterior tongue position), the overall articulation is very similar.

For both the palatal and palatalized lateral, the maximum height across the three tongue sensors is recorded by the TB sensor at the end of the phone. Similar to the palatal approximant (see Figure 3.22), the palatal and palatalized lateral demonstrate a pivot motion near the TM sensor where there appears to be little change in tongue position throughout the duration of both phones. However, there is more movement at other points in the tongue as compared to the dynamic models produced for the palatal approximant.

Despite the pronounced similarity between the palatal and palatalized lateral, some small differences are evident. Tongue position during the palatalized lateral is slightly more an-

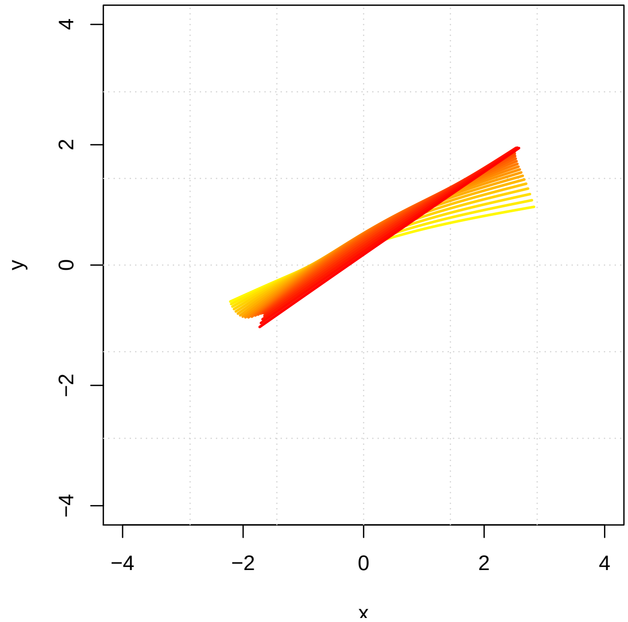

Figure 3.21: Dynamic model of /ʎ/, yellow to red represents first to last sample. Tongue oriented anterior (left) to posterior (right). Axes are scaled.

terior than the palatal lateral. Additionally, the ending position of the TB sensor is higher and more retracted for the palatal lateral in comparison to the palatalized lateral. Lastly, the palatalized lateral demonstrates more variation in tongue tip height, while the tongue tip height during the palatal lateral remains comparatively static.

The anterior tongue shape for the alveolar lateral (see Figure 3.23) is convex, in comparison to the concave shape present in the palatal or palatalized sounds. The alveolar lateral approximant is articulated with the tongue tip pointed up towards what can be presumed to be the alveolar ridge. This would be consistent with the description of an alveolar speech sound. The model illustrates an apical gesture with the tongue tip pointed only slightly upwards, corresponding with the almost neutral tongue blade angle reported in Section 3.3.3.

There is very little tongue movement when articulating the alveolar lateral, at least with regards to the midsagittal section of the tongue anterior. A comparison of the palatal (Figure 3.21) and palatalized lateral (Figure 3.22) to the palatal approximant (Figure 3.20) and the alveolar lateral approximant (Figure 3.23) reveals potential articulatory motivations for why *yeísmo* occurs (i.e. the merger of the palatal lateral to the palatal approximant); the palatal lateral is quite obviously more similar to the palatal approximant than the alveolar

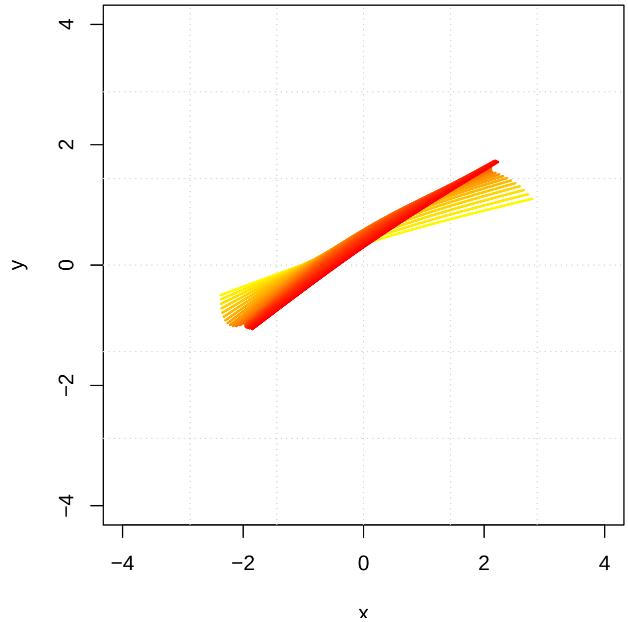

Figure 3.22: Dynamic model of [lʲ], yellow to red represents first to last sample. Tongue oriented anterior (left) to posterior (right). Axes are scaled.

lateral approximant in terms of anterior tongue articulation.

While the overall shape of the alveolar nasal stop (see Figure 3.24) is similar to the alveolar lateral, there are some distinct differences. The convex nature of the anterior tongue shape present in the alveolar lateral approximant is much less evident in the alveolar nasal; in fact, the anterior tongue shape for the alveolar nasal is almost entirely linear. However, like the alveolar lateral approximant, the alveolar nasal demonstrates very little movement across the production of the phone. It is not possible to determine from these models whether the observed differences is an effect of manner, i.e. a difference resulting from the lateral or nasal quality of the sound.

A comparison of the palatal (see Figure 3.25) and palatalized nasal (see Figure 3.26) reveals some particularly interesting observations in contrast to the lack of differences between the palatal and palatalized lateral approximant. In particular, there are clear articulatory differences between the palatal and palatalized nasal, which were not evident in the palatal and palatalized lateral, such as the presence of concavity in the palatalized nasal throughout the production of the sound and the lack thereof in the palatal nasal. There is much more movement in general during the production of the palatalized nasal in comparison to the

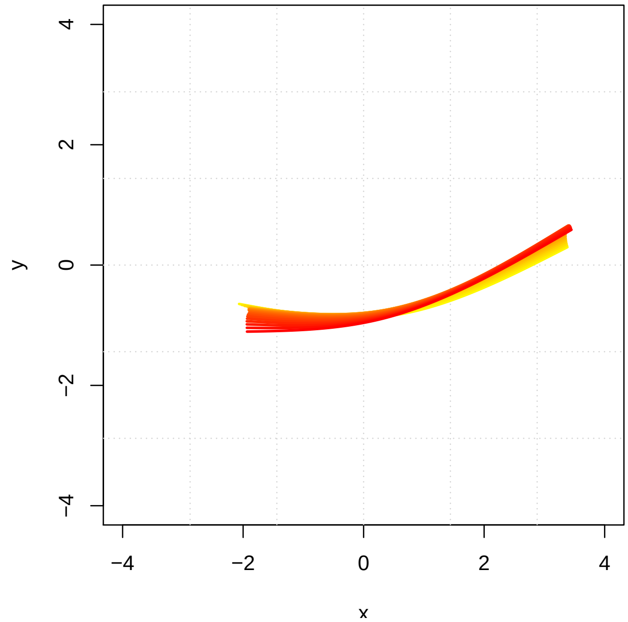

Figure 3.23: Dynamic model of /l/, yellow to red represents first to last sample. Tongue oriented anterior (left) to posterior (right). Axes are scaled.

palatal nasal.

Additionally, the palatalized nasal resembles the palatal and palatalized lateral, while the palatal nasal resembles the palatal approximant. Like the palatal and palatalized lateral, the palatalized nasal is articulated with the front of the tongue lowering and the area near the TB sensor rising as the sound is articulated. Like the palatal approximant, the palatal nasal is articulated with the majority of movement occurring near the TB sensor and the maximum TB height occurring approximately a third of the way into the duration of the phone.

The resemblance of the palatal nasal to the palatal approximant and the palatalized nasal to the palatal and palatalized lateral is evident also in the previous sections; a similar grouping is present in Figures 3.10 - 3.12 of the angle formed between the first and second sensor, as well as in the average phone duration (see Table 4.1). However, it is still unclear why a clear distinction between the palatalized and palatal nasal exists while the same distinction is not present between the palatalized and palatal lateral. Given the closer similarity of the palatal nasal to the palatal approximant, especially in comparison to the palatal and palatalized lateral, accounts of the palatal nasal merging with the palatal approximant should be

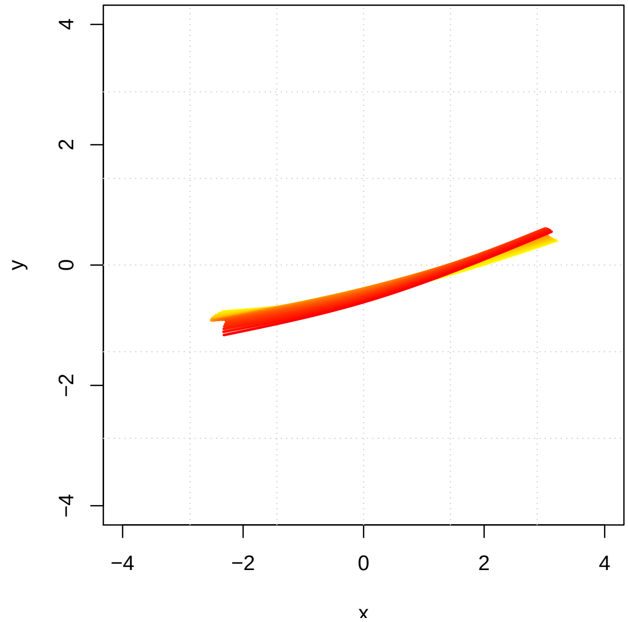

Figure 3.24: Dynamic model of /n/, yellow to red represents first to last sample. Tongue oriented anterior (left) to posterior (right). Axes are scaled.

much more frequent. Mere articulatory similarity, at least with respect to the shape of the tongue anterior, cannot be the complete explanation as to why *yeísmo* occurs.

Dynamic anterior tongue models by speaker

Given the relatively low number of speakers and large range in age (20 to 55 years of age, see Section 3.2.1), it would be interesting to study the variation occurring across speakers. This section will present a short investigation into dynamic anterior tongue models created from the data of each individual speaker, while making reference to potential social and geographic factors. Again, all the speakers in this study were female.

Speaker 1 was 36 years old at the time of data collection and had lived the majority of her life in São Paulo state - a total of 28 years. She had moved to Illinois only one year prior to the time of data collection.

Like the consolidated data, the palatalized lateral is slightly more anterior than the palatal lateral, which is particularly noticeable at the at the onset of the palatalized lateral. While the ending position appears to be similar for all the sensors, the palatalized lateral

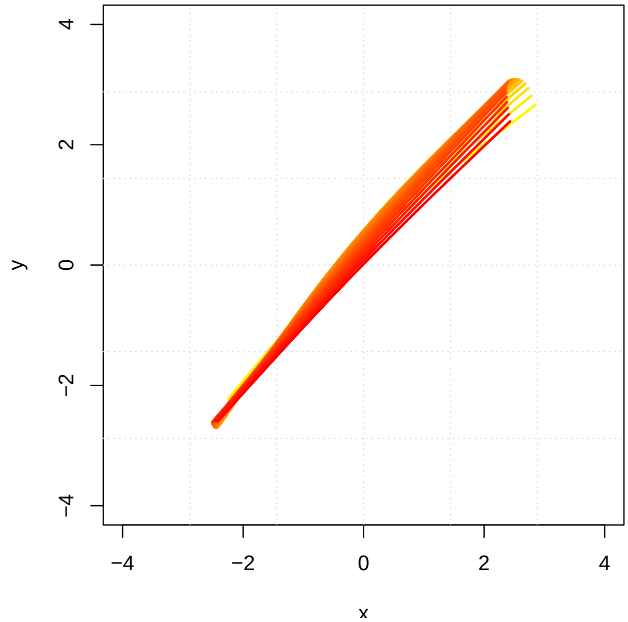

Figure 3.25: Dynamic model of /ɲ/, yellow to red represents first to last sample. Tongue oriented anterior (left) to posterior (right). Axes are scaled.

demonstrates a slightly higher overall sensor maximum position than the palatal lateral (occurring maybe halfway through the phone). Both sounds are articulated with the majority of the movement occurring during the earlier half of the phone duration.

There are some striking differences between Speaker 1 and the dynamic models created from the consolidated data, particularly with regards to comparisons between the palatalized and palatal lateral with the palatalized nasal, as well as the palatal nasal and palatal approximant. Both sets of comparisons were far more similar in the consolidated data. Speaker 1 demonstrates a noticeably steeper tongue blade angle for the palatalized nasal in comparison to the palatal and palatalized lateral, though maintaining certain resemblances such as a similar TT sensor position. Additionally, like the palatal and palatalized lateral, the palatalized nasal is articulated with the majority of movement occurring during the first half of the phone - the tongue anterior barely moves during the later half.

The dynamic model of the palatal nasal produced from Speaker 1 bears a great deal of resemblance to the palatal nasal model created from the group data. There is little movement at the tip and an evenly timed arching motion at the TB sensor. However, the palatal approximant produced by Speaker 1 does not demonstrate the same resemblance;

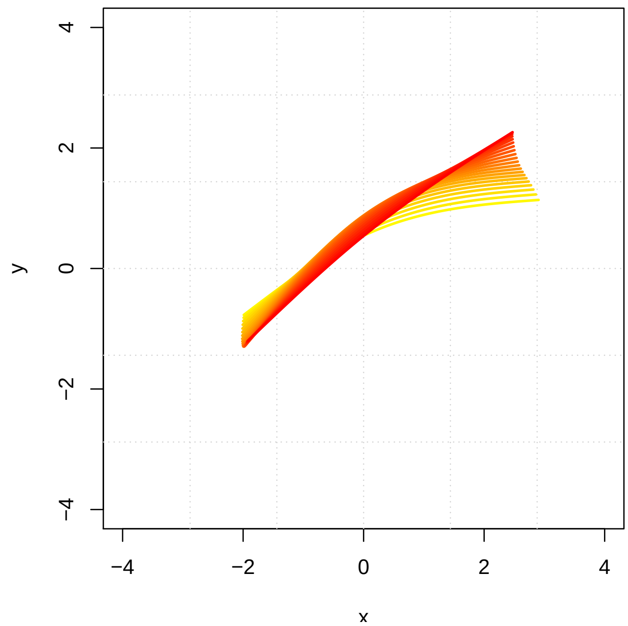

Figure 3.26: Dynamic model of [n^j], yellow to red represents first to last sample. Tongue oriented anterior (left) to posterior (right). Axes are scaled.

the TB sensor begins at a higher and more anterior position, with an overall convex tongue shape at the onset of the sound.

The alveolar lateral approximant is also different, lacking the higher TB sensor position observed in the consolidated data. It does, however, maintain a raised TT position that would presumably create a closure at the alveolar ridge that is characteristic of an alveolar lateral approximant. The alveolar nasal also has a flatter tongue angle in comparison to the model of the alveolar nasal from the combined data. Both the alveolar nasal and lateral produced by Speaker 1 are similar to their combined counterparts, demonstrating much less movement in comparison to the other sounds included in this study.

Speaker 2 was 47 years old at the time of data collection. She had resided in São Paulo City for 29 years and Champaign-Urbana for 18 years. It is immediately noticeable that Speaker 2 demonstrates much less tongue movement for all seven phones, especially in comparison to Speaker 1. The overall shape and position of the tongue appears to be similar to models created from Speaker 1 and the consolidated data, though there are some definite differences on closer investigation.

One particular difference between Speaker 2 and the consolidated data can be observed

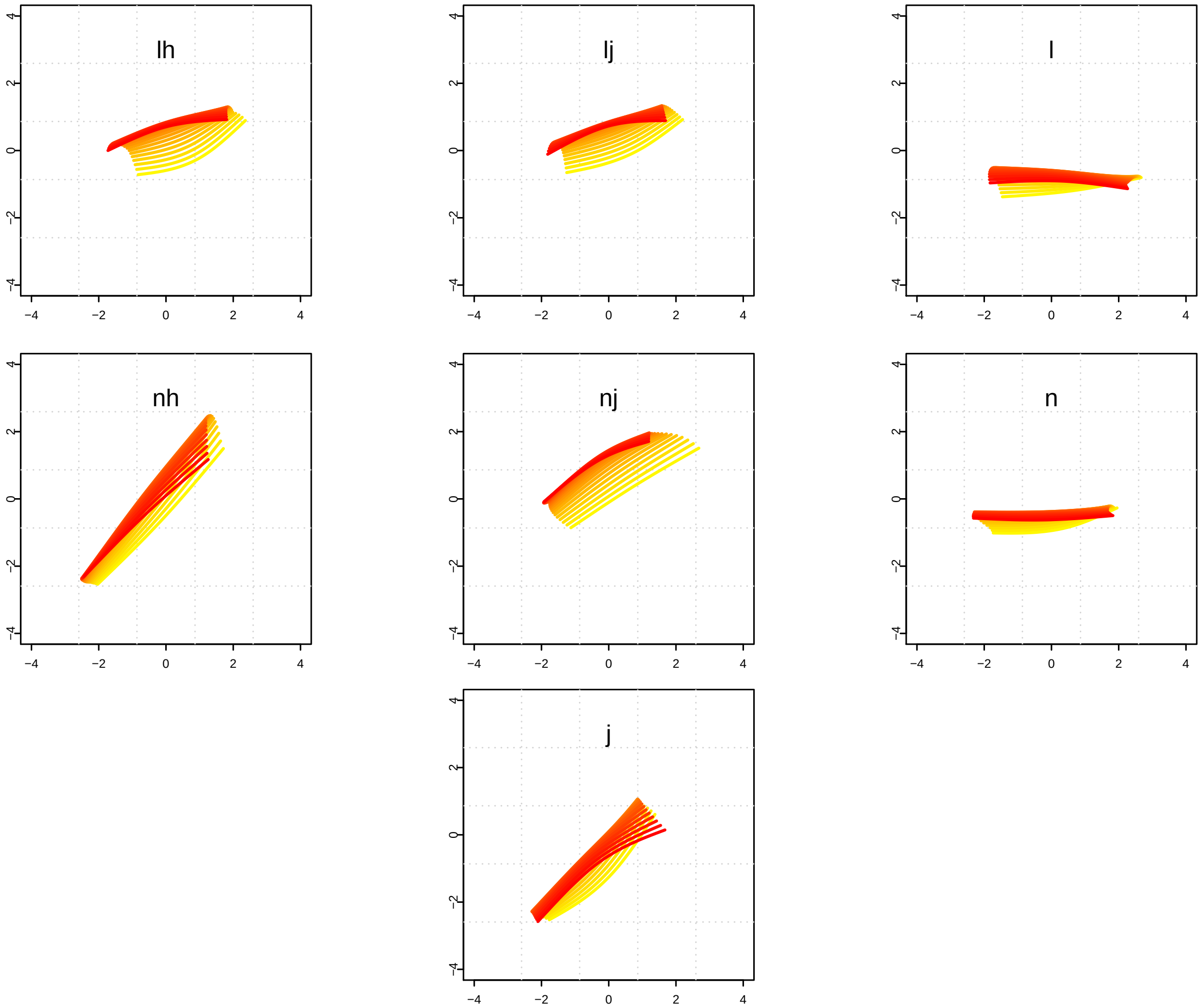

Figure 3.27: Dynamic anterior tongue models for Speaker 1.

when comparing the palatal and palatalized lateral. Unlike in the consolidated data or with Speaker 1, Speaker 2 articulates the palatal lateral with a more anterior tongue tip position. By comparison, the palatalized lateral is articulated with a more retracted and higher tongue tip position, though both the palatal and palatalized lateral are produced with the TB sensor in similar positions.

Here, the palatalized nasal bears a much greater resemblance to the palatal lateral, especially when examining the position of the tongue tip. In fact, the palatalized nasal is articulated with a slightly more anterior tongue tip position in comparison to the palatal lateral.

Speaker 2 articulates the alveolar nasal and lateral approximant with a much more pro-

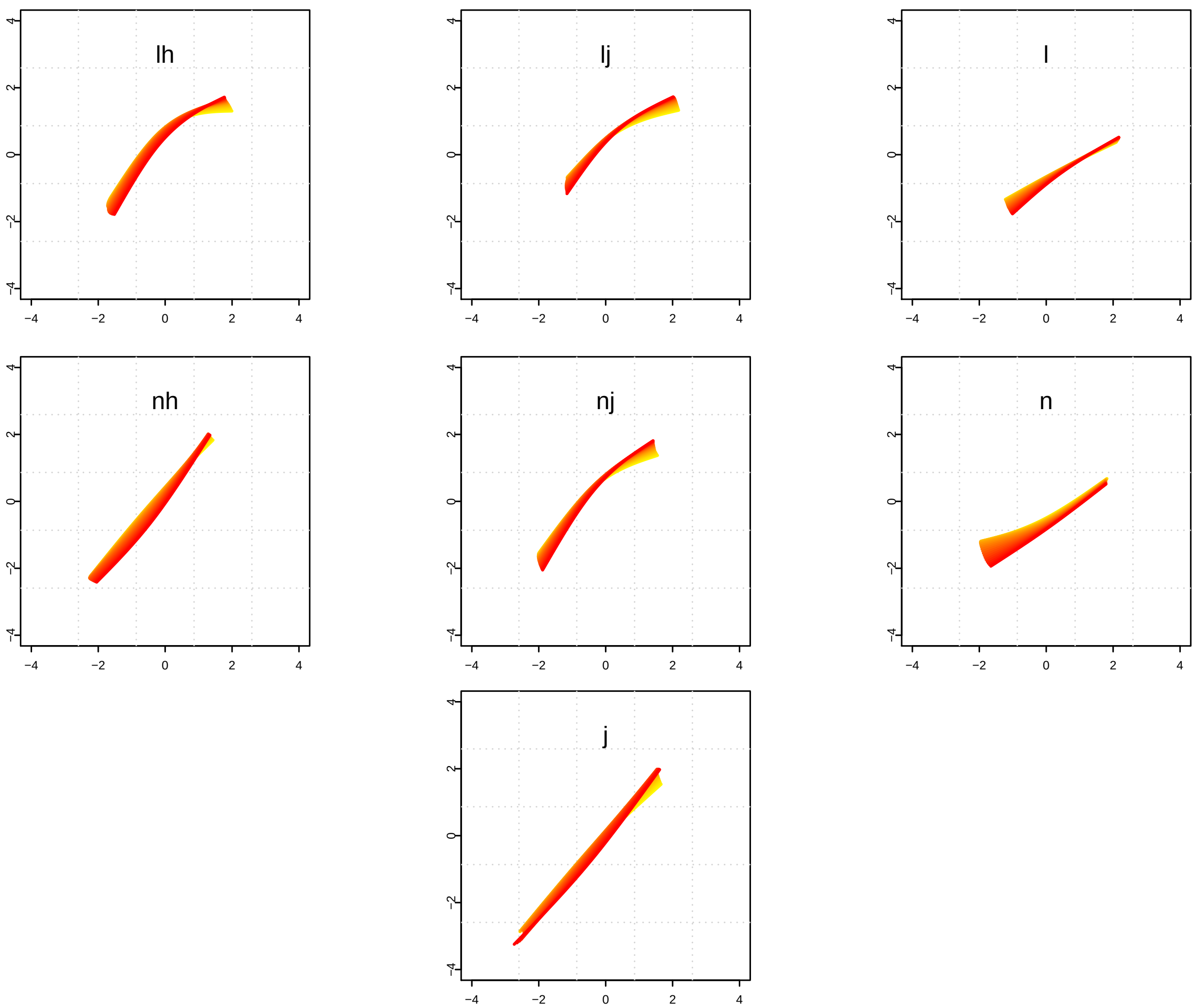

Figure 3.28: Dynamic anterior tongue models for Speaker 2.

nounced tongue angle in comparison to both Speaker 1 and the consolidated data. Unlike Speaker 1 but similar to the consolidated data, Speaker 2 lowers her tongue tip as the phone progresses for both alveolar sounds. Again, neither the alveolar nasal nor alveolar lateral approximant are articulated with as similar of a tongue shape and movement as expected from the literature; the alveolar nasal is articulated with a noticeably more anterior tongue tip position and a larger range in tongue tip height across the duration of the phone.

Also interesting to note is that the palatal approximant is not quite as similar to the palatal nasal as is observed in the consolidated data. Speaker 2 demonstrates slightly more movement when producing the palatal nasal, particularly around the TM sensor. The palatal approximant is articulated with a much more anterior and lower tongue tip position as well, though both phones have similar TB positions throughout the production of either phone.

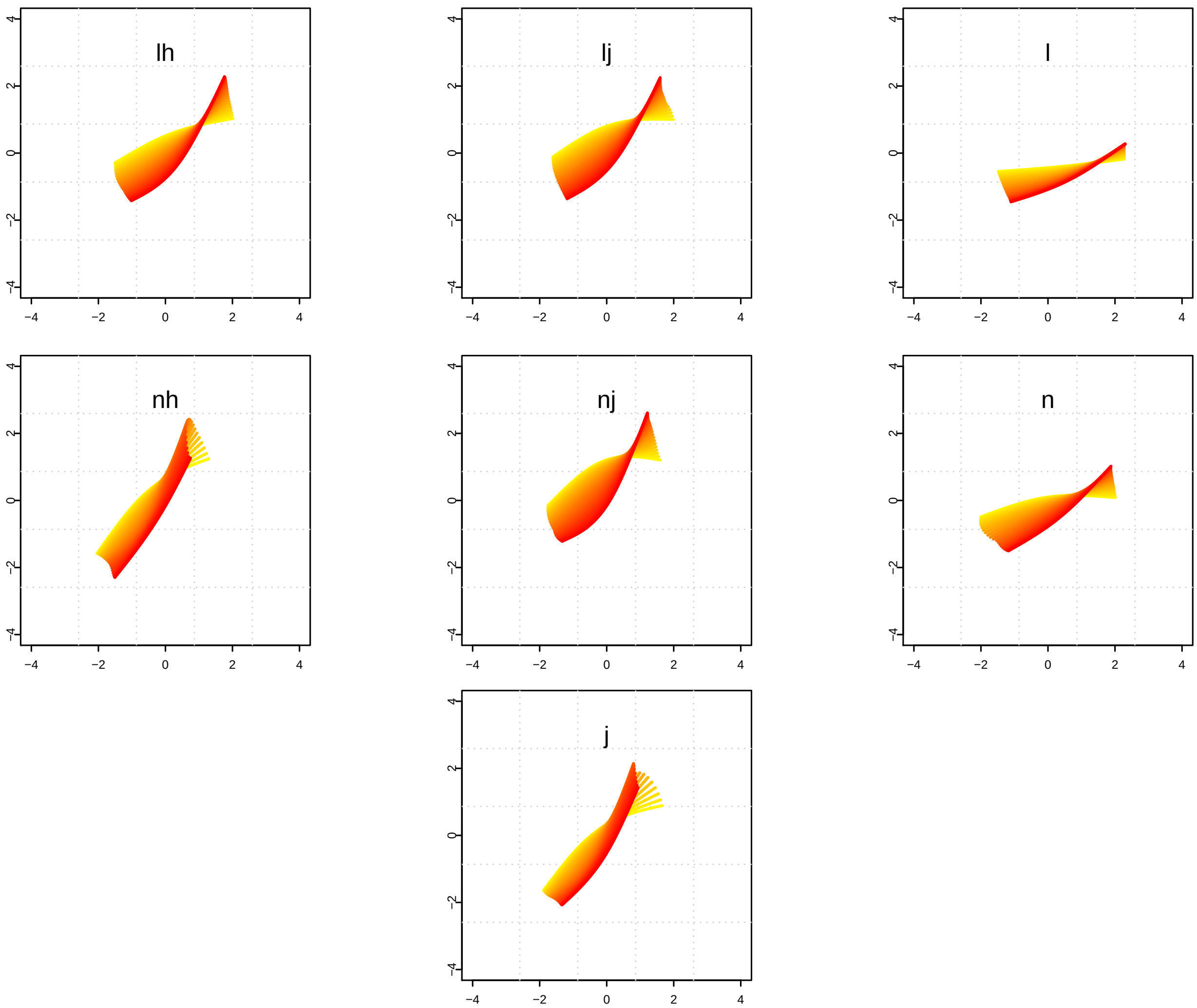

Figure 3.29: Dynamic anterior tongue models for Speaker 3.

Speaker 3 was 55 years old at the time of data collection. She lived in São Jose dos Campos for 30 years, followed by another 25 years in Champaign-Urbana. This speaker self-reported her accent as “Rio”, purportedly because she had been born in Rio de Janeiro. In general, Speaker 3 demonstrates more movement during the articulation of all seven sounds, similar to Speaker 1.

Again, like Speaker 1, Speaker 3 produces the palatalized lateral with a slightly more anterior tongue position. Other than this small difference in frontness, models of the palatal and palatalized lateral appear very similar to one another. However, unlike the previous two speakers, Speaker 3 articulates the palatal and palatalized lateral with a concave tongue shape that graduate changes to convex. Models of the palatalized nasal are also similar to the palatal and palatalized lateral, with the palatalized nasal being noticeably more anterior than both the palatalized and palatal lateral.

As in all the previous analyses, the palatal nasal bears a greater resemblance to the palatal approximant as opposed to the palatalized nasal. Similar to the consolidated data and Speaker 1, the palatal nasal begins and ends at similar positions after demonstrating a rise in tongue height at all three sensor positions. Unlike the consolidated data however, there are easily noticeable differences in how Speaker 3 articulates the palatal nasal and the palatal approximant. Speaker 3 does not articulate the palatal approximant with the same TB height at the onset and end of the sound, unlike the palatal nasal. In comparison to the consolidated data, Speaker 3 articulates both the palatal nasal and approximant with much more movement at the TT and TM sensor, though movement at the TB sensor appears similar.

There are also prominent differences in how Speaker 3 articulates the alveolar lateral approximant and the alveolar nasal. The alveolar nasal is articulated with a final TB position that is much higher than the onset, especially when compared the alveolar lateral approximant. There is also more lowering and retraction of the tongue tip apparent in the alveolar nasal that is not present during the articulation of the alveolar lateral approximant.

Speaker 4 was 47 years old at the time of data collection. She lived in São Paulo city for 19 years before moving to the United States. This speaker spent several summers abroad in Canada or the United States prior to her move. The dynamic models created from her speech are similar to Speaker 2, revealing less movement in comparison to the other speakers. However, there are notable differences - where Speaker 2 articulates the alveolar lateral approximant with a concave tongue shape and the alveolar nasal with a convex tongue shape, Speaker 4 articulates the two sounds with the opposite shapes.

Differences between the palatal and palatalized lateral are more apparent in the models created for Speaker 4 than in the previous models. Interestingly, the length of the tongue is considerably longer for the palatalized lateral than the palatal lateral. The palatalized lateral demonstrates comparatively less movement around the tongue tip, with a far more anterior TT sensor position. However, TB sensor position is relatively similar between the

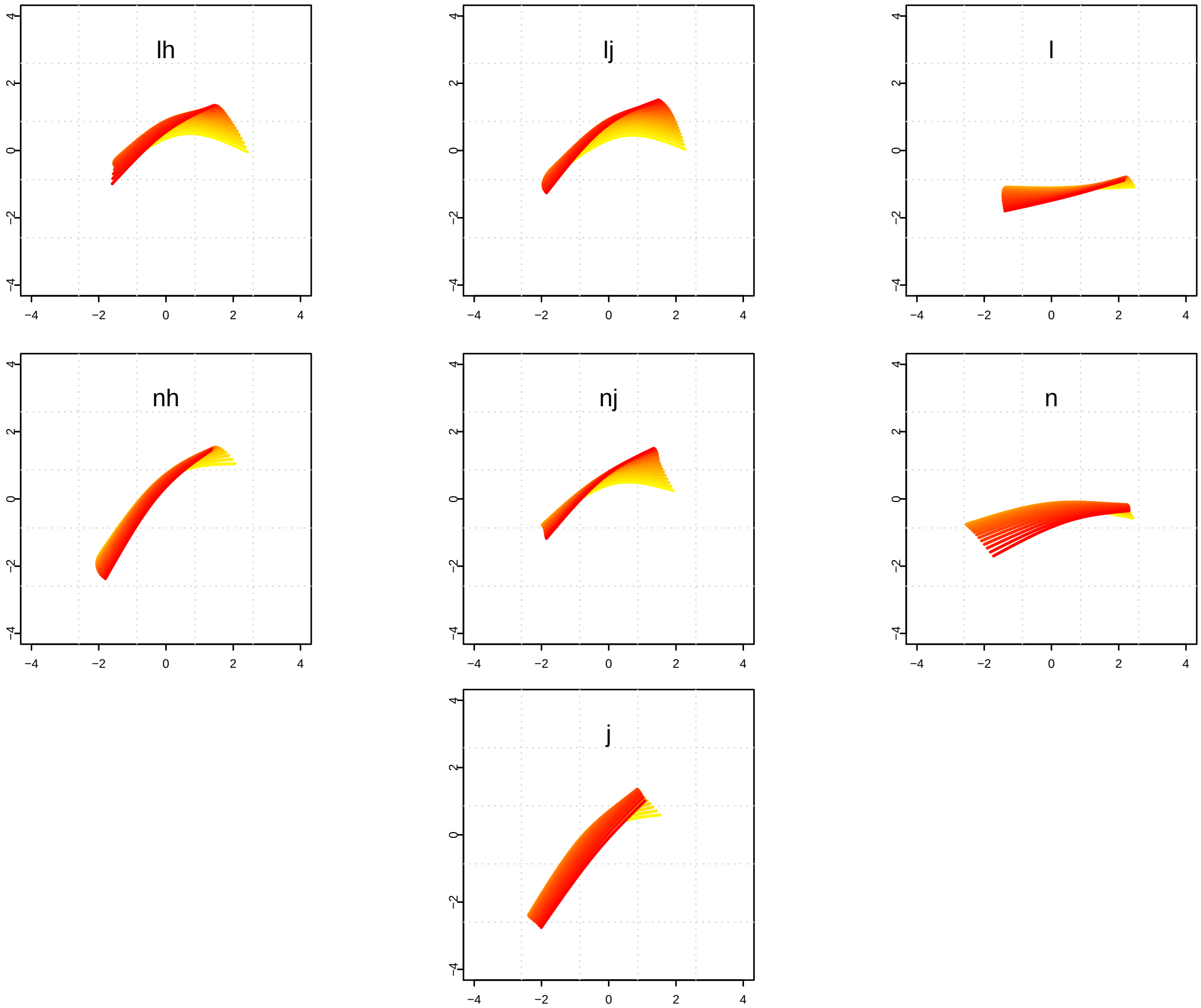

Figure 3.30: Dynamic anterior tongue models for Speaker 4.

palatal and palatalized lateral. Again, like in the previous models, the palatalized nasal is articulated with a similar overall tongue shape but with a slightly more anterior tongue position at all three tongue sensors.

There are larger differences between the palatal nasal and palatal approximant in Speaker 4, especially when compared to the consolidated data. In relation to the palatal nasal, the palatal approximant lowers and retracts more, though the initial tongue shape appears similar. The initial overall height of the palatal nasal is also much greater across all three sensors than the palatal approximant.

Interestingly, Speaker 4 produces the palatal and palatalized nasal with the most similarity, relative to the previous speakers and the consolidated data. Both sounds begin with a concave tongue shape that becomes less concave as time progresses, as a result of the tongue back rising. The length of the palatalized nasal in comparison to the palatal nasal is also striking; the palatal nasal is noticeably longer than the palatalized nasal, indicating that the tongue is able to stretch and compress to accommodate different articulations.

As noted above, the shape of the tongue for the alveolar sounds is different from the previous three speakers. The alveolar constriction appears to be achieved early on in the phone, with the tongue tip lowering for the majority of either phones' duration. The majority of the movement occurs at the tongue tip, with the tongue back remaining rather immobile.

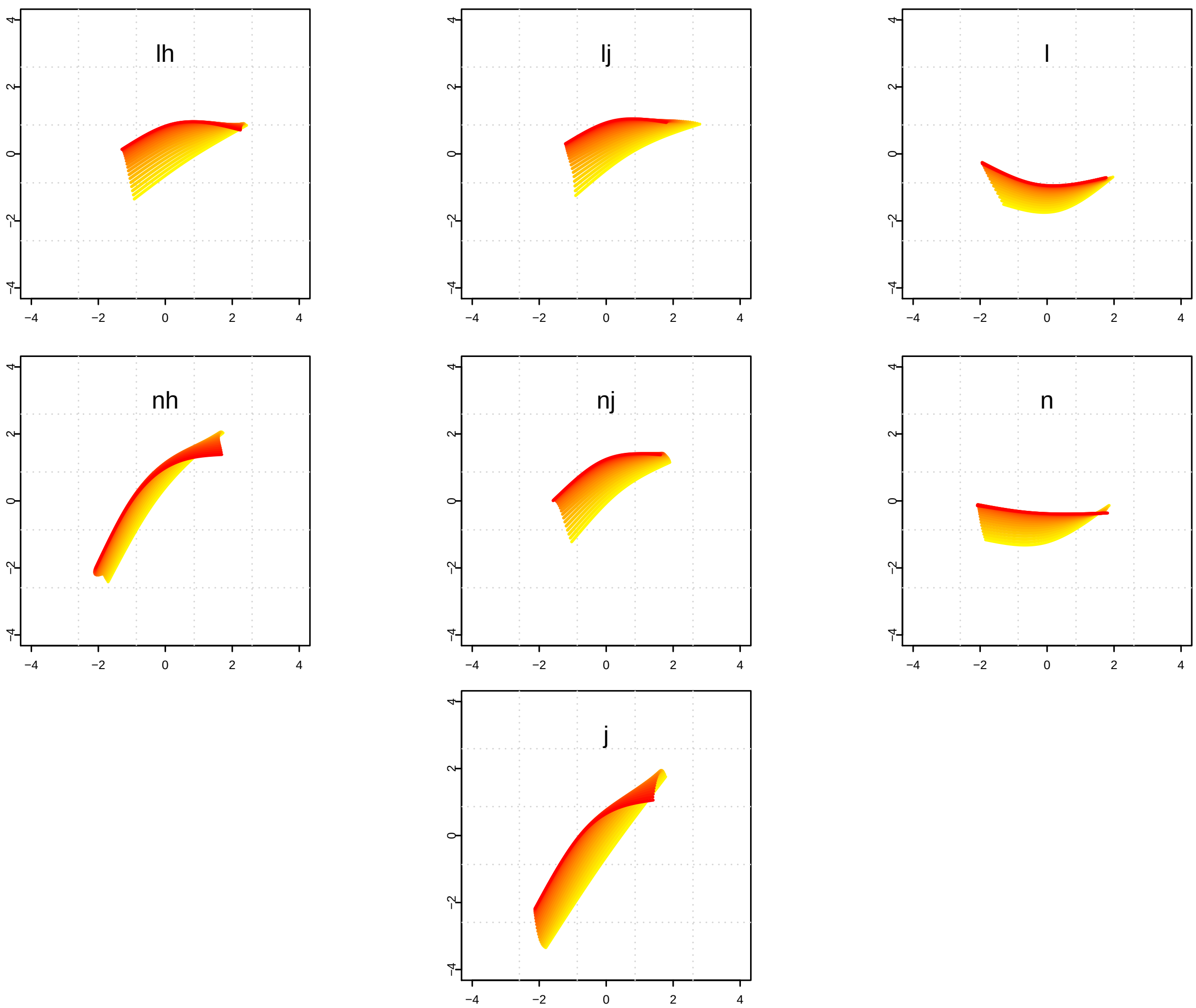

Figure 3.31: Dynamic anterior tongue models for Speaker 5.

Speaker 5 was 20 years old at the time of data collection. She had lived in São Paulo state for 19 years and Champaign-Urbana for one year. This speaker also provided the acoustic stimuli for the perception study included in the appendix (see Section A). Her voice was recorded separately for the perception study; the acoustic and articulatory data presented in this chapter and the following chapter were not collected from the same data acquisition session.

Unlike the consolidated data and majority of the previous speakers, Speaker 5 produces five of the seven speech sounds (excluding the palatal approximant and palatal nasal) with the maximum tongue position occurring at the end of the sound. This indicates that Speaker 5 completes the articulatory gesture at the end of the following sounds: /ʎ/, [l^j], /l/, /n/, [n^j]. The palatal nasal and palatal approximant are produced with the maximum tongue position occurring at approximately the first third of the phone duration, similar to the consolidated data.

Unlike the previous speakers, the differences between the palatal lateral and palatalized lateral are much less distinct; there is no visibly obvious evidence of increased tongue tip anteriority in the palatalized lateral, though the back of the tongue appears to be slightly more retracted for the palatalized lateral. There is also evidence of the tongue lowering at the TB sensor and rising at the TM sensor at the end of the palatal lateral, though not in the palatalized lateral.

As observed previously in the above analyses, there is a pronounced similarity between the palatalized nasal and palatal and palatalized lateral. The palatalized nasal is slightly more anterior at the tongue tip in comparison to the palatal and palatalized lateral. The palatalized nasal is also considerably more anterior at the TB sensor and demonstrates greater height at the TM and TB sensor, creating a steeper angle between the TT and TM sensor. All three sounds are similar in that the majority of the movement occurs at the tongue tip, while the TB sensor remains relatively immobile.

There are also observable similarities between the palatal nasal and palatal approximant; as mentioned above, both sounds are the only sounds to be articulated with the maximum tongue height occurring at about a third of the way into the phone by Speaker 5. While the two sounds are similar, they are not identical – the palatal approximant demonstrates greater variation in tongue height across time in comparison to the palatal nasal, which is particularly evident at the tongue tip. Additionally, the TB sensor moves forward as the palatal approximant is articulated, while the same sensor retracts during the palatal nasal. While the palatal approximant begins with the tongue tip much lower than the palatal nasal, both sounds end with a very similar final overall tongue position.

As in the above models, the alveolar lateral approximant is articulated with a much more pronounced convex curve in the tongue anterior in comparison to the alveolar nasal, though both sounds are articulated with similar final (and maximum) tongue tip positions. In comparison, the shape of the tongue during the alveolar nasal is comparatively neutral and becomes increasingly more so as the sound progresses, resulting in a nearly linear final tongue shape. Like the palatal lateral, palatalized lateral, and palatalized nasal, the two alveolar sounds are articulated with the greatest variation at the tongue tip and an immobile TB sensor.

Speaker 6 was 21 years old at the time of data collection. She had lived in São Paulo state for 20 years and Champaign-Urbana for one year. Despite the similarity in age to Speaker 5, the models created for Speaker 6 bear a greater resemblance to those of Speaker 2 and Speaker 4.

Like Speaker 2, Speaker 6 articulates the seven sounds with relatively less movement. As such, tongue movement throughout the phone duration appears to be quite static. For Speaker 6, this is particularly evident in models of all the sounds except for the palatal and palatalized lateral; these two sounds are articulated with a greater range of tongue position.

Change in tongue shape over time observed in Figure 3.32 more greatly resembles that of Speaker 4 than Speaker 2. Both Speaker 6 and Speaker 4 produce the palatal and palatalized sounds with a lower tongue back position that gradually rises until the maximum height is achieved at the end of the sound. This movement is in conjunction with a tongue tip position that continues to lower until the end of the sound.

