Abstract

Perception and Production of Timing in Non-native Speech: Russian Palatalization

Leandro José Bolaños

2017

It is well known that adults experience difficulty in perceiving and producing certain phonological contrasts not present in their native language. Adults also find difficulty in learning the specific timing of non-native articulatory gestures and contextual differences present in the language. This dissertation investigates the perception and production of Russian contrasts involving palatalized consonants in varying contexts by native speakers of American English lacking any experience in Russian, a group of native English speakers with varying degrees of Russian experience and a native Russian speaking group. Of interest are the effects of syllable position and palatalization on speakers' performance in perception and production. The framework of Articulatory Phonology (e.g. Browman & Goldstein 1986, 1992) and the Perceptual Assimilation Model (e.g. Best et al. 2001) are explored to account for differences in timing between English and Russian with respect to palatalization, and to subsequently make predictions on English speakers' perception as well as their production of the different timing property present in Russian palatalization. This dissertation also investigates the relationship between perception and production, specifically on the phenomenon of palatalization in non-native speech. In addition to group differences I also investigate individual speaker differences with perception and production and the relationship between them.

Participants from the three groups participated in a series of perception and production experiments involving Russian palatalized stops which vary in place of articulation (labials, coronals) and syllable position (onset, coda). The experiments consisted of three perception tasks (categorization, discrimination of natural speech, and discrimination of synthetic speech) and one production task. I show that the Perceptual Assimilation Model (PAM) is largely borne out by investigating the relation between categorization and discrimination. I also show that invoking Articulatory Phonology allows us to examine the relative differences in timing between the two languages and account for certain perception and production patterns for non-native speakers of a language. I also show that speakers differ both at the group level but also within groups we see different strategies and patterns in perception and production for different speakers. This dissertation therefore contributes new data to the body of work of perception and production in non-native speech.

Perception and Production of Timing in Non-native Speech: Russian Palatalization

A Dissertation
Presented to the Faculty of the Graduate School
of
Yale University
in Candidacy for the Degree of
Doctor of Philosophy

by
Leandro José Bolaños

Dissertation Director: Stephen Anderson

May 2017

Contents

List of Figures

List of Tables

Acknowledgments

This dissertation would not have been possible without the support and friendship of many different people. I will start by formally thanking everyone who has had the most direct impact in helping me along the way in producing the work presented here. First I want to thank Jelena Krivokapić, my first advisor at Yale. Thank you for the countless discussions about far flung ideas ranging from lenition of voiced stops in Spanish, the status of trills and taps and nasal assimilation, and finally to the topic presented in this thesis. It seemed like each week I wanted to write about something completely different. Jelena spent many hours year after year listening to my sometimes crazy ideas, putting up with less than readable drafts of papers and chapters, and guiding me through the process of writing a dissertation. I am eternally grateful for her guidance and mentorship. I also want to thank Stephen Anderson, current advisor for my dissertation for providing me very useful feedback on this and previous projects and iterations of this project through the years. I am very grateful for all the help Steve has given me through the years. I wish to thank Ryan Bennett for providing extremely useful and pointed feedback on previous drafts, and for his insightful comments on particular aspects of this work which have greatly broadened my understanding of the phenomenon. I also want to thank and acknowledge the entire faculty of the linguistics department at Yale University. The classes I took with Jelena Krivokapić, Darya Kavitskaya, Gaja Jarosz, Raffaella Zanuttini, Ashwini Deo, Steve Anderson, Erich Round, Dianne Jonas, Matt Wolf and others were all extremely enlightening and engaging. I learned very much from everyone on the faculty and from others whose courses I did not have the pleasure of formally taking (Bob Frank, Claire Bowern, Larry Horn) thank you as well.

Everyone who taught a class I took, gave me feedback on a talk, or applauded or congratulated me (and others) on completion of milestones, you have my deep gratitude. I always felt like a part of the community at Yale and very much appreciate the tireless efforts of faculty members in helping students progress and succeed. I also want to thank Pam Beddor from the University of Michigan, who taught me so much about phonetics and science as an undergrad at the University of Michigan. I also want to thank participants during poster sessions at the 2013 Acoustical Society of America meeting for useful feedback on previous versions of this work.