However, Speaker 6 articulates the palatal and palatalized lateral differently from Speaker 4. Figure 3.32 illustrates that Speaker 6 has a more anterior initial tongue tongue tip position during the palatal lateral, in comparison to the palatalized lateral. In general, it appears that the palatal lateral is relatively more anterior, which is particularly evident when comparing the position of the posterior portion of the tongue model. However, as the palatal lateral is articulated, Speaker 6 retracts the tongue tip. This results in the palatalized lateral

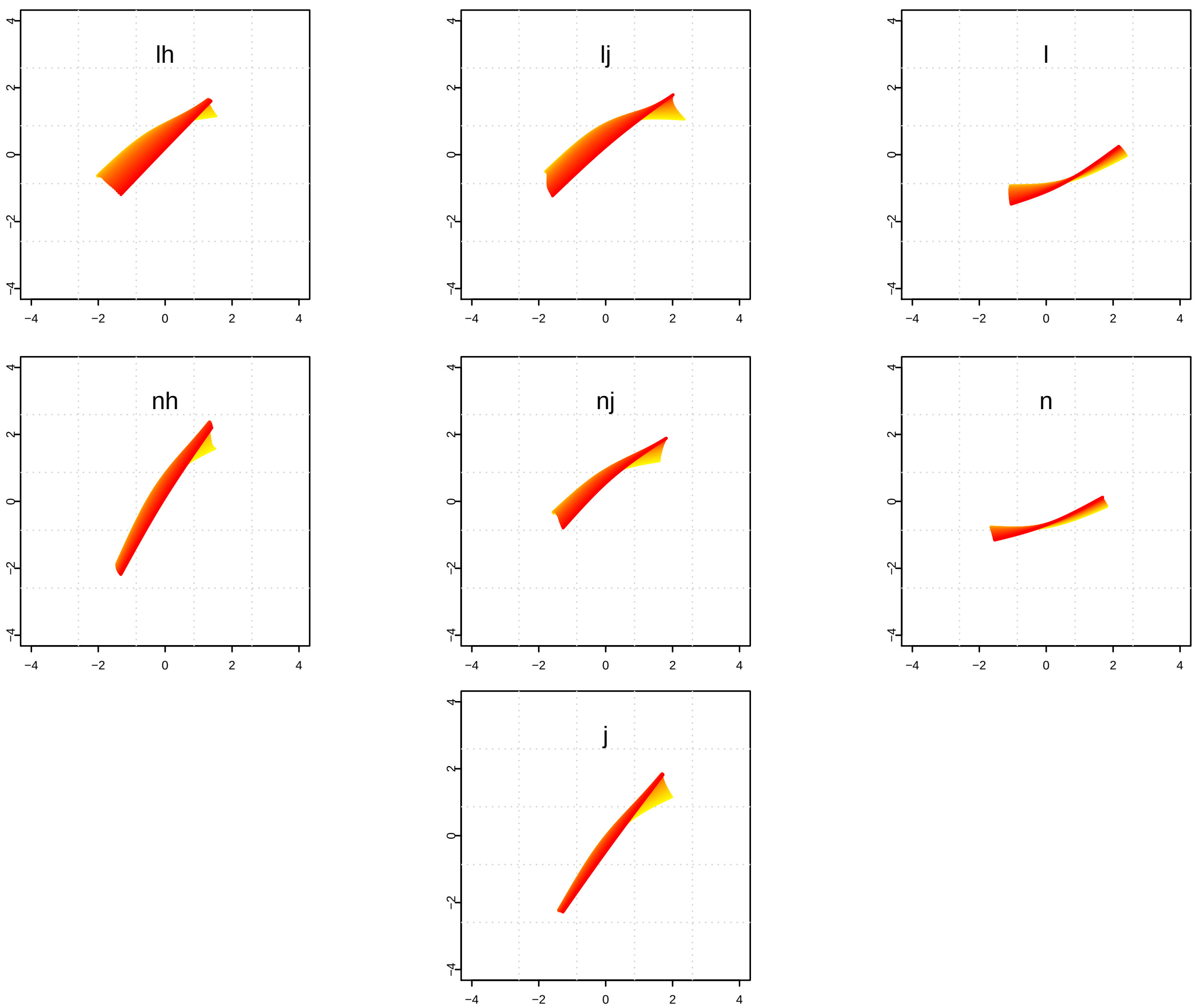

Figure 3.32: Dynamic anterior tongue models for Speaker 6.

ultimately demonstrating a comparatively more anterior tongue tip position at the end of the sound, despite not having moved a great deal with regards to frontness.

The palatalized nasal produced by Speaker 6 is more similar to the palatal lateral than the palatalized lateral in terms of distance between the TT and TB sensor; the palatalized lateral is considerably longer with regards to this distance. The overall characteristic movement during the palatalized nasal however is similar to the palatal and palatalized lateral: the tongue back raises as the tongue tip lowers. Like the palatal and palatalized lateral, the palatalized nasal is articulated with a more moderate tongue angle in comparison to the palatal nasal and palatal approximant. This is also true of the models created from the consolidated data.

The palatal nasal and palatal approximant are again quite similar, though the palatal approximant is articulated with a relatively more retracted and lower tongue back. Since the tongue tip position is similar for both sounds, this results in a larger tongue degree angle for the palatal nasal. Both sounds are articulated with less movement at the tongue tip in comparison to the palatal and palatalized lateral and palatal nasal. In this aspect, Speaker 6 is more similar to Speaker 1.

The alveolar nasal and alveolar lateral approximant are very similar to one another when produced by Speaker 6, especially in comparison to the previous speakers. With the exception of frontness (the alveolar nasal is more anterior than the alveolar lateral at all three sensors), the overall shape of the two sounds are very similar. Both the alveolar nasal and alveolar lateral approximant are articulated with a slightly convex tongue shape at the onset of the phone; as the tongue tip lowers, the shape becomes increasingly linear.

Speaker 7 was 32 years old at the time of data collection. She was born in Campinas and continued to reside in São Paulo state for 28 years before moving to North America. The models for Speaker 7 are most similar to those of Speaker 3, who is not in the same age group of Speaker 7. Note that Speaker 3 self-reportedly speaks with a “Rio” accent.

The similarity between Speaker 7 and Speaker 3 is particularly evident when comparing the palatal and palatalized lateral and palatal nasal. Both speakers produce these sounds with a relatively high tongue tip and low tongue back, which respectively lower and raise as the sound progresses. This results in a section of the tongue, close to the tongue back, that seems to stay immobile throughout the duration of the phone for both speakers. However, while the tongue shape remains concave throughout the duration of these three sounds when produced by Speaker 7, the tongue moves from a concave to convex shape when articulated by Speaker 3.

While Speaker 3 articulates the palatalized lateral with a visibly more anterior overall tongue position in comparison to the palatal lateral, this difference in anteriority is not obvious in Speaker 7. However, Speaker 7 does articulate the palatalized lateral with a slightly higher overall tongue position across time relative to the palatal lateral, as well as a relatively more retracted tongue back position at the onset.

Like Speaker 3, Speaker 7 articulates the palatalized nasal with a relatively more anterior

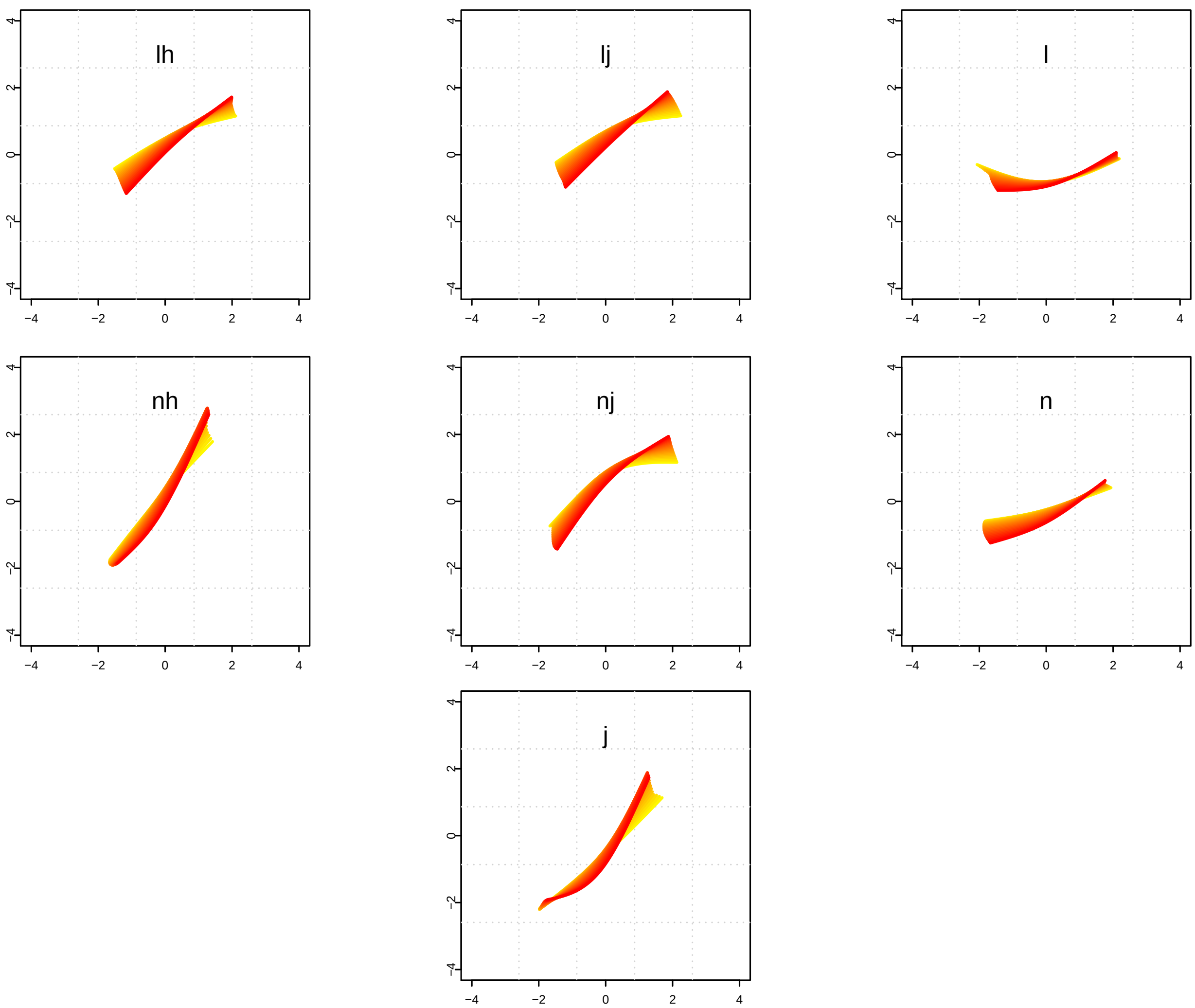

Figure 3.33: Dynamic anterior tongue models for Speaker 7.

tongue tip position. Unlike Speaker 3 however, the tongue tip is considerably lower during the palatalized nasal in comparison to the palatal and palatalized lateral. Additionally, while the tongue back position during Speaker 7's production of the palatalized nasal remains very similar to the palatalized lateral, the section of the tongue near the TM sensor is noticeably higher. This creates a larger arch in the curvature of the tongue shape during the palatalized nasal in comparison to the palatal and palatalized lateral.

As expected, the palatal nasal is more similar to the palatal approximant than the palatalized nasal, with a much higher tongue back position and lower tongue tip. This results in the steeper tongue angle observed in Section 3.3.3. There are a few differences between the palatal nasal and palatal approximant however. While the tongue tip position is similar to the palatal approximant, the palatal nasal is articulated with an obviously higher tongue back position. Additionally, towards the end of the palatal approximant, the front portion of the tongue begins to flatten out and creates a bend in the part of the tongue near the tongue tip.

The alveolar nasal and alveolar lateral approximant are articulated similar to the other speakers, with the maximum tongue position occurring near the beginning of the phone. Once the articulatory target is reached, i.e. once the tongue tip touches the alveolar ridge, the tongue tip lowers until the end of the sound. Movement occurs primarily at the tongue tip, with very little movement towards the tongue back. Like many of the above speakers and like the consolidated data, the alveolar lateral approximant is articulated with a lower TM sensor, resulting in a small bend in the tongue shape. In contrast, the tongue remains relatively linear during the production of the alveolar nasal.

Speaker 8 was 31 at the time of data collection. She lived in Rio de Janeiro for six years, after which she lived in São Paulo for 15 years. The remaining years were divided between Spain, Aracaju (Sergipe state), and São Caetano (São Paulo state). She had only recently moved to Champaign-Urbana. Her accent was self-reported as São Paulo "countryside Portuguese". Speaker 8 was in the same age group (30-39 years) as Speaker 1, Speaker 7, and Speaker 9; of those speakers, Speaker 8 is most similar to Speaker 7.

Speakers 1 through 7 were consistent in pronouncing [l^j], /ʎ/, and [n^j] similar to one another; this trend is not present in Speaker 8. The articulation of the palatalized lateral is more similar to the palatalized nasal while the palatal lateral is more similar to the palatal nasal (at least with regards to the shape of tongue, though not the tongue angle). As the palatalized lateral and palatalized nasal are articulated, the tongue becomes increasing linear, while the tongue anterior develops a slight convex curvature towards the end of the palatal lateral and palatal nasal. These differences suggests that Speaker 8 does distinguish between the palatal and palatalized lateral, though it is unknown whether listeners would also be able to distinguish between the two pronunciations.

While the articulation of the palatal lateral by Speaker 8 is more similar to the palatal nasal than the palatalized lateral, the palatal nasal and the palatal approximant are still more similar to one another. This follows the pattern observed in the previous speakers and

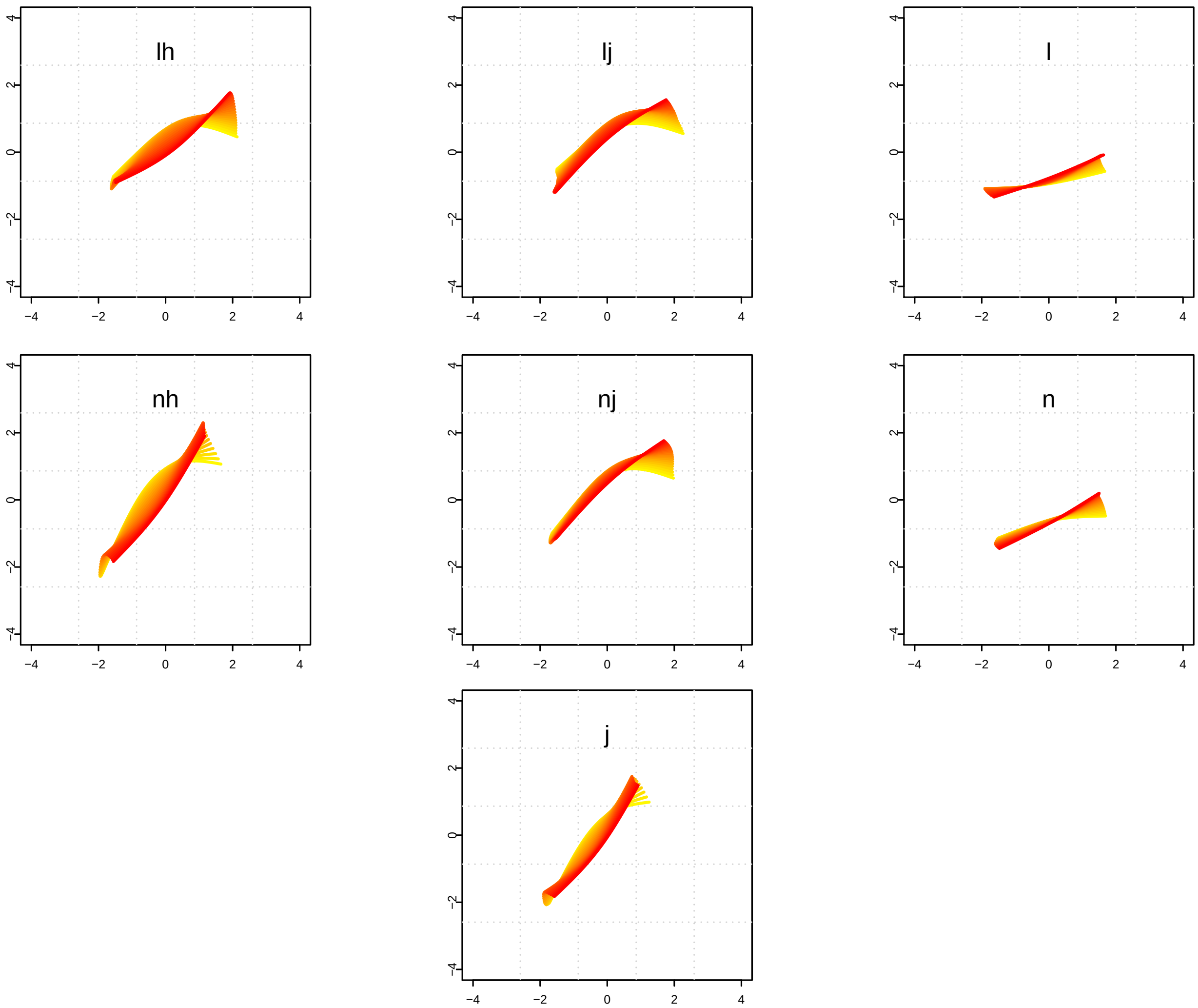

Figure 3.34: Dynamic anterior tongue models for Speaker 8.

the consolidated data. However, there are some obvious differences between the palatal nasal and palatal approximant. The change in tongue height is much less during the production of the palatal approximant in comparison to the palatal nasal, in addition to a comparatively larger distance between the TT and TB sensor during the palatal nasal.

The concavity that is present at the onset of all five palatal or palatalized sounds suggests that Speaker 8 articulates the palatal gesture by pressing the tongue blade up and against the hard palate. The tongue back then rises continuously across time as the tongue mid lowers, indicating that the palatal gesture is begin released. While the tongue tip and tongue mid remain relatively stable for the palatalized lateral and nasal, the tongue mid region of the tongue drops noticeably during the production of the palatal lateral, nasal, and approximant.

It is unclear why the palatal sounds release the palatal gesture in a different manner from the palatalized sounds, though it presents additional evidence supporting the observation that Speaker 8 differentiates between palatalized and palatal sounds.

Like the previous speakers and the consolidated data, Speaker 8 pronounces the alveolar lateral approximant and alveolar nasal with a flat tongue shape and nearly neutral tongue angle. There is also evidence of the tongue tip bending upwards during the production of the alveolar lateral approximant, which is also observed in previous speakers. However, the majority of movement during these two sounds occurs towards the tongue back at the beginning of the phone.

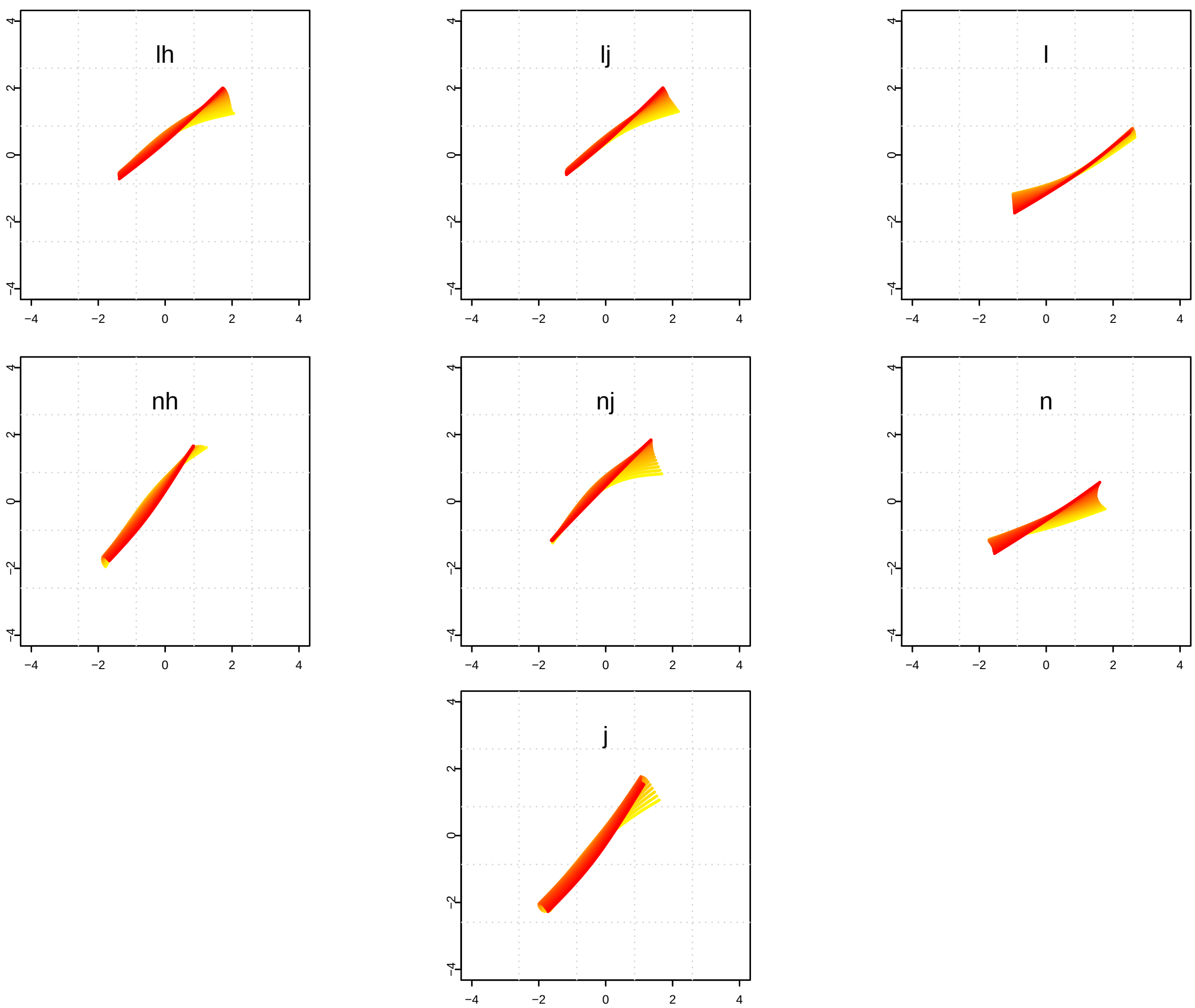

Figure 3.35: Dynamic anterior tongue models for Speaker 9.

Speaker 9 was 30 years old at the time of data collection. She had lived in various regions of São Paulo state; the lengthiest stay was in Ibitinga for 18 years. Her accent was described as “more pronounced”. Out of all 10 speakers, Speaker 9 was the only speaker that was not currently residing in Champaign-Urbana; her current residence at the time was Bauru, Brazil.

Excluding Speaker 1, Speaker 5, and Speaker 10, Speaker 9 follows the general trend demonstrated in the previous speakers; the tongue appears pressed up against the palate at the onset of the palatal and palatalized lateral and palatalized nasal, after which the gesture is gradually released. This is illustrated by a comparatively lower tongue back position at the phone onset, which gradually rises as the phone progresses.

Like Speaker 5, the differences between the palatal and palatalized lateral are not very obvious. When articulated by Speaker 9, the tongue tip during the palatal lateral is slightly more anterior and lower than the palatalized lateral. There also appears to be a little more movement at the tongue tip and tongue mid during the palatal lateral. However, these differences are not as pronounced as it is in other speakers, especially in comparison to Speaker 6, who is in the same age group as Speaker 5 (20-29 years).

The differences between articulation of the palatalized nasal and the palatal and palatalized lateral by Speaker 9 are more visible than the differences between the palatal and palatalized lateral. The entire tongue anterior has very clearly shifted forward and down for the palatalized nasal, with a larger arch in the tongue body at the onset. As with the previous speakers and the consolidated data, Speaker 9 articulates the palatalized nasal similar to the palatal and palatalized lateral, with a clear divide between the palatalized nasal and palatal nasal.

As expected, the palatal nasal is more similar to the palatal approximant, though the palatal approximant shows much more tongue back raising. The tongue back during the palatal approximant is much lower than the palatal nasal, approximating the same height as the tongue back during the palatal and palatalized lateral and palatalized nasal. In contrast, the palatal nasal is articulated with a much higher onset tongue body. There is more movement during the palatal approximant than the palatal nasal, with the tongue lowering more overall during the palatal approximant.

Speaker 9 articulates the alveolar lateral approximant and alveolar nasal according to expectations; there is a slight bow in the tongue tip region during the alveolar lateral approximant and a comparatively more neutral tongue shape during the alveolar nasal. Interestingly, the target does not appear to be the same; the tongue tip is considerably more anterior during the alveolar nasal. This is similar to all the other speakers, except for Speaker 8 and Speaker 10, who do not demonstrate any difference in tongue tip anteriority during the two alveolar sounds.

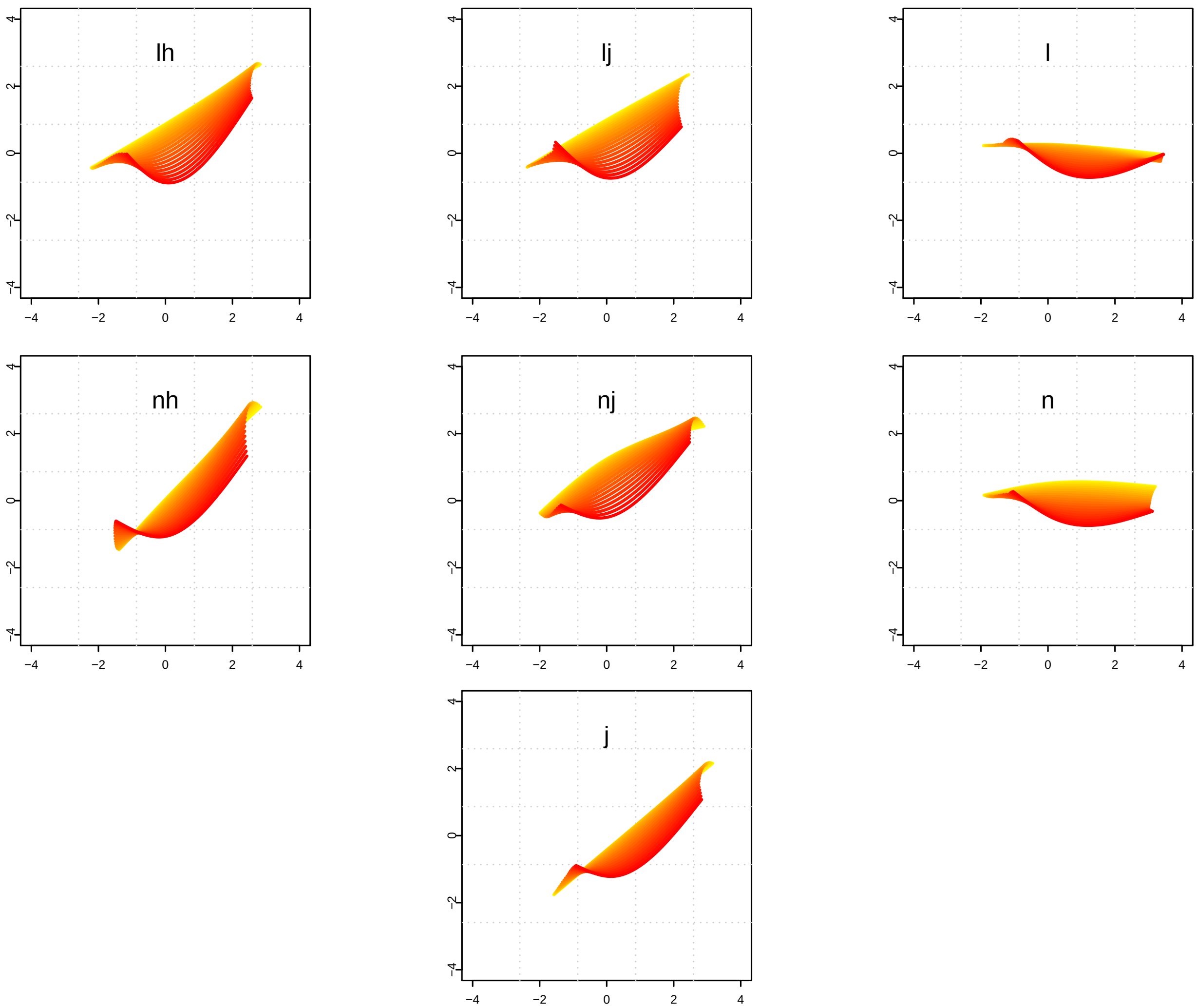

Figure 3.36: Dynamic anterior tongue models for Speaker 10.

Speaker 10 was 20 years old at the time of data collection. She lived in São Paulo city for 19 years before moving to Champaign-Urban for one year. With regards to age, Speaker 10 is in the same group as Speaker 5 and Speaker 6 (20-29 years). However, there are notable differences between Speaker 10 and the other members of her age group. Speaker 10 demonstrates much more variation in tongue position across time in general. Additionally, while Speaker 10 and Speaker 6 both achieve the target articulation near the onset of the phone, Speaker 5 achieves the maximum tongue position near the end of the phone.

Out of all 10 speakers, Speaker 10 demonstrates the greatest amount of movement once the target or maximum tongue position is achieved. This movement away from the articulatory target could be interpreted as the tongue relaxing to a more natural or neutral position. The final tongue position for all seven sounds varies with regards to the reported TB sensor height, but all the sounds culminate with a raised tongue tip.

Speaker 10 produces the palatal and palatalized sounds with the greatest amount of similarity in comparison to the other nine speakers. There are small differences in anteriority and height, but the overall shape of the tongue is very similar across time for all five palatal and palatalized sounds. Collectively, the palatal and palatalized sounds are articulated with a relatively flat and high tongue shape that gradually lowers across time. As the sounds are produced, a section of the tongue near the tongue tip begins to lower while the tongue tip remains relatively immobile (or even moving upwards), resulting in a deep bend at the front of the tongue. Unlike the previous speakers, Speaker 10 produces the palatal and palatalized sounds with relatively similar tongue degree angles.

Like several of the previous speakers and the consolidated data, Speaker 10 articulates the palatalized lateral with a relatively more anterior tongue in comparison to the palatal lateral. The back of the tongue is noticeably lower during the palatalized lateral, especially as the phone is articulated. In contrast to the palatal lateral, the final tongue tip position is much higher during the palatalized lateral.

Of all 10 speakers, Speaker 10 produces the palatalized nasal the most similar to the palatal nasal, though there are obvious differences in tongue height (the tongue tip is noticeably lower and the tongue back higher for the palatal nasal, resulting in a larger tongue degree angle). The palatalized nasal is the only palatal or palatalized sound to be articulated with a concave tongue shape at the onset; the other palatal or palatalized sounds have a flat initial tongue shape.

The palatal nasal and palatal approximant are articulated by Speaker 10 with a larger tongue degree angle than the other palatal or palatalized sounds, with the palatal nasal demonstrating the largest angle. Of all the palatal sounds, the palatal nasal has the highest tongue back position. While the tongue tip is initially more anterior and higher during the palatal nasal in comparison to the palatal approximant, the tongue tip for both sounds is very similar at the culmination of both phones.

With regards to the alveolar sounds, Speaker 10 produces the two phones with a final convex tongue shape, similar to the palatal or palatalized sounds. However, the alveolar lateral approximant and alveolar nasal have a much more neutral tongue blade angle. The final tongue back height is similar for the two sounds, but the alveolar nasal begins with a higher initial TB and TM sensor. The alveolar lateral approximant is articulated with a relatively more retracted tongue back, suggesting that Speaker 10 possesses a so-called "dark" or velarized /l/.

3.3.6 Understanding the linguopalatal relationship

Situating a model of the tongue against a model of the palate will assist in providing a visual understanding of the relationship between the tongue anterior and the hard palate with regards to the production of palatal and palatalized sounds. While only the palates of two speakers will be considered here for analysis, this section seeks to provide a preliminary description of the linguopalatal relationship in palatal and palatalized laterals and nasals that may be informative to future researchers who wish to perform a similar analysis using methodologies more suitable for obtaining an accurate image of the hard palate, i.e. ultrasound or MRI.

A trace of the palate was performed for nine of the 10 speakers by instructing the speaker to trace the roof of their mouth with the tip of their tongue, starting from as far back in the oral cavity as was comfortable. In doing so, the TT sensor should provide a reasonable approximation of the hard palate. Note, since the examination of suitable palate traces was necessarily post-hoc as observation of the traces was conducted in Matlab (MATLAB, 2014) following data processing, it was not possible to ask participants to continue making attempts until an acceptable trace had been obtained. The palate traces for two speakers (Speaker 4 and 10) were found suitable for analysis; the results from their data will be presented here.

As in Section 3.3.5, a natural spline of 2 degrees was used to model both the tongue and the palate (see Figures 3.37 and 3.38). Separate models were built for the palate and tongue contour of each speaker. The blue contour represents the palate, with light gray to dark gray representing dynamic changes in the tongue contour from the first to last sample point of each phone. There are a total of 21 discrete sample points in time (see Section 3.3.4 for how the sample points were obtained). The plot is oriented such that the oral cavity anterior is on the left and the posterior is on the right.

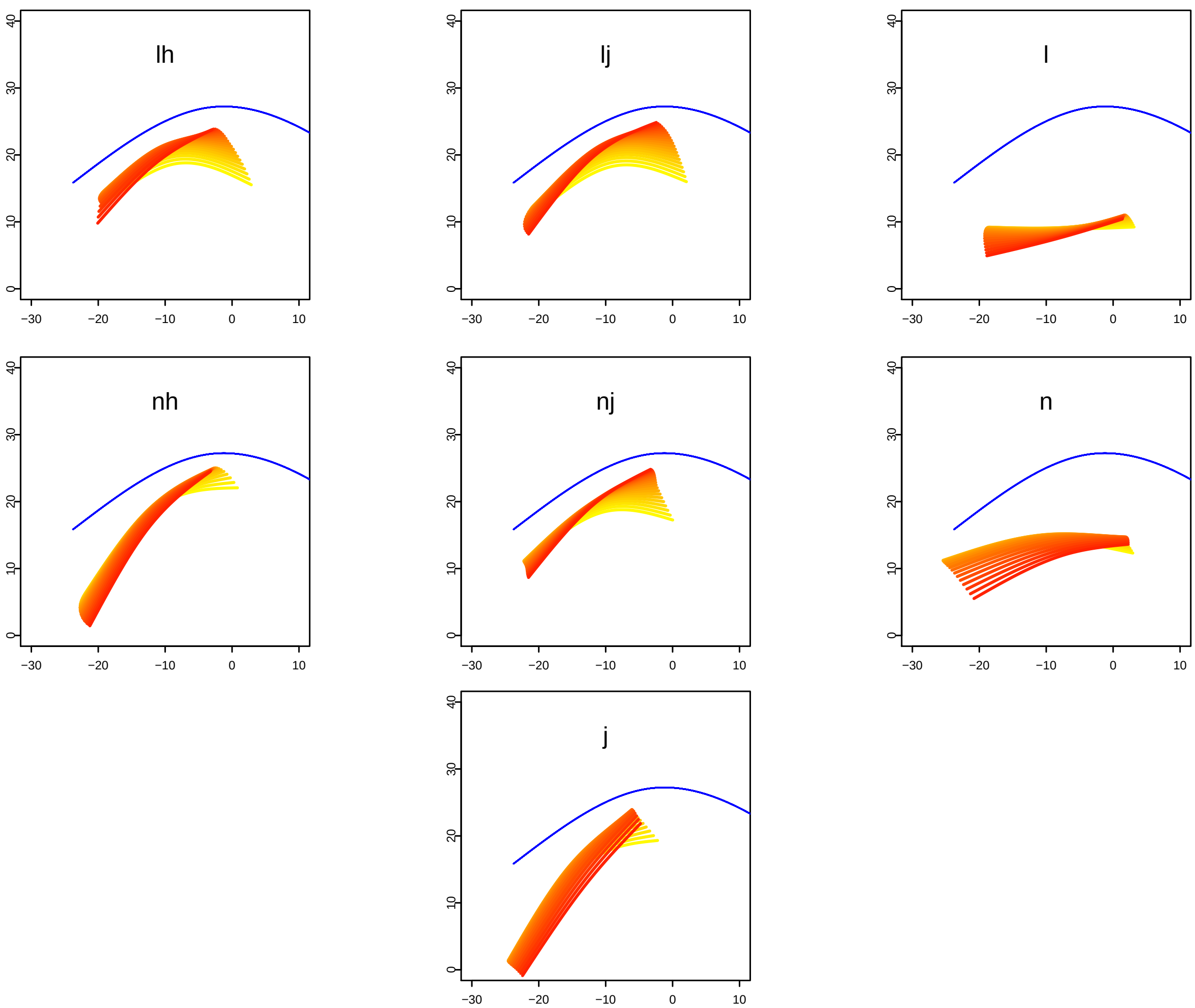

Figure 3.37: Speaker 4: Dynamic models of the tongue at 21 points in time. Yellow to red represents first to last sample point, blue represents hard palate. Plot orientated with tongue anterior on left, posterior on right. Axes units in mm.

Figure 3.37 reveals clear similarities in how Speaker 4 articulates palatal and palatalized sounds; a high tongue body position and a lowered tongue tip is easily observable for all five palatal and palatalized sounds studied here. There are, however, subtle differences: 1) the shape of the tongue anterior does not change greatly for the palatal approximant and palatal nasal across time, and 2) the tongue shape remains roughly the same, simply lowering across time. In contrast, the palatal and palatalized lateral and palatalized nasal are articulated with a domed tongue shape for the first half of the phone. As the sounds are articulated, the posterior portion of the observable tongue shape rises while the anterior half lowers, pivoting around the middle.

Likewise, the two alveolar sounds /l/ and /n/ bear an expected resemblance to one another, with a level anterior tongue position and space between the hard palate and the middle of the tongue. Both sounds appear to pivot around the position of the TB sensor. However, the alveolar nasal has a higher and more forward tongue position, with an arched tongue anterior that contrasts with the bowing apparent in the alveolar lateral.

The models created for Speaker 4 are similar to the models built over the combined data (see Section 3.3.5), with some minor differences. The models of the alveolar sounds are more similar to each other in the combined data, with bowing in both sounds. While the alveolar nasal is also more forward than the alveolar lateral, the change in height is not present as it is in the models here for Speaker 4. The pivoting around the tongue tip that is apparent in the models of the palatal nasal and palatal approximant in the combined data is not observable here. In contrast, the pivoting motion around what is likely the TM sensor for the palatal and palatalized lateral and palatalized nasal in Speaker 4's data is not present in the models built for the combined data; there is greater movement of the middle area of the anterior tongue for those models.

The palate for Speaker 10 was raised slightly by 2.5 mm to accommodate the tongue models (see Figure 3.38). The models built for Speaker 10 are visibly different from Speaker 4, with the most notable difference observable in the palatal and palatalized sounds. Figure 3.38 indicates that instead of the domed anterior tongue shape that characterized palatal and palatalized sounds in Figure 3.37 and in Section 3.3.5, these sounds are articulated by Speaker 10 with a an increasingly bowed tongue anterior. Despite the change in shape, the tongue

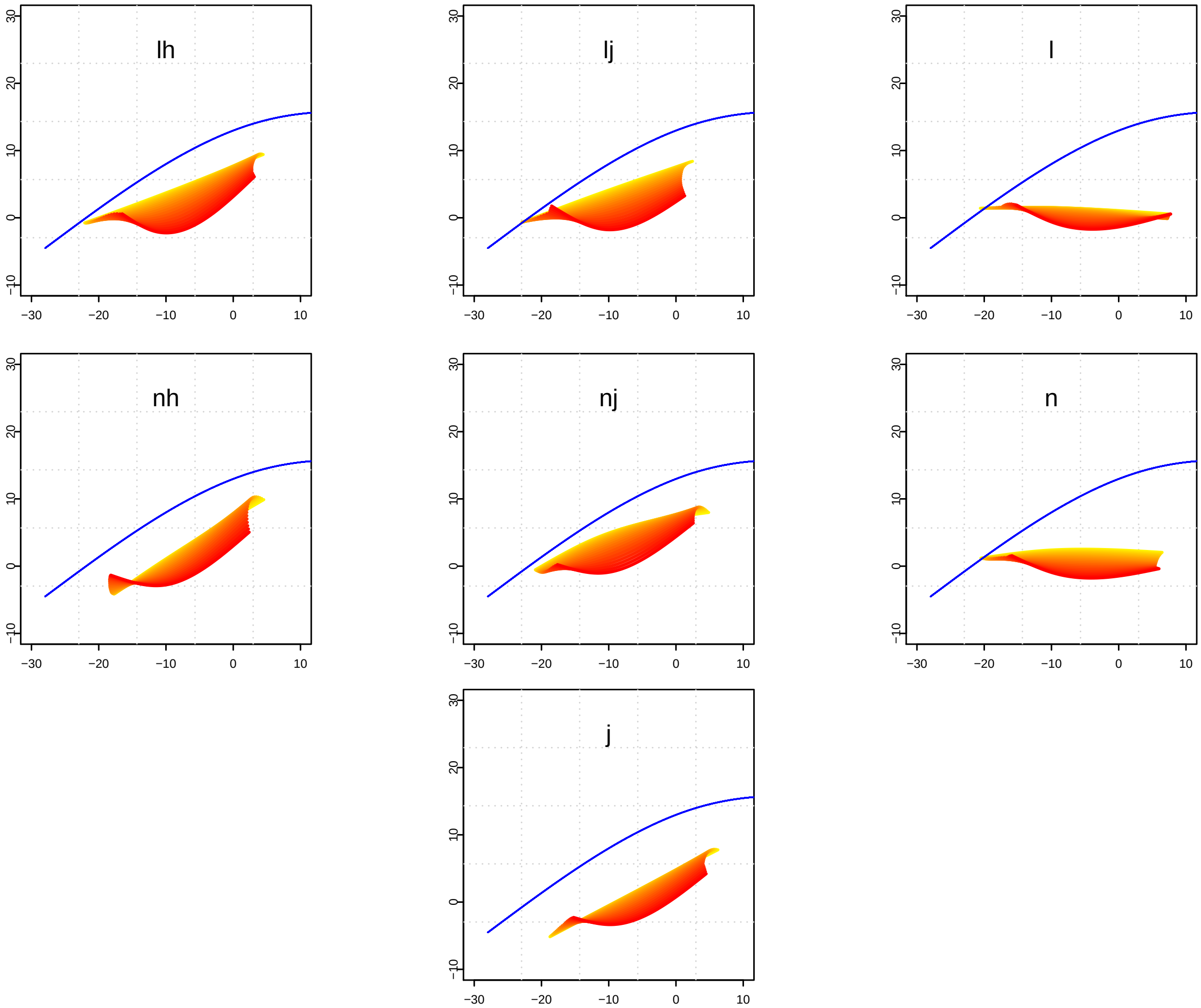

Figure 3.38: Speaker 10: Dynamic models of the tongue at 21 points in time. Yellow to red represents first to last sample point, blue represents hard palate. Plot orientated with tongue anterior on left, posterior on right. Axes units in mm.

tip appears to maintain contact with the alveolar ridge. Contrasting with Speaker 4, the models of for the alveolar lateral and alveolar nasal are nearly identical, with a obvious contact between the tongue tip and the alveolar ridge.

Similarities are also present however, with striking resemblances between the palatal and palatalized lateral and palatalized nasal, as well as between the palatal nasal and palatal approximant. Like the models from Section 3.3.5, the palatal nasal and palatal approximant appear to pivot around the tongue tip. The previously observed pivot point near the TM sensor is not, however, present in Speaker 10.

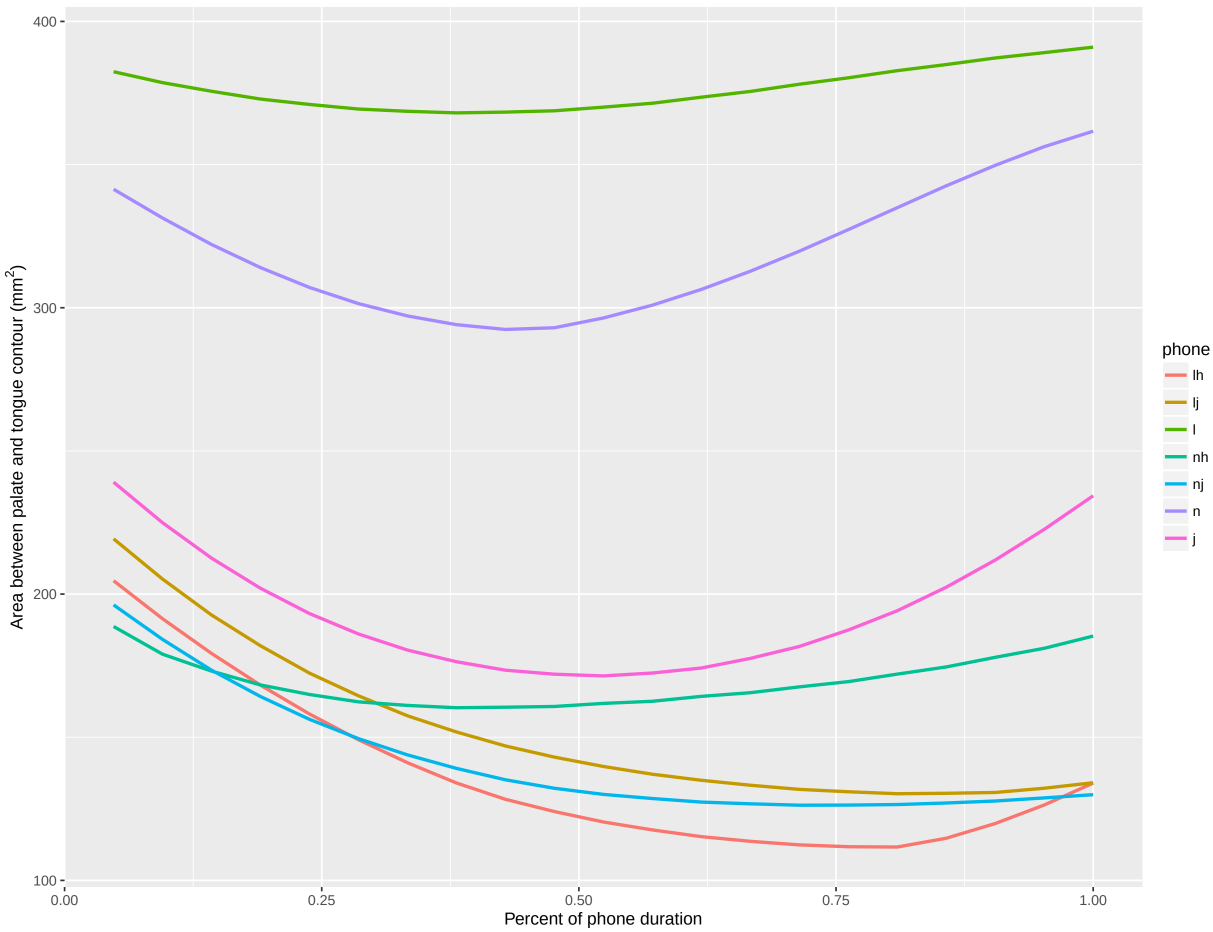

Figure 3.39: Speaker 4: Change in area between palate and tongue contour in mm^2 throughout duration, by phone.

A comparison of how the area between the palate and the anterior portion of the tongue as it changes across time was conducted as a means of quantifying the difference between the sounds. The area between the palate and each tongue contour was calculated in R (R Core Team, 2015) using the function *integrate*() from the package *stats* (R Core Team, 2015). Figures 3.39 and 3.40 illustrate the change in area between the palate and tongue anterior across time.

There are expected similarities between certain sounds, especially given the results presented in the above sections; for both speakers, we find that the palatal and palatalized lateral and palatalized nasal pattern together, while the alveolar sounds and remaining palatal sounds are paired together. However, the actual change in area for the palatal and

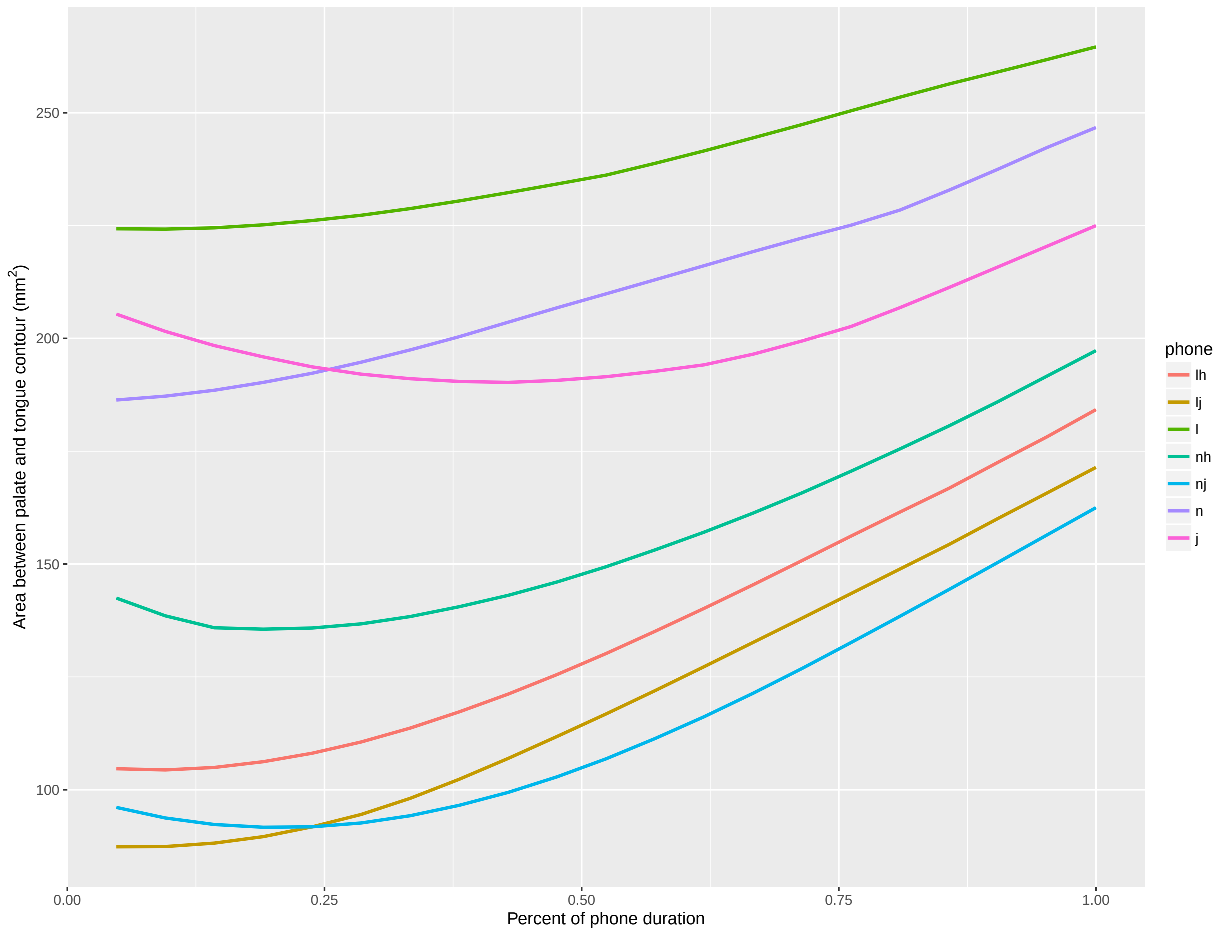

Figure 3.40: Speaker 10: Change in area between palate and tongue contour in mm^2 throughout duration, by phone.

palatalized sounds is different between the two speakers: Speaker 4 achieves the palatal gesture towards the end of the phone (indicated by declining area in Figure 3.39) while Speaker 10 completes the gesture towards the beginning of the phone (evident by the initially lower areas in Figure 3.40).

Since the models of the tongue shapes produced by Speaker 4 resemble the models created for the combined data of all 10 speakers, the general trend is towards achieving the palatal gesture at the end of the phone. Whether this trend holds true for the larger population of BP speakers will need to be tested through a more extended study. While both speakers claim São Paulo city as their hometown, it is possible that a difference in age (Speaker 10 is 20 years old, while Speaker 4 is nearly 30 years older) may have a small effect on speech patterns, though the effect is unlikely to be strong.

Though the timing for when the gesture is achieved is different for these two speakers, there are some clear similarities in terms of overall proximity to the palate. Summing the area between the palate and the tongue anterior across phone duration (see Table 3.2) provides a clear indication of these similarities. The two alveolar sounds have the largest combined area, followed by the palatal approximant and palatal nasal. The lowest combined areas belong to the palatal and palatalized lateral and palatalized nasal. While the motivation for grouping the palatal approximant with the palatal nasal and the palatal/palatalized lateral with the palatalized nasal is not as clear here (the difference in the summed area is not large), it is strongly supported by the findings presented in the previous sections.

	ʎ	lʲ	l	ɲ	nʲ	n	j
Speaker 4	2,876	3,201	7,908	3,561	2,979	6,703	4,090
Speaker 10	2,838	2,539	5,031	3,281	2,419	4,441	4,206

Table 3.2: Combined area between palate and tongue anterior across phone duration. Areas given in mm^2.

It is unsurprising that the two non-palatal sounds included here have the largest combined area; however, it is unexpected that the palatal nasal and palatal approximant have the next largest combined area, especially given that the palatal nasal and palatal approximant demonstrate the largest tongue blade angle (see Figures 3.10 - 3.12), indicating the strongest laminal articulation. While a laminal articulation may be characteristic of stereotypical palatal sounds (of which the palatal approximant is one), the extent of palatal contact or constriction may not need to be extensive when observing simply the anterior oral cavity. That is to say, at least with regards to the first centimeters of the tongue, strongly laminal sounds demonstrate less palatal contact or constriction. This corresponds with previous research (Recasens et al., 1993; Recasens and Pallarès, 2001; Recasens and Espinosa, 2006) that finds that the palatal approximant and palatal nasal is articulated further back on the palate than the palatal lateral.

3.4 Discussion

This section concludes the first portion of the study of the articulation and acoustics of the BP palatal lateral. I provide a detailed analysis of the articulation of the BP palatal lateral through a battery of experimental measures; these measures were carried out in comparison to the sounds /l, ɲ, n/ and [l^j, n^j]. I find that the articulation of the palatal lateral better resembles the palatalized lateral and nasal than the palatal approximant. In short, the BP palatal lateral can be described as articulated with a (1) -31 degree tongue blade angle with respect to the occlusal plane at the beginning of the phone, (2) concave anterior tongue shape that gradually flattens out over time, and (3) a palatal constriction that is achieved towards the end the of the phone.

Contrary to expectations that *yeísmo* might be the result of articulatory similarity between the palatal lateral and the palatal approximant, the two sounds do not exhibit any particularly remarkable similarities. Additionally, there are only minimal differences between the articulation of the palatal and palatalized lateral, with both differences existing only in the tongue tip: 1) the tongue tip during the production of the palatalized lateral is significantly higher than the palatal at the middle of the phone, though only slightly so, and 2) the tongue tip during the first half of the palatal lateral is significantly more anterior than the palatalized lateral. The implications of these differences will be discussed in the following section.

While this chapter was unable to uncover more or larger differences in the articulation of the palatal and palatalized lateral, there is a strong and consistent pattern that is present throughout the results of the statistical analysis. Throughout the multiple measures conducted in this chapter, the palatal and palatalized lateral are typically grouped together with the palatalized nasal, while the palatal nasal and approximant form one group and the alveolar nasal and lateral form another. Potential reasons for why these sounds were consistently found to be grouped together will be discussed in the sections below.

3.4.1 Palatal versus palatalized lateral

In the introduction, the palatal lateral was defined as possessing the following articulatory characteristics: (1) high front tongue body, (2) laminal articulation, (3) lingual contact along the hard palate extending up to two to three times as far as /l/, (4) a lateral air stream channel on one or both sides of the tongue. While the presence of a lateral air stream could only be assumed due to methodological limitations, the first three characteristics are confirmed and supported by the evidence presented in this chapter. Establishing the basic definition of the palatal lateral in BP was essential due to the limited information available on the BP palatal lateral - while there have been recent studies on the palatal lateral in European Portuguese (Martins et al., 2008; Teixeira et al., 2012), it cannot be assumed that these two varieties will be identical.

Figures 3.6 - 3.8 (illustration of sensor position distribution by phone) indicate that the BP palatal lateral is higher and fronter than the BP alveolar lateral (a baseline for this study) for all three sensors at three points of the phone duration (i.e. beginning, middle, and end). The difference in tongue height is particularly pronounced at the TM and TB sensor (see Figure 3.7 and Figure 3.8), with frontness particularly pronounced at the TM sensor (see Figure 3.8). By observation alone, there does not appear to be any significant differences between the palatal and palatalized lateral in these figures. This observation is supported statistically at least for the TM and TB sensor in the SS-ANOVA results (see Figure 3.13 - 3.15 and Figure 3.17 - 3.18), which indicate that there is no evidence of significant difference between the palatal lateral and palatalized lateral; note, this finding is in terms of tongue height or frontness with regards to the portion of the tongue that encompasses 2 cm to 3 cm behind the tongue tip. As mentioned in the previous section, Figure 3.14 and Figure 3.16 indicate that there is evidence of significant difference in the tongue tip during the production of the palatal and palatalized lateral; respectively, the figures indicate that the palatalized lateral is significantly higher than the palatal lateral at the middle of the phone, while the tongue tip is significantly more anterior during the first half of the palatal lateral. However, these differences, while significant, appear minor: the trajectory of the sensors during the production of these two sounds maintain close proximity.

Contrary to observations of a neutral tongue angle in x-ray tracings of the Russian palatal lateral (Fant, 1960:163), the BP palatal and palatalized lateral demonstrate a distinctly negative tongue angle similar to observations of the European Portuguese palatal lateral, which is articulated with a high tongue body and the tongue tip touching the lower teeth (Teixeira et al., 2012:322), as well as x-ray images of the Spanish palatal lateral before it underwent *yeísmo* (Straka, 1965). There is also no difference found between the palatal and palatalized lateral regarding the presence of a laminal articulation; both are articulated laminally and statistical analysis does not indicate a significant difference in terms of the tongue blade angle (see Figure 3.10 - 3.12). While the tongue blade degree does become moderately steeper as either phone progresses (a progression that is also visible in the dynamic models in Figures 3.20 - 3.26, which are based off interpolated data), the lack of significant difference remains.

The absence of evidence of significant difference between the tongue angle of the palatal and palatalized lateral is interesting, given the significant differences in height and frontness of the tongue tip during the production of these two sounds. This can be easily resolved by postulating the hypothetical configuration as illustrated in Figure 3.41; since the TT sensor is simultaneously more anterior and lower during the first half of the palatal lateral, the resulting angle created by the TT and TM sensor is the same for both the palatalized and palatal lateral.

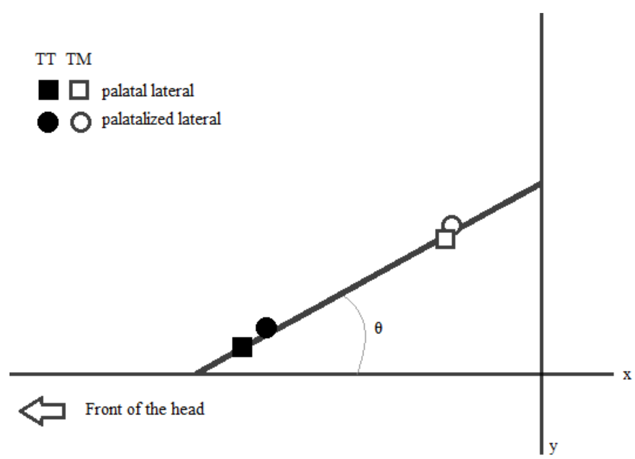

Figure 3.41: Illustration of possible TT and TM configuration resulting in a more forward TT for the palatal lateral, while maintaining a similar tongue angle degree with the palatalized lateral.

While actual physical contact with the hard palate could not be measured, models of the anterior portion of the hard palate permitted the calculation of total area between the anterior tongue contour and the palate for two speakers. These calculations indicate that there may be two to three times less area between the modeled palate and tongue contour for both the palatal lateral and palatalized lateral in comparison to the alveolar lateral, providing further support that palatal sounds demonstrate two to three times more linguopalatal contact than their alveolar counterparts (Keating, 1988:87). Based on palatograms of the historical Spanish palatal lateral, Navarro Tomás (1968) claims that the area of linguopalatal contact is nearly identical for the palatal lateral, palatal nasal, and palatal approximant. While the reported combined area is clearly different for these three sounds in BP, they are certainly more similar than the alveolar sounds; the summed area for the palatal approximant is about 0.4 times larger than the palatal lateral, while the alveolar sounds have a summed area nearly two to three times as big as the other palatal and palatalized sounds.

These models of the linguopalatal relationship also indicate that while not the majority, observations of tongue tip contact with the teeth and a flatter tongue body shape in the Russian palatal lateral (Fant, 1960:163) may be true for some speakers of BP; Speaker 10 (see Figure 3.38) appears to exhibit tongue tip contact with may be the alveolar ridge or even the teeth as well as a flatter anterior tongue body in comparison to Speaker 4 (see Figure 3.37) and models of the tongue shape for the pooled participant information (see Figures 3.20 - 3.26). These observations of palatal contact, while potentially speaker-dependent, can be applied to the palatalized lateral as well; there again does not appear to be a difference between the palatal and palatalized lateral.

Unlike observations of the historical Spanish palatal lateral made by Lipski (1989), we do not find that the palatal lateral in BP is composed of an alveolar lateral approximant component and a palatal component. Neither do we find that this is the case for the palatalized lateral or the palatalized nasal (in this case, an alveolar nasal component and palatal approximant component). With regards to the articulation, if the palatalized lateral and palatalized nasal were composed of distinct consonant and glide elements, we would expect to find that these sounds begin with an articulatory configuration resembling the corresponding alveolar sound and culminate at a position approximating the palatal approximant. However, this is not the case, as can be clearly observed in Figure 3.13 - Figure 3.15 and Figure 3.16 - Figure 3.18).

As mentioned above, findings from the two speakers indicate that the palatal and palatalized lateral and palatalized nasal demonstrate the greatest proximity, followed by the palatal nasal, the palatal approximant, the alveolar nasal stop, and then the alveolar lateral approximant. However, an EPG study on Catalan consonants by (Recasens, 1984b) indicates that the alveolar nasal has the least amount of tongue dorsum contact, followed by the palatal lateral, palatal nasal, and then the palatal approximant, while other EPG studies on Catalan (Recasens et al., 1993; Recasens and Pallarès, 2001; Recasens and Espinosa, 2006) also find that the palatal lateral is articulated further forward than the palatal nasal. A potential explanation for why the results from this dissertation do not contradict previous research is provided in Figure 3.5.

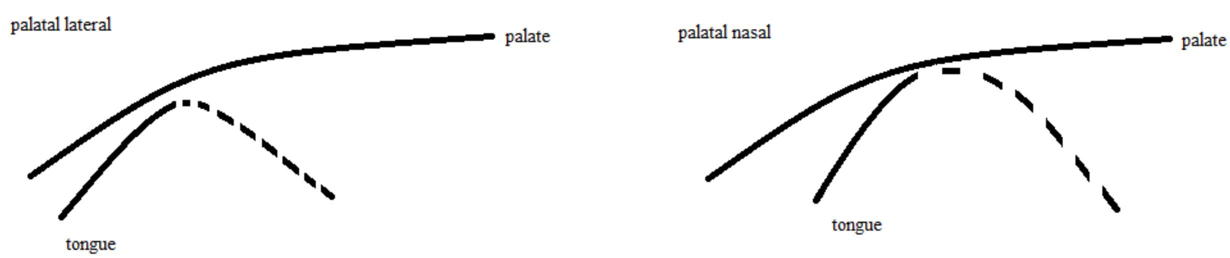

Figure 3.42: A simulation of the palatal lateral and palatal nasal, with dotted line indicating the portion of the tongue not captured by EMA sensors.

The linguopalatal configuration presented in Figure 3.42 accounts for the findings here and by previous EPG studies on Catalan palatal consonants (Recasens, 1984b; Recasens et al., 1993; Recasens and Pallarès, 2001; Recasens and Espinosa, 2006). Due to the nature of EMA sensor placement, only the anterior portion of the tongue blade was captured. MR images of the palatal lateral in EP (Teixeira et al., 2012) indicate that this portion of the tongue blade is pressed up against the upper teeth and alveolopalatal region of the roof of the mouth. As Figure 3.42 indicates, there is no additional palatal contact following the extent of the tongue contour captured by EMA. In the case of the palatal nasal, MR images of the EP palatal nasal (Martins et al., 2008) corroborates EPG findings by demonstrating extended linguopalatal contact and a more posterior place of contact. However, the linguopalatal contact includes the part of the tongue extending past the portion captured by the EMA. In conjunction with the steeper anterior tongue angle when articulating the palatal nasal, this results in our measure reporting a larger summed area for the palatal nasal than the palatal lateral.

Given the tendency in the literature to use the terms *palatal* and *palatalized* lateral interchangeably, the evidence presented here suggests that this may in fact, be an acceptable practice; neither sound demonstrates large articulatory differences and the difference in terminology is appropriate for distinguishing phonemic status as opposed to articulatory differences. That is to say, referring to the palatalized lateral in Ukrainian (Zilyns'kyj, 1979) would essentially be the same as referring to the sound as a palatal lateral, when considering only how the sound is articulated. The results reported in this chapter indicate that findings from the study of velars comparing phonemic palatal gestures and palatalization occurring as a result of articulation (Keating and Lahiri, 1993) do not appear to hold true for laterals. Unlike the reported findings on velars, more extreme fronting for the palatal lateral in comparison to the palatalized lateral is not found.

If it is indeed the case that the slight differences observed between the articulation of the palatal and palatalized lateral are truly minor, then differences are not expected for the acoustics and perception of these two sounds. If so, then distinguishing between the palatal and palatalized lateral in BP is redundant. While the presence of minimal pairs and different orthographical representations might serve as sufficient reason for the maintenance of the palatalized lateral as a distinct entity, this reason alone is not sufficient for preventing sound change. Sound changes resulting in the creation of multiple homophones is a relatively common process that occurs cross-linguistically and orthography does change to match pronunciation, though this process occurs more slowly.

Future studies should incorporate methodologies that can be used to reliably model the palate, as well as provide greater insight into the posterior region of the tongue (particularly the upper pharyngeal region). It is possible that larger distinctions between the palatal and palatalized lateral do exist, but in other dimensions not available for study using EMA. If so, then a study of the acoustics (see Chapter 4) and the perception (see Chapter 5) of the palatal lateral will be able to reveal insight into these differences. Subtle distinctions between the palatalized and palatal lateral are apparent and are reported in Section 4.3.4 and A.3.1.

3.4.2 Comparisons against similar sounds

The inclusion of [ɲ, n^j, j, l] were also included in this study to facilitate a discussion of how the palatal and palatalized lateral is articulated in comparison to other, similar sounds. It was expected that patterns observed between the palatal and palatalized lateral would also be observed in the palatal and palatalized nasal, which according to classic definitions, should differ only with regards to manner. However, the evidence presented in this study does not support this prediction; there are consistent significant differences between the palatal and palatalized nasal not found in the palatal and palatalized lateral. Additionally, previous predictions that *yeísmo* may be due to articulatory similarities between the palatal lateral and palatal approximant are not supported by the evidence presented in this study; results indicate that the BP palatal lateral bears a greater articulatory resemblance to the palatalized nasal than the palatal approximant.

While the palatal lateral does indeed possess a high front tongue body in comparison to the alveolar lateral approximant, the baseline, the lateral sounds (i.e. [ʎ, l^j, l]) demonstrate a more retracted tongue position in comparison to their nasal counterparts. This is apparent in Figures 3.16 - 3.18 comparing change in x-position over time by sensor and Figures 3.6 - 3.8 illustrating sensor position distribution by phone, and consistent with reports of an additional retracted tongue root articulation in the EP alveolar lateral (Cunha et al., 1985; Andrade, 1999; Oliveira et al., 2011). The more posterior position observed in the palatal lateral in comparison to the palatal nasal and palatal approximant directly contradicts findings from EPG studies on palatal consonants in Catalan (Recasens et al., 1993; Recasens and Pallarès, 2001; Recasens and Espinosa, 2006), which report that the palatal lateral in Catalan is articulated at a more anterior place than the palatal stop and palatal nasal.

It is unclear whether the addition of a retracted tongue root is responsible for the lack of distinction between the palatal and palatalized lateral here. If this line of questioning is to be pursued, future study of languages with a palatal lateral that is not articulated with a retracted tongue root is suggested. For example, a language like Breton (Elmar, 1970) might be appropriate, as the laterals in this languages have not been previously reported as velarized. Most notably, while results from fronted velars (Keating and Lahiri, 1993) suggest that the palatal nasal and palatal lateral ought to be more extremely fronted in comparison to their palatalized counterparts, only the palatal nasal is observed to be more fronted than its palatalized counterpart (i.e. the palatalized nasal).

Figures 3.13 - 3.15 indicate that the palatal and palatalized lateral and palatalized nasal have a higher overall anterior tongue position than the palatal nasal and palatal approximant. The position of all three tongue sensors is consistently more anterior for the palatal nasal (see Figures 3.16 - 3.17) than the palatal and palatalized lateral. This is true of the palatalized and alveolar nasal as well, corresponding to the observation made in Section 3.3.2. Figures 3.17 and 3.18 also illustrates similar changes in x-position across time for all palatal and palatalized sounds at the TM and TB sensors; palatal and palatalized sounds are articulated with a more anterior position at the phone culmination than phone onset, while the two alveolar sounds culminate with roughly similar anteriority as at the onset.

In contrast, Figure 3.16 demonstrates completely different movement in the tongue tip, lending evidence for articulatory independence between these two points of the tongue (roughly speaking, the tongue tip and the tongue blade). However, while the TB sensor appears to be the most indicative of palatalization (more consistent tongue raising is apparent at this sensor in Figures 3.6 - 3.8), there is also more variability. This may be indicative of the relative precision that this portion of the tongue might have in achieving specific articulatory targets.

A laminal articulation is found in all of the palatal and palatalized sounds; the palatal approximant and the palatal nasal are articulated with the steepest tongue blade angle, followed by the palatal and palatalized lateral and palatalized nasal. The change in angle is not large, apparent from calculations of tongue blade angles in Figures 3.10 - 3.12 and dynamic models of the tongue in Figures 3.20 - 3.26. The palatalized nasal is again more similar to the palatal and palatalized lateral, though not identical, with all three sounds articulated with a tongue blade angle that is comparatively less steep with respect to the palatal approximant and palatal nasal. While there is a smaller overall change in degree across the phone duration observed for the palatal nasal and palatal approximant, the palatal and palatalized lateral and palatalized nasal demonstrate a change in angle of more than 10 degrees, resulting in a steeper tongue blade angle. This is also observed in the SS-ANOVA results presented in Figures 3.13 - 3.15, with the largest change in tongue height observed from TM to TB for the palatal nasal and palatal approximant. An SS-ANOVA of the tongue blade angle (see Figure 3.19) confirms observations made from Figures 3.10 - 3.12; the palatal approximant and palatal nasal become less steep as the phone progresses, while the other five sounds are articulated with a steeper tongue blade angle throughout the phone duration.

With regards to the linguopalatal approximation, the same groupings are found again. The alveolar nasal and lateral demonstrate the largest summed area between the palate and tongue contour, followed by the palatal nasal and palatal approximant. The palatal and palatalized lateral and palatalized nasal demonstrate the smallest summed area, indicating that these sounds are articulated with the greatest proximity to the palate. Both Recasens and Espinosa (2006) and Martins et al. (2008) report a palatoalveolar place of articulation for the palatal lateral and a palatal place of articulation for the palatal nasal, i.e. the palatal lateral is articulated at a more anterior location. Their findings are supported by Figures 3.37 and 3.38, which illustrate what appears to be tongue tip contact with an anterior section of the palatal trace (it was impossible to identify the exact position of the alveolar ridge) for the palatal and palatalized lateral and palatalized nasal, while the narrowest constriction occurs further back (approximately around where the TB sensor was placed) for the palatal nasal and palatal approximant.

It is perhaps unsurprising that the palatal approximant and palatal nasal possess larger summed areas than the palatal lateral, given that the palatal constriction for the former two sounds reportedly occurs further back on the palate than the latter (Recasens et al., 1993). Since the palatal trace primarily captures the more anterior portion of the palate, the smaller summed area for the palatal and palatalized lateral (articulated at the palatoalveolar region rather than truly the palate) is expected. This is true also of the palatalized nasal by extension of its continued similarity with the palatal and palatalized lateral, as well as with regards to the amount of linguopalatal approximation.

Given a methodology such as ultrasound or MRI that is able to capture more of the palatal contour, the summed area between the palate and tongue will of course change depending on the chosen region of interest. EMA is not the ideal method for capturing palatal contact, yet this measure does provide useful information regarding the extent of linguopalatal approximation with regards to the anterior tongue blade, which can then inform future imaging studies that wish to better capture this relationship. While EPG is typically the method of choice for capturing linguopalatal contact (Recasens et al., 1993; Recasens and Espinosa, 2006; Shosted et al., 2012b), this method is also limited, as it is unable to capture tongue movement that approximates but does not directly contact the palate.

The reason for the similarity between the palatal and palatalized lateral may be related to the nature of a lateral sound, given that clear articulatory distinctions exist for the palatal and palatalized nasal. It is apparent that it is not simple articulatory similarity that causes the palatal lateral to merge with the palatal approximant, as the palatal nasal is far more similar to the palatal approximant than is the palatal lateral. The similarity between the palatal nasal and palatal approximant is supported by EPG evidence, which indicates that the palatal nasal in BP is commonly articulated as an approximant (Shosted et al., 2012b). Despite this finding, the palatal nasal is not typically associated with a sound merger like *yeísmo* and is found in nearly seven times the languages than the palatal lateral (Maddieson and Precoda, 1991). The distinction between the oral and nasal manner of articulation appears to play an important role in the differences observed here; lateral airflow may be more susceptible to being interchanged with central airflow than nasal airflow.

There is the potential for frequency to play an interacting factor in the clear articulatory distinction made by speakers between the palatal and palatalized nasal. If the palatal and palatalized nasal appear more frequently than their lateral counterparts in daily use, there may be more motivation to maintain a distinction between the two. However, a search of the São Carlos corpus (composed of Brazilian texts, with 32.5 million unique words and 42.9 million total words) reveals similar frequency rates for the palatal lateral (digraph *lh* occurs in 0.023% of unique words and 0.674% of total words) and palatal nasal (digraph *nh* occurs in 0.020% of unique words and 0.682% of total words) in BP (Linguateca, 1999). Frequency rates for the palatalized lateral and nasal were more difficult to obtain as these sounds only occur when in unstressed syllables. To obtain frequency rates for the palatalized lateral and nasal, the corpus search specifications were set such that there was an adjacent stress mark (i.e. acute, circumflex, and tilde) to the orthographic representations *li* and *ni*, with all vowels possible on either side of *li* and *ni*. This excluded words with more than three syllables where the palatalized lateral or nasal occurs in the antepenultimate syllable; it is unknown how many words this may have excluded. The frequency of words that satisfied these conditions was similar for both the palatalized lateral (digraph *li* occurs in 0.002% of unique words and 0.088% of total words) and the palatalized nasal (digraph *ni* occurs in 0.002% of unique words and 0.1006% of total words). Results from the corpus search indicate that the palatal and palatalized lateral appear in BP at similar rates as the palatal and palatalized nasal. This presents clear evidence that frequency is not the motivating factor for why BP speakers do not maintain a clear articulatory difference between the palatal and palatalized lateral, while doing so for the palatal and palatalized nasal.

3.4.3 Individual speaker variation

Analysis of individual speaker variation in Section 3.3.5 suggests that age is not the dominant contributing factor with regards to the similarity of the palatal and palatalized lateral in BP. Speaker 8 demonstrates the largest and most noticeable difference in tongue shape between the two sounds, yet she is only 31 years old. Speaker 5 articulates the palatal and palatalized lateral nearly identically; at 20, she is only 11 years younger than Speaker 8.

Only three speakers under the age of 30 participated, with just one of the three speakers (Speaker 5) articulating the palatal and palatalized lateral without clear differences in tongue height or anteriority. In a separate recording section, Speaker 5 produced the acoustic stimuli for the perception task included in the appendix; results indicate that listeners (both native speakers of BP and naive native speakers of English) are unable to distinguish between the palatal and palatalized lateral as produced by this speaker. Speaker 6 and Speaker 10 are also in their early twenties, yet articulatory models for Speaker 6 group her with Speaker 2 and Speaker 4, both of whom are in their late 40s. Speaker 10 is unique amongst the 10 speakers; articulatory models for Speaker 10 do not resemble any of the other speakers. Additionally, while Speaker 5 reaches maximum tongue position towards the end of the phone, Speaker

10 and Speaker 6 do so near the phone onset. The above observations indicate that there is a great deal of variation in articulatory strategies even among speakers of the same age range.

Similarly, there was a great deal of variation amongst the three speakers in the highest age range. Speaker 2, Speaker 3, and Speaker 4 were between the age of 45 and 55 years, yet the articulatory strategies were different across all three speakers. Speaker 2 demonstrates a much lower and anterior tongue tip during the production of the palatal lateral in comparison to the palatalized lateral; this effectively extends the length of her tongue blade during the palatal lateral. However, Speaker 3 and Speaker 4 do the opposite: both speakers produce the palatalized lateral with a more anterior tongue tip in comparison to their own productions of the palatal lateral. In terms of overall tongue shapes, Speaker 2 and Speaker 4 are similar, while Speaker 3 demonstrates noticeably different tongue shapes.

These observations indicate that at least for now in BP, the instability of the palatal lateral is not linearly transmitted down to each new generation of BP speakers; in other words, if articulatory instability during the production of the palatal lateral is what results in the historical loss of the palatal lateral, the effect of instability is likely a spontaneous event as opposed to a process that occurs gradually. Note that these observations may be confounded by regional differences, as Speaker 8 self-identified as speaking with a so-called "countryside" accent. Future research may wish to focus on speakers from a specific city or region in São Paulo state. Additionally, given the small sample size presented in this study, it is unknown whether a larger sample size would confirm that older speakers distinguish between the palatal and palatalized lateral at a higher rate than younger speakers.

3.5 Conclusion

This chapter establishes the articulatory characteristics of the palatal lateral in comparison to other, similar sounds. While clear differences in the shape of the anterior tongue blade were found between palatal and palatalized nasals, only a few subtle differences were found for the palatal and palatalized lateral. Given that the palatal lateral and palatal approximant are articulated with significantly different anterior tongue blade configurations, articulatory similarity as a motivating factor for *yeísmo* is not supported by the evidence presented here.

Consistent groupings were observed across measures; the palatal and palatalized lateral was typically similar to the palatalized nasal, the palatal nasal to the palatal approximant, and the alveolar nasal with the alveolar lateral. The palatal nasal and palatal approximant demonstrate the largest tongue blade angle, followed by the palatal and palatalized lateral and palatalized nasal, with the alveolar lateral and alveolar nasal articulated with the smallest (or most neutral) tongue blade angle. Dynamic models of the anterior tongue reveal that with regards to overall tongue blade shape over time, the palatal nasal and palatal approximant are both articulated with an immobile tongue tip and tongue body fronting, while the palatalized and palatal lateral and palatalized nasal exhibit tongue body fronting as the tongue tip retracts and lowers. Both the alveolar lateral and alveolar nasal demonstrate tongue tip lowering while the tongue body remains relatively immobile.

A few potential reasons for the rareness of the palatal lateral were discussed, with the presence of a lateral airstream channel as the most likely contributing factor. The findings reported here lay the necessary groundwork for comparing the articulation of the BP palatal lateral to not only palatal laterals in other languages, but palatalized laterals as well. It is the first step towards creating a comprehensive understanding of the palatal lateral in BP, as well as the behavior of the palatal lateral in other languages in which it is attested.

Chapter 4

STUDY 2: ACOUSTICS

4.1 Introduction

This study will provide a description of the acoustics of the palatal lateral in Brazilian Portuguese (BP). A comparison to similar sounds in BP will also be made and correlations between the acoustics and articulatory events will be identified.

4.2 Methodology

Acoustic data was simultaneously collected using a Countryman Isomax E6 head-mount microphone during the EMA study. For additional information, please refer to Section 3.2.1 regarding participants and Section 3.2.2 regarding materials.

4.3 Results

As described in Section 3.3, data annotation was conducted in Praat (Boersma and Weenink, 2014) by visually identifying the beginning and end of each phone. An abrupt reduction in amplitude was taken to indicate a constriction in the airflow, marking the boundary of the consonant and distinguishing it from the surrounding back vowels. Each phone was embedded in a nonsense word that contained the target consonant in a syllable-initial and intervocalic position; i.e. [ˈpa.Ca], where ‘C’ represents the target consonant (please see Section 3.2.2 for more information on the constriction of materials). Each word was presented in the carrier phrase *Diga* ___ *para nós* [ʤi.gɐ ___ pa.rɐ nɔʃ] ‘Say ___ four times’. To facilitate quick identification of the VCV sequence enclosing the target consonant, each nonsense word began with a bilabial stop and was superseded by a word beginning with a bilabial stop.

With regards to Section 4.3.2 - 4.3.5, nasal sounds were excluded from the analysis of formant frequency as nasality introduces an unwanted ambiguity, such that it is unclear whether changes in the formant frequencies are a result of the oral configuration or nasalization (see Engwall et al. (2006), Carignan et al. (2011), and Shosted et al. (2012a) for more on the effect of nasalization on the acoustic signal and articulation). Data from all 10 BP speakers was included for analysis in Section 4.3.2 - 4.3.5, with a total of 13,320 items and an average of 3,330 repetitions per phone.

Measurements of the formant frequencies for [ʎ, lʲ, l, j] were obtained automatically via a modified Praat (Boersma and Weenink, 2014) script (Kawahara, 2010)[1] at the beginning, middle, and end of each phone using the default Praat settings (i.e. the number of predicted formants was five, with a maximum formant (Hz) of 5,500 and window length of 25 ms). While the rate of error with regards to correct formant identification was not calculated, box plots of formant distribution (see Figures 4.2- 4.4) indicate that the data that falls within the first and third quartile is tightly clustered around the median, indicating that the rate of error is likely sufficiently small.

While normalization presents major advantages, particularly by reducing or removing the effect of vocal tract size, normalized formant frequencies cannot be compared against non-normalized frequencies reported by previous studies. In order to better compare the current findings with previous findings in the literature, the data was not normalized in the analyses of formant frequency distribution in Section 4.3.2 and Section 4.3.3 of this chapter. However, to reduce the effect of speaker-specific physiological characteristics, analyses of the relationship between the articulations and corresponding acoustic output incorporated normalized acoustic and articulatory data in Section 4.3.4 and Section 4.3.5.

This section will begin by cataloging general acoustic characteristics of the palatal lateral (as well as similar sounds [l, j, and lʲ]) and culminate with a model of the relationship between the articulatory data from Chapter 3 and acoustic data presented in this chapter. Note that throughout the figures in this chapter, orthographic representations *lh*, *lj*, *nh*, and *nj* will be

[1] Included in appendix.

used to refer respectively to the sounds [ʎ, l^j, ɲ, n^j] in figures for plotting convenience.

4.3.1 Average phone duration

Since phone duration is not affected by nasality, all seven phones elicited during the EMA study from 10 native speakers of BP in Chapter 3 were included for analysis in this section. A total of 23,301 items were included, with an average of 3,329 repetitions per phone.

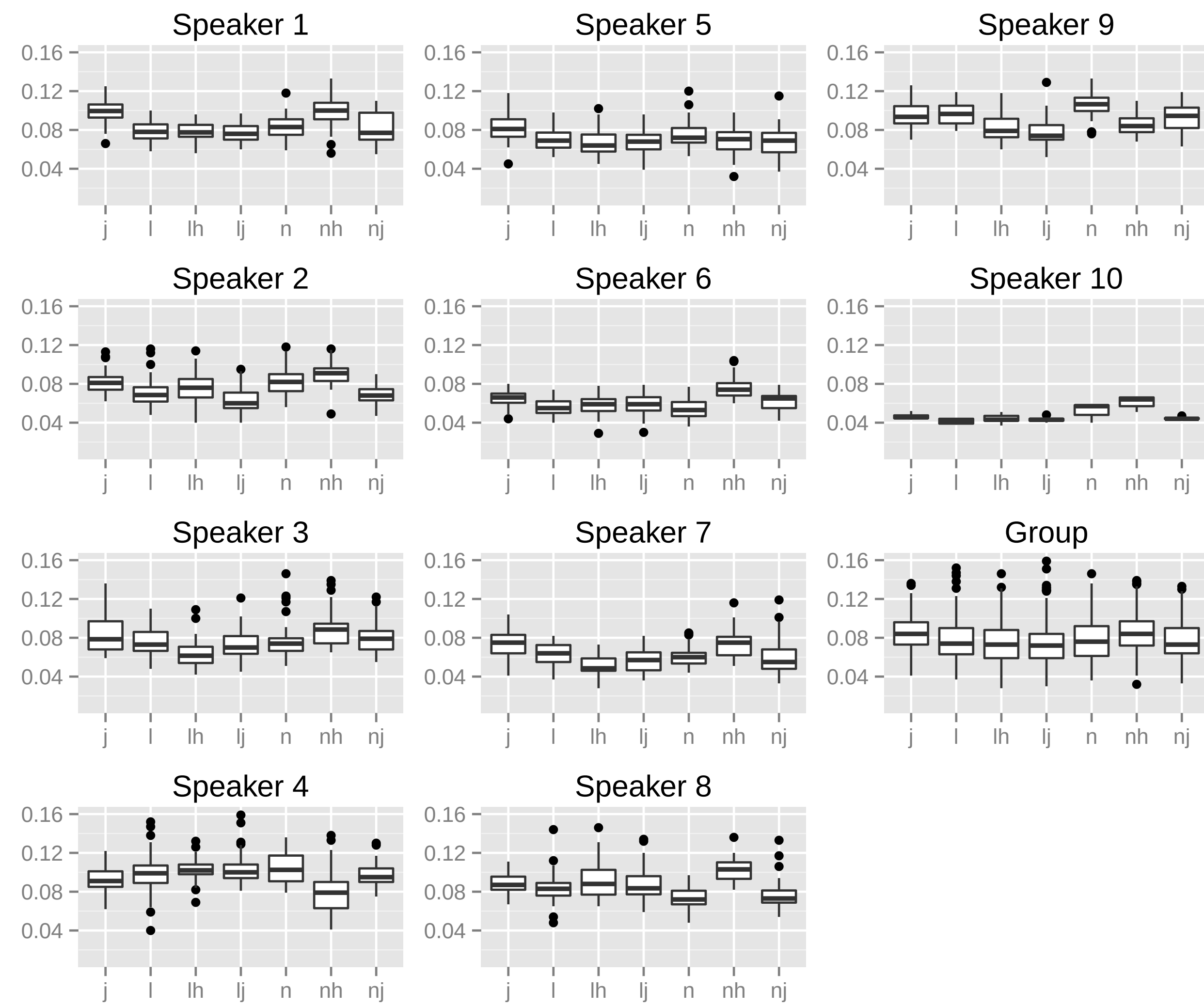

Figure 4.1: Distribution of phone durations by speaker and group. Time given in seconds.