The people I'm about to thank at this point may not have had as direct an influence on the work in this thesis, but without them I would not have had the support and friendship necessary to complete a PhD. First, my cohort, thank you to Michael Freedman and E-Ching Ng for being great friends and colleagues during the whole time at Yale. Thanks also to Argyro Katsika, Erich Round, Gregg Castellucci, Dolly Goldenberg, Christian DiCanio and other members of the Yale and Haskins communities for useful feedback on previous versions of this work and for your friendship as well. Thanks also to Doug Whalen, Aude Noiray and Khalil Iskarous who took me on as an RA at Haskins Laboratories early in my grad school career and who also taught me more about the beauty and rigor of scientific discovery (in speech). Finally, thanks to anyone who came out to play volleyball on Sundays, it also made life at Yale more fun.

Finally, I want to thank my family. My parents, Lourdes and Enrique who always taught us to work hard, treat others with respect, honor our backgrounds and best of all - do what you want to do in life. The support given to me (and my siblings) by my parents is unquestionably one of the most important factors in my completion of this thesis and in everything else I've done of value in my life. I want to thank my siblings Lila, Enrique, Alberto and Gabriel for just being there for me. I also want to thank my grandfather who taught my father the same values my parents taught me and who also taught me values of honesty and hard work. Finally I want to thank Sabina. Without her I'd be a lonely, lost fool, and with her this dissertation really came to life.

Thanks to everyone who taught me, lectured me, corrected me, confronted me and made me laugh these past years, you've all made me a better person. And of course any shortcomings in this dissertation are solely my own responsibility - this is the most important part.

Chapter 1 Introduction

This dissertation investigates the perception and production of non-native timing in speech, and specifically in articulatory gestures. This dissertation is set within several different areas of study in linguistics. By investigating non-native speech the studies reported here contribute to the body of knowledge in the cross-linguistic and second language acquisition literature. By investigating the specific properties of the particular language and phenomenon studied here, specifically Russian palatalization, the dissertation contributes to the literature on this topic, and given the nature of Russian palatalization to be explained in subsequent chapters, this dissertation is crouched within the framework of Articulatory Phonology. By investigating performance in production and perception by individual speakers and their differences this dissertation makes contributions to the body of knowledge on individual speaker differences in speech perception and production. More specifically I investigate the perception and production of Russian palatalized consonants by speakers of English who differ in their experience with Russian. Following I describe the goals of the dissertation. Goals

The goals of this dissertation are outlined here and motivated further in upcoming sections. The goals are both empirical and theoretical. They are empirical in that a relatively understudied phenomenon - palatalization as a secondary articulation in the context of non-native perception and production will be investigated. They are theoretical in that the results will have bearings on prominent theories of speech perception, non-native perception, second language acquisition (SLA) as well as the link between perception and production in the domain of non-native speech. The main goals for this dissertation are outlined and elaborated below.

The first goal is to investigate perception and production of an example of non-native gestural timing, namely palatalization. This is important because within non-native and second language speech perception and production studies, the timing of articulators or gestures is seldom studied (although see e.g. Zsiga 2011). Indeed most such studies investigate segmental contrasts which differ between languages. Timing is an often overlooked property of speech, and as such this dissertation investigates non-native timing of gestures.

The second goal is to test the Perceptual Assimilation Model (PAM) of non-native speech and introduce new data into this model, namely non-native perception of gestural timing (palatalization). Specifically I attempt to test the PAM with new kinds of data, namely non-native speech differing primarily in timing of articulatory gestures. This study also introduces different syllable positions (onset and coda) to be tested, of which PAM studies normally test onsets. The results and findings of this dissertation partially support the PAM.

The third goal is to investigate the extent of a link or correlation between perception and production of non-native timing. The following more specific questions arise to be addressed: 1) What is the relation between perception and production of palatalization by non-native speakers? 2) Are non-native speakers better able to perceive the palatalization contrast, or produce the contrast? These two questions represent different levels of inquiry, the first represents the general question of whether there is a relation between perceptual and production ability (see e.g. ; Fowler 2003, Fowler & Galantucci 2005, Beddor 2015), and the second represents the more specific issues, namely if non-native speakers are better at perceiving or producing the contrasts.