Box plots of the phone durations for each speaker (see Figure 4.1) indicates a general trend for lateral sounds to be produced with shorter durations, similar to the palatalized nasal. In general, the palatal approximant and the palatal nasal appear to have the longest

Palatal	Plain	Palatalized
/ʎ/ 74.4	/l/ 76.8	[lʲ] 73.8
/ɲ/ 84.4	/n/ 78.9	[nʲ] 75.7
/j/ 84.4		

Table 4.1: Average duration by phone, in milliseconds.

durations. This holds true when looking at the average phone duration for the pooled data (see Table 4.1); the palatal approximant and the palatal nasal both have the longest average phone lasting 84.4 milliseconds, with the shortest average phone lasting 73.8 milliseconds recorded for the palatalized lateral.

While the palatal approximant and palatal nasal appear similar with regards to the length of phone duration, the palatal lateral does not follow this trend. T-tests reveal that there is no evidence of significant difference between the durations for the palatal approximant and palatal nasal ($p=0.99$, Bonferroni-adjusted significance level=0.017). Both the palatalized nasal and lateral have a shorter average duration than their plain and palatal counterparts, however, T-tests reveal that while the duration for the palatal nasal and palatalized nasal are significantly different ($p<0.001$), no evidence of significant difference is found in terms of phone length between the palatal lateral and palatalized lateral ($p=0.64$).

4.3.2 Formant frequency distribution

The following box plots highlight the similarities between the palatal and palatalized lateral. Notches on the box plots indicate confidence intervals around the median; overlapping notches suggest that there is strong evidence that the medians are not different.

The similarity between the first formant frequency of the palatal and palatalized lateral is evident (see Figure 4.2), with only one to two Hz difference in median formant frequency throughout the phone. Overlapping notches at all three sampled points during the phone

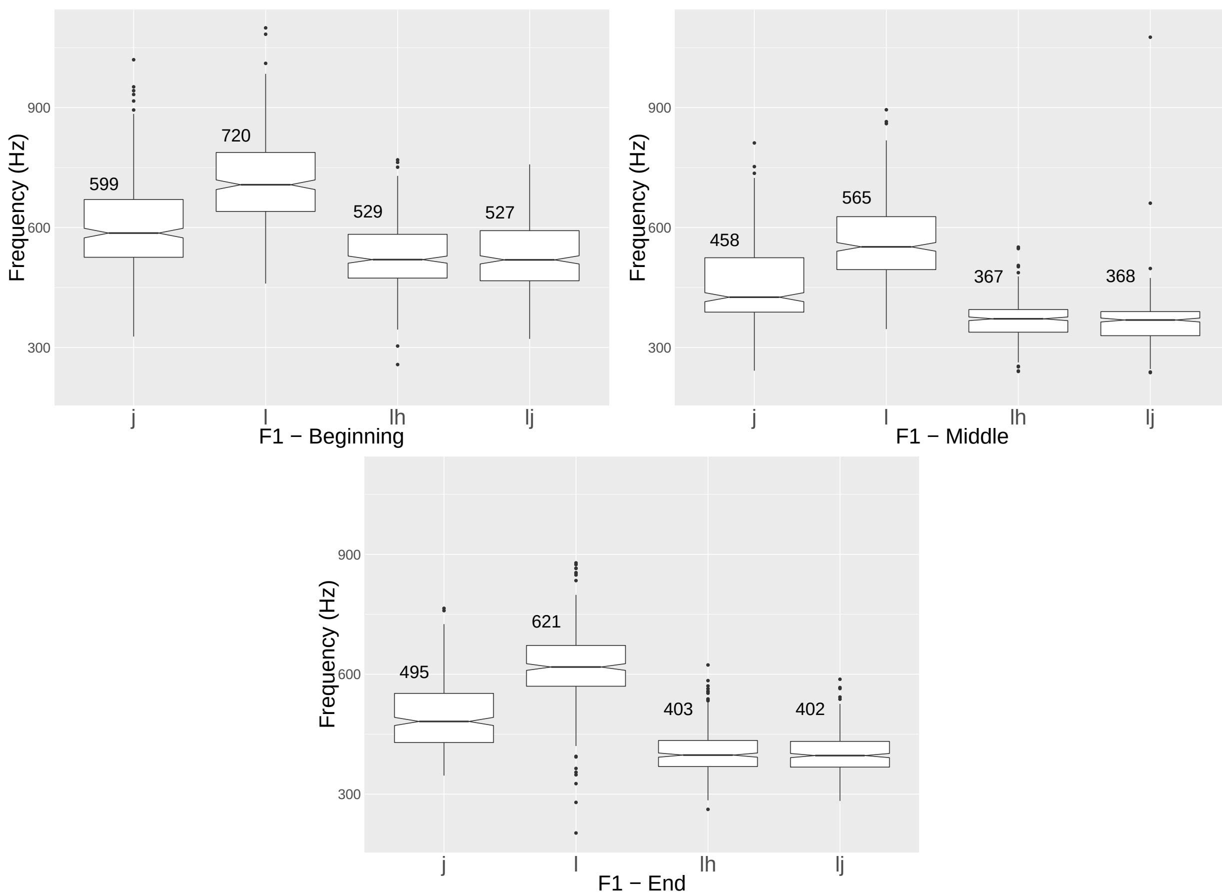

Figure 4.2: Distribution of F1 by phone, measured at the beginning, middle, and end of the phone.

indicate that there is no evidence of significant difference between the first formant frequency of the palatal and palatalized lateral. These two sounds consistently have the lowest F1, followed closely by the palatal approximant. Though similar, the palatal approximant is significantly different from the palatal and palatalized lateral. As the sound articulated the furthest back in the oral cavity, the alveolar lateral has the highest first formant frequency. For all four phones, the maximum median F1 is observed at the beginning of the phone and the minimum median F1 at the end. Given that F1 is supposedly inversely correlated with tongue height (Ladefoged and Johnson, 2014), this finding indicates that as F1 decreases during the production of the palatal and palatalized lateral, palatal approximant, and alveolar lateral approximant, tongue height should rise. Referring back to the dynamic models in Section 3.3.5, this finding corresponds to an increase in TB sensor height for the alveolar lateral approximant and the palatal and palatalized lateral, but not the palatal approximant, for which TB sensor height reaches a maximum at the midpoint of the phone.

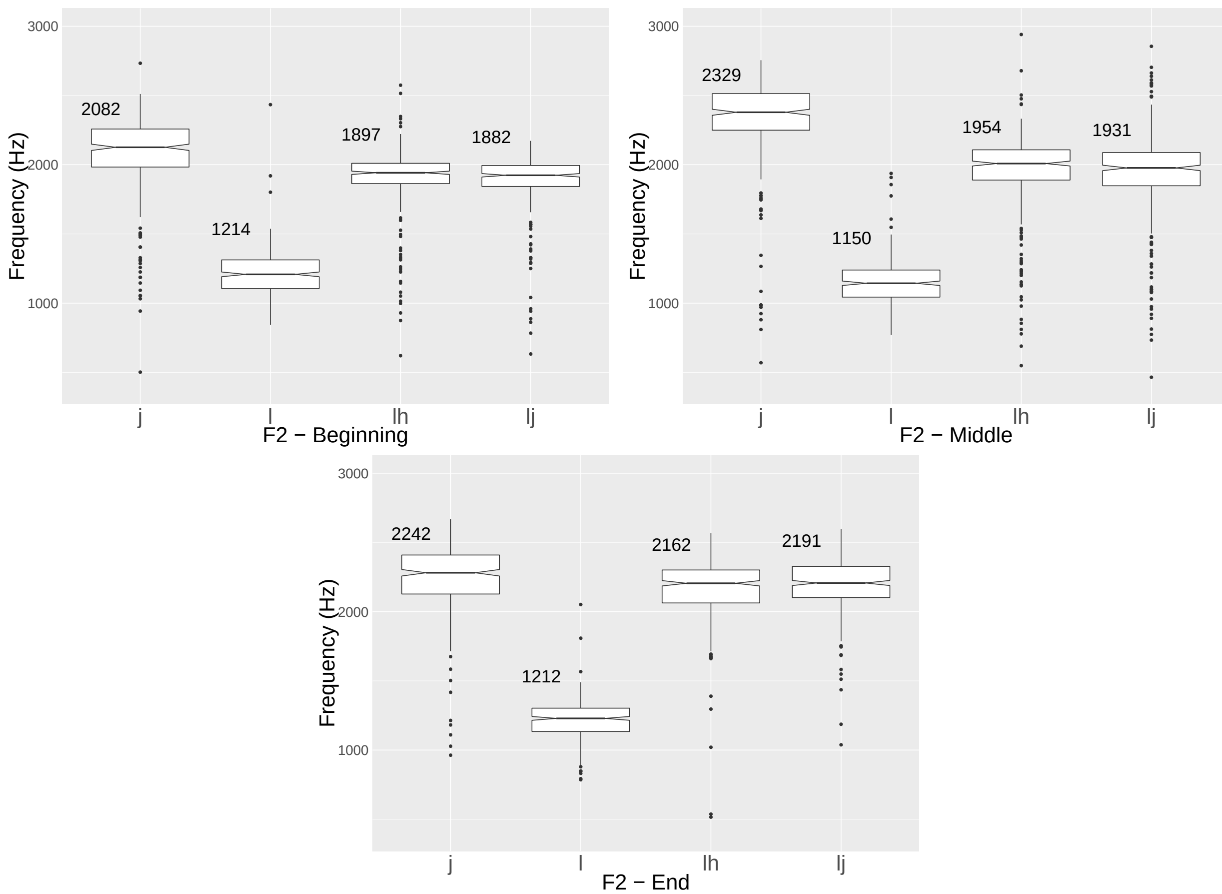

Figure 4.3: Distribution of F2 by phone, measured at the beginning, middle, and end of the phone.

Overlapping notches indicate that throughout the phone, there is no evidence that the median second formant frequency is significantly different between the palatal and palatalized lateral (see Figure 4.2). Again, though similar to the palatal and palatalized lateral, the palatal approximant second formant frequency is significantly higher. All three palatal/-palatalized sounds have a high second formant frequency, as expected for a sound articulated with a high anterior tongue position (Fant, 1960). In contrast, the alveolar lateral has a much lower median F2.

The second formant frequency increases throughout the palatal and palatalized lateral, while the palatal approximant achieves a maximum F2 and the alveolar lateral achieves a minimum F2 at the midpoint of the phone. As the phone progresses, the distance between the first and second formant frequency increases in the palatal and palatalized lateral and palatal approximant as the sound progresses, while the distance decreases for the alveolar lateral.

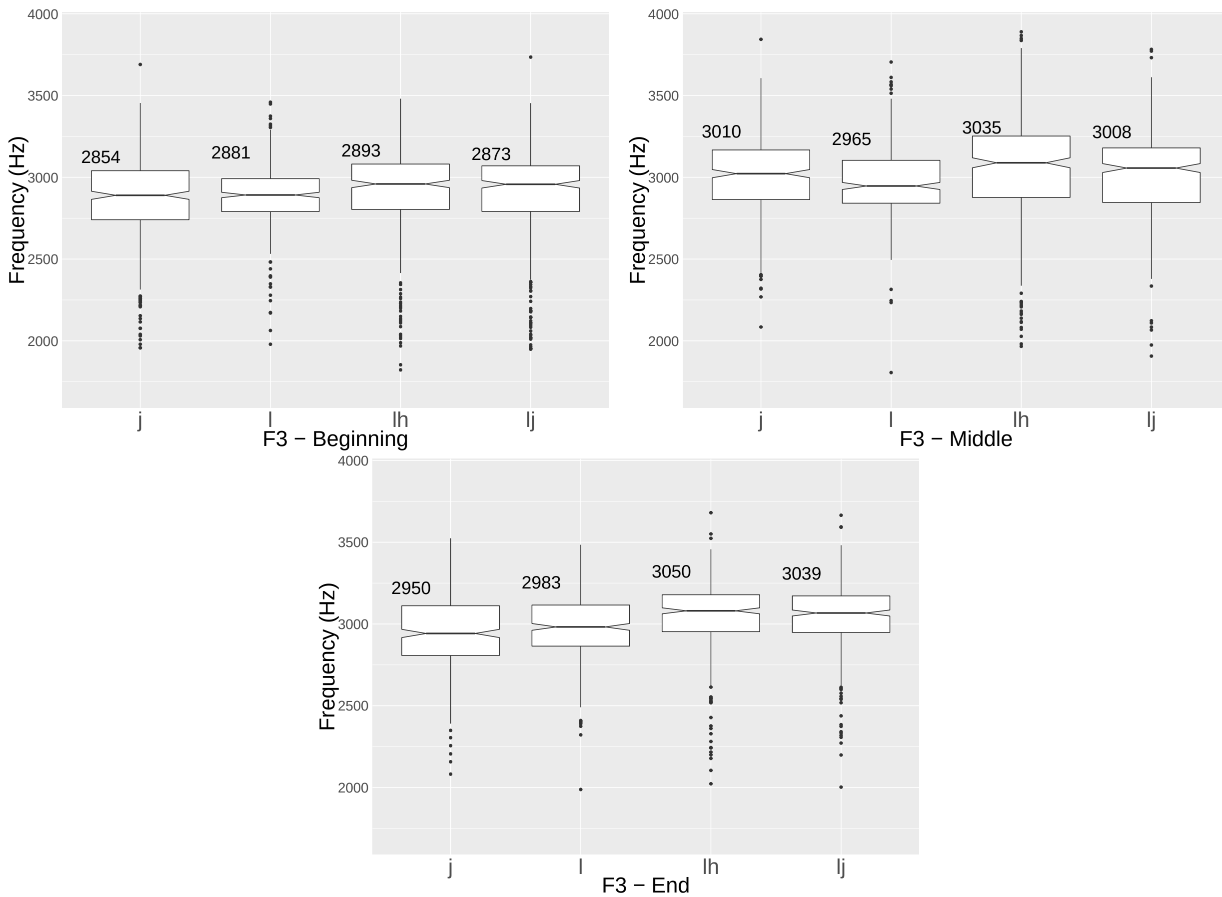

Figure 4.4: Distribution of F3 by phone, measured at the beginning, middle, and end of the phone.

Unlike the first and second formant frequencies, large differences are not observed in median third formant frequencies for all four phones (see Figure 4.4). Once again, there is no evidence that the palatal and palatalized lateral are significantly different from one another, though significantly higher than the palatal approximant and alveolar lateral at the beginning, middle, and end of the phone. The median third formant frequency is significantly different between the alveolar lateral and palatal approximant at the middle of the phone, but not at the beginning or end.

This section establishes the acoustic similarity between the palatal and palatalized lateral, which both demonstrate a low F1 and high F2 (with an approximate difference of 1,600 Hz at the phone midpoint) that is characteristic of sounds articulated with the high front tongue position (Fant, 1960). While the palatal approximant closely resembles the palatal and palatalized lateral, it is significantly different from these two sounds. As expected, the alveolar lateral approximant has a relatively high F1 and low F2 (an approximate difference of 600 Hz at the phone midpoint), which is characteristic of sounds articulated with a low retracted tongue body (Fant, 1960).

4.3.3 Smoothing Spline ANalysis of VAriance (SS-ANOVA)

A Smoothing Spline ANalysis of VAriance (SS-ANOVA) was used to compare the distribution of the first through third formant frequency for [ʎ, lʲ, l, j]. The same data from Section 4.3.2 was used for analysis here. Formants were extracted in the same manner as in Section 4.3.2, using the same Praat script (Kawahara, 2010) and settings.

SS-ANOVA requires splines to be of the same length, so formant trajectories were resampled to 21 samples per phone using the MATLAB (2014) function *resample*(). The following figure (Figure 4.5) was created using R package *ggplot2* (Wickham, 2009) and an R script by (Mielke, 2013)[2]. Overlapping confidence intervals indicate that there is no evidence that the two curves are significantly different for that portion of the curve. It is immediately apparent from Figure 4.5 that there is no evidence that the distribution of formant frequencies for the palatal and palatalized lateral is significantly different for the first three formants.

The first formant frequency for the palatal approximant begins around 500 Hz and is approximately equidistant from the first formant frequency for the alveolar lateral (approximately 100 Hz higher) and the palatal and palatalized laterals (approximately 100 Hz lower). Note that there is nearly complete overlap in confidence intervals for the first formant frequency of the palatal and palatalized lateral, which renders the two sounds nearly indistin-

[2] Included in appendix and available at *http://phon.chass.ncsu.edu/manual/tongue_ssanova.r*, accessed 08/30/2015.

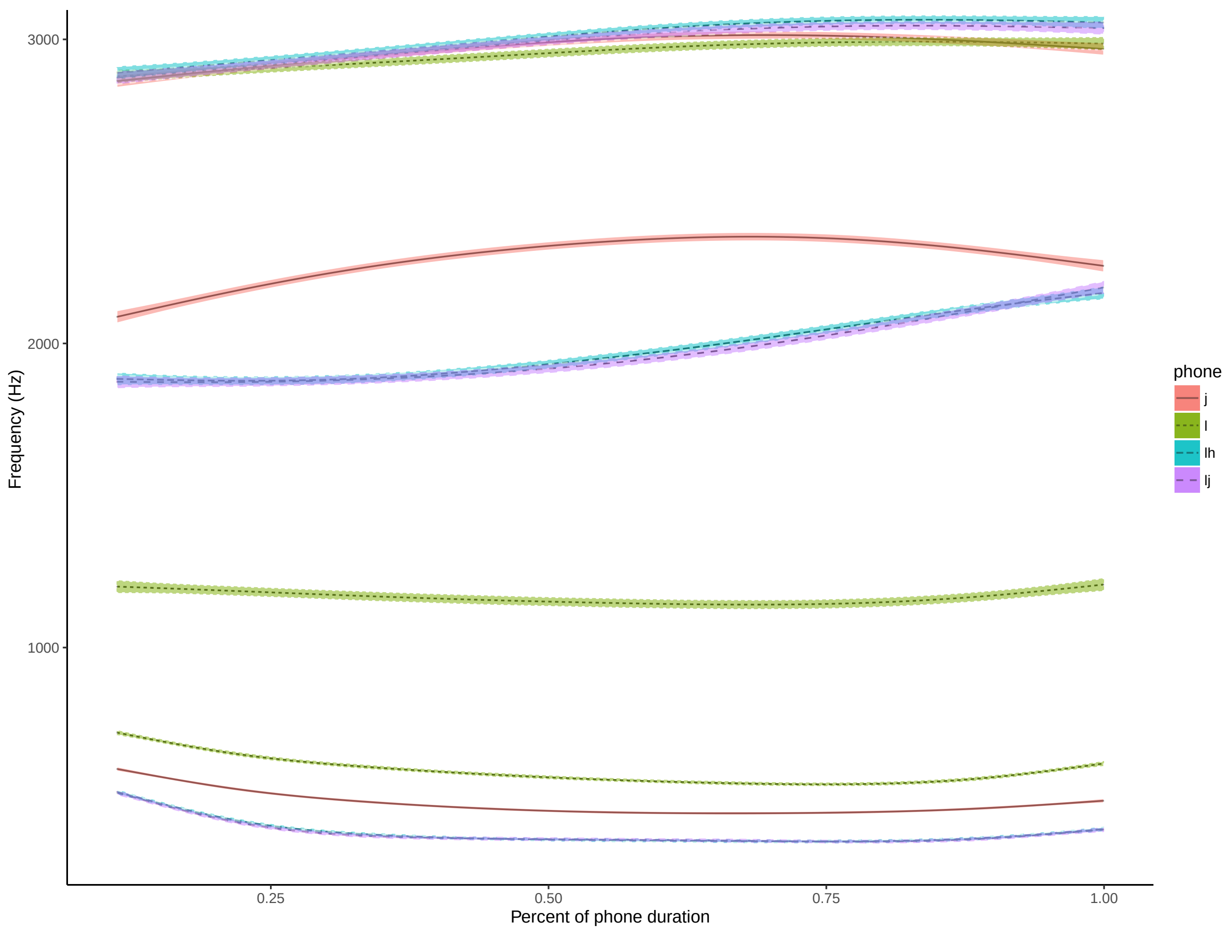

Figure 4.5: SS-ANOVA of the first three formant frequencies for the palatal lateral, palatalized lateral, alveolar lateral, and palatal approximant.

guishable in Figure 4.5. The first formant frequency for all four sounds begin at a relatively higher frequency, with a moderate decrease until about 25% into the duration of the sound; the same frequency is then maintained for the remainder of the phone duration. The decrease in Hz is slightly larger for the palatal and palatalized lateral, as it begins a bit closer to the palatal approximant before decreasing. As in the previous figures (see Figures 4.2 - 4.4), the overall change in the first formant frequency is minimal, no more than 150 Hz. Other than the palatal and palatalized laterals, there is no overlap in the confidence intervals.

While all four sounds were fairly similar in terms of the shape of the formant trajectory for the first formant, the second formant frequency reveals more obvious differences between the sounds. The overlapping second formant frequencies for the palatal and palatalized lateral are located around the 2000 Hz range; it is relatively stable until a third of the phone duration, after which it increases by several hundred Hz and culminates near the palatal approximant. The second formant frequency for the palatal approximant begins higher than the palatal and palatalized lateral, with a final frequency that is higher than the initial frequency. The lowest second formant frequency is observed for the alveolar lateral, at approximately 1250 Hz, with no large changes in frequency from start to end. In comparison, the formant frequencies for the palatal and palatalized lateral are modeled as a moderately convex curve while the palatal approximant is modeled as a concave curve, with a peak at two-thirds of the phone. Given that F2 is supposed to be positively correlated with tongue anteriority (Ladefoged and Johnson, 2014), similar curvatures should be observed in SS-ANOVAs of tongue anteriority (see Figure 3.16 - 3.18). Comparisons of the two sets of figures indicate that this is the case only for the palatal approximant.

The distance between the third formant frequencies of the four sounds is the smallest; confidence intervals calculated for all four sounds overlapped for the first quarter of the phone, with confidence intervals for the palatal and palatalized lateral overlapping with the palatal approximant until halfway through the phone. The highest overall frequency was observed for the palatal and palatalized lateral. From about a quarter until three-quarters of the phone duration, the alveolar lateral is lower than the palatal approximant, before the confidence intervals overlap at the end of the sound.

4.3.4 Correlation between tongue position and acoustics

One useful application of being able to simultaneously collect both articulatory and acoustic data is the opportunity to study the relationship between the two. Given the many-to-one problem (Maeda, 1990) (i.e. the dilemma of identifying which of many potential articulatory configurations has contributed to a single acoustic output), this is not a simple task. However, it would be very beneficial to field researchers interested in articulatory phonetics if a better understanding of possible responsible articulations can be established. Through such an understanding, field researchers would be able to discuss, with greater confidence, the articulatory implications of acoustic data collected from linguistic communities that are either hard to access or difficult to recruit for a laboratory setting. This section seeks to establish such an understanding of the relationship between tongue position and subsequent acoustics.

Classic descriptions (Ladefoged and Johnson, 2014) of the first two formant frequencies describe F1 as being negatively correlated with height and F2 positively correlated with frontness. The focus on the first two formant frequencies is due to their importance in listener identification of vowels, while the third formant frequency provides crucial perceptual cues for distinguishing between liquids /l, r/ and semi-vowels /w, j/ (O'Connor et al., 1957a; Espy-Wilson, 1992). F3 is generally associated with quality distinctions (Ladefoged and Johnson, 2014); a high F3 assists in the perceived naturalness of a synthesized alveolar lateral approximant (Fant, 1960) and while an incorrect F3 does not detract from the correct identification of voiced stop consonants, a correct F3 can augment the listener's perception of said sounds (Harris et al., 1958).

	TTx	TMx	TBx	TTy	TMy	TBy
F1	0.05	0.11	0.13	-0.41	**-0.55**	-0.51
F2	-0.18	-0.43	**-0.52**	0.02	0.39	**0.63**
F3	0.07	-0.04	0.07	0.04	0.15	0.19

Table 4.2: Correlations between the x- and y-position of the three tongue sensors with the first three formant frequencies. Strongest correlations indicated in bold print.

Correlation coefficients (see Table 4.2) were calculated in R (R Core Team, 2015) with the function *cor*() (the default method, Pearson, was selected) for [ʎ, lʲ, j, l] using the articulatory data from Section 3.3.2[3]. The corresponding acoustic information (i.e. the first three formants) were extracted using a Praat script; the same acoustic data used in the previous sections of this chapter is used to perform the analysis in this section (see Section 4.3 for more information on how the formants were extracted). Since the articulatory data in Section 3.3.2 only included sample points at the beginning, middle, and end of each phone, only the corresponding acoustic information was included in this section for analysis.

Both acoustic and articulatory data were normalized by speaker in R (R Core Team, 2015) using the function *scale*(). Figures illustrating the correlations were created using R (R Core Team, 2015) package *ggplot2* (Wickham, 2009). The strongest correlation between F1 and y-position was observed for the TM sensor, while the TB sensor had the strongest correlation with F2. None of the sensors demonstrated a strong correlation with F3.

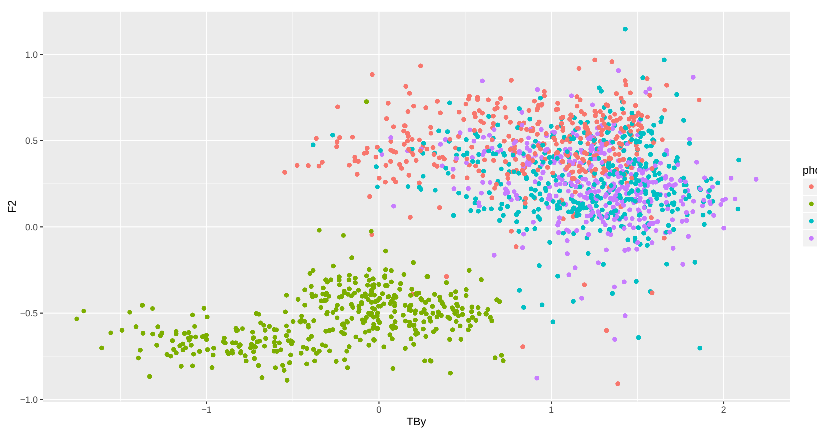

Figure 4.6: Distribution of TB height plotted against F2 by phone. Axes units are scaled.

True to traditional descriptions of the relationship between tongue position and formant frequency, there is a high correspondence between the x- and y-position of the TB sensor with F2 and F1 respectively. If the first and second formant frequency is known, one might be able to predict with a modest amount of accuracy where the midpoint (i.e. approximately three

[3]Non-interpolated articulatory data taken from the beginning, middle, and end of every phone.

centimeters from the tongue tip) of the tongue body is in the oral tract . It is unsurprising that this point of the tongue is perhaps the most highly correlated, as these correlations were calculated for predominantly palatal sounds.

Figure 4.7: Distribution of tongue height plotted against F1, by sensor (TT, TM, TB, from top to bottom) and phone. Axes units are scaled.

The correlation between height and frontness with F1 and F2 respectively is not consistent across sensors. The correlation between tongue tip frontness and F2 is weak, unlike the much stronger correlations observed for the TM and TB sensor. This indicates that tongue tip frontness is largely unrelated to F2. Surprisingly, there is a strong positive correlation between the y-position of the TB sensor with F2. Referring to Figure 4.6, it is evident why this might be the case; the height of TB is similar for all the palatal sounds but much lower for the alveolar lateral with a lower F2, resulting in a positive correlation between TB height and F2. This figure also illustrates the trend that has been discussed at length in Chapter 3: the palatal and palatalized laterals are overlapping, while the alveolar lateral and palatal approximant remain distinct (though the palatal approximant is occasionally similar to the palatal and palatalized laterals).

Figure 4.7 demonstrates the negative correlations reported in Table 4.2; as height increases, F1 decreases. Here again, the same pattern is noticed, with the palatal and palatalized laterals overlapping and a notable separation between the palatal and palatalized laterals and the other sounds. However, it becomes increasingly evident that within each phone, there does not seem to be as strong of a correlation between the first formant frequency and tongue height. While there is a general trend overall of F1 decreasing as height increases, this trend is not entirely apparent within each phone; the increase in height has little to no effect on the first formant frequency within a phone, particularly for the palatal and palatalized laterals.

Figure 4.8 illustrates the relationship between the second formant frequency and tongue frontness reported in Table 4.2. The reason for the weak correlation between tongue tip frontness and F2 is evident (see Figure 4.8, upper left), as the distribution of tongue frontness for all four sounds is within the same range of x-values. In contrast, TM and TB demonstrate a greater separation between the palatal/palatalized sounds and the alveolar lateral approximant, resulting in the moderately strong correlations reported above. The same trend observed in the previous figures continues, with the palatal and palatalized laterals completely overlapping while the other two phones are separate. In particular, the alveolar lateral is distinctly different from the palatal sounds. Interestingly, the distribution of palatal and palatalized laterals is much more spread out here (i.e. there is much more variation observed in these sounds), while the palatal approximant and alveolar lateral are clustered closer to the mean.

Given the low correlation between tongue frontness and the third formant frequency, Figure 4.9 is not surprising. Here, the correlation scores seem to be related to variability, with TB having the highest correlation score and least amount of variability. The overlapping distribution of palatal and palatalized laterals and separation from the palatal approximant

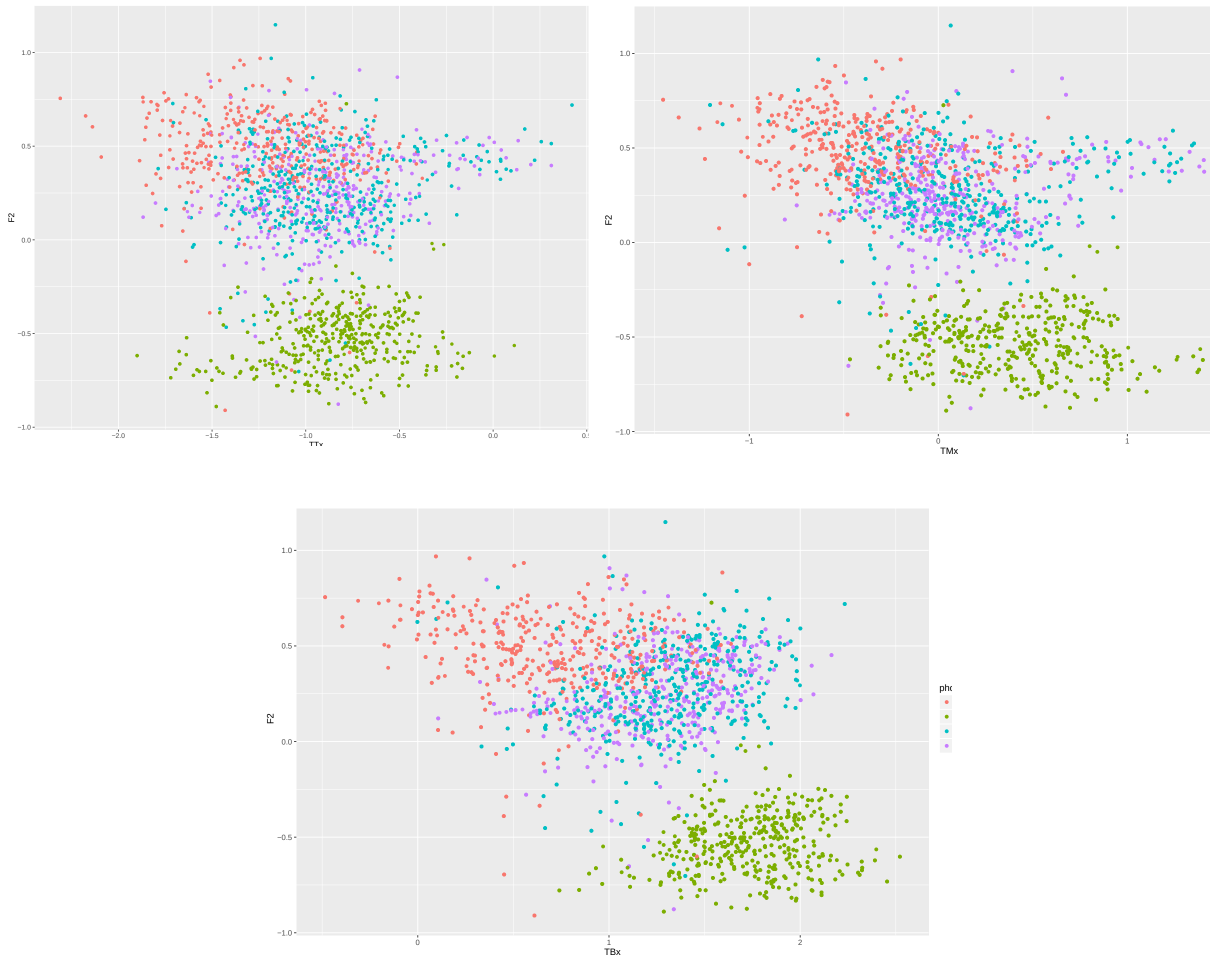

Figure 4.8: Distribution of tongue frontness (left to right in graph = front to back in oral cavity) plotted against F2, by sensor (TT, TM, TB, from top left, top right, bottom) and phone. Axes units are scaled.

and alveolar lateral is most evident in the TB sensor. In the TT sensor, it is difficult to distinguish between the different phones, while only the palatal approximant is distinguishable for the TM sensor.

In Figures 4.6 - 4.9, something becomes increasingly apparent. The figures presented here indicate that within phones, there is not as strong of a relationship between tongue position and formant frequency. As a result, the correlations were recalculated by phone. The results are indicated in Table 4.3 - 4.5. It is clear from all three figures that the strength of the correlations drops when calculated over individual phones.

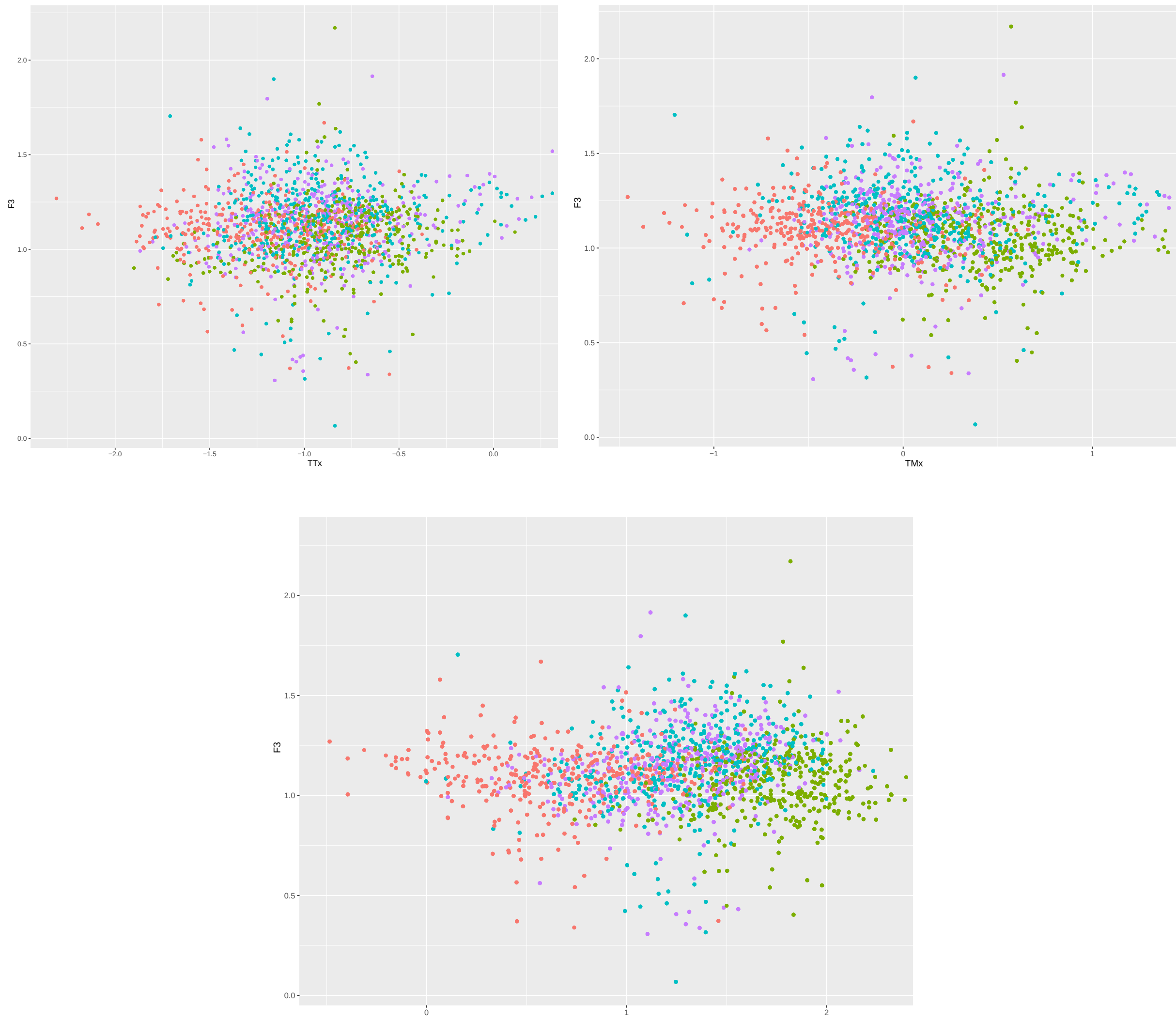

Figure 4.9: Distribution of tongue frontness (left to right in graph = front to back in oral cavity) plotted against F3, by sensor (TT, TM, TB, from top left, top right, bottom) and phone. Axes units are scaled.

There are still some moderate correlations between the formant frequencies and sensor positions however, particularly for the palatal approximant (see Table 4.3). Table 4.3 indicates that the similarity between the palatal and palatalized lateral observed in Figures 4.6 - 4.9 is also represented in the relationship between F1 and tongue height. These two sounds however have the lowest correlation coefficients reported in comparison to the palatal approximant and alveolar lateral approximant.

Except for the palatal approximant, the other three sounds demonstrate relatively stronger

	/ʎ/	$[l^j]$	/l/	/j/
TT	-0.12	-0.15	-0.28	-0.33
TM	-0.10	-0.14	-0.24	-0.36
TB	0.02	-0.01	-0.07	-0.32

Table 4.3: Correlation coefficients for F1 by sensor height.

correlations at the TT and TM sensor, with regards to the relationship between sensor height and F1. This indicates that the tongue back (defined here as three centimeters from the tongue tip) does not play a role in the production of F1 for these sounds. In contrast, the correlation coefficients between sensor height and F1 are relatively consistent across the three sensors for the palatal approximant, indicating that this entire area of the tongue is important for F1 control during the production of the palatal approximant. This result is interesting as it suggests that with regards to F1 and tongue height, the palatal and palatalized lateral are more similar to the alveolar lateral approximant.

	/ʎ/	$[l^j]$	/l/	/j/
TT	0.12	0.16	0.17	-0.16
TM	-0.04	0.14	-0.07	-0.22
TB	0.22	0.22	0.13	-0.15

Table 4.4: Correlation coefficients for F2 by sensor frontness.

None of the sounds demonstrate strong correlations with regards to the relationship between tongue frontness and the second formant frequency (see Table 4.4). In particular, the second formant frequency associated with the palatalized lateral has a positive correlation with the x-position for all three sensors, which suggests that as the tongue moves backwards in the mouth, F2 rises. This is clearly the opposite of the predicted relationship, since F2 is expected to decrease as the tongue retracts. However, reassessing Figure 4.8 reveals that the expected relationship between F2 and tongue anteriority holds true when considering the distribution of all four phones as a whole (though more so for TM and TB than TT)

There is a notable difference in correlation scores between the palatal and palatalized lateral; the correlation scores are similar for the two sounds with regards to the TT and TB sensor, but not the TM sensor - the palatalized lateral has a positive correlation coefficient of 0.14, while the palatal lateral has a negligible negative correlation coefficient of -0.04. The difference in correlation coefficients indicates that the relationship between the articulation and acoustics is not the same for both sounds at the TM sensor (about two centimeters from the tongue tip). In particular, speaker correlation coefficients for the palatalized lateral indicate that there is a weaker linear relationship between sensor position and formant frequency for the palatal lateral in comparison to the palatalized lateral. This suggests that there may be more variability at this part of the tongue for the palatal lateral, in comparison to the palatalized lateral, which is somewhat surprising, as palatalized sounds are generally expected to demonstrate more variability by comparison. In this aspect however, the palatal lateral and alveolar lateral approximant are similar, as both sounds demonstrate a weaker negative correlation between TM frontness and F2.

	/ʎ/	[lʲ]	/l/	/j/
TT	0.03	0.11	0.17	-0.04
TM	-0.04	0.11	-0.10	-0.07
TB	0.24	0.26	0.05	-0.01

Table 4.5: Correlation coefficients for F3 by sensor frontness.

Given the already low correlations reported between the third formant frequency and tongue frontness in Table 4.2, it is unsurprising that the correlations are so low when calculated by phone (see Table 4.5). It is interesting to note however that the palatalized lateral has the strongest correlations out of the four phones, while the alveolar lateral again demonstrates a very slight positive correlation. Interestingly, the correlation scores for the palatal and palatalized lateral do not resemble one another.

While the correlations are high when the phones are combined, overall there is not a very strong relationship between the x- and y-position of the sensors and the acoustics when divided by phone. Comparatively however, correlation coefficients reveal that the relationship between sensor height and F1 is relatively stronger than sensor frontness and F2. The palatal approximant demonstrates the strongest correlations out of the four sounds investigated in this section and maintains a moderate correlation between the F1 and tongue height; the correlation between sensor position and acoustics is comparatively stronger for the TM sensor (placed two centimeters from the tip of the tongue). This finding is the reverse of the three lateral sounds, which is also indicative of the role in which this section of the tongue plays in the production of a palatal approximant.

4.3.5 Generalized Additive Models (GAMs)

This section seeks to understand how well the first two formant frequencies can be used to predict sensor position. Unlike linear regression models or SS-ANOVAs, smooth estimates in Generalized Additive Models (GAMs) are nonparametric and are better suited for attempts to fit unbalanced data; given the absence of a linear relationship between formant frequencies and sensor position (see Section 4.3.4), a GAM was fitted to the data in order to better understand the predictive power of the acoustics for the palatal and palatalized lateral, alveolar lateral, and palatal approximant. The third formant frequency was excluded from analysis here, as Section 4.3.4 indicates that the third formant frequency is not strongly correlated with sensor position.

Using R (R Core Team, 2015), the *gam*() function from the *mgcv* package (Wood, 2000) was used to build the models and the *visreg*() function from the *visreg* package (Breheny and Burchett, 2016) was used to plot the model fits. Four models were built, two to predict y- and x-position based on the first and second formant frequency respectively and another two to predict the first and second formant frequency based on the y- and x-position. The focus here will be on how much of the variance in the data the model is able to explain, which will be used to evaluate the predictive power of tongue sensor position as opposed to the first and second formant frequency.

Figure 4.10 illustrates fits of the two models with formant frequency as the predictor. In the first model (top row), sensor y-position is the dependent variable, while F1 is a smooth term with sensor identity as a by variable and phone identity as a separate independent factor. A summary of this model reveals that 28% of the deviance in the data is explained by the model; when modeling linguistic data, especially given the few predictors included, this is considered a moderately satisfactory figure (as a reference point, typically a number equivalent to or larger than 60% is considered quite high). The estimated degrees of freedom (EDF) is greater than 1 for each parameter; an EDF of 9.0, 8.5, and 9.0 for parameters TT, TM, and TB respectively indicates a high degree of non-linearity in the relationship between

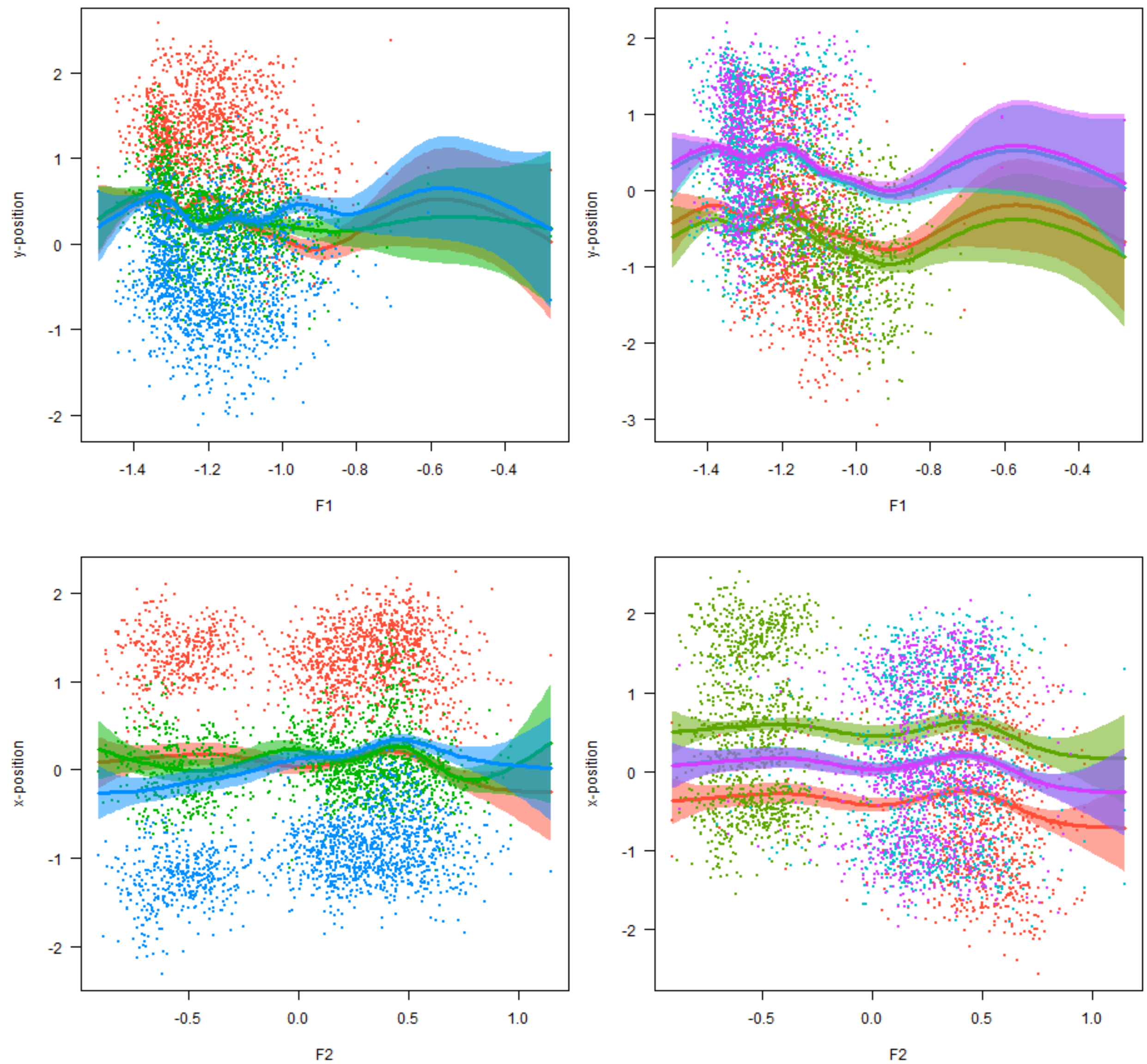

Figure 4.10: Fit of GAMs. Independent variable on x-axis (formant frequency), smooth term on y-axis (sensor position). Plotted by tongue sensor identity (left) and by phone identity (right).

the predictor parameter and dependent variable.

The second model (Figure 4.10 bottom row) is built with the x-position as the dependent, with F2 as a smooth term and sensory identity as a by variable and phone identity as an independent factor. The deviance explained by this model is 7.48%, with an EDF of 6.6, 7.7, and 6.2 for TT, TM, and TB respectively. The low deviance explained for the second model indicates that the second formant frequency is not a good predictor of tongue sensor

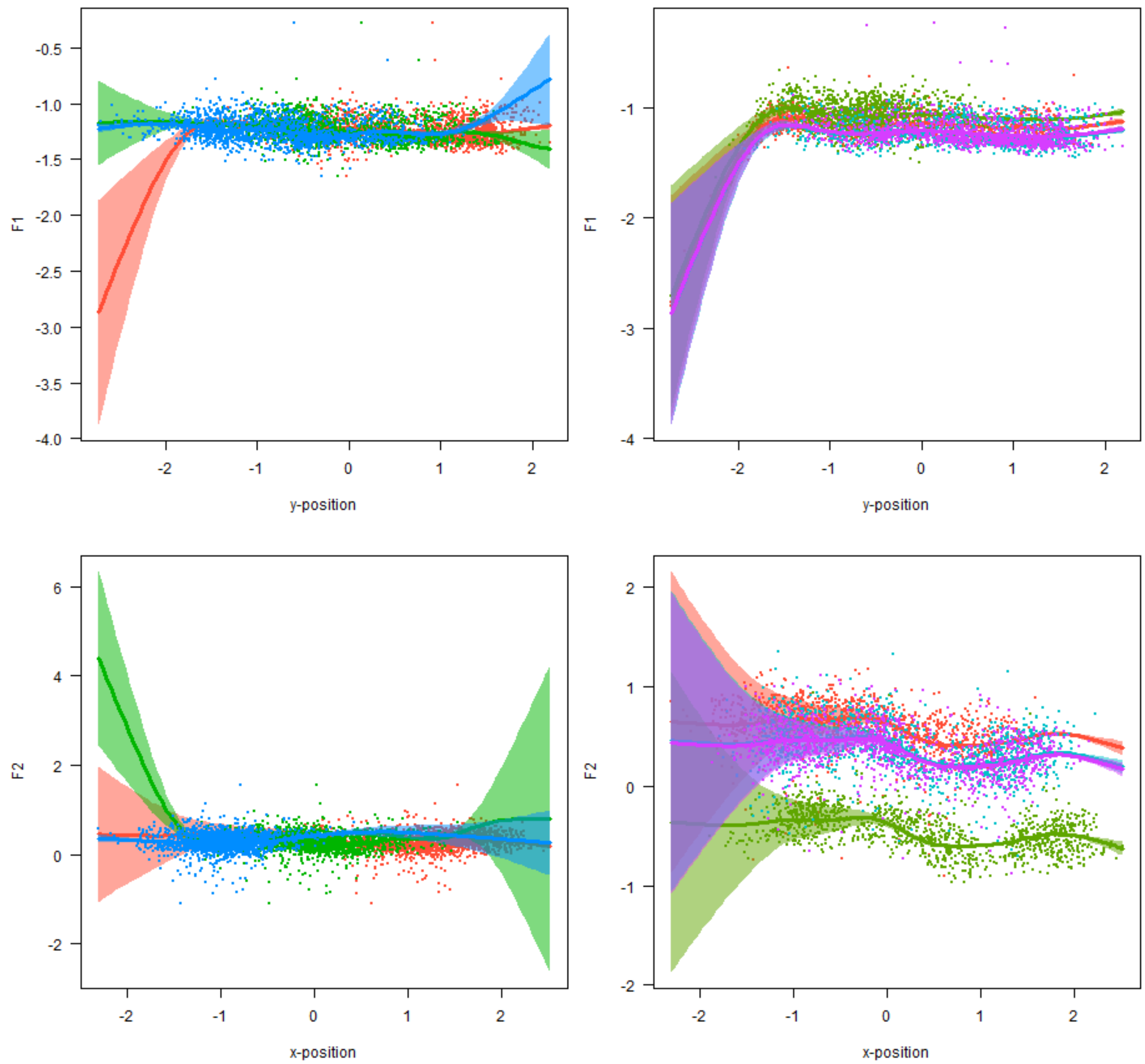

Figure 4.11: Fit of GAMs. Independent variable on x-axis (sensor position), smooth term on y-axis (formant frequency). Plotted by tongue sensor identity (left) and by phone identity (right)

x-position respectively for the four approximant sounds studied in this chapter, while the first formant frequency can be considered a satisfactory predictor of tongue sensor y-position. An EDF greater than 1 for all three parameters in both models indicates that there is a high degree of non-linearity between tongue sensor position and the first and second formant frequency.

The third and fourth model (Figure 4.11 top and bottom row, respectively) reverse the direction of the prediction; F1 is the dependent while y-position is the smooth term and likewise for F2 and x-position. The third model (where F1 is the dependent and y-position is the predictor) accounts for 40.2% of the deviance in the data. When y-position is the predictor, the explained deviance goes up by 12.2%. An EDF of 8.1, 7.4, and 8.8 is reported for TT, TM, and TB respectively.

The fourth model (where F2 is the dependent and x-position is the predictor) accounts for much more deviance in the data; the deviance explained is 76.9%, capturing 69.42% more of the information in the data. This is particularly high, as explained deviance of linguistic data is not typically expected to approach 100%. An EDF of 5.1, 8.5, and 6.5 is reported for TT, TM, and TB respectively. Like the first two models, the third and fourth model demonstrate a non-linear relationship between the first and second formant frequency and tongue sensor position.

This section finds that sensor position is a better predictor of F1 and F2 overall, as opposed to the reverse. However, the y-position of tongue sensors performs only slightly better at predicting F1 than the reverse, explaining just 12.2% more of the deviance in the data. Comparatively, x-position of tongue sensors is a much better predictor of F2 than the reverse, capable of explaining 69.42% more deviance. This indicates that while F2 is a poor predictor of tongue anteriority, F1 and tongue height are equally capable at predicting one another. Regardless, given the ability of the first and second formant frequency to explain just 28% and 7.48% of the deviance in the y-position and x-position of the three tongue sensors respectively, field researchers interested in making articulatory claims based purely on acoustic data recorded for the palatal lateral, palatalized lateral, alveolar lateral, and palatal approximant are advised against doing so.

4.4 Discussion

This study investigates the acoustics of the BP palatal lateral, facilitated by comparisons to the palatalized lateral, alveolar lateral approximant, and palatal approximant. The corresponding nasal counterparts were not included in Sections 4.3.2 - 4.3.5 of this chapter due to the known effect of nasality on acoustics (Engwall et al., 2006; Carignan et al., 2011; Shosted et al., 2012a), effects not predictable given the articulatory methods utilized here.

4.4.1 Acoustic characteristics

Results indicate that the palatal and palatalized lateral have a low F1 and high F2 of approximately 350 Hz and 1900 Hz respectively, with no evidence that formant frequencies are significantly different between the two speech sounds; this indicates that while EMA is unable to capture the entire tongue contour, large differences in the vocal tract beyond the third EMA sensor (TB) are unlikely to have occurred between the palatal and palatalized lateral as the effects would be evident in the first and second formant frequency. Compared to formant values calculated for the palatal lateral as produced by a male speaker of Russian (Fant, 1960:167), the values observed here are about 200 Hz higher. Similarly, the values reported here for the palatal approximant and alveolar lateral in BP are about 200 to 600 Hz higher than the formant frequencies reported for the same sounds in English (Stevens, 2000:516,526). This discrepancy between the current findings and previous literature is likely due to the effect of gender. Like the articulatory results from Chapter 3, this study does not find a significant difference between the acoustics of the palatal lateral and palatalized lateral; SS-ANOVA results indicate that there is no evidence that the first through third formant frequencies are significantly different from one another.

The F2 values of the palatal sounds reported in this study fall within the range of F2 values (1,600 to 2,400 Hz) reported by Recasens (1984b) for a single male speaker of Catalan; the BP palatal lateral, palatalized lateral, and palatal approximant have a respective median F2 of 1,954 Hz, 1,931 Hz, and 2,329 Hz at the middle of the phone. While Ladefoged and Maddieson (1996) suggest that laminal laterals are characterized by close proximity between the first and second formant frequency (i.e. the acoustic consequences of a low back tongue body position), this study finds this only for the alveolar lateral, which has a median difference of 585 Hz between the first and second formant frequency at the middle of the phone. This is consistent with observations that the alveolar lateral approximant in Portuguese is “dark” (Strevens, 1954; Andrade, 1999), i.e. articulated with a retracted tongue position (Ladefoged and Johnson, 2014). However, contrary to the prediction made by Ladefoged and Maddieson (1996), articulatory results from Chapter 3 (see Section 3.3.3) indicate that the alveolar lateral is articulated with an apical gesture, while the laminal laterals (i.e. the palatal and palatalized lateral) have a large respective F1 and F2 difference of 1,587 Hz and 1,563 Hz respectively (see Section 4.3.2). This suggests that the palatal and palatalized laterals are articulated with a high front tongue body, which is consistent with articulatory results in Chapter 3. The findings in this chapter and in Chapter 3 do not support the suggestion that laminal laterals are articulated with contiguous first and second formant frequencies; contiguity between the first and second formant frequencies is not expected for a sound produced with a high front tongue body.

Models of the lateral side branch indicate that the resulting lateral antiresonance should affect frequencies which lie in the range of 3,500 to 4,000 Hz, with the lowest possible lateral antiresonance extending down to 2,200 Hz (Stevens, 2000)[4]. Since antiresonances dampen overlapping formant frequencies (Johnson, 2011), lateral antiresonance is expected to attenuate the third formant frequency for the three lateral sounds included in this study. Given that the third formant frequency is associated with quality or naturalness in the alveolar lateral approximant (Harris et al., 1958; Fant, 1960; Ladefoged and Johnson, 2014) and is crucial for distinguishing between liquids and semi-vowels (O'Connor et al., 1957a; Espy-Wilson, 1992), attenuation of the third formant frequency may crucially affect listeners ability to correctly and consistently identify the palatal lateral. As a result, the lateral antiresonance may be a strong factor in the development of *yeísmo*.

Even though Recasens et al. (1993) define "duration" in articulatory terms (i.e. occlusion), they report a mean duration for the Catalan palatal nasal that is almost identical to the acoustically-defined durations reported here (84.4 ms), with only a 0.8 ms difference. Reported mean duration for the Italian palatal nasal is of course much longer (nearly double), due to the geminate nature of the sound. In contrast, later results from another study on Majorcan Catalan (Recasens and Espinosa, 2006) report durations that directly contradict the findings here; they find that the palatal lateral is consistently longer than the palatal nasal across syllable position in an /a/ vowel context, while we find that the palatal nasal is

[4]Note that the frequency of the antiresonance can vary quite significantly as a result of vowel environment or speaker-dependent oral tract specifications (Stevens, 2000).

significantly longer than the palatal lateral in the same vowel context (84.4 ms as opposed to 74.4 ms). Both Recasens et al. (1993) and Recasens and Espinosa (2006) define duration in articulatory terms (i.e. amount of electrodes activated on an artificial palate), which should correspond to a decrease in amplitude in the acoustics (how phone duration was defined here).

Since the palatal nasal and palatal approximant are articulated with significantly longer durations than the palatal lateral, the shorter duration may affect the amount of linguistic information a listener is able to ascertain regarding the identity of the speech sound. Shorter durations may also result in a higher occurrence of linguistic undershoot, which would further increase the listener's difficulty in correctly identifying the intended speech sound. Note that the observation by (Ladefoged and Maddieson, 1996) that laminal articulations may result in slower transitions in comparison to apical articulations is not borne out in this study, as the palatal and palatalized lateral and palatalized nasal - all three laminal sounds, have the shortest phone durations of all the sounds studied here. The combination of attempting an articulation that takes longer to achieve in conjunction with a shorter overall constriction duration may also aggravate the issue of articulatory undershoot, resulting in a more vowel-like articulation as opposed to an approximant. The role of duration in the typological rarity of the palatal lateral will be revisted in the discussion section of Chapter 5 (see Section A.4.2), which will focus on the perception of the BP palatal lateral by native and non-native speaker.

4.4.2 Mapping the relationship between articulation and acoustics

Correlations between sensor position and formant frequencies by phone reveal another subtle difference between the palatal and palatalized lateral; the palatal and palatalized lateral have similar correlation scores between y- and x-position with F1 and F2 respectively for two sensors (TT and TB), but not with regards to the TM sensor. This finding suggests that there may be more variation at this section of the tongue for the palatal lateral. However, both the palatal and palatalized lateral demonstrate rather weak relationships between the articulation and acoustics; correlation scores across sensor for both sounds do not go above 0.22.

According to predictions made by the acoustic theory of speech production (Fant, 1960; Stevens, 2000), one would expect a stronger relationship between the first two formant frequencies and sensor position. However, this expectation is supported only in the broader context when calculating the correlations for all four of the sounds combined (see Table 4.2), indicating that within-phone differences in tongue position do not demonstrate a significant effect with regards to the resulting acoustic signal.

While generative phonology (Jakobson et al., 1952; Chomsky and Halle, 1968) describes speech sounds in terms of categorical terms (i.e. [+high], [+consonantal]), articulatory studies on the relationship between vowel "height" (as defined in phonological theory) and the physical tongue position have found that the articulatory configurations do not correspond to categorical expectations (Russell, 1970; Wood, 1975; Ladefoged, 1996). These same studies found that while the acoustics were faithful to expectations for the vowels, the paired articulatory results demonstrated inter-speaker variation; i.e. different speakers utilized different articulatory strategies to produce the same acoustic result.

The many-to-one issue is present in Section 3.3.6, where speaker-specific models of the tongue reveal idiosyncratic strategies for articulating the palatal and palatalized sounds. This is supported by findings from the GAMs in Section 4.3.5, which indicate that articulation is a stronger predictor for the acoustics than the reverse. This is particularly true of the relationship between tongue frontness and the second formant frequency; sensor x-position is able to explain more than 76.9% of the F2 data, which is 69.42% more than the reverse. Similarly, sensor y-position performs moderately better than F1, explaining an additional 12.2% of the deviance in the data; F1 is able to explain 28% of the sensor height data, while sensor height explains 40.2% of the F1 data. The distinction between the first and second formant frequency is also captured in the individual correlations (see Tables 4.3 - 4.4), which reports stronger correlations overall between F1 and sensor height than F2 and sensor frontness. These findings indicate that the idiosyncratic articulatory configurations employed by the speakers to produce the palatal and palatalized lateral in this study are more diverse in terms of tongue frontness, but not tongue height.

As a result, one might be able to state with some degree of confidence that an increase in F2 during the production of a palatal lateral is likely due to lowering of the tongue blade anterior. However, even when squaring the highest correlation coefficient observed, the highest possible R-squared value is less than 0.30. At best, the second formant frequency is only able to capture less than 27% of the variation in tongue position. Therefore, acoustic studies concerning the palatal and palatalized lateral should carefully consider this dichotomy when discussing associated articulations. Regarding the findings presented above, changes in the acoustics cannot be used to indicate changes in the articulations with sufficient confidence.

4.5 Conclusion

Evidence from the correlations and GAMs indicate that the acoustics alone are unable to account for differences in the palatal and palatalized lateral. As a result, the study of the articulations and their relationship to the acoustic output becomes all the more important. While formant frequencies reported for the palatal approximant are more similar to the palatal and palatalized lateral than the alveolar lateral, the findings here are unable to explain why the palatal lateral merges with the palatal approximant and not the reverse. To answer this question, a study of how the palatal lateral is perceived was conducted (see Appendix A). With regards to the existence of a phonemic contrast between the palatal and palatalized lateral, articulatory and acoustic evidence from Chapter 3 and Chapter 4 do not support the expectation that the palatal lateral and palatalized lateral are two separate phonemes with distinctive articulatory or acoustic characteristics.

Chapter 5

GENERAL DISCUSSION

A detailed investigation and comparison of the articulation (Chapter 3), acoustics (Chapter 4), and perception (appendix) of the palatal lateral against similar speech sounds in BP was completed. This section integrates the results from the three studies in a codified discussion of the implications for the rarity of the palatal lateral, while situating the findings in the context of previous literature.

5.1 Major findings

Major findings of this dissertation include the discovery that there is only minimal evidence of significant differences between the palatalized lateral and palatal lateral with regards to how the two sounds are articulated. Additionally, there is no evidence of significant differences between the palatalized lateral and palatal lateral with regards to the resulting acoustics. A comparison of the formant frequencies of the palatal lateral, palatalized lateral approximant, palatal approximant, and alveolar lateral approximant reveal that there is no evidence that the palatalized and palatal lateral are significantly different with regards to the first three formant frequencies in both static and dynamic measures. There is a greater resemblance between the palatal approximant and the palatal lateral (as well as the palatalized lateral) than the alveolar lateral approximant, though the resemblance is not significant.

Given that the articulation of the palatalized nasal resembles the palatalized lateral and the palatal nasal resembles that of the palatal approximant, the presence of only minimal differences between the palatal lateral and palatalized lateral suggests that the palatal lateral in Brazilian Portuguese is likely simply a palatalized lateral. This is supported by results from the perception study included in the appendix (see Appendix A), which found that even

Articulatory study	**Acoustics study**	**Perception study**
Minimal evidence of significant difference between TT position for the /ʎ/ and [l^j] with regards to midsagittal tongue blade region.	No evidence of significant differences found in F1, F2, and F3 between /ʎ/ and [l^j].	Higher confusion rates between [l^j] and /j/ than /ʎ/ and /j/ for both groups when SNR is 0 and -3 dB.
[n^j] resembles /ʎ/ and [l^j] with regards to midsagittal tongue blade region.	/ʎ/ and [l^j] manifest a low F1 (lower than /l/ and /j/) and high F2 (approaching /j/).	Effect of native language: When SNR is 0 and -3 dB, True Positive rates for /ʎ/ and [l^j] are comparable for both groups (chance), while rates for /ɲ/ and [n^j] are relatively high for BP and chance for English.
/ɲ/ resembles /j/ with regards to midsagittal tongue blade region.		Lateral/oral resonance not as salient as nasal resonance.

Table 5.1: Major findings from each study included in this dissertation.

native speakers of Brazilian Portuguese were unable to identify the palatal and palatalized lateral at better than chance.

The most distinct difference between the palatal and palatalized lateral appears during the perception experiment. In the quiet condition and when SNR is raised to -3 dB, the palatalized lateral demonstrates higher confusion rates with the palatal approximant in comparison to the confusions between the palatal lateral and palatal approximant. This is true for both the BP- and English-speaking listeners; i.e. both groups demonstrate higher confusion rates between [l^j] and /j/ and lower confusion rates between /ʎ/ and /j/. Note however that the difference in confusion rate is not large; the biggest difference during the first two conditions is about 10% and 15% respectively for the BP and English group. There does not appear to be an effect of native language with regards to discriminating the palatal and palatalized lateral; the True Positive rates observed for the palatal and palatalized nasal are comparable between the two participant groups in all five SNR conditions. In contrast, familiarity with the palatal and palatalized nasal results in higher True Positive rates for the first two SNR conditions; while the BP group is able to identify the palatalized and palatal nasal with relatively low confusion, the English group identifies the two sounds at nearly chance.

The difference reported above from the perception study does not appear to be the result of differences in the posterior oral cavity, as large differences in the portion of the tongue not captured by EMA would result in observable differences in the first three formant frequencies, which we do not find. Furthermore, while EMA did not capture large differences in the tongue blade shape with regards to the palatalized and palatal lateral, our measures were able to successfully demonstrate a significant difference between the palatalized and palatal nasal; this suggests that EMA is capable of capturing the difference between a palatalized and palatal sound. Arguably, the palatalized lateral and palatal lateral should be considered separate sounds, as the former occurs as the result of coarticulation with a following high front vowel and the latter is already present in the language as a member of the BP phone inventory. However, we do not find stable measurable acoustic or articulatory differences between the palatalized and palatal lateral. It is possible that the perceptual difference between the palatalized and palatal lateral might be due to differences in the size and shape of the lateral airstream channels, as the acoustic model of speech production demonstrates that small changes in the side branch greatly affect the frequency range of the antiresonance formed by this side branch (Fant, 1960; Stevens, 2000). Regardless, even if large significant differences in the size and shape of the side branch exist between the palatalized lateral, these differences do not seem to assist listeners in the correct identification of the palatalized and palatal lateral.

While only minimal significant differences between the palatalized lateral and palatal lateral were found, the two sounds exhibited a few trends in the perception data that might be interesting to pursue as the focus of future study. Results from the perception study indicate a subtle difference between the perception of the palatalized and palatal lateral; these findings indicate that in comparison to the palatal lateral, the palatalized lateral is easily misheard as the palatal approximant, even when compared to the palatal lateral. This phenomenon does not appear in native speakers until noise is added to the signal; it is present in the non-native speakers even in the quiet condition.

Findings from this dissertation suggest that the palatal and palatalized lateral should be treated as the same, at least with regards to discussions of typology. Several studies have presented substantial information on the articulation (Catalan (Recasens, 1984b; Recasens et al., 1993; Recasens and Pallarès, 2001; Recasens and Espinosa, 2006; Recasens and Rodríguez, 2015), EP (Martins et al., 2008, 2010; Teixeira et al., 2012), Russian (Fant, 1960; Kochetov, 2005; Proctor, 2009), Spanish (Straka, 1965; Navarro Tomás, 1968),) and acoustics (BP (Silva, 1999), Corrientes Spanish (Colantoni, 2004), Catalan (Recasens, 1984b), Italian (Vagges et al., 1978), Russian (Fant, 1960:167)) of the palatal lateral, however there has yet to be a careful study of the differences between a palatal lateral and a palatalized lateral; given the detailed and comprehensive investigation of the palatal and palatalized lateral presented here in this dissertation, it is suggested that the findings could be generalized for all languages which can claim to possess a palatal lateral. The differences between the palatal and palatalized lateral are small enough that the effect of phonemic status is largely limited to minor differences in confusion patterns.

However, the presence of only minimal significant differences (i.e. at the tongue tip during the first half of the phone) between the palatal lateral and palatalized lateral is unexpected, especially when compared to results reported here for the palatal and palatalized nasal which find that the two sounds are articulated with a significantly different anterior tongue shape throughout the entire duration of the phone and are readily discriminated by listeners who are native speakers of BP. The production of the palatalized and palatal nasal (at least with regards to the midsagittal aspect of the anterior tongue blade) were expected to mirror the production of the palatalized and palatal lateral; findings from this study do not support this expectation. Instead, the production of the palatalized nasal resembles the palatalized and palatal lateral, while the palatal nasal strongly resembles the palatal approximant. Large and consistent significant differences found in the articulation and acoustics of the palatalized and palatal nasal but not the palatalized and palatal lateral are mirrored in the perception results obtained from BP speakers during the quiet condition; the palatalized and palatal nasal are identified without too much trouble while the palatalized and palatal lateral are confused for one another at nearly equal rates. Given that palatographic evidence from Spanish (Navarro Tomás, 1968) illustrates nearly identical linguopalatal contact patterns between the palatal nasal and palatal lateral, the palatal nasal and palatal lateral were expected to be articulated similarly here as well.

A few possibilities for why the palatal lateral and the palatal nasal are different are considered, the first of which is the difference in manner. It is well known that a nasal sound is produced by opening the velopharyngeal port, while a lateral sound is produced with one or both sides of the tongue lowered to create a lateral airstream (Ladefoged and Johnson, 2014). With the addition of a palatal place of articulation, the nasal sound coordinates two different articulators (i.e. the velopharyngeal port and tongue blade) while the lateral sound coordinates different muscles within the same articulator (i.e. the tongue blade). The coordination of creating the palatal place of articulation and lateral airstream may prevent the palatal lateral from achieving the same tongue shape as the palatal nasal; i.e. the steeper tongue blade angle that is characteristic of a palatal approximant as well. Instead, it is articulated with a more moderate tongue blade angle, resembling the palatalized nasal instead.

Another possibility considers whether it is unusual that we actually find a difference between the palatal and palatalized nasal. The articulatory similarity between the palatal nasal and the palatal approximant observed in this dissertation corresponds to an EPG study by Shosted et al. (2012b), who find that BP speakers approximate the palatal nasal in the majority of elicitations, while PS speakers consistently produce the sound with complete occlusion. Paired with our findings, this indicates that while the BP palatal nasal has undergone a nasal version of *yeísmo*, the palatal lateral has not; this dissertation clearly demonstrates that while some varieties of Portuguese may be experiencing *yeísmo* (Azevedo, 2005), speakers of BP from São Paulo state do not pronounce the palatal lateral as an approximant. It is possible that if the studies in this dissertation were reproduced with speakers that articulate both the palatal lateral and palatal nasal with complete occlusion, statistical analysis will indicate no evidence of significant differences between the palatal lateral and palatal nasal, with the possibility of the lack of evidence extended to the palatal and palatalized nasal as well.

5.2 Factors contributing towards *yeísmo*

This dissertation placed a special emphasis on the merger of the palatal lateral with the palatal approximant and the relationship of this sound change with the rarity of the palatal lateral. Out of 28 languages that are known to possess or have possessed the palatal lateral, five languages were identified as having merged the palatal lateral with the palatal approximant, i.e. Basque (Hualde and Bilbao, 1992), French (Dauzat, 1899), Hungarian (Benko and Imre, 1972), Italian (Bladon and Carbonaro, 1978), and Spanish (De los Heros Diez Canseco, 1997). In particular, the palatal lateral in Spanish developed at two separate points in time and yet was lost both times (Lapesa and Pidal, 1942; Menéndez Pidal, 1950; Lipski, 1989; Penny, 2000; Hualde et al., 2005; Pharies, 2007; Zampaulo, 2015). It was hoped that understanding the mechanism behind *yeísmo* might provide additional insight into the rarity of the palatal lateral.

Initial proposals that the palatal lateral merges with the palatal approximant due to articulatory similarity are not supported by our findings; articulatory results indicate that the palatal approximant bears greater resemblance to the palatal nasal than the palatal lateral. Instead, evidence from the perception study in Appendix A indicates that noisy environments cause listeners to incorrectly identify the palatal lateral as a palatal approximant. Results indicate that the addition of speech-shaped noise increases the perception of a palatal approximant when listeners who are native speakers of BP are presented with a set of sounds (i.e. [ʎ, lʲ, l, ɲ, nʲ, n, d, dʲ, j]). Of these sounds, the palatalized lateral, followed by the palatal lateral, are the most likely to be misheard as a palatal approximant. The effect of noise is especially emphasized when the sounds are presented to English-speaking listeners, who misinterpret the palatalized and palatal lateral as a palatal approximant even in the quiet condition. *Yeísmo* is effectively the by-product of signal perturbation; when the acoustic signal of a palatalized or palatal lateral is disrupted, listeners are unable to accurately reconstruct the intended interpretation.

These findings support the claim made by Colantoni (2004), who argues that in Spanish, glide-like transitions from the palatal lateral to the following vowel result in the perception of a palatal approximant. While providing a plausible explanation for *yeísmo*, this study does not include a perception study to conclusively determine whether listeners do indeed perceive a palatal approximant in the transitions. Perception results from this dissertation successfully demonstrate that listeners do indeed perceive the palatal (and palatalized) lateral as a palatal approximant at a higher rate as compared to similar sounds including the palatal nasal and the alveolar lateral approximant. This is likely related to nasal resonance being more perceptually salient than oral resonance (Miller and Nicely, 1955).

I argue here that *yeísmo* occurs as a result of the palatal lateral's vulnerability to noise disruptions. When a listener hears the palatal lateral in the context of speech-type noise, an occurrence that commonly occurs during daily life, the disruption of the signal causes the listener to misinterpret the palatal lateral as a central palatal approximant; this characteristic of the palatal lateral directly contributes to the rarity of the sound.

5.3 Factors contributing towards the rarity of the palatal lateral

If the palatal or palatalized lateral is commonly misinterpreted as a palatal approximant, then perhaps the palatal and palatalized lateral are less perceptually salient in comparison to other palatal sounds. This would likely be a strong contributor towards the rarity of the palatal lateral. Findings from the perception study support this hypothesis; even in the quiet condition, d-prime scores indicate that both native speakers of BP and English are considerably less sensitive when discriminating the palatal and palatalized lateral from the other sounds included in the study. The comparable performance of both the BP group and English group also presents substantial support for the proposal that the palatal and palatalized lateral are typologically less perceptually salient. The overall difficulty that listeners have correctly identifying a palatal lateral contributes directly to the rarity of the palatal lateral.

A factor for why the palatal and palatalized lateral are less perceptually salient may be related to manner requirements. Recasens et al. (1993) claim that the production of a lateral airstream forces the palatal lateral to be articulated at a more anterior position than the palatal nasal and palatal approximant in Spanish. The more anterior place of articulation is said to contribute to lower linguopalatal contact, which is correlated with lower resistance to coarticulation with adjacent vowels (Recasens, 1984b). According to prototype theory (Rosch, 1978; Mervis and Rosch, 1981), greater articulatory variation (both inter- and intra-speaker) could cause difficulties for the maintenance of a distinct prototypical palatal lateral, which would in turn contribute to the rarity of the palatal lateral. However, proponents of exemplar theory (Goldinger, 1996; Pierrehumbert, 2002) have demonstrated that the brain is quite capable of retaining fine phonetic detail. Regardless, both interpretations place the onus of phone maintenance on the listener. I argue here that the rarity of the palatal lateral is largely due to the listener's inability to accurately recover the palatal lateral during speech; the evidence presented in this dissertation supports the theory of listener-driven sound change as proposed by Ohala (1993). This dissertation presents strong evidence that the resonance that results from nasal airflow is more salient than lateral or even central airflow; this is reflected in the d-prime scores and confusion matrices in the perception study at the third SNR condition, when the SNR is raised to -6 dB. During the first two conditions (i.e. the quiet condition and when the SNR is -3 dB), both groups of participants are nearly equally sensitive to the alveolar lateral approximant and alveolar nasal, with both groups slightly more sensitive to the alveolar lateral approximant in the first condition. It is during the third SNR condition when it becomes especially apparent that nasal resonance is particularly robust against the perturbation of speech-weighted noise for both participant groups: (1) d-prime scores remain positive for the alveolar nasal while the scores drop below zero for the alveolar lateral approximant, and (2) the alveolar nasal maintains a True Positive rate of 76.8% and 69.0% for the BP and English group respectively while the other sounds drop to below 20%.

The relative weakness and strength of lateral and nasal resonance respectively is reflected in cross-linguistic sound change, which suggests that the replacement of the alveolar lateral approximant with an alveolar nasal is a common occurrence. Cantonese speakers tend to pronounce the syllable-initial alveolar lateral approximant in English as an alveolar nasal (Chan and Li, 2000) and sinitic loanwords containing an alveolar lateral approximant were typically borrowed into Korean as an alveolar nasal (Kang, 2012); neither example is a result of coarticulation with adjacent nasal sounds. There is evidence that nasal harmony can result in the nasalization of an alveolar lateral approximant in languages like Yoruba (Ladefoged,

1964) and Kwa (Hyman, 1972), where a following nasal vowel causes an underlying alveolar lateral approximant to be nasalized (i.e. /l/>[n]). The alveolar nasal is also more frequent in the language database UPSID than the alveolar lateral approximant; the alveolar nasal is found in 202 of the languages contained within UPSID (i.e. 44.7% of the languages in UPSID), while the alveolar lateral approximant is found in 174 languages (35.8% of the languages in UPSID) (Maddieson and Precoda, 1991).

Substantial support is provided here for the role of the listener in the rarity of the palatal lateral. Perception results from this dissertation indicate that listeners (both BP-speaking and English-speaking) manifest the greatest difficulty when identifying the palatal and palatalized lateral. SS-ANOVA of the articulation and acoustics indicate that within a single given vowel context, the palatal and palatalized lateral are articulated with more or less the same degree of variability as other palatal sounds, including the palatal nasal and palatal approximant. Lastly, perception evidence suggests that lateral resonance is less salient than nasal resonance, contributing to misinterpretation and therefore loss of the palatal lateral in languages undergoing *yeísmo*.

5.3.1 Implications for the phonetic representation of palatal sounds

In Chapter 1, I argue that the phonemic status of a sound must be considered. Specifically, a palatalized sound should be a sound that occurs as a result of a phonetically palatalizing environment (e.g. occurring as a result of proximity to a glide or a high front vowel), while a palatal sound should present evidence of phonemic contrast. While the palatal and palatalized lateral in BP do not demonstrate strong differences with regards to the articulation or acoustics, the palatal and palatalized nasal do; this provides further evidence of significant differences between palatalized and palatal sounds (See also Keating and Lahiri (1993)), lending additional support to the argument that the phonemic status of a sound is crucial towards correctly representing and categorizing speech sounds.

Following the above argument, the "plain" and "palatalized" contrast in Russian should be corrected to a "plain" and "palatal" contrast, as the so-called palatalized sounds are contrastive and are therefore palatal sounds. Separate IPA symbols are not necessary to represent these sounds, as there are already pre-existing IPA symbols that are used to indicate a stop, nasal, or lateral that is articulated along or at the palate. In particular, the IPA symbols for palatal and palatalized sounds should be kept distinct, as it is necessary for consistent and responsible linguistic reporting.

Given that the IPA is organized in an articulatory manner (i.e. assigning categories based upon the manner and place of articulation), one must be circumspect in the reliance on language databases such as UPSID (though an indispensable resource to linguists), which typically draws upon field research conducted using acoustic or even impressionistic data. If the basis for the phonetic representation is articulatory in nature, then sounds to be represented based on those phonetic symbols should be identified and categorized based on their articulatory characteristics as well. Naturally, detailed articulatory research is often difficult to conduct when in the field, however, it bears mentioning that these classifications should be approached with a certain amount of caution.

In sum, this dissertation does not argue for a change in the current phonetic representation of speech sounds. However, the findings here indicate that phone classification is an aspect which must be conducted in a manner that is thorough and consistent cross-linguistically. If not, the identification and comparisons of rare sounds across languages increases in complexity.

5.4 Future research

Given that lateral resonance is identified as one of the major contributors towards the rarity of the palatal lateral, a study of how variability in the lateral airflow channel(s) affects acoustics would be relevant for understanding the precise role of lateral resonance during perception. In Appendix A, it was proposed that the perceived perceptual difference between the palatal and palatalized lateral may be due to listeners attending to differences in how the lateral antiresonance affects the acoustics. A more detailed description of the shape and size of the lateral channel(s) using MRI would be relevant for understanding how changes in the lateral channel(s) affect how listeners perceive the palatal lateral. Future research should reference a study by Teixeira et al. (2012), who have modeled the lateral channel(s) during the production of the EP palatal lateral from 3D MRI articulatory data. While there is no evidence as of yet whether the palatal lateral in EP is articulated the same as the BP palatal lateral, further study of the lateral channel(s) during the production of the palatal and palatalized lateral in both EP and BP would be useful for improving acoustic models of the palatal lateral. More accurate acoustic models will also contribute to our understanding of how palatalization affects the size and shape of the lateral channel(s) - and therefore, explain why listeners demonstrate higher levels of confusion between [lʲ] and /j/ in comparison to /ʎ/ and /j/.

A reproduction of the articulatory study in this dissertation comparing the palatal lateral in BP and EP using either EPG or EMA would be relevant for future research on dialectal variation in Portuguese. Shosted et al. (2012b) find evidence of approximation in BP productions of the palatal nasal and Teixeira et al. (2012) find that the EP palatal lateral is produced with complete occlusion, both of which correspond to the observations of the BP palatal nasal and palatal lateral in this dissertation. An exploration of the differences between the BP and EP palatal lateral may reveal why developments in the palatal nasal are not mirrored in the palatal lateral, despite the two sounds differing supposedly only in terms of manner.

While the perception study included in the appendix compared the perception of the palatal lateral and the palatal nasal in a large-scale identification task as a means of identifying major trends in the perception of the palatal lateral and similar sounds, future studies should be conducted on a smaller set of sounds. An experiment paradigm that forces participants to identify a contrast (e.g. a forced-choice task (Fechner et al., 1966)) is advised, as such a task is better suited for quantifying a listener's ability to distinguish between the palatal and palatalized lateral. In particular, listener confusion between the palatal and palatalized lateral may have obscured how well (or poorly) listeners distinguish the palatal lateral from the palatal nasal and palatal approximant.

Chapter 6

CONCLUSION

This dissertation presents a detailed and holistic description of the palatal lateral in BP through a three-part study of the articulation, acoustics, and perception. In particular, it fills the current knowledge gap regarding the perception of the palatal lateral, while presenting novel information of the production of the palatal lateral in BP. Comparisons to the palatal and palatalized nasal reveal that contrary to expectations, the production and perception of the palatal and palatalized lateral do not mirror their nasal counterparts. Major findings from this dissertation include the discovery that the palatal and palatalized lateral demonstrate only minor differences with regards to articulation or acoustics, which is reflected in listeners' inability to distinguish between the two sounds in an identification task included in the appendix. With regards to the rarity of the palatal lateral, this dissertation concludes that lateral resonance causes the palatal lateral to be vulnerable to perceptual confusions, especially when perceived in an acoustically noisy environment. The relatively higher frequency of the palatal nasal cross-linguistically is explained as a result of nasal resonance being a more salient perceptual cue than lateral resonance.

Details of the shape of the anterior tongue and dynamic models of tongue position over time were provided in Study 1 with regards to the palatal lateral, palatalized lateral, palatal nasal, palatalized nasal, alveolar lateral approximant, alveolar nasal, and palatal approximant. Findings indicate that there are only minor differences between the palatal lateral and palatalized lateral, which are articulated similarly to the palatalized nasal; all three sounds are produced with a high front tongue body position and a tongue blade angle that is more neutral in comparison to the palatal approximant and palatal nasal. Furthermore, the palatal nasal is in many regards articulated the same as the palatal approximant; in comparison to the other sounds included in this study, the palatal nasal and palatal approximant are articulated with the most anterior tongue position and the steepest tongue blade angle. Likewise, the alveolar nasal and alveolar lateral approximant are also articulated similarly; the two sounds manifest a neutral tongue blade angle and low tongue body position. Results from Study 2 reveal that there is no evidence that the formant frequencies reported for the palatal and palatalized lateral are significantly different, indicating that if larger differences in the tongue shape not captured by EMA (e.g. the tongue root) exist, there is no evidence that the difference has a significant effect on perception.