This second issue is related to the question of whether perception must precede production in L2. A prominent second language acquisition model, the Speech Learning Model (e.g. Flege 1999) posits among other things, that production errors in L2 have a basis in perception, that is, when learners fail to adequately produce speech sounds in an L2, they fail to adequately perceive them (Flege 1999). However, it has also been found that perception does not need to precede production in L2 (Sheldon and Strange 1982). This dissertation will attempt to address the relation and extent of perception and production in L2 in order to address this issue. We will ask, then, whether non-native speakers perceive a non-native contrast correctly and whether this has any bearing on their production of the non-native contrast. A related goal that is expanded on below looks at this same question at the individual speaker level.

The fourth goal is to investigate the effect of learning L2 experience in perception and production of palatalization through the recruitment of several groups of speakers with varying knowledge of the target language (Russian). Studying the effect of learning is important as there is much debate and inquiry into whether and how adults are able to learn the phonological, phonetic, or any other aspect of a foreign language (Piske, MacKay and Flege 2001). The critical period hypothesis (e.g. see Birdsong 1999) has been posited partly to explain why adults rarely attain native-like pronunciation levels in non-native languages. In some cases, adults can learn a nonnative phonology, but struggle with reaching native-like pronunciation. On the other hand young children can and do acquire a second language natively. Older children generally exhibit better pronunciation skills in L2 than older adults learning an L2. So, it is generally seen that in acquiring a second language, the younger the learner, the better the chances of acquiring native-like pronunciation (e.g. see Major 2001). In the SLA literature, this factor is known as Age of Learning (AOL) (Major 2001; Piske, MacKay and Flege 2001).

Other factors such as L1 and L2 use, length of residence in an L2 environment, and individual speaker properties such as motivation are known to affect non-native speech (Major 2001; Piske, MacKay and Flege 2001). Piske and colleagues (2001) investigated L1 use and found it to be a significant contributing factor to perceived foreign accent; greater L1 use is correlated with higher degrees of foreign accent. Their study found that of all the factors mentioned above, the most significant contributors to perceived foreign accent are AOL and L1 use, while other factors have not been found to correlate with perceived foreign accent as consistently. In sum, the relevant factor to be investigated here for the purposes of Goal 4 is general experience in the L2.

The fifth goal of this dissertation is related to the fourth, to examine individual differences in the link between perception and production in non-native speech. Individual differences in speech perception and production have garnered much recent interest in the literature (e.g. Schertz, Cho and Lotto 2015, Franken et. al. 2015). A major focus of this dissertation is to investigate more deeply the behavior of individual speakers who are naïve to the second language (i.e. have never had exposure to the target language) and those who are learners of the target language. A prediction here is that there will be more variability between speakers who are less proficient in the target language in both perception and production, and also within speakers. We could also expect to see differences in the relation between perception and production. Recent work has shown that individual speakers use different cues for a non-native sound differently in perception versus production. Schertz, Cho and Lotto found that cue usage of L1 Korean speakers of English voiceless stops did not always correspond between perception and production and that individual strategies differed in perception and production. Individual speakers' relative use of the cues in production was more homogenous than in perception.

In sum, this dissertation aims to investigate how speakers of English differing in experience with Russian perceive and produce phonological phenomena which differ in timing from their native language. This dissertation tests a theory of non-native perception and further investigates how perception and production interact at different levels of learning. This dissertation is organized as follows. In Chapter 2, a phonetic and articulatory description of palatalization both in Russian and in other languages is given. In Chapter 3, relevant theories of nonnative and second language speech perception are introduced. Chapter 3 then introduces hypotheses for perception and production of the specific Russian consonants. Chapter 4 introduces the experiments to test the hypotheses. Chapter 5 to Chapter 9 describe in detail each perception and production experiment. Finally Chapter 10 contains analysis and discussion of the studies.

In the next section, in order to further explain non-native timing of gestures, as well as to ground this dissertation within theoretical work, the framework of Articulatory Phonology (Browman & Goldstein 1986; 1992; 1995) will be described.

1.1. Background

1.1.1. Theoretical background and framework: Articulatory Phonology

Articulatory Phonology (AP) is a theory of speech production which takes articulatory gestures as basic phonological units. Gestures in Articulatory Phonology are defined as "abstract characterizations of articulatory events, each with an intrinsic time or duration" (Browman & Goldstein 1992). This intrinsic duration, however, is an abstract duration, not a physical one. According to this definition, gestures have both abstract, phonological properties and physical, phonetic properties. In the sense that gestures are both abstract and physical, AP assumes that linguistic (or discrete) and physical (or continuous) properties of speech are high and low (respectively) dimensional descriptions of a single system. In this way, AP attempts to bridge the gap between phonology and phonetics by considering the two domains as two sides of the same coin.

The physical and temporal properties of articulatory gestures allow for AP to account for timing in speech. First, given that gestures have temporal properties, they exist in time and have features such as onset and offset. For example, the constriction parameters of the tongue tip closure of the first consonant in /tap/ result in an activation period for that gesture which spans a period of time. Likewise, the other gestures in /tap/, such as the vocalic tongue body gesture are associated with an activation period. Along with the activation periods of gestures, relative timing between gestures is a lawful consequence of the model. Let us consider in more detail how the model accounts for utterances at and above the segmental level.

A formal way of implementing gestures in Articulatory Phonology is through the model of Task Dynamics (Saltzman & Munhall 1989). Task Dynamics has been developed as a formal system of modeling and describing skilled actions such as reaching for an object, drinking from a cup, or speech. Task Dynamics has been used to model speech production by equating the production of speech as a series of task-driven skilled actions. These actions are modeled as formations and releases of constrictions in the vocal tract. The tasks are modeled using second-order dynamical equations, hence the name of the model. In speech these tasks themselves are accomplished through the movement of articulators (e.g. lips, jaw for [b]), which in turn are governed by context-independent tract variables. Each constriction type (e.g. [b]) is represented by typically two tract variables, one representing the location of constriction in the vocal tract and another representing the degree of constriction. For example, tongue tip constriction location (TTCL), and degree (TTCD) are two related tract variables which together help specify certain gestures (e.g. [t] or [d]). In the Task Dynamics model, each gesture is associated with one or more tract variables. Articulatory gestures are then more accurately defined as a “dynamical system with a characteristic set of parameter values” (Browman & Goldstein 1995).

An utterance is composed of a combination of a set of gestures, which can overlap in time. During the planning phase of an utterance, gestures are combined into gestural scores. Gestural scores specify the durations and the relative timing of activations of the gestures for a given utterance. Gestures are thought to be regulated via a coupled oscillator model, which can result in different modes of coordination between gestures in an utterance. For example, the most common modes of coordination are in-phase and anti-phase (e.g. Goldstein, L., Byrd, D. & E. Saltzman. 2006; Marin & Pouplier 2010). This means that gestures which are coordinated in-phase to each other are timed to be synchronous, while gestures coordinated in an anti-phase mode are timed to be sequential. More specifically on the mechanisms of coordination between gestures, several different points on a gesture’s activation trajectory have been proposed as landmarks used for coordination, among them are the onset of movement of the articulator towards the target; the target (or point in time the target is reached); the c-center (or mid-point of gestural plateau); the release (the start of the offset of target or movement of the articulator away from target); and finally the release offset (point in time at which the active control of the gesture ends) (Gafos 2002). The target in the AP model is the goal, or maximum constriction of the gesture. For example, the closure of the TT constriction gesture in /t/ can be considered the target of the gesture.

The degree of overlap between gestures in an utterance depends on the relative activation, which landmarks (onset, target, offset) of one gesture coincide with which target of another, as well as factors such as speech rate, context and style.

While gestures are invariantly specified phonological units, they are at the same time systems of articulatory parameters which specify the movement of individual articulators which, in the end, can and do influence each other and produce contextually varying articulatory and acoustic outputs which have been the focus of decades of phonological and phonetic research.