The perception study included in the appendix presents completely new information regarding the perception of the palatal lateral. Results indicate that while there is no evidence that the acoustics are significantly different and only minor differences in the articulation of the palatal and palatalized lateral, both native and non-native listeners confuse the BP palatalized lateral with the palatal approximant at slightly higher rates than the BP palatal lateral. This finding suggests that a palatalized lateral is more susceptible to *yeísmo* than a palatal lateral. Regardless, both the palatalized and palatal lateral are mistaken as the palatal approximate at higher rates than the other sounds included in the study, lending support for the role of the listener as the main contributing factor for the occurrence of *yeísmo*. While the palatalized and palatal lateral are significantly different from the palatal approximant with regards to both articulation and acoustics, the findings here suggest that lateral resonance is not a salient perceptual cue; this results in listeners inaccurately recovering the palatalized and palatal lateral approximant as a palatal central approximant, and thereby providing the conditioning environment for the palatal and palatalized lateral to merge with the palatal approximant. Additionally, large significant differences observed in the articulation of the palatalized nasal and palatal nasal are reflected in the perception of the two sounds by the BP-speaking listeners (and to a lesser degree, the English-speaking listeners as well); confusion matrices of the BP-speaking listeners' responses to the palatal and palatalized nasal reveal relatively low levels of confusion when identifying the two sounds in the quiet condition.

The investigation of potential factors resulting in the rarity of the palatal lateral greatly assists our understanding of the distribution of speech sounds across the world's languages, a topic which is particularly relevant for modeling language sound systems and for language documentation. Additionally, understanding the effect of the listener in sound change is particularly informative for both historical linguists attempting to reconstruct prototypes of long-dead languages, as well as linguists interested in modeling speech perception. Through the careful documentation of the articulatory characteristics of the tongue blade and resulting acoustics, this dissertation provides a platform from which cross-linguistic comparisons of the palatal lateral can be conducted.

Appendix A

PERCEPTION STUDY

A.1 Introduction

This study seeks to understand how the palatal lateral, a rare speech sound in human languages, is perceived. Compared to other palatal sounds, the palatal lateral is a rare sound amongst the world's languages (see Table 2.1). A potential explanation for its rarity may derive from perceptual sources, e.g. the palatal lateral may be vulnerable to signal disruptions in noisy environments, or acoustic similarity to the palatal approximant may cause the palatal lateral to be mistakenly identified as the palatal approximant. To test this, listeners in this study were administered an identification task via an on-line link to investigate the perception of highly similar sounds such as the palatalized lateral approximant [l^j] and palatal lateral /ʎ/ in Brazilian Portuguese (BP). Listeners were also presented with other palatal (/ɲ/ and /j/) and palatalized sounds ([n^j] and [d^j]) from BP, in order to determine whether the palatal lateral is more susceptible to identification errors in comparison to other palatal sounds. By exploring the perceptual similarity of the palatal lateral to the palatal approximant, results from this study may have implications for our understanding of why *yeísmo* occurs (i.e. the replacement of the palatal lateral with the palatal approximant). This study also included the alveolar sounds /l/, /n/, and /d/ to establish a baseline for the perception of non-palatal sounds.

During the identification task, the speech sounds are presented first in a quiet condition, followed by increasingly noisy environments. The addition of speech-shaped noise (i.e. noise where the long-term average spectrum (LTAS) is similar to the LTAS of speech) is used to induce perception errors. Confusion matrices will be used to analyze the direction and degree of confusions between sounds. The incremental addition of noise is expected to affect sounds differently (Miller and Nicely, 1955; Phatak and Allen, 2007; Toscano and Allen, 2014); if the palatal lateral is a rare sound as a result of being insufficiently salient, the accurate perception of the palatal lateral is expected to degenerate faster than the other sounds included in this study.

An identification task was chosen as the ideal methodology for this study. While a forced-choice task is one of the favored tools for exploring phone categories, this methodology assumes that listeners already have categories for these sounds. To avoid this bias and to enable the exploration of listener confusions, an identification task was preferred. Forced-choice tasks can also become cumbersome when more than one sound contrast is investigated. Here, the inclusion of several similar sounds is made possible because of the identification task, which also results in a more accurate imitation of the complexity of day-to-day speech.

Two groups of participants are included in this study: (1) listeners who are native speakers of BP from São Paulo state and (2) listeners who are native speakers of American English. The BP group is the primary target group, while native speakers of English are included as the control group. Since the English phone inventory does not include a palatal lateral, the English group should be completely unfamiliar with the palatal lateral as a phone category. Responses recorded from the English group will be used as a baseline measure against which the BP group can be compared, providing a means for understanding how the palatal lateral is perceived by listeners whose native language includes the palatal lateral or does not include the palatal lateral as a phone category.

A.2 Methodology

A.2.1 Participants

25 (11 male, 14 female) native speakers of BP were recruited for this task. 20 (11 male, 9 female) native speakers of American English were recruited as a control group. Native speakers of American English were recruited through Amazon Mechanical Turk (MTurk) (see (Heer and Bostock, 2010) for more on the use of MTurk in experimental design), while native speakers of BP were recruited by word of mouth. Only participants from São Paulo state were

included, in order to control for dialect variation; speakers from other states in Brazil might not distinguish between the palatal lateral and the palatal approximant (Azevedo, 2005), which is a crucial aspect of this study. All native speakers of BP from São Paulo state were eligible to participate, regardless of their current location. English speakers who had received any kind of exposure to Portuguese were not eligible to participate. Eligibility for inclusion or exclusion was self-reported by participants, who answered a short questionnaire about their language background. Portuguese speakers received $10, while English-speakers received $6 per MTurk compensation standards.

The participants in this chapter will be referred to as *listeners* who are *speakers* of a language. The term *listeners* will be used when discussing participants' responses and when referring to participants in general, particularly since this study is interested in the participants' capacity as listeners as opposed to speakers. The term *speaker* will be used when discussing the participant's native language, e.g. *native speakers of English*, since participants are typically not referred to as *native listeners* in perception literature.

A.2.2 Materials

The word list (see Table A.1) was produced by a single female speaker in her early twenties from São Paulo state. This speaker also participated in the articulatory study; she is identified in Chapter 3 as Speaker 5. While the acoustics were simultaneously recorded during the articulatory study, the level of background noise was deemed unsuitable for a perception study. As a result, the test items used in this study were re-recorded with an AKG C-520 head-mounted microphone in a sound-attenuating booth in the Phonetics & Phonology lab in the Foreign Language Building at the University of Illinois at Urbana-Champaign.

The word list was composed of two-syllable nonsense words, with stress on the first syllable. The second syllable contained the target speech sound in an intervocalic, syllable-initial position. The list contained nine words and were repeated twice, for a total of 18 individual test items.

Speech-weighted white noise was added to these items at varying levels of Signal to Noise Ratio (SNR). Noise was added in order to simulate real-world conditions and to investigate

Palatal	Plain	Palatal-ized
[ˈpa.ʎa] *bálha*	[ˈpa.la] *pála*	[ˈpa.lʲa] *pália*
[ˈpa.ɲa] *pánha*	[ˈpa.na] *pána*	[ˈpa.nʲa] *pánia*
[ˈpa.ja] *páia*	[ˈpa.da] *páda*	[ˈpa.dʲa] *pádia*

Table A.1: List of nonsense words for perception task. Phonetic representation in brackets [], orthographic representation in italics.

which sounds are more perceptually robust, i.e. less affected by noise. Given that unweighted white noise unevenly masks higher frequency speech more than lower frequency speech (Phatak and Allen, 2007:2314), speech-weighted white noise was created by applying a Linear Predictive Coding (LPC) filter to white noise generated by the function *randn*() in MATLAB (2014). Each item had noise added at five different SNR levels: quiet (no added noise), -3 dB, -6 dB, -9 dB, and -12 dB. Each file was then uploaded to SoundCloud as a wav file, in preparation for use with Qualtrics.

A.2.3 Procedures

The study was distributed to native speakers of English via MTurk and to native speakers of BP via word of mouth. The study was hosted by Qualtrics, an online survey software and insight platform. Each participant provided electronic informed consent. Participants filled out a language background questionnaire in which they specified information such as their age, native language, previous history of speech or hearing disorders, and knowledge of other languages. Checks were placed in the beginning of the study, e.g. "Do you have any hearing problems" and "Are you a native speaker of English/Portuguese?". Failing these checks disqualified participants from continuing, effectively preventing listeners who did not fit the appropriate profile from participating in the study.

Prior to the start of the study, participants were asked to move to a quiet room, turn off music, close unrelated tabs in their browser, wear headphones, and adjust audio settings to a comfortable level. Participants were also advised to avoid taking the study in Internet Explorer as the audio may not play properly.

The English-speaking participants were given a short practice session and test to determine eligibility for the study. During the practice session, participants were first presented with an interactive version of Table A.1; clicking on the written word would play a recording of the pronunciation. This was to allow the English-speaking participants to become familiarized with Portuguese orthography and the pronunciation. They were allowed to take as much time as necessary, but were advised to spend no more than five minutes on this section. A mock version of the study was administered when the participants were ready, consisting of recordings of each of the nine words played in random order. Participants received feedback on their selections during and only during the mock study.

Following the mock study, participants were given a short test to confirm that they had understood the directions. Both English- and Portuguese-speaking participants participated in the test. The three target sounds included in this test were [dʲ], /j/, and /l/. Since these sounds should be easily discernible for speakers of both languages, the test served to identify participants who did not meet baseline requirements for the study, e.g. those who did not have the audio adjusted to an appropriate level or those who may have some form of auditory impairment due to hearing loss.

During the study, participants were presented with a screen as in Figure A.1. Each screen contained nine words to select from and an embedded audio file. The audio played automatically as the page loaded. While participants were able to replay the audio if necessary, they were advised to make their selection as quickly as possible. After making their selection, participants clicked on the » button to move to the next page.

Each block of 18 unique test items (9 words x 2 repetitions) was presented five times before the SNR was raised, for a total of 450 presentations (18 unique items x 5 presentations x 5 SNR). Items were randomized within each block. The study took approximately one hour to complete.

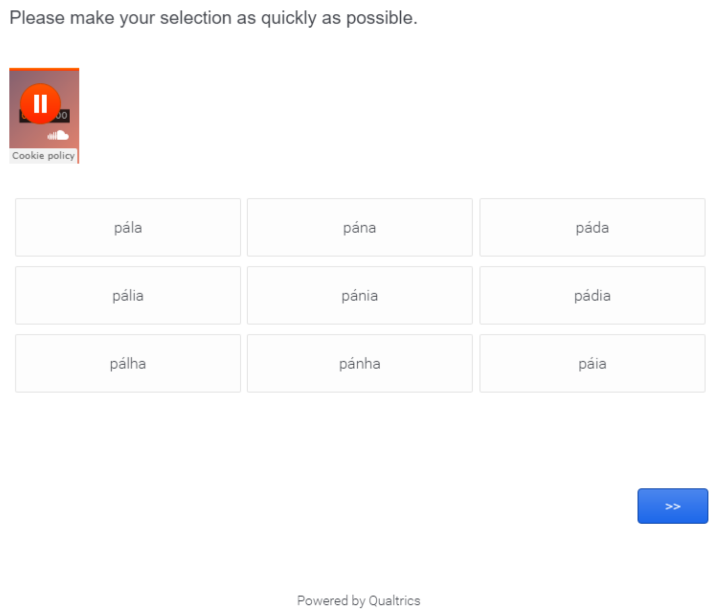

Figure A.1: Reproduction of what was presented to a participant during the study.

A.3 Results

Additional precautions for ensuring that participants were engaging in the study in a manner approximating that of controlled laboratory setting also included monitoring the length of time required by each participant to complete the study. None of the English-speaking participants required more than the expected completion time of approximately an hour, while a few of the Brazilian Portuguese-speaking participants did require a significantly longer completion time. However, each participant who recorded a longer completion time also personally reported to the researcher any difficulties which may have resulted in the discrepancy, indicating that the additional time required was not a result of distraction. Typically, the difficulties were all related to slower internet speeds in Brazil which disrupted or delayed audio streaming, especially towards the end of the study. However, this issue would have been present even if the study had been administered in a controlled laboratory setting in Brazil, as some of the participants reported having taken the study while using a campus internet connection. These issues do not appear to have significantly affected the outcome of the study however; further discussion of any potential effects is presented in Section A.5.

A.3.1 Confusion matrices

A confusion matrix is a useful tool for visualizing how listeners (both native and non-native) perceive and classify similar speech sounds (see Miller and Nicely (1955) and Phatak and Allen (2007) for additional examples of confusion matrices used in speech-related research). It is typically used to depict how well a classification model (here, the human brain) performs. Interpretation of confusion matrices often refer to the predicted versus actual class; here, these terms refer to the selections made by the participants versus the speech sound's actual identity.

$$ConfusionRate = \frac{ActualFrequency}{ObservedFrequency} \tag{A.1}$$

where:

ActualFrequency = # of times sound A identified as sound B.

ObservedFrequency = # of times sound B selected.

The rates within the confusion matrices (see Figures A.2 - A.6) were calculated by dividing the actual frequency by the observed frequency, with the actual frequency defined as the number of times sound A was identified as sound B and the observed frequency defined as the number of times sound B was selected. Calculating confusion rates is preferable over accuracy rates as confusion rates indicate where listeners are confusing two categories with each other. For example, listeners may accurately identify *palha* 100 times out of the 100 times that *palha* appears. However, listeners are selecting *palha* a total of 200 times, which indicates that listeners are actually confusing another sound with *palha* 50% of the time.

Confusion matrices were built for each language group (left: BP, right: English) by SNR level. The vertical axis represents the actual class (i.e the phone that was produced by the speaker) and the horizontal axis represents the observed class (i.e. the phone that was perceived by the listener). A heat map of the confusion rates was incorporated for ease of interpretation using R (R Core Team, 2015) package *ggplot2* (Wickham, 2009), with 0.0% represented by white and 100.0% represented by dark blue. The script used to created the confusion matrices was based on an R script by (Agrawal, 2011)[1]. The primary focus of

[1]Script included in appendix.

interest in the matrices here are the values running from the lower left hand corner to the upper right hand corner of the matrix. These values represent the True Positive rate: i.e., of the number of times that participants classify a sound as some sound X, how often is the classification actually correct?

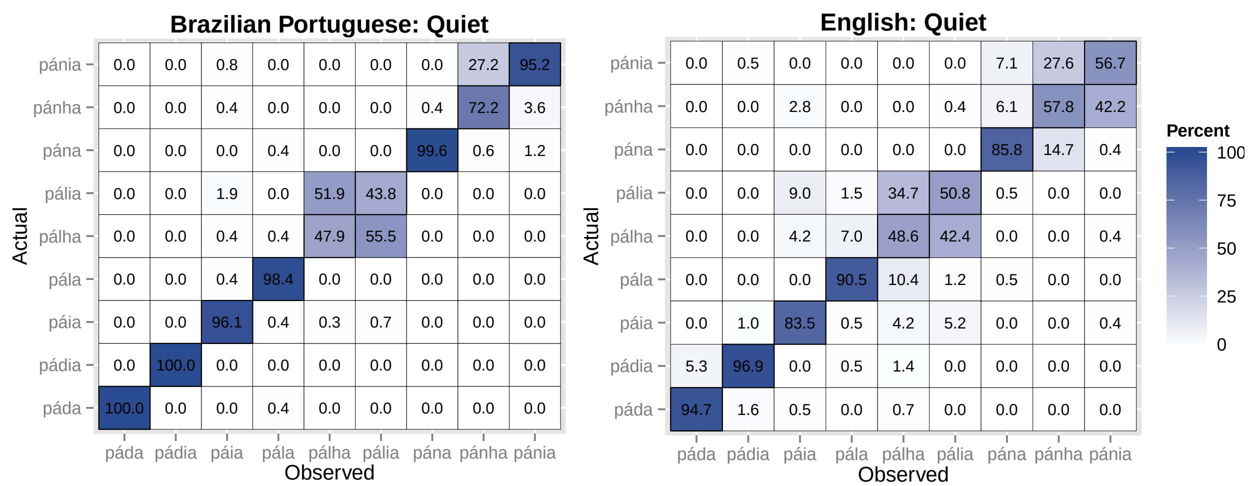

Figure A.2: Confusion matrix of participants' responses with no noise added. Refer to equation (A.1) for the calculation of confusion rates.

Figure A.2 indicates that the BP listeners (left) confused the palatal and palatalized lateral with one another at nearly equal rates, suggesting that speakers of BP are unable to discern between the two. This contrasts with the confusion rates observed for the palatal and palatalized nasal. The BP listeners performed much better when identifying the palatal and palatalized nasal; the palatal nasal was confused for the palatalized nasal only 27.2% of the time, while the palatalized nasal was confused for the palatal nasal just 3.6% of the time (though both sounds were occasionally confused with the plain alveolar nasal). Even in the quiet condition, there is already a small amount of confusion with (and only with) the palatal approximant (0.3% and 0.7% for *palha* and *palia* respectively), i.e. when *palha* or *palia* were chosen, the actual word was *paia* 0.4% and 0.7% of the time respectively. The fact that these confusions even occur in the quiet condition lends support for *yeísmo* occurring as a result of perceptual ambiguity between the palatal approximant and the palatal lateral. Though the participants correctly identified the palatal approximant 96.1% of the total times that *paia* was selected, participants also incorrectly identified palatalized nasals, palatal nasals, palatalized laterals, palatal laterals, and plain alveolar lateral approximants as a palatal approximant. The sounds that were the least confused were the plain and palatalized alveolar stops, with both sounds identified with perfect accuracy.

In the quiet condition (i.e. no noise), the English-speaking participants (see Figure A.2, right) demonstrated more confusion than the native-speaker participants; none of the words are identified with perfect accuracy. It is interesting to observe that differences in the responses between the two groups of listeners arise even prior to the addition of noise, indicating that native speaker familiarity does affect a listener's ability to discriminate between the palatal and palatalized contrasts. Similar to the BP listeners, the English listeners confused the palatal and palatalized laterals with other speech sounds around 50% of the time. The largest of these confusions were between the plain lateral and the palatal lateral: 10.4% of the time, they confused plain alveolar laterals for the palatal lateral, and the palatal lateral for a plain alveolar lateral 7.0% of the time. In contrast to the BP listeners, the English listeners had much higher confusion rates for the palatal and palatalized nasals. They identified palatalized nasals as palatal nasals 27.6% of the time and plain alveolar nasals as palatal nasals 14.7% of the time. Of all the times they chose the palatalized nasal, they chose correctly only 56.7% of the time, 42.2% of the time they incorrectly identified a palatal nasal as a palatalized nasal.

The comparable amount of difficulty experienced by the English-speaking group when discriminating the palatal and palatalized contrast for both the nasals and laterals highlights the dichotomy in the BP-speaking group when discriminating between the palatal and palatalized contrast for the laterals (poor discrimination) and nasals (good discrimination). Though both sounds are present as phones in English, the palatal approximant and the alveolar lateral approximant both demonstrate more generalized confusion, especially when compared to the BP listeners. Even without the addition of noise, these sounds seem to be more easily confused. There are also low levels of confusion of palatal approximants for other palatal or palatalized and lateral sounds. The plain lateral approximant was occasionally identified as a palatal lateral, a palatalized lateral, or a plain nasal, while listeners occasionally identified a palatal lateral, palatalized lateral, palatal approximant, and palatalized stop as the plain lateral approximant. Again, the plain and palatalized alveolar stops are the most robust against confusion, with 94.7% and 96.9% True Postive rates respectively.

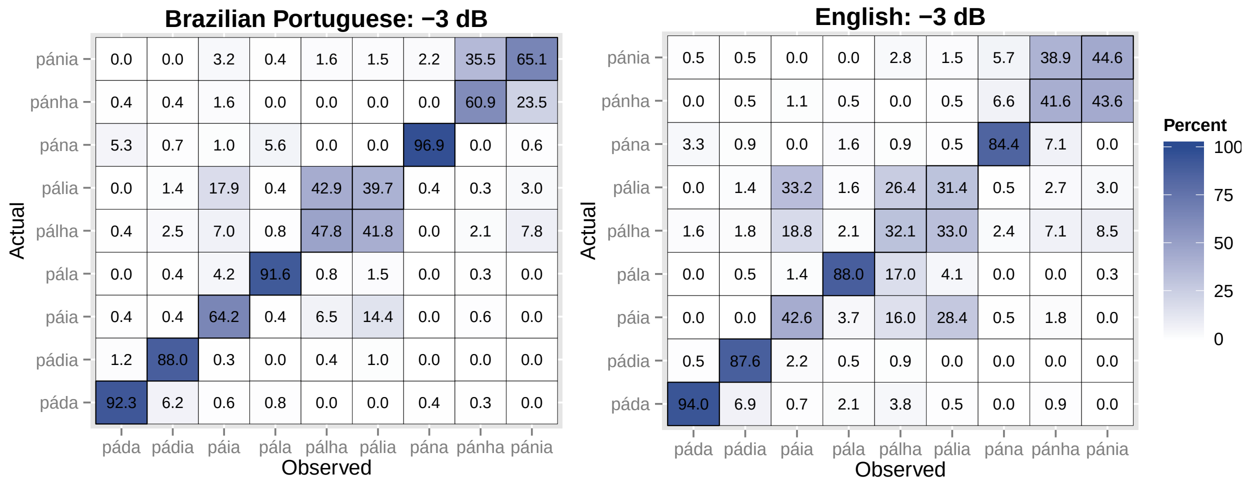

Figure A.3: Confusion matrix of participants' responses with an SNR of -3 dB. Refer to equation (A.1) for the calculation of confusion rates.

Figure A.3 reveals that the addition of -3 dB of noise increased the rate of confusion for palatal and palatalized laterals with the palatal approximant, resulting in the True Positive rate dropping 31.9% for BP listeners and 40.9% for English listeners. This increase in confusion is apparent in both participant groups and is stronger in one direction; palatalized and palatal laterals are misidentified as a palatal approximant at higher confusion rates as opposed to vice versa. This directly corresponds to the literature on historical sound changes related to the palatal lateral; palatal laterals are typically replaced with palatal approximants and not vice versa. In contrast, the plain and palatalized alveolar stop, alveolar lateral, and alveolar nasal maintain relatively high True Positive rates. Though the True Positive rate drops for the palatal and palatalized laterals in both groups, it is only by less than 10% and 20% for BP and English listeners respectively.

Additionally, the BP listeners now identify the palatal and palatalized nasals at nearly equal True Positive rates, in comparison to when the sounds were presented without noise (see Figure A.2, left). However, even with noise added, the BP listeners still perform better than the English-speaking participants by approximately 20%. We also begin to see a small amount of confusion with the palatal and palatalized laterals for both groups, particularly in the case for *palha*, where palatal and palatalized laterals are misheard as a palatal or palatalized nasal. This observation does not extend to when the palatal and palatalized nasals are misheard as palatal or palatalized laterals; this confusion occurs less than 3% of the time for both groups. The unbalanced direction of confusion lends support to the idea that the rarity of the palatal lateral may be due to a lower ability to be accurately perceived.

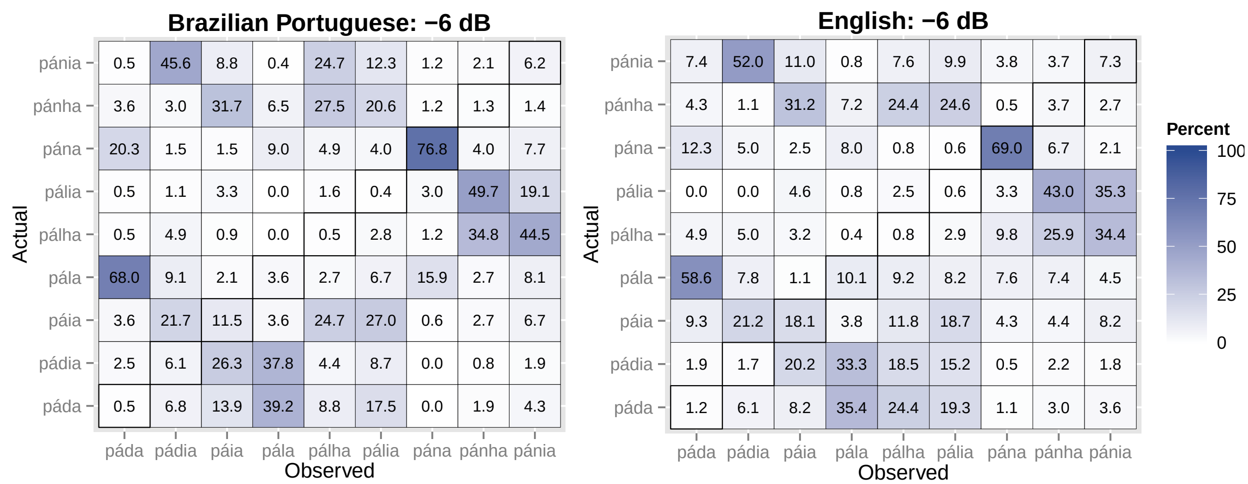

Figure A.4: Confusion matrix of participants' responses with an SNR of -6 dB. Refer to equation (A.1) for the calculation of confusion rates.

When the SNR is increased to -6 dB (see Figure A.4), there is a drastic change in the confusion patterns for both groups. We no longer see the distinct diagonal of high True Positive rates, with rates dropping to as low as 0.4% and 0.6% for the BP group (left) and the English group (right), respectively. In general, there is a large reduction in the True Positive rates except for the alveolar nasal, which maintains levels of 76.8% and 69.0% for BP and English listeners respectively. Additionally, the added noise results in the alveolar lateral approximant becoming confused as the plain alveolar stop at high rates (68.0% and 58.6% of the time for BP and English listeners respectively). While the reverse is also observed, participants from both language groups are nearly equal in confusing the plain and palatalized plosive as the alveolar lateral approximant. For both groups, the palatalized nasal was misheard as a palatalized alveolar stop for approximately 50% of the time that *padia* was chosen. The palatal approximant was also incorrectly identified as a palatalized alveolar stop at relatively high levels; approximately 21% for both groups.

For the BP listeners (Figure A.4, left), there are clusters of higher confusion rates between the palatal and palatalized laterals and nasals (though less so between the palatalized nasal and lateral). The same clusters can be observed in the English group (Figure A.4, right), though with lower confusion rates for the palatalized nasal misheard as either the palatal or palatalized lateral. However, the BP listeners confuse the palatal and palatalized laterals for a palatal approximant 24.7% and 27.0% of the time respectively, while the English listeners do so only 11.8% and 18.7% of the time respectively. The English listeners incorrectly identify the plain and palatalized alveolar stops as palatal and palatalized laterals at nearly equal confusion rates.

In the next SNR condition, the resemblance between the two groups becomes even more evident. While there are differences between the two sets of confusion matrices - sufficient to illustrate that the two groups do not behave identically - the extent of the similarity does not appear to be accidental. In this particular condition (see Figure A.4), even when the distinctive diagonal of high True Positive rates is lost, the matrices still mirror one another in their confusions rates. In particular, similarly high confusion rates of the alveolar lateral approximant for the alveolar stop are observed, as well as a high True Positive rate for the alveolar nasal (the only phone to still demonstrate a high True Positive rate).

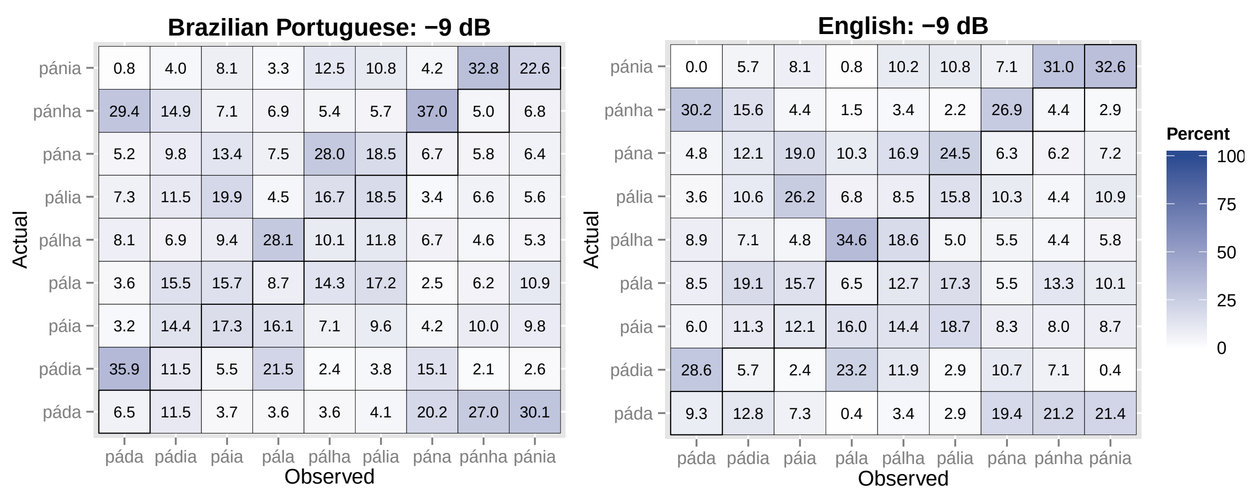

Figure A.5: Confusion matrix of participants' responses with an SNR of -9 dB. Refer to equation (A.1) for the calculation of confusion rates.

As more noise is added to raise the SNR to -9 dB (see Figure A.5), similar patterns of confusion develop again in both groups. For both the BP and English listeners, the palatal and palatalized nasal tends to be predominantly confused with either the palatalized nasal or the plain alveolar stop. The same bimodal confusion can be found for the alveolar stop, where the palatal nasal and palatalized alveolar stop are the most commonly misidentified as a plain alovelar stop. Likewise, the palatal nasal and alveolar stop are the two most commonly misidentified as a plain alveolar nasal. Lastly, palatal laterals, palatalized alveolar stops, and palatal approximants are the most commonly confused for the plain lateral approximant (in that order) for both the BP and English listeners.

Oddly enough, we find that the alveolar nasal and both the palatalized and non-palatalized alveolar lateral approximant demonstrate relatively high levels of confusion with the palatal approximant. Given that the alveolar nasal and palatal approximant do not share any articulatory similarities, it is unexpected that the alveolar nasal would demonstrate levels of confusion with the palatal approximant comparable to confusion levels reported for the palatalized and non-palatalized alveolar lateral approximant, which both share more articulatory similarities to the palatal lateral. Additionally, results from Chapter 2 indicate that there is no evidence of significant difference between the formant frequencies reported for the palatalized and palatal lateral, both with regards to dynamic and static measures; given that the palatalized lateral is acoustically more similar to the palatal lateral than the alveolar lateral approximant, participants were expected to confuse the palatalized and palatal lateral with other sounds at similar rates. However, this expectation is not supported by the perception data.

Note that the confusion rate for the palatalized lateral with the palatal approximant is much higher for the English listeners as compared to the BP listeners (26.2% versus 19.9% respectively), with a respective True Positive rate of 12.1% versus 17.3%. Native speakers of languages with a palatal lateral in their phoneme inventory may be slightly less likely to confuse palatal or palatalized laterals with a palatal approximant. However, since the English listeners demonstrate much higher confusion rates when inaccurately identifying palatal and palatalized laterals as a palatal approximant (13.4% and 18.7% respectively, compared to 7.1% and 9.6% for the BP group), it may be that non-native speakers find this contrast equally confusing regardless of the direction of confusion.

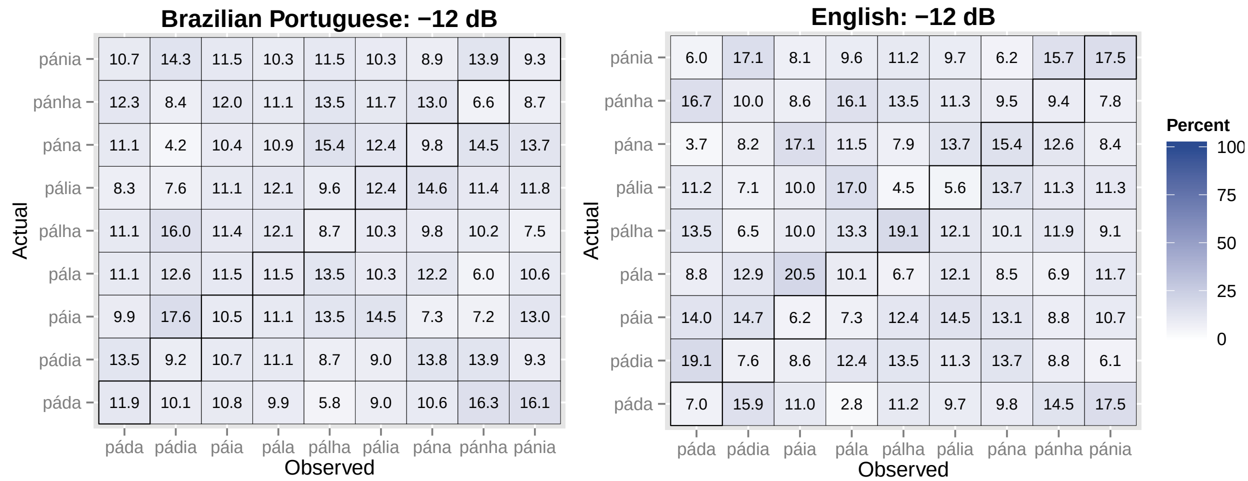

Figure A.6: Confusion matrix of participants' responses with an SNR of -12 dB. Refer to equation (A.1) for the calculation of confusion rates.

The last condition (see Figure A.6) was included to test the threshold of hearing capacity by raising the SNR to -12 dB. At this SNR, it is very difficult to distinguish speech from noise, which should provide a baseline comparison for when participants are simply guessing. Many participants reported that they were unable to identify anything during this condition and this is evident in Figure A.6, where the confusion rates are roughly the same across the board. In this SNR condition, the English listeners seem to demonstrate relatively greater confusions for certain sounds, particularly *pala* for *paia*. However, it is unwise to draw significant conclusions from such small differences in rates, especially since the stimuli in this condition are nearly inaudible due to the level of noise added.

A.3.2 Fleiss' Kappa Statistic

The Fleiss' kappa statistic was calculated for both the BP listeners and the English listeners, per SNR condition (see Table A.2). This statistic shows the degree of agreement between participants after subtracting the effect of agreement by chance, regarding the word selections that were made. While there is a suggested interpretation of kappa values (Landis and Koch, 1977), interpretation of a kappa statistic less than 1 (perfect agreement), greater than 0 (zero agreement), or less than 0 (patterned disagreement, not by chance) is somewhat arbitrary. However, since it is a normalized measure, the kappa statistic can be used to make relative comparisons between the degree of agreement in the native and non-native group.

SNR	Brazilian Portuguese	English
Quiet	0.812	0.646
-3 dB	0.602	0.514
-6 dB	0.349	0.291
-9 dB	0.0874	0.136
-12 dB	-0.0102	0.0157

Table A.2: Fleiss' kappa scores calculating agreement within both participant groups for each SNR condition.

High agreement for the BP listeners is expected in the quiet condition and this is found to be true. In contrast, the English listeners demonstrate lower agreement even in the quiet condition, with a kappa score that is similar to that of BP listeners when presented stimuli with a -3 dB SNR. By the time noise is raised to a -6 dB SNR, both groups are nearly equal with regards to within listener agreement. When -9 dB noise is added, there is very little agreement within each group of participants. The final condition results in negative agreement amongst the BP listeners; the low agreement within both groups indicates that there are few conclusions that can be drawn from the addition of -12 dB of noise.

A.3.3 d-prime measure

One method of measuring how discriminable one category is from another (also known as "sensitivity") utilizes signal detection theory to calculate a statistic referred to as d' (Macmillan and Creelman, 1991). This statistic takes the z-transform of the hit rate (e.g. the proportion of *palia* responses to the item *palia*) and subtracts the z-transform of the false alarm rate (e.g. the proportion of *palia* responses to other items). This effectively removes the effect of a listener's response bias, i.e. a listener's tendency to favor selecting a particular item. The larger the d', the easier the word (referred to as the "signal") is detected. A negative d' would indicate that the false alarm rate is larger than the hit rate.

When confronted with a hit rate of one or a false alarm rate of zero, a standard correction was first performed. For a hit rate of one, the hit rate is corrected using the formula 1 - $1/(2n)$, where n is the total number of items in the category of interest (e.g. the total number of times *palia* appears). For a false alarm rate of zero, the false alarm rate is corrected using the formula $1/(2n)$, where n is the maximum amount of times a listener could incorrectly identify other items as the correct item (e.g. the number of times a listener could incorrectly identify items as *palia*). With this method of correction, the largest possible d' is 6.36, indicating the highest level of sensitivity.

This statistic can also be used to describe the perceptual distance between sounds, which is especially relevant for investigating whether the rareness of the palatal lateral is due to the listener's difficulty in discerning it from another similar and more perceptually salient sound. The d' was calculated for the responses to each individual nonsense word (responses to each word were grouped by SNR condition and native language prior to calculation), which provides the perceptual distance of a single sound from the remainder of the sounds included in its group. Comparisons of the d' prime scores will be made in this section. Given the complete loss of discriminability in the -12 SNR condition (see Figure A.6), d' was not calculated for the this last SNR condition. Calculating the d' will provide a means of understanding how easily the two language groups were able to discriminate each stimulus in a given SNR condition.

	Quiet		-3 SNR		-6 SNR		-9 SNR	
	BP	Eng	BP	Eng	BP	Eng	BP	Eng
páda	6.13	3.31	3.67	3.04	-1.36	-1.00	-0.33	-0.12
pádia	6.36	3.40	4.04	2.93	-0.36	-0.90	0.02	-0.35
páia	4.83	2.66	2.45	1.34	0.02	0.37	0.37	0.06
pála	5.53	2.94	3.71	2.72	-0.64	-0.06	-0.17	-0.33
pálha	1.77	1.20	1.44	0.70	-1.31	-1.09	-0.05	0.31
pália	1.08	1.56	1.07	0.81	-1.50	-1.28	0.40	0.21
pána	5.34	2.78	3.81	2.70	2.08	1.92	-0.26	-0.34
pánha	3.53	1.35	2.41	0.95	-1.14	-0.54	-0.46	-0.45
pánia	2.98	1.85	1.72	1.49	-0.33	-0.28	0.56	0.96

Table A.3: d' scores for each word by group (BP versus English) and SNR condition.

Some obvious differences between the BP and English group are automatically apparent even in the quiet condition (see Table A.3). There is a larger spread in d' scores for the BP group, ranging from the highest possible d' score of 6.36 reported for *pádia* to a low d' score of 1.08 for *pália*, which is even lower than the d' score of 1.56 reported for the same word by the English group. The equally low scores for the palatal and palatalized lateral in both groups (1.77 and 1.08 versus 1.20 and 1.56 for /ʎ/ and [lʲ] in BP and English, respectively) indicate that regardless of native language, listeners had a great deal of trouble distinguishing these two sounds from the other sounds present in this study. For the remaining items, d' scores for the BP group are approximately double the d' scores calculated for the English group. As a whole, native language has a strong effect in this study.

Despite overall differences in the absolute scores, the distribution of relatively high versus low scores is similar in both groups. The highest sensitivity was reported for the alveolar stops, the second highest sensitivity for the alveolar lateral approximant and alveolar nasal, followed by the palatal approximant, and then the palatal and palatalized nasals. This pattern indicates that while the absolute perceptual distances may be affected by native language, the relative distances may be reflective of cross-linguistic characteristics regarding the ease of discrimination of the speech sounds examined in this study.

In the second SNR condition, d' scores reported for the BP group are nearly halved for all items except those containing the palatal and palatalized lateral. However, a SNR of -3 dB does not seem to greatly affect the BP listeners' ability (or inability) to discriminate these two sounds from the others. In comparison, while there is an overall reduction in sensitivity by the English listeners, it is not as dramatic as the reduction observed in the BP listeners. Note that the BP d' scores resemble those of the English listeners in the quiet condition; the BP participants demonstrate greater perceptual distances in the quiet condition but are easily affected by noise.

When the SNR is raised to -6 dB, the sensitivity is roughly the same for both the English and BP listeners. While the d' is negative for the majority of the words, it is interestingly positive for two words: *páia* and *pána*. An explanation for why d' is higher for *pána* can be found in Figure A.4, which shows that *pána* is the only word that remains resilient to the addition of more noise. As for *páia*, while the True Positive rate was low, the hit rate was relatively high and the false alarm rate was similar to the false alarm rate for the other words.

Since listeners made correct selections of *páia* more often than for other words, this resulted in *páia* being more perceptually distinctive (i.e. having a higher d') despite demonstrating similar confusion rates.

In the following SNR condition, raising the SNR to -9 dB severely reduces the perceptual distinctiveness of all words for both groups. There are no clear outliers in this condition. While some words improve in terms of d' (in particular, *pália* for both groups), it seems unlikely that the addition of noise aided participants in correctly identifying the word. Referring back to Figure A.5, the spread of confusion rates paints a similar picture: the increase in SNR rendered the sounds relatively similar to one another, with only a few sounds demonstrating slightly better True Positive rates.

A.4 Discussion

In this chapter, an identification task is administered to two groups of listeners: the target group, native speakers of BP, and the control group, native speakers of American English. The focus of this task is to understand how the palatal lateral is perceived. Similar sounds such as the palatalized lateral and the palatal approximant were included in this study in order to investigate whether the rareness of the palatal lateral may be a result of this sound being easily confused with similar speech sounds. The findings lend support for the expectation that the palatal lateral is easily confused with the palatal approximant, providing an explanation for the prevalence of *yeísmo* and rareness of the palatal lateral cross-linguistically.

A.4.1 The effect of noise on the perception of the palatal lateral

Confusion between the palatal lateral and palatal approximant is apparent even during the quiet condition (see Figure A.2). The confusion goes in both directions for both groups and includes the palatalized lateral as well, i.e. palatal and palatalized laterals are inaccurately identified as a palatal approximant and vice versa for both language groups. Of the palatal and palatalized sounds, excluding the palatalized stop, the palatal approximant has the lowest confusion rate. This is a possible indicator for why the palatal lateral merges with the palatal approximant and not the reverse during *yeśmo*.

Nativeness makes a clear contribution, resulting in obvious differences between the confusion matrices of the native (BP) and non-native-group (English) for the first two conditions. Across the board, the rate of confusion of course is higher for the English group. This group also demonstrates lower cross-listener agreement (see Table A.2) for the first three SNR conditions. Even in the absence of noise, the English listeners were unable to achieve similarly high levels of cross-listener agreement for sounds present in their native language. It is likely that the native language of the speaker producing the test items contributed to this uncertainty, since the English group is unable to discriminate sounds present in their native language at high levels of sensitivity (see Table A.3). However, the finding that listeners struggle when presented with non-native contrasts is not a novel observation; extensive literature on second language acquisition in adults has already established the difficulty of learning non-native phonemic contrasts (Flege, 1991; Guion et al., 2000; Ingram and Park, 1998; Aoyama and Guion, 2007).