We ground this dissertation in Articulatory Phonology for the following reasons. First, the notion of gestural overlap inherent in AP will be relevant[1] because the phenomenon we are investigating – palatalization in Russian, is very well described and modeled in Articulatory Phonology since it involves coordination of multiple gestures at the segmental level. Second, with regards to non-native speech perception and production, we will make claims about possible assimilation and production strategies by native speakers of American English of Russian consonants. In some cases, it is possible that English speakers could assimilate the single segment Russian speech sounds to multi-segment English categories. In the next chapter I introduce a prominent model of non-native speech perception and how it relates not only to perception of speech at the segmental level, but also with the interaction of segments and gestures and Articulatory Phonology provides insights into how such cross-linguistic perceptual assimilation might interact with articulatory gestures.

---

[1] For a review of recent developments of the AP model, including the specification of intergestural timing, see Goldstetin, L., Byrd, D. & E. Saltzman (2006), Marin, S. & M. Pouplier (2010).

Chapter 2 Perception and SLA

In this chapter I describe some general theories and previous studies on both native and non-native speech perception. I first describe the phenomenon of Categorical Perception (CP) to simply set the framework on how some basic aspects of speech perception work. A major reason why the phenomenon of CP is introduced here is because in this dissertation we are dealing with perception of phonological contrasts by speakers which speak as a native language that lacks such a contrast.

The importance of Categorical Perception studies for this dissertation is that such studies have shown that speech can be organized in a discrete set of phonological categories, while continuously varying in a phonetic property. The classic example is VOT, where VOT itself is a continuous variable ranging from negative VOT, to positive VOT. The phonological categories along a VOT continuum are, however, discrete and differ between languages both in absolute number of categories (e.g. 2 for Spanish, 3 for Thai), and boundaries between categories (e.g. zero VOT, positive VOT for English; negative VOT, zero VOT for Spanish). In the AP framework, VOT is analyzed as the coordination or timing of gestures, for zero and positive VOT this usually refers to two gestures, a lingual gesture (e.g. LA, TT, TB) for a stop, and a glottal opening gesture (GL) for voicelessness[2]. A negative VOT corresponding to a voiced stop is analyzed as simply lacking a GL gesture. A zero VOT configuration (English initial /b, d, g/ or Spanish initial /p, t, k/)

---

[2] In Articulatory phonology, not all consonants have a glottal gesture. Voiced consonants lack any glottal gesture altogether, which corresponds to the mechanism of voicing. Only voiceless consonants (and hence consonants that can be analyzed in the VOT paradigm) have a glottal gesture, which accounts for the voicelessness.

corresponds to a simultaneous coordination of lingual and GL gestures, while a positive VOT (e.g. aspiration) corresponds to a sequential coordination of lingual and GL gestures, where the GL gesture lags with respect to the lingual gesture.

It is likewise possible to analyze the speech gestures involved in Russian palatalization in the same way. As stated above, Russian palatalization shall be analyzed as an example of a difference in articulatory timing with respect to English. In the same way that glottal and lingual gestures exhibit different coordination patterns to yield differing VOT values, so are lingual and tongue body (TB) gestures coordinated in different ways in Russian and English to yield differences in palatalization. For example, differences in coordination between the two types of gestures (e.g. TT and TB) can be argued to yield 2 categories in Russian, while the notion of which phonological categories these two gestures can represent in English is not as clear. The specifics on coordination of gestures in Russian and English will be explored in detail in Chapter 3. This notion of category is important for non-native speech perception, which is the focus of the next section.

2.1. L2 and non-native speech

Comparing any two languages yields different relations between the speech sounds present in each language and how those speech sounds constitute contrasts in each language. Speech sounds in one language are often similar or seemingly identical to speech sounds in another. For instance, Spanish voiceless stops /p/ /t/ and /k/ are very similar phonetically and phonologically to French voiceless stops. Phonetically, they have similar articulatory properties, such as constriction degree and location, as well as VOT. Phonologically they are similar in that they have similar distributional properties such as sonority sequencing and phonotactics, and they contrast with each other. Other speech sounds might be similar phonologically but distinct phonetically from sounds in another language. For example, Spanish /r/ is similar phonologically to the Standard French rhotic phoneme in terms of distribution and phonotactic properties, but phonetically these two phonemes are realized distinctly in the two languages; an alveolar trill in Spanish versus a uvular trill or fricative in French. In either of these examples just described, a speaker of one language learning the other as a second language (L2) could perceive the speech sounds as being members of already existing native langue (L1) categories. Assimilation of an L2 speech sound to the L1 phonological system occurs when the listener perceives and incorporates that speech sound within an existing L1 category.

Phonological contrasts present in an L2 which are absent in L1 have been described as the most difficult for non-native speakers. For example, Japanese lacks a liquid (/l/-/r/) contrast, meaning that Japanese makes no phonemic distinction between a lateral /l/, and rhotic. However, it is now well known that not all non-native contrasts are perceived equally as poorly. Some non-native contrasts are perceived well by speakers (Best et al 1988).

Best et al. (1988) performed a study to address the question of why infants lose the ability to discriminate initially discriminable non-native contrasts. This question is related to the question of why adults are worse at distinguishing some contrasts over others. Their study evaluated two proposals, 1) the linguistic proposal and 2) the psycho-acoustic proposal. The linguistic proposal states that the reorganization of perception abilities by adulthood occurs due to specific linguistic experience. Linguistic experience is both phonological and allophonic, such as members of a non-native phonological contrast, or if one or more members of the non-native phonological contrast exhibit allophones that occur in the native language non-contrastively.

The psycho-acoustic proposal (also called auditory proposal) states that the decline in perception abilities occurs due to a lack of auditory experience with sounds that do not occur in native phonological contrasts.

Best and her colleagues (1988) tested these proposals by studying click sounds from the Zulu language. They argued that Zulu clicks do not occur either phonologically, or as allophones of any English category. Thus, according to the psychoacoustic proposal, English speakers should discriminate the clicks poorly as the acoustic properties of the clicks are not present in English. The linguistic proposal also predicts poor discrimination, because the clicks are absent phonologically and phonetically from the English inventory. The reason for testing clicks is that they are not predicted to assimilate to English phonological or phonetic categories, and as such, different perception patterns were argued to arise. 18 Zulu click minimal pairs were tested on native English speakers. Contrary to either of the 2 above proposals, subjects were found to perform well on discrimination tests. Comparing previous discrimination data for non-native consonants that assimilated in different ways to native phonology such as the Hindi retroflex versus dental stop contrast /t-t/ and the Thompson /k'/ - /q'/ ejective contrast, Best and her colleagues concluded that Zulu clicks did not perceptually assimilate to the speakers’ L1. They argued that Zulu clicks were phonetically distinct enough from any English phone and as a result they failed to assimilate to any English category. They also argued that the failure to assimilate the Zulu clicks to any English category resulted in the English subjects employing psychoacoustic and phonetic discrimination strategies, as opposed to phonemic perception. That is, Thompson and Hindi contrasts were assimilated to English categories, and Zulu clicks were not. This study and others (Polka 1992) have shown that speakers with varying language backgrounds perceive different non-native contrasts differently, and the Perceptual Assimilation Model (e.g. Best, Sithole and McRoberts 1988; Best, Hallé and Pardo 2007; Best and Tyler 2007; Hallé, Best and Levitt 1999; Best, McRoberts and Goodell 2001; Best and Strange 1992; Best 1995) was developed to model the perception of non-native contrasts.

2.2. Perceptual Assimilation Model

The finding described above (Best et al 1988), namely that non-native speech sounds are perceptually assimilated to native phonological categories when possible is a major insight of the Perceptual Assimilation Model (PAM). Likewise, the possibility exists that speakers could fail to assimilate non-native speech sounds to any L1 categories, as happened with the Zulu clicks described above.