Additional effects of nativeness are apparent in the first two conditions; non-native listeners have much greater difficulty identifying palatal and palatalized nasals, which is a distinction not present in English. The BP group does well when presented with the palatal and palatalized nasal in the quiet condition, with True Positive rates of 72.2% and 95.2% respectively, in comparison to the English group (57.8% and 56.7% respectively). Despite the ease of identification that previous familiarity with the sound contrasts provides, the effect is largely confined to the quiet condition. In the quiet condition, BP participants demonstrate very low confusion rates with regards to the sounds that are mistakenly identified as the palatal approximant (i.e. all the lateral sounds and the palatal and palatalized nasal), while the English participants demonstrate more pronounced confusion rates for the palatal and palatalized lateral in contrast to the lower confusions for the palatal and palatalized nasal. When noise is added, this same pattern is observed in the BP group, with the likelihood of the palatal and palatalized lateral being identified as a palatal approximant notably higher than that of the palatal and palatalized nasals.

The contrast between the palatal and palatalized lateral against the palatal and palatalized nasal is reflected in the d' scores; the palatal and palatalized lateral have the lowest sensitivity for both language groups, while the palatal approximant is much more easily discriminated from the rest of the phones. This discrepancy in sensitivity is repeated for the palatal and palatalized lateral but not the palatal and palatalized nasal up until the third SNR condition. Additionally, the extremely low sensitivity scores reported for the palatal and palatalized lateral even in the context of no added noise for native speakers (1.77 and 1.08, respectively) illustrates a simple reason for why this sound is so rare: it is difficult for listeners to correctly identify.

These findings indicate that while linguistic familiarity with the contrasts included in this study can reduce confusion in optimal auditory contexts, the addition of noise or other types of distractions present in daily life makes it more likely for the palatal and palatalized lateral to be misheard as a palatal approximant. The same is not found for the palatal and palatalized nasal, which is likely why the palatal nasal is found in 31.26% of the languages in the UPSID language database and the palatal lateral in only 4.43% (Maddieson and Precoda, 1991). Additionally, given that the palatalized lateral suffers higher confusion rates with the palatal approximant both groups (see Figure A.3), this provides a potential explanation of frequent historical accounts of so-called palatal laterals resulting from palatalizing contexts undergoing *yeísmo* (Dauzat, 1899; Maiden, 1995a; Pharies, 2007).

There is evidence here that there are subtle acoustic differences with regards to how these two sounds are perceived. The palatalized lateral is more likely to be mistakenly perceived as a palatal approximant by native speakers when uttered in an acoustic context that is less than optimal, which may be related to findings from Chapter 4, where the first two formant frequencies of the palatalized lateral were found to have a lower correlation to the position of the tongue anterior than the palatal lateral (see Section 4.3.4). Note, however, that higher confusion does not equate to less sensitivity in listeners; d' scores (see Table A.3) indicate here that sensitivity remains similar between the palatal and palatalized lateral regardless of SNR.

Referring back to the corpus frequency results mentioned in the discussion section of Chapter 3, words containing the digraph *lh* (representing the palatal lateral) are much more common than words containing *li* (specifically for contexts where the digraph is pronounced as a palatalized lateral, i.e. in an unstressed syllable); frequency rates of words containing the digraphs that represent the palatal and palatalized lateral are respectively 0.674% and 0.088% of total words in the São Carlos corpus (Linguateca, 1999). There are some difficulties when obtaining the frequency rate of the palatalized lateral, since the digraph *li* is pronounced as the CV sequence [li] when stressed (please refer to Section 3.4.2 for more information on how this frequency rate was calculated); even though the frequency rate for the palatalized lateral may not be exact, the corpus frequency results suggest that the higher confusion rates for the palatalized lateral are reflected in the lower frequency rate of the digraph *li* in comparison to the digraph *lh*.

A review of the literature on sound mergers involving the palatal lateral finds only a single language, Hungarian (Benko and Imre, 1972), that underwent a historical merger between the palatal lateral and the alveolar lateral approximant. An explanation for the infrequent merger of the palatal lateral with the alveolar lateral approximant (as compared to the palatal approximant) can be found in the results from the first two conditions of the identification task. While the palatal and palatalized lateral are indeed mistaken for the alveolar lateral approximant in both the BP and English group, the combined confusion rate of the palatal and palatalized lateral is considerably higher for the palatal approximant than for the alveolar lateral. The difference is more pronounced in the second SNR condition, indicating that the addition of noise increases the confusion in the direction of the palatal approximant and results in a lower likelihood for the palatal lateral to merge with the alveolar lateral approximant cross-linguistically.

The addition of more than -3 dB of noise results in a definite departure from the norm as depicted in the confusion matrices shown in Figures A.2 and A.3 of the first two conditions. Each additional increase in SNR after the first two conditions results in vastly different confusion matrices. The direction of confusions change and the strength of some confusions even reverse, demonstrating how different levels of noise can affect the speech signal. This study finds that intelligibility begins to break down when the SNR is raised to -6 dB, which corresponds to findings from Miller and Nicely (1955), who report that listeners are unable to discriminate sounds by place of articulation when the SNR is -6 dB. Intelligibility is lost entirely when the SNR is raised to -12 dB, representing the threshold of hearing capacity.

While there is always the possibility for inattention or boredom to affect participant choices (particularly as the sounds becomes progressively harder to identify), this study was designed specifically to elicit responses that were guided by subconscious rather than conscious decision, reducing the effect of boredom. Additionally, while this task took approximately one to two hours to complete, breaks were built into the design so that participants could refresh themselves as needed.

Despite differing only in terms of manner, the alveolar nasal is decidedly far less affected by -6 dB of noise than the alveolar lateral, with True Positive rates of approximately 70% for both groups in contrast to just 3.6% and 10.1% for the BP and English group respectively. As noted in the discussion of the articulatory study in Chapter 3, this finding lends further support for the possibility that the susceptibility of the palatal lateral towards historical mergers is due to the lateral nature of the sound. Specifically, lowering the side of the tongue to create a lateral airstream may affect the articulatory stability of the palatal lateral; Recasens (1984b) finds that the palatal lateral in Catalan is articulated with less linguopalatal contact than the palatal approximant and palatal nasal (which he attributes to lateral manner requirements), with lower linguopalatal contact linked to higher rates of coarticulation with surrounding vowels. Higher rates of coarticulation could plausibly contribute to the rarity of the palatal lateral.

It is also possible that the palatal lateral, particularly in comparison to the palatal nasal and palatal approximant, is rare because lateral resonance[2] is less salient than the resonance that results from nasal or central airflow. Simply put, nasal resonance renders the sound more audible or noticeable to listeners, especially when in comparison to lateral sounds articulated with only lateral (oral) airflow. A similar study (Miller and Nicely, 1955) of perception confusions finds that listeners are particularly adept at identifying the alveolar nasal even when the acoustic signal is disrupted with noise. Confusion matrices indicate that among 16 English consonants (i.e. /p, t, k, f, θ, s, ʃ, b, d, g, v, ð, z, ʒ, m, n/), the bilabial nasal /m/ and alveolar nasal /n/ have the highest True Positive rates; out of four SNRs (0, -6 dB, -12 dB, -18 dB), the alveolar nasal has the highest True Positive nasal except for

[2] I use the term *lateral resonance* here to refer to the oral resonance that occurs during the production of lateral sounds.

when the SNR is -6 dB, when the bilabial nasal has the highest True Positive rate. Miller and Nicely (1955) find that the percent of information transmitted as a function of SNR is higher for nasality than that of duration or affrication, which the authors use to explain why nasal sounds are discriminable at SNRs of even -12 dB when place of articulation is lost at SNRs less than -6 dB. While Miller and Nicely (1955) do not include lateral sounds in their study, findings from the perception study included here indicate that the alveolar nasal is far more resistant to the effects of noise than the alveolar lateral approximant.

The results reported in this chapter are comparable to Miller and Nicely (1955), with place of articulation discriminable at an SNR of -3 dB but not at -6 dB and the alveolar nasal consistently demonstrating low confusion rates (i.e. high True Positive rates). Given that the previous study (Miller and Nicely, 1955) was conducted with English consonants, this suggests that nasal sounds are particularly robust against perceptual confusion cross-linguistically. If nasal sounds are inherently easier to perceive even in noisy environments, this may explain why the palatal nasal is more frequent in the world's languages in comparison to the palatal lateral, despite the two sounds differing only in terms of manner.

The perception of palatalization is affected when the SNR is raised to -9 dB. Palatalized stops are incorrectly identified as a non-palatalized stop with a confusion rate of approximately 30% for both groups, while non-palatalized stops are identified as a palatal or palatalized nasal at similar confusion rates. Similarly, palatal laterals are incorrectly identified as an alveolar lateral at approximately 30% confusion rates for both groups as well. Note that the same is not observed for the palatalized lateral, which demonstrates low confusion rates with the alveolar lateral and higher confusion rates with the palatal approximant. This difference in the perception of the palatal and palatalized lateral provides an explanation for why languages like Spanish only briefly possess a palatal lateral before the sound merges to a palatal approximant (Penny, 2000; Pharies, 2007; Zampaulo, 2015). It is likely that the historical sound typically described as a palatal lateral was actually a palatalized lateral approximant and thus more likely to be confused with a palatal approximant. The evidence here indicates that a palatalized lateral is more likely to merge with a palatal approximant than a palatal lateral, potentially contributing to the instability of the palatal lateral as a phone category in historical Spanish.

The palatal and palatalized lateral are largely similar to one another in terms of their True Positive rates and the sounds that are mistaken for them; discriminating between the two sounds is roughly chance level even in optimal auditory conditions. This finding coincides with findings from the articulatory study in Chapter 3 and acoustic study in Chapter 4, which found only minor differences between the articulation and the acoustics of the palatal and palatalized lateral. The predominant difference between the two sounds emerges only when observing the rates at which the palatal and palatalized lateral are mistaken for other sounds, particularly the palatal approximant.

A.4.2 Additional evidence for *yeísmo*

In the discussion of the acoustics results (see Section 4.4.1), it was suggested that the shorter duration of the palatal and palatalized lateral may affect the listener's ability to accurately retrieve the necessary information regarding phone identity from the acoustic signal. However, sensitivity measures indicate that while BP participants were better at identifying the palatal lateral than the palatalized lateral in the first two SNR conditions, the reverse was found for the English participants. If duration truly is the main contributing factor, similar d' scores to the palatalized lateral should be found for the palatal and palatalized lateral. Instead, the palatal and palatalized lateral are much more poorly discriminated than the palatal and palatalized nasal for the BP listeners, while d' scores are roughly the same between the palatal and palatalized nasals and laterals for the English group. In fact, during the second condition, the English listeners identify the palatalized nasal with higher sensitivity than all the other palatal and palatalized sounds, including the palatal approximant. Together, this indicates that duration is not the sole factor for the rareness of the palatal lateral or the frequent occurrence of *yeísmo*, though it may still have a contributing influence.

It was also suggested that acoustic similarity may be a possible explanation for why *yeísmo* occurs and certainly, observations of the modeled formant frequencies for the palatal approximant and both the palatal and palatalized lateral support this theory. However, it is unable to account for why the palatal lateral merges with the palatal approximant and not the reverse. The perception results here, as discussed in the previous section, indicate why this is the case: while the two sounds are acoustically similar, listeners are better at identifying the palatal approximant even when noise is added up to -6 SNR.

A.5 Future research

There are a few directions in which future research may proceed. An expanded version of this study using the acoustics recorded during the EMA experiment would provide insight with regards to how cross-speaker articulatory differences affect perception and phone identification. Additionally, the speaker chosen to produce the stimuli in this study was also the same speaker that demonstrated the least articulatory differences between the palatal and palatalized lateral in Chapter 3, a side-effect that was unintentional. Given that by-speaker analysis found that articulatory similarity between the palatal and palatalized lateral might be an effect of age, future iterations of this study may wish to include speakers sampled from a larger range of ages.

Note that some participants reported requiring a longer time to complete the study. The longer completion times were an unforeseen effect of presenting acoustic stimuli in a wav format, which overloaded the browser cache for some of the BP participants. Pilot versions of this study were conducted several times using multiple browsers, internet connections, and operating systems with no ill-effects, with the exception of Internet Explorer (which participants were asked to avoid before attempting the study). However, since the pilot studies were conducted in the United States, the effect of unstable internet connections was not fully tested. Future online studies in regions with slower or undependable internet access may wish to incorporate acoustic stimuli in an MP3 format, to avoid issues with audio streaming. Firefox is also recommended as a preferred browser; several participants reported occasional delays in acoustic streaming when using Chrome.

A.6 Conclusion

A model of the listener's role in the rarity of the palatal lateral is presented in this study. Both the BP and English group demonstrated similar levels of confusion when asked to identify the palatal lateral and palatalized lateral; the two sounds are identified at roughly chance even when no noise has been added. Additionally, the findings here support the prediction that acoustic similarity can cause the palatal lateral to be misinterpreted as a palatal approximant; d' results indicate that the direction of *yeísmo* is due to listeners identifying the palatal approximant with greater sensitivity, even when the signal is perturbed with noise. This is especially true of the palatalized lateral, which is more susceptible to such confusions at a higher rate than the palatal lateral. Cumulatively, these results suggest that the cross-linguistic rarity of the palatal lateral is due to the listener's difficulty in accurately identifying the sound when the acoustic signal is perturbed by the presence of even minimal noise.

Appendix B

MATLAB SCRIPTS

B.1 EMA data processing

The following suite of scripts (Wong and Hermes, 2015) were developed together with Zainab Hermes, a colleague from the Department of Linguistics at the University of Illinois. The scripts are also available online: https://www.drop box.com/s/02lxp0u1eiyb2yv/Data_processing_02212015.zip?dl=0.

```
%Written by Nicole W. Wong and Zainab Hermes, Fall 2014 at UIUC
   . Updated
%10/12/2014.

%Rotates the data to the occlusal plane, corrects for head
%movement (translation), and applies a butterworth filter for
%smoothing. Outputs corrected files into the same folder.

%Give the folder path and specific bite plate file name. Also
   give an
%arbitrary acquisition name to calculate the cut off frequency.
    Give the
%sampling rate. Give the index for one of the cheek reference
   sensors.

%Requires four functions:
```

```
%1. reorient_data
%2. compute_rot_matrix
%3. my_butter
%4. replace_NaNs


function [] = my_data_processing(directory_path,BP_filename,
   cut_off,SR,cheek_ref)

%Turns the warning off for the butterworth filter. The input is
    correct.
warning('off','signal:filtfilt:ParseSOS');

%Create a pop-up box that reads the following:
Message = 'You may want to set the cut off slightly higher, as
   the Butterworth filter requires the signal to be forward and
    reverse filtered, which produces a lower cut off and
   sharper roll off.' ;
waitfor(msgbox(Message,'Determine Cut Off Frequency for
   Butterworth Filter'))

%Determine cut off frequency from an arbitrary file and sensor.
%Select and save the desired cut off
cut_off = importdata(strcat(directory_path,'\',cut_off));
co_data = cut_off.data(:,80);   %TM
my_fft(co_data,SR);
[x,y] = ginput(1);
co = round(x);        %We want the cut off to be an integer
```

```
co_num = sprintf('%0.2f', co); %Force cut off to 2 decimal
   places.
co_nom = sprintf('%0.2f',co*2/SR); %Calculate norm. cut off, to
    2 dec. places.
close all

%Test the filter and determine whether it's accurate.
%If not, quit and rerun the script to reselect.
%Note the output dialog, for when you need to report
   methodology.
my_butter(co_data,SR,co,1);

choice = questdlg('Is the filter approriate?', 'Cut Off
   Frequency', 'Yes',...
     'No','Yes');
Handle response
switch choice
    case 'Yes'
        close all
        waitfor(msgbox(strcat('Selected cut off frequency = ',
           co_num,...
            '. Normalized cut off frequency = ',co_nom,'.'),'
               Selected Cut Off Frequency'));
    case 'No'
        msgbox('Please re-run the script and select another cut
            off frequency.')
        return
end
```

```
%For all the files within the folder path, determine which are
   .tsv
%Read all the files in the folder. Find which files to read and
    take.
%Save all the file names to a matrix. Strip the last three
   digits off the
%file names and store it.
a = dir(directory_path);
avector = length(a');
j = 1;

for i = 1:avector
    [p,q,r] = fileparts(a(i,1).name);

    if strcmp(r, '.tsv') == 1
        mynames{j,1} = q;
        myacq{j,1} = str2double(q(10:12)); %Changes depending
           on your file name
        j=j+1;
    end

end

%Head rotation. The function compute_rot_matrix computes the
   rotation
%matrix based on the selected BP acquisition. Then it saves the
    rotaiton
%matrix to variables R and trans, so that it can be applied to
   the other
```

```
%acquisitions.
BP_filename = strcat(directory_path,'\',BP_filename);
[R,trans] = compute_rot_matrix(BP_filename);

%Rotate and save the Biteplate data so that it can be used to
%apply translation later (my_trans). reorient_data takes the R
%and trans calculated by compute_rot_matrix above and applies
   it
%to the data (BP_filename) and then reorient data also requires
%that you specify how many sensors will be rotated.

BP_rot = reorient_data(R, BP_filename, trans, 12);
my_trans = BP_rot(100, cheek_ref) - BP_rot(100, 105:107);

%Loop through all the .tsv files in the given directory and do
   the
%following: correct for head movement, reorient to the occlusal
    plane, and
%filter the data.

for ii = 1:length(mynames')

filename = strcat(directory_path,'\',mynames{ii,1},'.tsv');

%Rotate to the occlusal plane. Apply reorient_data only to 11
%sensors, since the last BP sensor is not in range.
output_data = reorient_data(R, filename, trans, 11);

%Replace all the NaNs
```

```
output_data = replace_NaNs(output_data);

%Butterworth filtering and head movement correction. Loop
   through
%all the sensors in the reorientated data. The number of
   sensors
%may change depending on how many sensors you collected data
   from
%during your experiment. Here, there are 11 sensors, not
%including the last BP sensor (12).
c = -2;
d = 1;

%Identify the cheek reference data first, so that it doesn't
   get rewritten.
ref = output_data(:,cheek_ref);

    for iii = 1:11
        index = iii*8+c;
        indices = index:index+2;

        %Correct for head movement by subtracting the cheek
           movements from
        %each given sensor. Delta = B = data ready to be
           filtered.
        A = repmat(my_trans, length(output_data(:,1)), 1);
        delta = ref-A;
        B = output_data(:,indices)-delta;
```

```
        %Apply butterworth filter to each individual data
           series
        co = 10;
        buttfilt1 = my_butter(B(:,1),SR,co,0);
        buttfilt2 = my_butter(B(:,2),SR,co,0);
        buttfilt3 = my_butter(B(:,3),SR,co,0);

        %Move the outputs so that front-back movement is x, up-
           down
        %movement is y, and left-right movement is z. The
           titles x,y,z are
        %arbitrary designations anyways.
        buttfilt = horzcat(buttfilt2,buttfilt3,buttfilt1);
        output_data(:,indices) = buttfilt;

        c = c+1;
    end

%Save file to original folder as text files
newfile = strcat(directory_path,'\',mynames{ii,1}, '_rcf.txt');
dlmwrite(newfile,output_data,'delimiter','\t');

end

clf
close all
```

B.1.1 reorient_data.m

```
%Written by Zainab Hermes @ UIUC 2014. Edited by Nicole Wong,
   11/2014.

function output_data = reorient_data(R, filename, w, sens_num);

% filename is the bite plate
%R is output from compute_rot_matrix
%w is the trans output from compute_rot_matrix

my_data = importdata(filename);
output_data = my_data.data;

[a b] = size(output_data);
trans = zeros(a, 3);
trans(:,3) = w;

% Loop through only 11 if you don't need the BP3

sensor_indices = [6:8, 15:17, 24:26, 33:35, 42:44, 51:53,
   60:62, 69:71, 78:80, 87:89, 96:98, 105:107];
c = -2;
for i = 1:sens_num
    index = i*8+c;
    indices = index:index+2;
    A = output_data(:, indices);
    B = A*R - trans;
    output_data(:, indices) = B;
    c = c+1;
end
```

B.1.2 compute_rot_matrix.m

```
%Written by Zainab Hermes @ UIUC 2014. Edited by Nicole Wong,
   12/2014.
%Requires function replace_NaNs

function [R, trans] = compute_rot_matrix(filename)

%filename is the bite plate acquisition

bite_data = importdata(filename);
bite_data.data = replace_NaNs(bite_data.data);

% bite plate sensors are 6 & 7 & 11 i.e. xyz coordinates are in
    columns 60-62 and
% 69-71 and 105 - 107
% read the xyz coordinates for midpoint sample

[rows, ~] = size(bite_data.data);
index = round(rows/2);

p6 = bite_data.data(index, 60:62);
p7 = bite_data.data(index, 69:71);
p11 = bite_data.data(index, 105:107);

% compute bite plane equation coefficients using p6, p7, and
   origin (0,0,0)
nor = cross(p6 - p7, p6 - p11);
a = nor(1);
b = nor(2);
```

```
c = nor(3);
d = - dot(nor, p6);

% compute angle between bite plate plane and xy plane (
   horizontal plane)
theta = acosd( abs(nor(3)) / sqrt( nor(1)^2 + nor(2)^2 + nor(3)
   ^2));

% compute unit vector of intersection between two planes
inter = cross(nor, [0 0 1]);
umag = sqrt(inter(1)^2+inter(2)^2+inter(3)^2);
u = inter/umag;
ux = u(1);
uy = u(2);
uz = u(3);

% now compute rotation matrix
R =  [cosd(theta)+ux^2*(1-cosd(theta)) ux*uy*(1-cosd(theta))-uz
   *sind(theta) ux*uz*(1-cosd(theta))+uy*sind(theta); ...
      uy*ux*(1-cosd(theta))+uz*sind(theta) cosd(theta)+uy^2*(1-
         cosd(theta)) uy*uz*(1-cosd(theta))-ux*sind(theta); ...
      uz*ux*(1-cosd(theta))-uy*sind(theta) uz*uy*(1-cosd(theta)
         )+ux*sind(theta) cosd(theta)+uz^2*(1-cosd(theta))];


w = p11*R;
trans = w(3);
```

B.1.3 my_butter.m

```
%Written by Nicole W. Wong @ UIUC Fall 2014
%Prior to this script, plot the Fast Fourier Transform of the
   data
%and determine where the cutoff should be.

%Input the cuttoff to apply a 2nd order Butterworth low-pass
   filter. The
%function will normalize the cut-off frequency and plot the
   filtered data
%superimposed over the unfiltered data. It will also plot the
   FFT of the
%filtered and unfiltered data, to verify that only the desired
   frequencies
%were affected.

%Also give the function the sampling rate of the data. There is
    an option
%(plotopt) to plot the filtered signal or not. 1=yes, 0=no.

function buttfilt = my_butter(data,fs,cutoff,plotopt)

%Determine where the cutoff should be. Remember that you may
   want
%to set the cut off slightly higher, as the Butterworth filter
%requires the signal to be forward and reverse filtered (using
%filtfilt). The first application of the filter cause a phase
%shift (to the right), so the second application is required to
%reverse it. The double application of the filter creates a
   lower
```

```
%cut-off and sharper roll off.

% Cutoff frequency is that frequency where the magnitude
   response of the
% filter is the square root of 1/2. For butter, the normalized
   cutoff
%frequency Wn must be a number between 0 and 1, where 1
   corresponds to the
%Nyquist frequency, pi radians per sample. (from the Matlab
   help page)

%To calculate the normalized cutoff frequency: Cutoff Freq * 2
   / SR

%Creates a second order low-pass filter that passes all
   frequencies
%below 20 Hz
[z,p,k] = butter(2,(cutoff*2/fs),'low');
[sos,g] = zp2sos(z,p,k);

%Applies the filter to the data
buttfilt=filtfilt(sos,g,data);
warning('off','signal:filtfilt:ParseSOS');

%Option of plotting. If plotopt==0, does not plot. If plotopt
   ==1, plot.
if plotopt==1

%Plot the fourier transform of the filtered data.
```

```
my_fft(buttfilt,fs);

%Plots the filtered data against the unfiltered data.
figure('units','normalized','outerposition',[0 0 1 1]);
subplot(2,1,1)
plot(buttfilt,'r')
subplot(2,1,2)
plot(data, 'b');
end
```

B.1.4 replace_NaNs.m

```
%By Zainab Hermes @ UIUC 2014

function no_NaNs = replace_NaNs(inputdata)

c = -2;
d = 1;

no_NaNs = inputdata;

for j = 1:11
    index = j*8+c;
    indices = index:index+2;
    A = inputdata(:, indices);
    a1 = A(:,1);
    a2 = A(:,2);
    a3 = A(:,3);

    not_a_nan_a1 = 0;
```

```
not_a_nan_a2 = 0;
not_a_nan_a3 = 0;

b4_1 = a1;
b4_2 = a2;
b4_3 = a3;

for i = 1:length(a1);
    if( isnan( a1(i) ) )
        a1(i) = not_a_nan_a1;
    else
        not_a_nan_a1 = a1(i);
    end
end
% -------


for i = 1:length(a2);
    if( isnan( a2(i) ) )
        a2(i) = not_a_nan_a2;
    else
        not_a_nan_a2 = a2(i);
    end
end
% -------


for i = 1:length(a3);
    if( isnan( a3(i) ) )
```

```
            a3(i) = not_a_nan_a3;
        else
            not_a_nan_a3 = a3(i);
        end
    end
    % ----------------------



    B = horzcat(a1, a2, a3);
    no_NaNs(:, indices) = B;
    c = c+1;

end
```

B.2 Tongue blade angle

```
%Written by Nicole W. Wong @ UIUC 2015

%Imports an entire struct and finds the maximum degree of
   apical
%laminality for all target consonants in the struct.

%requires:
%(1) my_apic_lam.m
%(2) my_tongue_contour.m

%It outputs the angle between the line formed by the TT and TM
   and the
%horizontal plane at the point when the tongue is the most
%forward. It also outputs the point at which the tongue is the
```

```
%most forward (point = in seconds, timepoint = in samples].

function [angleTTxTM] = my_apic_lam_all(bounds_name,WAVE_struct
   ,plotopt)
my_bounds = importdata(bounds_name);
bounds = my_bounds.data;
text = my_bounds.textdata;
text = text(2:length(text),:);

%For every line in bounds,
for i = 1:length(bounds)

    %find the matching acquisition in WAVE_struct
    for ii = 1:length(WAVE_struct.sweeps)
        if strncmp(text{i,2},WAVE_struct.sweeps{ii,1}.filename
           ,11) == 1

            if isempty(text{i,4}) == 1
                continue
            else

            %Identify the variables TT, TM, and time as neeeded
                for my_apic_lam.
            TT = WAVE_struct.sweeps{ii,1}.Data{1,11}.SIGNAL;
            TM = WAVE_struct.sweeps{ii,1}.Data{1,10}.SIGNAL;
            time = WAVE_struct.sweeps{ii,1}.Time;

            %Calculate the angle for the sensors identified
               here at the
```

```
%point in time for each line in bounds.
[angle1,point1,timepoint1]=my_apic_lam(TT, TM,
   bounds(i,:),time);

%Put the following values into a matrix: 1) The
   stimulus, 2)
%The target, 3) The angle, 4) The EMA sample when
   the angle
%was calculated, 5) The time at when the angle was
   calculated.
angleTTxTM{i,1} = WAVE_struct.sweeps{ii,1}.Stimulus
   ;
angleTTxTM{i,2} = text{i,4};
angleTTxTM{i,3} = angle1;
angleTTxTM{i,4} = point1;
angleTTxTM{i,5} = timepoint1;

if plotopt == 1
    %For that same point at which the tongue tip
       angle
    %was calculated, plot the tongue contour and
       label it.
    %First identify the additional inputs necessary
    TD = WAVE_struct.sweeps{ii,1}.Data{1,7}.SIGNAL;
    UL = WAVE_struct.sweeps{ii,1}.Data{1,5}.SIGNAL;
    LL = WAVE_struct.sweeps{ii,1}.Data{1,6}.SIGNAL;
    NOSE = WAVE_struct.sweeps{ii,1}.Data{1,3}.
       SIGNAL;
    %Make the name of the figure
```

```
                    name = strcat(WAVE_struct.sweeps{ii,1}.Stimulus
                       ,'-',text{i,4});
                    %Now plot
                    my_tongue_contour(TT,TM,TD,UL,LL,NOSE,
                       timepoint1, name)
                end
                end
            end
        end
        end
end
```

B.2.1 my_apic_lam.m

```
%Written by Nicole W. Wong @ UIUC 2015


%This function finds the angle in degrees between two lines: 1)
    one line
%that connects the Tongue Tip sensor and the Tongue Mid-Body
   Sensor and 2)
%a horizontal line that runs through the Tongue Mid-Body Sensor
   . Finding
%the angle is a means of determining whether a gesture is
   apical or
%laminal, under the assumption that if the tongue tip is
   pointed upwards,
%the angle between the two lines should be positive. Likewise,
   if the
%gesture is laminal, the angle between the two lines should be
   negative.
```

```
%It takes as input the Tongue Tip data, Tongue Medial data, and
    bounds.
%It also needs the time for each sample point.

%It outputs the angle between the line formed by the TT and TM
   and the
%horizontal plane at the point when the tongue is the most
   forward. It
%also outputs the point at which the tongue is the most forward
    (point =
%in seconds, timepoint = in samples].

function [degree,point,timepoint] = my_apic_lam(TT,TM, bounds,
   time)

%Find the sampling rate
timevector = length(time);
SR = timevector/time(timevector,1);

%Convert the bounds (in seconds) to EMA samples.
L = round(bounds(1,1)*SR);
R = round(bounds(1,2)*SR);

%Extract the TT and TM data contained only for this time
   segment.
TT = TT(L:R,:);
TM = TM(L:R,:);
```

```
%Find apex of gesture based on the max x-value for the TT (i.e.
    most forward movement of the tongue tip).
[r]=find(TT(:,1)==min(TT(:,1)));
[r]=r(1,1);

%Convert [r] back to correct sample point
timepoint = L+(r-1);
point = time(timepoint,1);

%Find x and y for TT (x1,y1) and TM (x2,y2) at the apex of the
   gesture
x1 = TT(r,1);
y1 = TT(r,2);

x2 = TM(r,1);
y2 = TM(r,2);

%Calculate the angle in degrees
degree = atand((y1-y2)/(x2-x1));
```

B.2.2 my_tongue_contour.m

```
%Written by Nicole W. Wong @ UIUC 2015

%Plots the tongue contour by drawing a line that connects all
%three points. Takes as input the tongue tip data (TT), tongue
%medial data (TM), and the tongue dorsum data (TD). 'r'
%represents the specific EMA sample that will be plotted.
```

```
function [] = my_tongue_contour(TT,TM,TD,UL,LL,NOSE,timepoint,
   name)

% Plot
figure('name',name)
hold on

% Set the aspect ratio
set(gca,'DataAspectRatio',[1 1 1])
axis square

% Plot the tongue contour in green, lips in blue, and nose in
   black for
% reference.

plot(TT(timepoint,1), TT(timepoint,2),'.g','MarkerSize',5)
plot(TM(timepoint,1), TM(timepoint,2),'.g','MarkerSize',10)
plot(TD(timepoint,1), TD(timepoint,2),'.g','MarkerSize',15)
plot(UL(timepoint,1), UL(timepoint,2),'.b','MarkerSize',15)
plot(LL(timepoint,1), LL(timepoint,2),'.b','MarkerSize',15)
plot(NOSE(timepoint,1), NOSE(timepoint,2),'.k','MarkerSize',15)
hold off
```

B.3 Re-sampling tongue sensor trajectories

```
%Written by Nicole W. Wong @ UIUC 2015

%Script to resample sensor trajectories to an equal number of
%samples per phone.
```

```
%Preallocate for speed
% max_pt = zeros(length(bounds),8);
traject_x = [1.1 2.1 3.1 4.1 5.1 6.1 7.1 8.1 9.1 1.2 2.2 3.2
   4.2 5.2 6.2 7.2 8.2 9.2 1.3 2.3 3.3 4.3 5.3 6.3 7.3 8.3 9.3
   ];
traject_y = [1.1 2.1 3.1 4.1 5.1 6.1 7.1 8.1 9.1 1.2 2.2 3.2
   4.2 5.2 6.2 7.2 8.2 9.2 1.3 2.3 3.3 4.3 5.3 6.3 7.3 8.3 9.3
   ];

%Loop through all the files
for i = 1:length(bounds)
    %Identify the corresponding file.
    filename = strcat(directory_path,'\',strtrim(bound_labels(i
       ,2)),'_rcf.txt');
    file = importdata(filename{1,1});

    %Find the sampling rate
    time = file(:,1);
    timevector = length(time);
    SR = timevector/time(timevector,1);

    %Convert the bounds (in seconds) to EMA samples.
    L = round(bounds(i,1)*SR)-2; %Take one extra sample to the
       L and R to help pad.
    R = round(bounds(i,2)*SR)+2;
    %Identify and save the TT, TM, TD data. Extract the data
       contained only for this time segment.
    TT = file(L:R,87:88);
    TM = file(L:R,78:79);
```

```
    TD = file(L:R,51:52);

    %Take the x and y vector for TT-TD. Resample them and save
       the values after removing the padded samples.
    TT_resamps = resample(TT,13,length(TT));
    TTx_resamp = TT_resamps(3:length(TT_resamps)-2,1);
    TTy_resamp = TT_resamps(3:length(TT_resamps)-2,2);
    TM_resamps = resample(TM,13,length(TM));
    TMx_resamp = TM_resamps(3:length(TM_resamps)-2,1);
    TMy_resamp = TM_resamps(3:length(TM_resamps)-2,2);
    TD_resamps = resample(TD,13,length(TD));
    TDx_resamp = TD_resamps(3:length(TD_resamps)-2,1);
    TDy_resamp = TD_resamps(3:length(TD_resamps)-2,2);
    resampsx = [TTx_resamp',TMx_resamp',TDx_resamp'];
    resampsy = [TTy_resamp',TMy_resamp',TDy_resamp'];


    %Put all the resampled trajectories into one place.
    traject_x = vertcat(traject_x,resampsx);
    traject_y = vertcat(traject_y,resampsy);


    %Save the labels in a separate array.
    labels{i,1} = bound_labels(i,2);
    labels{i,2} = bound_labels(i,4);
    labels{i,3} = filename{1,1}(46:47);
end
```

Appendix C

PRAAT SCRIPTS

C.1 Formant frequency

The following script by Kawahara (2010) is the original script used to obtain formant frequencies in this dissertation. Some modifications were made (not shown here) to obtain formant frequencies at other sample points besides the midpoint. The script was accessed in September 2011 and is available online: http://user.keio.ac.jp/ kawahara/scripts/get_formants_midpoint.praat.

```
# This Praat script will get F1, F2, and F3 at the midpoints of
    all the intervals of the all files in the specified folder.
# Version: 3 Feb 2010
# Author: Shigeto Kawahara
# Input: TextGrid and wav in the same directly. They must have
   the same name.

form Get F1, F2, F3
        sentence Directory ./
        comment If you want to analyze all the files, leave
           this blank
        word Base_file_name
        comment The name of result file
        text textfile result.txt
endform
```

```
# Write-out the header

fileappend "'textfile$'" soundname'tab$'intervalname'tab$'F1'
   tab$'F2'tab$'F3'tab$'
fileappend "'textfile$'" 'newline$'

#Read all files in a folder
Create Strings as file list... wavlist 'directory$'/'
   base_file_name$'*.wav
Create Strings as file list... gridlist 'directory$'/'
   base_file_name$'*.TextGrid
n = Get number of strings

for i to n
clearinfo

#We first extract a formant tier
        select Strings wavlist
        filename$ = Get string... i
        Read from file... 'directory$'/'filename$'
        soundname$ = selected$ ("Sound")
        To Formant (burg)... 0 5 5500 0.025 50

# We now read grid files and extract all intervals in them
        select Strings gridlist
        gridname$ = Get string... i
        Read from file... 'directory$'/'gridname$'
        int=Get number of intervals... 1
```

```
# We then calculate F1, F2 and F3

for k from 1 to 'int'
        select TextGrid 'soundname$'
        label$ = Get label of interval... 1 'k'
        if label$ <> ""

                # calculates the mid point
                vowel_onset = Get starting point... 1 'k'
                vowel_offset = Get end point... 1 'k'
                midpoint = vowel_onset + ((vowel_offset -
                   vowel_onset) / 2)

                # get the formant values at the midpoint
                select Formant 'soundname$'
                f_one = Get value at time... 1 'midpoint' Hertz
                    Linear
                f_two = Get value at time... 2 'midpoint' Hertz
                    Linear
                f_three = Get value at time... 3 'midpoint'
                   Hertz Linear
                mid = 'midpoint'
                resultline$ = "'soundname$''tab$''label$''tab$
                   ''f_one''tab$''f_two''tab$''f_three''tab$''
                   midpoint''tab$'"
                fileappend "'textfile$'" 'resultline$' '
                   newline$'
        endif
```

```
endfor

fileappend "'textfile$'"


endfor

# clean up

select all
Remove
```

Appendix D

R SCRIPTS

D.1 SSANOVA

The following script by Mielke (2013) was used to create the SSANOVAs used in this dissertation. The script was accessed in November 2014 and is available online at the following location: http://phon.chass.ncsu.edu/manual/tongue _ssanova.r.

```
################################################
# tongue_ssanova.r                                revised
   October 22, 2013
# Jeff Mielke
# functions for SSANOVA comparisons of tongue traces in polar
   coordinates using gss
################################################
#
# BASIC COMMAND TO GENERATE AN SSANOVA PLOT (IF 'phone' IS THE
   NAME OF YOUR FACTOR)
#  ss <- polar.ssanova(data, 'phone')
#
# BASIC COMMAND TO PLOT THE RAW DATA
#  show.traces(data)
#
# TO PLOT TO FILE, SEPARATING BY TWO DIFFERENT FACTORS (COLUMNS
    IN YOUR DATA FRAME):
```

```
#  cairo_pdf('my_ssanova_pdf.pdf', h=4.5, w=5, onefile=T)
#    ss.by.C <- polar.ssanova(data, 'consonant')
#    ss.by.V <- polar.ssanova(data, 'vowel')
#  dev.off()
#
# TO HIGHLIGHT RAW DATA FOR THE LEVEL ('I'):
#  show.traces(data, c('I'))
#
# DATA FILE SHOULD BE ORGANIZED LIKE THIS (MULTIPLE COLUMNS CAN
    BE USED INSTEAD OF word):
#
# word,token,X,Y
# dog,1,307,262
# dog,1,311,249
# dog,1,315,240
# dog,2,308,261
# dog,2,311,250
# dog,2,314,249
# cat,1,307,240
# dog,2,311,250
# dog,2,314,259
# ...
#
################################################
#
# polar.ssanova() ARGUMENTS (ALL OPTIONAL EXCEPT data):
#
#          data: your tongue tracings (minimally including
   columns X and Y and a
```

```
#                 column with a factor)
#        data.cat: the factor to use to categorize the data (
   defaults to 'word')
#           scale: how much to scale the axis values (e.g. to
   convert from pixels to
#                 centimeters)
#   origin.method: how to choose the origin for calculating
   polar coordinates
#           debug: whether to generate the cartesian and non-
   transformed polar plots too
#        plotting: whether to plot anything (or just return the
   result of the test)
#            main: the main title for the plot
#         CI.fill: whether to indicate confidence intervals with
    shading (like ggplot)
#                 or with dotted lines (like the earlier
   SSANOVA code).
#                 Defaults to FALSE (dotted lines)
#        printing: if TRUE, different splines use different line
    types, so that the
#                 figure can be printed in black and white.
#            flip: whether to flip the Y values (useful for
   plotting data from images
#                 in cartesian coordinates, but ignored if
   using polar coordinates)
# cartesian.only: used by cart.ssanova()
#        is.polar: if TRUE, the data is already in polar
   coordinates
#
```

```
################################################
#
#  cart.ssanova() SAME AS polar.ssanova() BUT DOESN'T USE POLAR
   COORDINATES
#
################################################
#
#   show.traces() ARGUMENTS (ALL OPTIONAL EXCEPT data):
#
#           data: your tongue tracings (minimally including
   columns X and Y and a
#                 column with a factor)
#       data.cat: the factor to use to categorize the tongues (
   defaults to 'word')
#   to.highlight: a list of factor levels to plot while muting
   the other levels
#        to.plot: a list of factor levels to plot, excluding
   the rest (defaults to all)
#    token.label: the factor to use to identify individual
   tokens (defaults to 'token')
#           flip: whether to flip the Y values (useful for
   plotting data from images)
#           main: the main title for the plot
#       overplot: whether to add the traces to an existing plot
#       is.polar: if TRUE, the data is already in polar
   coordinates
#         origin: used if the data is in polar coordinates
   already
#
```

```
################################################

library(gss)

#CONVERT POLAR COORDINATES TO CARTESIAN COORDINATES
make.cartesian <- function(tr, origin=c(0,0)){
    X <- apply(tr, 1, function(x,y) origin[1]-x[2]*cos(x[1]))
    Y <- apply(tr, 1, function(x,y) x[2]*sin(x[1])-origin[2])
    xy <- cbind(X, Y)
    return(xy)
}

#CONVERT CARTESIAN COORDINATES TO POLAR COORDINATES
make.polar <- function(data.xy, origin=c(0,0)){
    xy <- cbind(data.xy$X, data.xy$Y)
    all_r <- apply(xy, 1, function(x) sqrt((x[1]-origin[1])^2 +
        (x[2]-origin[2])^2))
    all_theta <- pi+apply(xy, 1, function(x,y) atan2(x[2]-
       origin[2], x[1]-origin[1]))
    data.tr <- data.xy
    data.tr$X <- all_theta
    data.tr$Y <- all_r
    return(data.tr)
}

#RESCALE DATA FROM PIXELS TO CENTIMETERS
us.rescale<-function(data, usscale, X='X', Y='Y'){
    data[,c(X)] <- data[,c(X)]*usscale
    data[,c(Y)] <- data[,c(Y)]*usscale
```

```
    data
}

#SELECT AN APPROPRIATE ORIGIN FOR THE DATA
select.origin <- function(Xs, Ys, method='xmean_ymin'){
    if (method=='xmean_ymin'){
        if (mean(Ys)>0){
            return(c(mean(Xs), max(Ys)*1.01))
        }else{
            return(c(mean(Xs), min(Ys)*1.01))
        }
    }
    if (method=='xmean_ymean'){
        return(c(mean(Xs), mean(Ys)))
    }
    return(c(mean(Xs), max(Ys)*1.01))
}

#PERFORM THE SSANOVA AND RETURN THE RESULTING SPLINES AND
   CONFIDENCE INTERVALS
#expand.grid + predict scheme based on http://www.ling.upenn.
   edu/~joseff/papers/fruehwald_ssanova.pdf
tongue.ss <- function(data, data.cat='word', flip=FALSE, length
   .out=1000, alpha=1.4){
    if (flip==TRUE){
        data$Y <- -data$Y
    }
    data$tempword <- data[,data.cat]
    #print(summary(lm(Y ~ tempword * X, data=data)))
```

```
    ss.model <- ssanova(Y ~ tempword + X + tempword:X, data=
       data, alpha=alpha)
    ss.result <- expand.grid(X=seq(min(data$X), max(data$X),
       length.out=length.out), tempword=levels(data$tempword))
    ss.result$ss.Fit <- predict(ss.model, newdata=ss.result, se
       =T)$fit
    ss.result$ss.cart.SE  <- predict(ss.model, newdata=ss.
       result, se=T)$se.fit
    #print(names(ss.result))
    #print(aggregate(ss.Fit ~ tempword, FUN=mean, data=ss.
       result))
    #print(aggregate(ss.cart.SE ~ tempword, FUN=mean, data=ss.
       result))
    ss.result$ss.upper.CI.X <- ss.result$X
    ss.result$ss.upper.CI.Y <- ss.result$ss.Fit + 1.96*ss.
       result$ss.cart.SE
    ss.result$ss.lower.CI.X <- ss.result$X
    ss.result$ss.lower.CI.Y <- ss.result$ss.Fit - 1.96*ss.
       result$ss.cart.SE
    names(ss.result)[which(names(ss.result)=='tempword')] <-
       data.cat
    ss.result
}

#PLOT THE SSANOVA RESULTS
plot.tongue.ss <- function(ss.result, data.cat, lwd=3, main='',
    CI.fill=FALSE, printing=FALSE, show.legend=T, plot.labels=c
   (main,'X','Y'),
```

```
                        overplot=FALSE, xlim=NULL, ylim=NULL
                           ){
n_categories <- length(levels(ss.result[,data.cat]))
Fit.palette <- rainbow(n_categories, v=0.75)
CI.palette <- rainbow(n_categories, alpha=0.25, v=0.75)

xrange = range(c(ss.result$X, ss.result$ss.lower.CI.X, ss.
   result$ss.upper.CI.X))
yrange = range(c(ss.result$ss.Fit, ss.result$ss.lower.CI.Y,
    ss.result$ss.upper.CI.Y))

if (is.null(xlim)){
    xlim <- xrange
}
if (is.null(ylim)){
    ylim <- yrange
}
if (!overplot){
    plot(0, 0, xlim=xlim, ylim=ylim,xlab=plot.labels[2],
       ylab=plot.labels[3], main=plot.labels[1], type='n')
}

if (printing){
    for (i in 1:n_categories){
        w=levels(ss.result[,data.cat])[i]
        subdata <- ss.result[ss.result[,data.cat]==w,]
        #if (CI.fill==TRUE){
            polygon(c(subdata$ss.upper.CI.X, rev(subdata$ss
               .lower.CI.X)),
```

```
                    c(subdata$ss.upper.CI.Y, rev(subdata$ss
                       .lower.CI.Y)),
                    col=CI.palette[i], border=F)
            #}else{
            #lines(subdata$ss.upper.CI.X, subdata$ss.upper.
               CI.Y, type='l', col=Fit.palette[i], lty=3)
            #lines(subdata$ss.lower.CI.X, subdata$ss.lower.
               CI.Y, type='l', col=Fit.palette[i], lty=3)
            #}
        lines(subdata$X, subdata$ss.Fit, type='l', col=Fit.
           palette[i], lwd=lwd, lty=i)
        }
    if (show.legend){
        #legend(xrange[1]+0.8*diff(xrange), yrange[1]+0.3*
           diff(yrange), c(levels(ss.result[,data.cat])),
           lwd=lwd, col=Fit.palette, lty=1:n_categories)
        legend(xlim[1]+0.8*diff(ylim), ylim[1]+0.3*diff(
           ylim), c(levels(ss.result[,data.cat])), lwd=lwd,
            col=Fit.palette, lty=1:n_categories)
    }
}else{
    for (i in 1:n_categories){
        w=levels(ss.result[,data.cat])[i]
        subdata <- ss.result[ss.result[,data.cat]==w,]
        if (CI.fill==TRUE){
            polygon(c(subdata$ss.upper.CI.X, rev(subdata$ss
               .lower.CI.X)),
                    c(subdata$ss.upper.CI.Y, rev(subdata$ss
                       .lower.CI.Y)),
```

```
                    col=CI.palette[i], border=F)
                }else{
                lines(subdata$ss.upper.CI.X, subdata$ss.upper.
                  CI.Y, type='l', col=Fit.palette[i], lty=3)
                lines(subdata$ss.lower.CI.X, subdata$ss.lower.
                  CI.Y, type='l', col=Fit.palette[i], lty=3)
                }
            lines(subdata$X, subdata$ss.Fit, type='l', col=Fit.
              palette[i], lwd=lwd)
            }
        if (show.legend){
            legend('bottomright', c(levels(ss.result[,data.cat
              ])), lwd=lwd, col=Fit.palette)
        }
    }
}

guess.data.cat <- function(data, data.cat){

}

#PLOT THE ORIGINAL DATA
show.traces <- function(data, data.cat='word', to.highlight=c
  (''), to.plot=c(''), token.label='token', flip=TRUE, main
  ='', overplot=FALSE, is.polar=FALSE, origin=c(0,0)){
    if (sum(!names(data)%in%c('token','X','Y'))==1 & !data.cat%
      in%names(data)){
        data.cat <- names(data)[!names(data)%in%c('token','X','
          Y')]
```

```
    warning(paste('Using column \"',data.cat,'" to group
       the data.\nTo avoid this warning, use "show.traces(
       data, \'',data.cat,'\')"',sep=''))
}
#print(data.cat)
show.cat <- function(data, data.cat, w, col){
    subdata <- data[data[,data.cat]==w,]
    subdata[,token.label] <- factor(subdata[,token.label])
    tokens <- levels(subdata[,token.label])
    for (t in tokens){
        token <- subdata[subdata[,token.label]==t,]
        lines(token$X,token$Y,col=col)
    }
}
if (flip){
    data$Y <- -data$Y
}
if (is.polar){
    data[,c('X','Y')] <- make.cartesian(data[,c('X','Y')],
       origin=origin)
}
categories <- levels(data[,data.cat])
n_categories <- length(categories)
trace.palette <- rainbow(n_categories, v=0.7)
ghost.palette <- rainbow(n_categories, v=0.7, alpha=0.1)
if (overplot==FALSE){
    plot(0,0,xlim=range(data$X), ylim=range(data$Y),xlab='X
       ',ylab='Y', main=main)
}
```

```
    for (i in 1:n_categories){
        w=levels(data[,data.cat])[i]
        if (w%in%to.plot >= mean(categories%in%to.plot)){
            if (w%in%to.highlight >= mean(categories%in%to.
               highlight)){
                show.cat(data, data.cat, w, col=trace.palette[i
                   ])
            }else{
                show.cat(data, data.cat, w, col=ghost.palette[i
                   ])
            }
        }
    }
    legend('bottomright', categories, lwd=1, col=trace.palette)
}

#CALCULATE AN SSANOVA IN POLAR COORDINATES AND THEN PLOT IT
   BACK IN CARTESIAN COORDINATES
polar.ssanova <- function(data, data.cat='word', scale=1,
   origin.method='xmean_ymin', debug=FALSE, plotting=TRUE, main
   ='',
                          CI.fill=FALSE, printing=FALSE, flip=
                             TRUE, cartesian.only=FALSE, is.
                             polar=FALSE, show.legend=TRUE,
                          plot.labels=c(main,'X','Y'), overplot
                             =FALSE, xlim=NULL, ylim=NULL, lwd
                             =3, alpha=1.4){
    if (sum(!names(data)%in%c('token','X','Y'))==1 & !data.cat%
       in%names(data)){
```

```
    data.cat <- names(data)[!names(data)%in%c('token','X','
       Y')]
    warning(paste('Using column \"',data.cat,'" to group
       the data.\nTo avoid this warning, use "polar.ssanova
       (data, \'',data.cat,'\')"',sep=''))
}
#if (flip==TRUE){
#    data$Y <- -data$Y
#}
data.scaled <- us.rescale(data, scale)
if (cartesian.only){
    ss.pol.cart <- tongue.ss(data.scaled, data.cat=data.cat
       , flip=flip, alpha=alpha)
    ss.cart <- ss.pol.cart
    ss.polar <- ss.pol.cart
}else{
    if (is.polar){
        #origin <- select.origin(data.scaled$X, data.
           scaled$Y, method=origin.method)
        origin <- c(0,0)
        print (origin)
        data.polar <- data.scaled
    }else{
        origin <- select.origin(data.scaled$X, data.
           scaled$Y, method=origin.method)
        print(paste('origin is',paste(origin)))
        print(summary(data.scaled$Y))
        data.polar <- make.polar(data.scaled, origin)
    }
```

```
    ss.polar <- tongue.ss(data.polar, data.cat=data.cat,
       alpha=alpha)
    ss.pol.cart <- ss.polar
    ss.pol.cart[,c('X','ss.Fit')] <- make.cartesian(ss.
       polar[,c('X','ss.Fit')], origin=origin)
    ss.pol.cart[,c('ss.cart.SE')] <- NA
    ss.pol.cart[,c('ss.upper.CI.X','ss.upper.CI.Y')] <-
       make.cartesian(ss.polar[,c('ss.upper.CI.X','ss.upper
       .CI.Y')], origin=origin)
    ss.pol.cart[,c('ss.lower.CI.X','ss.lower.CI.Y')] <-
       make.cartesian(ss.polar[,c('ss.lower.CI.X','ss.lower
       .CI.Y')], origin=origin)
}
if (plotting){
    if (debug){
        ss.cart <- tongue.ss(data.scaled, data.cat=data.cat
           , flip=T)
        plot.tongue.ss(ss.cart, data.cat, main=main, CI.
           fill=CI.fill, printing=printing, show.legend=
           show.legend, plot.labels=plot.labels, overplot=
           overplot, xlim=xlim, ylim=ylim, lwd=lwd)
        plot.tongue.ss(ss.polar, data.cat, main=main, CI.
           fill=CI.fill, printing=printing, show.legend=
           show.legend, plot.labels=plot.labels, overplot=
           overplot, xlim=xlim, ylim=ylim, lwd=lwd)
    }
    plot.tongue.ss(ss.pol.cart, data.cat, main=main, CI.
       fill=CI.fill, printing=printing, show.legend=show.
       legend, plot.labels=plot.labels, overplot=overplot,
```

```
            xlim=xlim, ylim=ylim, lwd=lwd)
    }
    return(ss.pol.cart)
}



#CALCULATE AN SSANOVA IN CARTESIAN COORDINATES (NOT ADVISED FOR
    ULTRASOUND DATA)
cart.ssanova <- function(data, data.cat='word', scale=1, origin
   .method='xmean_ymin', debug=FALSE, plotting=TRUE, main='',
                         CI.fill=FALSE, printing=FALSE, flip=
                            TRUE, show.legend=TRUE, plot.labels
                            =c(main,'X','Y'), overplot=FALSE,
                            xlim=NULL,
                         ylim=NULL, lwd=3, alpha=1.4){
    polar.ssanova(data=data, data.cat=data.cat, scale=scale,
       origin.method=origin.method, debug=debug, plotting=
       plotting, main=main,
                  CI.fill=CI.fill, printing=printing, flip=flip
                     , cartesian.only=TRUE, show.legend=show.
                     legend, plot.labels=plot.labels,
                  overplot=overplot, xlim=xlim, ylim=ylim, lwd=
                     lwd, alpha=alpha)
}
```

D.2 Dynamic model comparisons

Models of the tongue were created using R code by Stevenson (2013). The code was accessed in November 2015 and is available online: http://statistical-research.com/thats-smooth/?utm_source=rss&utm_ medium=rss&utm_campaign=thats-smooth.

```
library(graphics)
library(splines) # Used for the ns() function -- (natural cubic
    splines)
R = matrix(cbind(1,.99, .99,1),nrow=2)
U = t(chol(R))
nvars = dim(U)[1]
numobs = 1000
set.seed(1)
random.normal = matrix(rnorm(nvars*numobs,10,1), nrow=nvars,
   ncol=numobs);
X = U %*% random.normal
newX = t(X)
raw = as.data.frame(newX)
orig.raw = as.data.frame(t(random.normal))
names(raw) = c("response","predictor1")
raw$predictor1.3 = raw$predictor1^3
raw$predictor1.2 = raw$predictor1^2
fit = lm(raw$response ~ raw$predictor1.3)
plot(raw$response ~ raw$predictor1.3, pch=16, cex=.4, xlab="
   Predictor", ylab="Response", main="Simulated Data with
   Slight Curve")
abline(fit)
x = with(cars, speed)
y = with(cars, dist)
eval.length = 50
```

```
# This LOESS shows two different R function arriving at the
   same solution.
# Careful using the LOESS defaults as they differ and will
   produce different solutions.
fit.loess = loess.smooth(x, y, evaluation = eval.length,
                         29.
                         family="gaussian", span=.75, degree=1)
fit.loess2= loess(y ~ x, family="gaussian",
                  span=.75, degree=1)
## Set a simple 95% CI on the fit.loess model
new.x = seq(min(x),max(x), length.out=eval.length)
ci = cbind(
  predict(fit.loess2, data.frame(x=new.x)),
  predict(fit.loess2, data.frame(x=new.x))+
  predict(fit.loess2, data.frame(x=new.x), se=TRUE)$se.fit*
     qnorm(1-.05/2),
  predict(fit.loess2, data.frame(x=new.x))-
  predict(fit.loess2, data.frame(x=new.x), se=TRUE)$se.fit*
     qnorm(1-.05/2)
)
## Linear Model
fit = lm(y ~ x )
## Polynomial
fit.3 = lm(y ~ poly(x,3) )
## Natural Spline
fit.ns.3 = lm(y ~ ns(x, 3) )
## Smoothing Spline
fit.sp = smooth.spline(y ~ x, nknots=15)
```

```
plot(x,y, xlim=c(min(x),max(x)), ylim=c(min(y),max(y)), pch=16,
    cex=.5,
     ylab = "Stopping Distance (feet)", xlab= "Speed (MPH)",
        main="Comparison of Models"
     , sub="Splines")
## Add additional models on top of graph. It can get cluttered
   with all the models.
## LOESS with Confidence Intervals
matplot(new.x, ci, lty = c(1,2,2), col=c(1,2,2), type = "l",
   add=T)
## Linear
lines(new.x, predict(fit, data.frame(x=new.x)), col='orange',
   lty=3)
## Polynomial
lines(new.x, predict(fit.3, data.frame(x=new.x)), col='light
   blue', lty=4)
## Natural Spline
lines(new.x, predict(fit.ns.3, data.frame(x=new.x)), col='green
   ', lty=5)
## Smoothing Spline
lines(fit.sp, col='blue', lty=6)
## Kernel Curve
lines(ksmooth(x, y, "normal", bandwidth = 5), col = 'purple',
   lty=7)
legend("topleft",c("Linear","Polynomial","Natural Spline","
   Smoothing Spline","Kernel"),
       col=c('black','light blue','green','blue','purple'), lty
          =c(3,4,5,6,7), lwd=2)
```

D.3 Confusion matrices

The following R script by Agrawal (2011) was modified for our purposes to create the confusion matrices used in this dissertation. The script was accessed in August 2015 and is available online: https://ragrawal. wordpress.com/2011/05/16/visualizing-confusion-matrix-in-r/.

```
#generate random data
data = data.frame(sample(LETTERS[0:20], 100, replace=T),sample(
   LETTERS[0:20], 100, replace=T))
names(data) = c("Actual", "Predicted")

#compute frequency of actual categories
actual = as.data.frame(table(data$Actual))
names(actual) = c("Actual","ActualFreq")

#build confusion matrix
confusion = as.data.frame(table(data$Actual, data$Predicted))
names(confusion) = c("Actual","Predicted","Freq")

#calculate percentage of test cases based on actual frequency
confusion = merge(confusion, actual, by=c(&quot;Actual&quot;))
confusion$Percent = confusion$Freq/confusion$ActualFreq*100

#render plot
# we use three different layers
# first we draw tiles and fill color based on percentage of
   test cases
tile <- ggplot() +
  geom_tile(aes(x=Actual, y=Predicted,fill=Percent),data=
     confusion, color="black",size=0.1) +
```

```
  labs(x="Actual",y="Predicted")
tile = tile +
  geom_text(aes(x=Actual,y=Predicted, label=sprintf("%.1f",
     Percent)),data=confusion, size=3, colour="black") +
  scale_fill_gradient(low="grey",high="red")

# lastly we draw diagonal tiles. We use alpha = 0 so as not to
   hide previous layers but use size=0.3 to highlight border
tile = tile +
  geom_tile(aes(x=Actual,y=Predicted),data=subset(confusion, as
     .character(Actual)==as.character(Predicted)), color="black
     ",size=0.3, fill="black", alpha=0)

#render
tile
```

BIBLIOGRAPHY

Abondolo, D. (1998). Khanty. In Abondolo, D., editor, *The Uralic languages*, pages 358–86. London: Routledge.

Agrawal, R. (2011). Visualizing confusion matrix using HeatMap in R. https://ragrawal.wordpress.com/2011/05/16/visualizing-confusion-matrix-in-r/.

Aguilar, L. (1997). *De la vocal a la consonante*, volume 3. Universidade de Santiago de Compostela.

Ainsworth, W. (1968). First formant transitions and the perception of synthetic semivowels. *The Journal of the Acoustical Society of America*, 44(3):689–694.

Andrade, A. (1999). On /l/ velarization in European Portuguese. In *International Congress of Phonetic Sciences, San Francisco.*

Aoyama, K. and Guion, S. G. (2007). Prosody in second language acquisition. *Language experience in second language speech learning: In honor of James Emil Flege*, 17:281.

Arvaniti, A. (1999). Cypriot Greek. *Journal of the International Phonetic Association*, 29(02):173–178.

Austin, P. (1981). A grammar of Diyari, South Australia. *Cambridge Studies in Linguistics London*, (32).

Azevedo, M. M. (2005). *Portuguese: A linguistic introduction*. Cambridge University Press.

Barbosa, P. A. and Albano, E. C. (2004). Brazilian Portuguese. *Journal of the International Phonetic Association*, 34(02):227–232.

Bendor-Samuel, J. T., Association, I. L., et al. (1961). *The verbal piece in Jebero*. Linguistic Circle of New York.

Benko, L. and Imre, S. (1972). *The Hungarian language*. Mouton.

Berry, J. J. (2011). Accuracy of the NDI wave speech research system. *Journal of Speech, Language, and Hearing Research*, 54(5):1295–1301.

Bills, G.D., V. C. and Troike, B. (1969). *An introduction to spoken Bolivian Quechua.* University of Texas Press for Institute of Latin American Studies, Austin.

Bladon, R. A. and Carbonaro, E. (1978). Lateral consonants in Italian. *Journal of Italian Linguistics Amsterdam*, 3(1):43–54.

Blake, B. J. (1979). *A Kalkatungu grammar*. The Australian National University.

Boersma, P. and Weenink, D. (2014). *Praat: doing phonetics by computer [Computer program]*.

Bothorel, A. (1982). *Etude phonétique et phonologique du breton-parlé à Argol(Finistère-sud)*. PhD thesis.

Boyd-Bowman, P. (1952). Sobre restos de lleísmo en México. *Nueva Revista de Filología Hispánica*, 6(1):69–74.

Branks, T. H. and Branks, J. (1973). Fonología del Guambiano. *Sistemas fonológicos de idiomas colombianos*, 2:39–56.

Breen, G. and Henderson, J. (1990). The syllable in Arrernte phonology. *School of Australian Linguistics and Institute for Aboriginal Development, Alice Springs*.

Breheny, P. and Burchett, W. (2016). *visreg: Visualization of Regression Models*. R package version 2.2-2.

Browman, C. P. and Goldstein, L. (1995). Gestural syllable position effects in American English. *Producing speech: Contemporary issues*, pages 19–33.

Browman, C. P. and Goldstein, L. M. (1986). Towards an articulatory phonology. *Phonology*, 3(01):219–252.

Bubrikh, D. V. (1949). *Grammatika literaturnogo Komi Jazyka*. Leningrad State University, Leningrad.

Carignan, C. (2013). *When nasal is more than nasal: The oral articulation of nasal vowels in two dialects of French*. PhD thesis, University of Illinois at Urbana-Champaign.

Carignan, C., Shosted, R., Shih, C., and Rong, P. (2011). Compensatory articulation in American English nasalized vowels. *Journal of Phonetics*, 39(4):668–682.

Castellani, A. (1954). GL intervocalico in Italiano. *Lingua nostra*, 15:66–70.

Chan, A. Y. and Li, D. C. (2000). English and cantonese phonology in contrast: Explaining cantonese esl learners' english pronunciation problems. *Language Culture and Curriculum*, 13(1):67–85.

Chomsky, N. and Halle, M. (1968). *The sound pattern of English*. New York: Harper and Row.

Clements, G. N. and Sezer, E. (1982). Vowel and consonant disharmony in Turkish. *The structure of phonological representations*, 2:213–255.

Coates, H. and Elkin, A. (1974). *Ngarinjin - English Dictionary*, volume 1. Oceania Linguistic Monograph.

Colantoni, L. (2001). *Mergers, chain shifts and dissimilatory processes: palatals and rhotics in Argentine Spanish.* PhD thesis, University of Minnesota.

Colantoni, L. (2004). Reinterpreting the CV transition. In *Contemporary Approaches to Romance Linguistics: Selected Papers from the 33rd Linguistic Symposium on Romance Languages (LSRL), Bloomington, Indiana, April 2003*, volume 258, page 83. John Benjamins Publishing.

Cubberley, P. (2002). *Russian: A linguistic introduction.* Cambridge University Press.

Cunha, C., Cintra, L. F. L., et al. (1985). *Nova gramática do Português contemporâneo*, volume 2. Nova Fronteira Rio de Janeiro.

Dalston, R. M. (1975). Acoustic characteristics of English /w, r, l/ spoken correctly by young children and adults. *The Journal of the Acoustical Society of America*, 57(2):462–469.

Dauzat, A. (1899). Contribution a l'etude des articulations consonantiques. *La Parole*, 1:611–621.

Davidson, L. (2006). Comparing tongue shapes from ultrasound imaging using smoothing spline analysis of variance. *The Journal of the Acoustical Society of America*, 120(1):407–415.

De los Heros Diez Canseco, S. (1997). *Language variation: The influence of speakers' attitudes and gender on sociolinguistic variables in the Spanish of Cusco, Peru.* PhD thesis, PhD dissertation, University of Pittsburgh.

DiCanio, C. (2013). Visualizing vowel spaces in R: from points to contour maps. http://christiandicanio.blogspot.com/2013/10/visualizing-vowel-spaces-in-r-from.html.

Echeverria, M. S. and Contreras, H. (1965). Araucanian phonemics. *International Journal of American Linguistics*, 31(2):132–135.

Elmar, T. (1970). Grammaire structurale du Breton de L'Ile de Groix. *Heidelberg: Carl Winter Universitätsverlag.*

Engwall, O., Delvaux, V., and Metens, T. (2006). Interspeaker variation in the articulation of nasal vowels. *Proceedings of the 7th ISSP*, pages 3–10.

Espy-Wilson, C. Y. (1992). Acoustic measures for linguistic features distinguishing the semivowels /w j r l / in American English. *The Journal of the Acoustical Society of America*, 92(2):736–757.

Fairbanks, G. H. (1965). *Historical phonology of Russian*, volume 16. Deccan College, Post-graduate and Research Institute.

Fant, G. (1960). *Accoustic theory of speech production: with calculations based on X-ray studies of Russian articulations*. Mouton & Company.

Fechner, G., Adler, H., Howes, D., and Boring, E. (1966). *Elements of Psychophysics*. Number v. 2 in A Henry Holt edition in psychology. Holt, Rinehart and Winston.

Flege, J. E. (1991). Age of learning affects the authenticity of voice-onset time (VOT) in stop consonants produced in a second language. *The Journal of the Acoustical Society of America*, 89(1):395–411.

Gick, B., Campbell, F., Oh, S., and Tamburri-Watt, L. (2006). Toward universals in the gestural organization of syllables: A cross-linguistic study of liquids. *Journal of Phonetics*, 34(1):49–72.

Gim, S.-G. (1971). *Phonetics of Korean*. Hanul.

Goldinger, S. D. (1996). Words and voices: episodic traces in spoken word identification and recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22(5):1166.

Grandgent, C. H. (1927). From Latin to Italian. *Cambridge*, 10:24–52.

Gu, C. (2014). Smoothing Spline ANOVA Models: R package gss. *Journal of Statistical Software*, 58(5):1–25. http://www.jstatsoft.org/v58/i05/.

Guion, S. G. (1998). The role of perception in the sound change of velar palatalization. *Phonetica*, 55(1-2):18–52.

Guion, S. G., Flege, J. E., Liu, S. H., and Yeni-Komshian, G. H. (2000). Age of learning effects on the duration of sentences produced in a second language. *Applied Psycholinguistics*, 21(02):205–228.

Hanulíková, A. and Hamann, S. (2010). Slovak. *Journal of the International Phonetic Association*, 40(03):373–378.

Hardman, M. J. (1966). *Jaqaru: outline of phonological and morphological structure*, volume 22. Mouton.

Hardman, M. J. (1978). Jaqi: The linguistic family. *International Journal of American Linguistics*, 44(2):146–153.

Harris, K. S., Hoffman, H. S., Liberman, A. M., Delattre, P. C., and Cooper, F. S. (1958). Effect of third-formant transitions on the perception of the voiced stop consonants. *The Journal of the Acoustical Society of America*, 30(2):122–126.

Heer, J. and Bostock, M. (2010). Crowdsourcing graphical perception: Using Mechanical Turk to assess visualization design. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 203–212. ACM.

Honsa, V. (1975). Clasificación de los dialectos Españoles de América y la estructura de los dialectos de Colombia. In *Actas del III Congreso de Asociación Lingüística y Filología de la América Latina.*

Howard, L. (1967). Camsa phonology. *Phonemic systems of Colombian languages. Lomalinda, Colombia: Summer Institute of Linguistics.*

Hualde, J. I. (1991). *Basque phonology.* Routledge.

Hualde, J. I. and Bilbao, X. (1992). A phonological study of the Basque dialect of Getxo. *Anuario del Seminario de Filología Vasca "Julio de Urquijo".*

Hualde, J. I. et al. (2005). *The sounds of Spanish.* Cambridge University Press New York.

Hyman, L. M. (1972). Nasals and nasalization in Kwa. *Studies in African Linguistics*, 3(2):167–205.

IBGE (2010). Instituto Brazileiro de Geografia e Estatística: 2010 population census. http://www.ibge.gov.br/english/estatistica/populacao/censo2010/default.shtm.

Ingram, J. C. and Park, S.-G. (1998). Language, context, and speaker effects in the identification and discrimination of English /r/ and /l/ by Japanese and Korean listeners. *The journal of the Acoustical Society of America*, 103(2):1161–1174.

Iverson, G. and Sohn, H.-S. (1994). Liquid representation in Korean. *Theoretical Issues in Korean Linguistics*, pages 77–100.

Iverson, P., Hazan, V., and Bannister, K. (2005). Phonetic training with acoustic cue manipulations: A comparison of methods for teaching English /r/-/l/ to Japanese adults. *The Journal of the Acoustical Society of America*, 118(5):3267–3278.

Jakobson, R. (1941). *Child language: aphasia and phonological universals.* The Hague & Paris: Mouton.

Jakobson, R., Fant, G., and Halle, M. (1952). *Preliminaries to speech analysis.* Cambridge: MIT Press.

Johnson, K. (2011). *Acoustic and auditory phonetics.* John Wiley & Sons.

Joseph, B. D. and Tserdanelis, G. (2003). Modern Greek. In Roelcke, T., editor, *Variationstypologie/Variation Typology: Ein sprachtypologisches Handbuch der europäischen Sprachen in Geschichte und Gegenwart/A Typological Handbook of European Languages*, pages 823–836. Walter de Gruyter.

Kang, Y. (2012). The adaptation of English liquids in contemporary Korean: A diachronic study. *Catalan Journal of Linguistics*, 11:41–63.

Kawahara, S. (2010). get_formants_midpoint.praat. http://user.keio.ac.jp/k̃awahara/scripts/get_formants_ midpoint.praat.

Keating, P. and Lahiri, A. (1993). Fronted velars, palatalized velars, and palatals. *Phonetica*, 50(2):73–101.

Keating, P. A. (1988). Palatals as complex segments: X-ray evidence. *UCLA working papers in phonetics*, 69:77–91.

Knight, R.-A., Villafana Dalcher, C., and Jones, M. J. (2008). Cue switching in the perception of approximants: Evidence from two English dialects. *Journal of the Acoustical Society of America*, 123(5):3319.

Kochetov, A. (2004). Perception of place and secondary articulation contrasts in different syllable positions: language-particular and language-independent asymmetries. *Language and speech*, 47(4):351–382.

Kochetov, A. (2005). Phonetic sources of phonological asymmetries: Russian laterals and rhotics.

Koneczna, H. and Zawadowski, W. (1956). *Obrazy rentgenograficzne głosek rosyjskich*. PWN.

Kuhl, P. K. (2007). Is speech learning 'gated' by the social brain? *Developmental science*, 10(1):110–120.

Ladefoged, P. (1964). *A phonetic study of West African languages*. Cambridge: Cambridge University Press.

Ladefoged, P. (1988). Hierarchical features of the International Phonetic Alphabet. In *UCLA working papers in phonetics*, volume 70, pages 1–12.

Ladefoged, P. (1996). *Elements of acoustic phonetics*. University of Chicago Press.

Ladefoged, P. and Everett, D. (1996). The status of phonetic rarities. *Language*, 72(4):794–800.

Ladefoged, P. and Johnson, K. (2014). *A course in phonetics*. Cengage learning.

Ladefoged, P. and Maddieson, I. (1996). *The sounds of the world's languages*. Oxford Publishers.

Landis, J. R. and Koch, G. G. (1977). The measurement of observer agreement for categorical data. *biometrics*, pages 159–174.

Lapenda, G. (2005). *Estrutura Da Língua Iatê*. UFPE.

Lapesa, R. and Pidal, R. M. (1942). Historia de la lengua Española.

Lewis, M. P., editor (2009). *Ethnologue: Languages of the World*. SIL International, Dallas, TX, USA, sixteenth edition.

Liljencrants, J. and Lindblom, B. (1972). Numerical simulation of vowel quality systems: The role of perceptual contrast. *Language*, pages 839–862.

Lindblom, B. (1986). Phonetic universals in vowel systems. *Experimental phonology*, pages 13–44.

Lindblom, B. (1990). Explaining phonetic variation: A sketch of the H&H theory. In *Speech production and speech modelling*, pages 403–439. Springer.

Linguateca (1999). Núcleo Interinstitucional de Lingüística Computacional / São Carlos. In *Acesso a corpora / Disponibilização de corpora.* http://www.linguateca.pt/acesso/corpus.php/?corpus=SAOCARLOS/.

Lipski, J. M. (1989). Spanish yeísmo and the palatal resonants: Towards a unified analysis. *Probus*, 1(2):211–224.

Lipski, J. M. (1994). *Latin American Spanish.* Longman Pub Group.

Lotto, A. J., Sato, M., and Diehl, R. L. (2004). Mapping the task for the second language learner: The case of Japanese acquisition of /r/ and /l/. *From sound to sense*, 50(2004):C381–C386.

Lytkin, V. (1966). Komi-zyrjanskij jazyk. In Vinogradov, V., editor, *Jazyki Narodov SSSR*, volume 3, pages 281–299. Nauka, Moscow and Leningrad.

Macmillan, N. and Creelman, C. (1991). Detection theory: A user's guide (Cambridge UP, Cambridge).

Maddieson, I. and Precoda, K. (1991). UPSID-PC: The UCLA Phonological Segment Inventory Database.

Maeda, S. (1990). Compensatory articulation during speech: Evidence from the analysis and synthesis of vocal-tract shapes using an articulatory model. In *Speech production and speech modelling*, pages 131–149. Springer.

Maiden, M. (1995a). Evidence from the Italian dialects for the internal structure of prosodic domains. *Ansterdam studies in the theory and history of linguistic science: Series 4*, pages 115–115.

Maiden, M. (1995b). *A linguistic history of Italian.* Routledge.

Martínez-Celdrán, E., Fernández-Planas, A. M., and Carrera-Sabaté, J. (2003). Castilian Spanish. *Journal of the International Phonetic Association*, 33(02):255–259.

Martins, P., Carbone, I., Pinto, A., Silva, A., and Teixeira, A. (2008). European Portuguese MRI based speech production studies. *Speech Communication*, 50(11):925–952.

Martins, P., Oliveira, C., Silva, A., and Teixeira, A. (2010). Articulatory characteristics of European Portuguese laterals: A 2D & 3D MRI study. *FALA, Vigo, Spain.*

Mateus, M. H. and d'Andrade, E. (2000). *The phonology of Portuguese.* Oxford University Press.

MATLAB (2014). *version 8.3 (R2014a).* The MathWorks Inc., Natick, Massachusetts.

Menéndez Pidal, R. (1950). Orígenes del Español. *Estado lingüístico de la Península Ibérica hasta el siglo XI*, 3.

Mervis, C. B. and Rosch, E. (1981). Categorization of natural objects. *Annual review of psychology*, 32(1):89–115.

Mielke, J. (2013). tongue_ssanova.r. Functions for SSANOVA comparisons of tongue traces in polar coordinates using gss.

Miller, G. A. and Nicely, P. E. (1955). An analysis of perceptual confusions among some English consonants. *The Journal of the Acoustical Society of America*, 27(2):338–352.

Mosby (2013). *Mosby's Dental Dictionary.* Elsevier Health Sciences. Occlusal Plane.

Mouton, P. G. and Martos, I. M. (2012). The /ʎ/-/ɉ/ merger (*yeísmo*) in Central Spain: Advances since the ALPI. *Dialectologia: revista electrònica*, pages 23–42.

Navarro Tomás, T. (1968). Repertorio de estrofas Españolas.

O'Connor, J. D., Gerstman, L. J., Liberman, A. M., Delattre, P. C., and Cooper, F. S. (1957a). Acoustic cues for the perception of initial w, j, r, l in English. *Word-Journal of the International Linguistic Association*, 13(1):24–43.

O'Connor, J. D., Gerstman, L. J., Liberman, A. M., Delattre, P. C., and Cooper, F. S. (1957b). Acoustic cues for the perception of initial /w, j, r, l/ in English. *Word*, 13(1):24–43.

Oh, S. (2002). Cross-language blending of /l/ gestures by Korean-English bilingual children. In *26th Boston University Conference on Language Development (BUCLD)*, volume 2, pages 473–484.

Oh, S. and Gick, B. (2002). An ultrasound study of articulatory gestures of Korean /l/. In *Proceedings of the first international conference on speech sciences.*

Ohala, J. (1981). The listener as a source of sound change. *Papers from the Parasession on Language and Behavior*, pages 178–203.

Ohala, J. (1989). Sound change is drawn from a pool of synchronic variation. *Language change: Contributions to the study of its causes*, pages 173–198.

Ohala, J. J. (1993). Sound change as nature's speech perception experiment. *Speech Communication*, 13(1):155–161.

Oliveira, C., Martins, P., Marques, I., Couto, P., and Teixeira, A. (2011). An articulatory and acoustic study of European Portuguese /l/. In *International Congress of Phonetic Sciences, Hong Kong.*

O'Nolan, G. (1934). *The new era grammar of Modern Irish*. Educational Company of Ireland.

Penny, R. (2000). *A History of the Spanish Language*. Cambridge University Press.

Pharies, D. (2007). *A Brief History of the Spanish Language*. University of Chicago Press.

Phatak, S. and Allen, J. (2007). Syllable confusions in speech-weighted noise. *Acoustical Society of America Journal*, 121(4):2312–2326.

Pierrehumbert, J. (2002). Word-specific phonetics. *Laboratory phonology*, 7:101–139.

Proctor, M. I. (2009). *Gestural characterization of a phonological class: The liquids*. PhD thesis, Yale University.

Quilis, A., Esgueva, M., Araus, M. L. G., and Cantarero, M. (1979). Características de las consonantes laterales Españolas. *LEA: Lingüística Española actual*, 1(2):233–344.

R Core Team (2015). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org/.

Recasens, D. (1984a). Timing constraints and coarticulation: Alveolo-palatals and sequences of alveolar+[j] in Catalan. *Phonetica*, 41(3):125–139.

Recasens, D. (1984b). Vowel-to-vowel coarticulation in Catalan VCV sequences. *The Journal of the Acoustical Society of America*, 76(6):1624–1635.

Recasens, D. (2013). On the articulatory classification of (alveolo) palatal consonants. *Journal of the International Phonetic Association*, 43(01):1–22.

Recasens, D. and Espinosa, A. (2006). Articulatory, positional and contextual characteristics of palatal consonants: Evidence from Majorcan Catalan. *Journal of Phonetics*, 34(3):295–318.

Recasens, D., Farnetani, E., Fontdevila, J., and Pallarès, M. D. (1993). An electropalatographic study of alveolar and palatal consonants in Catalan and Italian. *Language and Speech*, 36(2-3):213–234.

Recasens, D., Farnetani, E., and Ni Chasaide, A. (1991). A study of the production and coarticulatory characteristics of palatal and palatalized consonants. *ACCOR Periodic Progress Report*, 2.

Recasens, D., Fontdevila, J., and Pallarès, M. D. (1995). A production and perceptual account of palatalization. *Bruce Connell & Amalia Arvaniti (eds.)*, pages 265–281.

Recasens, D. and Pallarès, M. D. (2001). Coarticulation, assimilation and blending in Catalan consonant clusters. *Journal of Phonetics*, 29(3):273–301.

Recasens, D. and Rodríguez, C. (2015). An investigation of lingual coarticulation resistance using ultrasound. In *Ultrafest, Hong Kong*.

Recasens, D. and Romero, J. (1997). An EMMA study of segmental complexity in alveolopalatals and palatalized alveolars. *Phonetica*, 54(1):43–58.

Repetti, L. and Tuttle, E. F. (1987). The evolution of Latin pl, bl, fl, and cl, gl in Western Romance. *Studi mediolatini e volgari*, (33):53–115.

Resnick, M. C. (1975). *Phonological variants and dialect identification in Latin American Spanish*. Walter de Gruyter.

Ristinen, E. K. (1960). An East Cheremis phonology. *American Studies in Uralic Linguistics*, 1:249–287.

Rohlfs, G. (1966). Grammatica storica della lingua Italiana e dei suoi dialetti. *Fonetica*, 1.

Rona, P. (1964). El problema de la división del Español Americano en zonas dielectales en presente y futuro de la lengua Española. *Madrid, SCIC*.

Rosch, E. (1978). Principles of categorization. In Rosch, E. and Lloyd, B. B., editors, *Cognition and categorization*, pages 27–48. Hillsdale, NJ: Lawrence Erbaum Associates.

Rousselot, J.-P. (1899). Les articulations Irlandaises etudiees a l'aide du palais artificiel. *La Parole*, 1:241–262.

Russell, G. O. (1970). *The vowel: Its physiological mechanism as shown by X-ray*. Ayer Publishing.

Scatton, E. A. (1975). *Bulgarian phonology*. Slavica Publishers.

Shosted, R., Carignan, C., and Rong, P. (2012a). Managing the distinctiveness of phonemic nasal vowels: Articulatory evidence from Hindi. *The Journal of the Acoustical Society of America*, 131(1):455–465.

Shosted, R., Hualde, J. I., and Scarpace, D. (2012b). Palatal complexity revisited: An electropalatographic analysis of /ɲ/ in Brazilian Portuguese with comparison to Peninsular Spanish. *Language & Speech*, 55(4).

Silva, A. H. (1999). Caracterização acústica de [R], [ɾ], [L] e [ʎ] nos dados de um informante paulistano. *Cadernos de estudos linguisticos*, (37):51–68.

Simonsen, H. G., Moen, I., and Cowen, S. (2008). Norwegian retroflex stops in a cross linguistic perspective. *Journal of Phonetics*, 36(2):385–405.

Sproat, R. and Fujimura, O. (1993). Allophonic variation in English/l/ and its implications for phonetic implementation. *Journal of phonetics*, 21(3):291–311.

Stevens, K. N. (2000). *Acoustic phonetics*. MIT Press.

Stevenson, C. W. (2013). That's smooth. http://statistical-research.com/thats-smooth/?utm_source=rss&utm_medium=rss&utm_campaign=thats-smooth.

Straka, G. (1965). *Naissance et disparition des consonnes palatales dans l'évolution du Latin au Français*. Centre de philologie et de littératures romanes de l'Université de Strasbourg.

Strevens, P. (1954). *Some observations on the phonetics and pronunciation of modern Portuguese*.

Tabain, M. (2009). An EPG study of the alveolar vs. retroflex apical contrast in Central Arrernte. *Journal of Phonetics*, 37(4):486–501.

Teixeira, A., Martins, P., Oliveira, C., and Silva, A. (2012). Production and modeling of the European Portuguese palatal lateral. *Computational Processing of the Portuguese Language*, pages 318–328.

Tereshchenko, N. (1966). Nganasanskij jazyk. In Vinogradov, V., editor, *Jazyki Narodov SSSR*, volume 3, pages 438–457. Nauka, Moscow and Leningrad.

Toscano, J. C. and Allen, J. B. (2014). Across-and within-consonant errors for isolated syllables in noise. *Journal of Speech, Language, and Hearing Research*, 57(6):2293–2307.

Uldall, E. (1954). Guarani sound system. *International Journal of American Linguistics*, 20(4):341–342.

Vagges, K., Ferrero, F. E., Magno-Caldognetto, E., and Lavagnoli, C. (1978). Some acoustic characteristics of Italian consonants. *Journal of Italian Linguistics Amsterdam*, 3(1):69–84.

Vago, R. (1974). *The sound pattern of Hungarian*. Georgetown University Press.

Valenzuela, P. M. and Gussenhoven, C. (2013). Shiwilu (Jebero). *Journal of the International Phonetic Association*, 43(1):97–106.

Venables, W. N. and Ripley, B. D. (2002). *Modern Applied Statistics with S*. Springer, New York, fourth edition. ISBN 0-387-95457-0.

Wickham, H. (2009). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York.

Wilkins, D. P. (1989). *Mparntwe Arrernte (Aranda): Studies in the structure and semantics of grammar*. PhD thesis, The Australian National University.

Wong, N. W. and Hermes, Z. (2015). *Matlab data processing suite for NDI Wave data*. https://www.dropbox.com/s/02lxp0u1eiyb2yv/Data_processing_02212015.zip?dl=0.

Wood, S. A. (1975). Tense and lax vowels-degree of constriction or pharyngeal volume? *Working papers, Phonetics Laboratory, Department of General Linguistics, Lund University*.

Wood, S. N. (2000). Modelling and smoothing parameter estimation with multiple quadratic penalties. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 62(2):413–428.

Yamada, R. A. and Tohkura, Y. (1992). The effects of experimental variables on the perception of American English/r/and/l/by Japanese listeners. *Perception & psychophysics*, 52(4):376–392.

Zampaulo, A. (2015). The evolution of the (alveolo)palatal lateral consonant in Spanish and Portuguese. In Haug, D. and Kristjánsson, E., editors, *Historical Linguistics 2013: Selected papers from the 21st International Conference on Historical Linguistics, Oslo, 5-9 August 2013*, Current Issues in Linguistic Theory, pages 69–88. John Benjamins Publishing Company.

Zhukova, A. (1980). *Jazyk palanskikh Korjakov.* Nauka, Leningrad.

Zilyns'kyj, I. (1979). A phonetic transcription of the Ukrainian language. *Cambridge, Mass.*