The PAM makes predictions on the ability by non-native speakers to perceive and discriminate non-native phonological contrasts based on the perceived similarity of the L2 phones to L1 categories (Best, Sithole and McRoberts 1988; Best, Hallé and Pardo 2007; Best and Tyler 2007; Hallé, Best and Levitt 1999; Best, McRoberts and Goodell 2001; Best and Strange 1992; Best 1995). Specifically, "PAM posits that non-native speech perception is strongly affected by listeners' knowledge (whether implicit or explicit) of native phonological equivalence classes, and that listeners perceptually assimilate non-native phones to native phonemes whenever possible, based on detection of commonalities in the articulators, constriction locations and/or constriction degrees used." (Best et al. 2001). This excerpt most succinctly describes the basis of the PAM. PAM posits that speakers will assimilate non-native phones to native L1 categories based on perceived articulatory similarities of the L2 phone to the closest L1 categories. It is articulatory properties of the non-native phones, not acoustic features, which the PAM argues listeners perceive when confronted with another language. In this way PAM makes claims about the objects of perception in non-native speech, specifically that they are articulatory gestures, and so PAM is a theory in the same spirit as direct realist theories of speech perception (e.g. Fowler 2003).

PAM provides two initial perceptual possibilities for people exposed to non-native speech sounds. The first and most common is that the non-native speech sounds perceptually assimilate to native phonological categories. In this case, a non-native speech sound is perceived as an instance of a native phonological category. As an example of perceptual assimilation consider a native speaker of American English hearing Spanish [l], and assimilating this speech sound to the English phonological category /l/. The second possibility is that the speech sounds fail to perceptually assimilate to native phonological categories. In such cases the speech sounds are not perceived as an instance of any native phonological category, an example being an English speaker hearing a Zulu click and failing to hear it as any instance of any English phonological category (as in Best et al. 1988).

Several additional possibilities arise within the two broad patterns described above. For phones that assimilate to native categories, such assimilation can range from good to poor. For instance, a non-native phone that is acoustically and phonetically (or articulatorily) very similar, or identical, to a native L1 phoneme will assimilate very well to that category. For example, Spanish and French voiceless stops, such as /t/, are very similar both phonologically and phonetically, thus a Spanish speaker who hears French [t] will most likely assimilate French [t] as a very good exemplar of Spanish /t/. Conversely, a non-native phone that is relatively similar along one or more acoustic or phonetic parameters to a native L1 category, but differs in another parameter can assimilate to the L1 category but less well. For example, the same

Spanish speaker who hears an instance of English syllable initial [t^h] could assimilate this phone to the Spanish phonological category /t/; however the listener could perceive it as worse exemplar of Spanish /t/ due to the phonetic difference between Spanish and English initial /t/ – namely longer aspiration in English.

For phones that do not assimilate to any native categories, the phones can be heard either as speech or non-speech sounds (Best et al. 2001). Given how individual phones may or may not assimilate to native L1 phonological categories, several patterns of assimilation of non-native contrasts are now described.

Based on the assimilation for individual non-native phones a non-native *contrast* of two phones can be assimilated in the following ways: 1) The two phones can be assimilated each to separate L1 categories constituting Two Category assimilation (TC). 2) The two phones can be assimilated equally well or poorly to a single L1 category constituting Single Category assimilation (SC). 3) Both may assimilate to a single L1 category, but one may assimilate as a better exemplar than the other, which is termed Category Goodness (CG). These CG instances are cases involving perceivable phonetic differences between the phones, where both are assimilated to the same phonological L1 category, but possible different phonetic categories. 4) It is also possible that one or both non-native phones fail to assimilate as any native L1 category which would constitute Uncategorized-Categorized (UC) or Uncategorized-Uncategorized (UU) patterns respectively (Best et al 2001). Finally one or more phones may be perceived as disparate enough to be assimilated as non-speech which would constitute Non-Assimilable (NA).

Based on the perceptual assimilation of individual phones, PAM makes explicit predictions on discrimination of contrasts. For TC assimilation patterns, discrimination by non-native listeners should be good; for SC patterns, discrimination should be poor; and for CG patterns, discrimination ability should be better than SC but worse than TC. The reasons for such predicted discrimination ability are as follows. For TC contrasts, the opposing phones exhibit a relation like that in the L1, that is, they each assimilate to a separate L1 category equally well. For CG patterns, the prediction is less precise because both phones assimilate to a single category, but one is noticeably worse than its pair. The CG pattern is a range, and different contrasts would fall under different regions within this range, some patterning closer to TC while others to SC. This is why the prediction for CG types is simply that they should be easier than SC but harder than TC. For SC types, the observation is that both opposing non-native phones should assimilate roughly equally well to a single native category, thus behaving together as one phonological and phonetic L1 category, and for this reason discrimination is predicted to be poor. Thus, the assimilation of phones to L1 categories determines the discrimination abilities of non-native contrasting phones according to PAM.

The predictions on discrimination of non-native contrasts the PAM makes are easily testable and ultimately arise from the perceptual assimilation (or lack thereof) of individual non-native phones to potential L1 categories. As stated above, perceived articulatory similarity of non-native phones to L1 categories is theorized to determine whether and how such phones perceptually assimilate to L1 categories. However, a principled way of determining perceived articulatory similarity between L1 and L2 is still elusive. PAM studies generally compare phones using articulatory data or phonetic descriptions in order to make predictions on perceptual assimilation. The phones are then compared, but it is difficult to determine with precision the strength of a perceptual assimilation relation for any two given phones. For instance, comparing Spanish /t/ to English /t/ involves a description of the articulatory gestures involved in each, and ideally, their phasing or timing. We know that English and Spanish /t/ both involve a TT constriction and a GL opening gesture, and we know that the phasing between TT and GL differs between the two. It is also true that for most Spanish speakers the TT closure is produced closer to the teeth than in English. By comparing the articulatory properties of these two phones, we now see three ways in which they differ. However, there is still no way to quantify the strength of the perceptual assimilation of English /t/ to Spanish L1 category /t/. We might predict that English /t/-/d/ contrast would perceptually assimilate to Spanish /t/-/d/ and as such constitute a TC assimilation pattern, but it is not possible to quantify how well or poorly the phones will assimilate to Spanish categories. Although it is possible to determine post hoc the assimilation patterns of different phones, it is indeed difficult to quantify the relations between L1 and L2, or indeed, any two phonological systems or even phones.[3]

2.3. Learning

The Speech Learning Model (SLM) was initially developed to model how experienced learners actively learn a second language (e.g. Flege 1987a, 1987b; Fox et al.1995; Piske et al. 2001; Flege 1990; Flege et al. 1994; Guion et al. 2000). The PAM was developed to account for behavior of speakers with no prior exposure to the non-native contrasts, however recently the PAM has been extended to include more experienced learners of an L2 (Best and Tyler 2007).

The SLM is a theory of speech perception and speech production, while PAM is a theory of perception. The SLM and PAM also differ in other ways described below. The main insight of the

---

[3] This is an area which could benefit from much future work. Developing a system to quantify the perceptual assimilation between languages would be extremely interesting.

SLM is what is known as equivalence classification. This notion states that learners classify different L2 phones based on their phonetic similarity/differences to phones in the L1. Thus, if a phone in the L2 is phonetically similar to a phone in the L1, this constitutes a "similar" phone in Flege's terminology. A similar phone should undergo equivalence classification and according to SLM, should not be learned as well as a "new" phone, which would not undergo equivalence classification. A new category would eventually be learned for the new phone, while the similar phone would remain in the established L1 category to which it is similar. Evidence for this has been reported, for example in L2 English learners of French who produced French /t/, /u/ (similar phones) and /y/ (new phone). Advanced English learners of French produced /y/ with more native-like characteristics than /t/ and /u/, thus seeming to confirm the hypothesis.

In this dissertation, specific hypotheses for learning will be tested, specifically by testing more similar (e.g. /t/, /p/) Russian phones[4] and more distinct, new (e.g. palatalized variants) Russian phones with speakers of differing Russian experience. The way that PAM has been extended to include L2 learning is described below. Best and Tyler (2007) compare their PAM for L2 learning (PAM-L2) to the established SLM for L2 learning. The SLM contains a series of postulates, which in turn yield hypotheses for SLA. Pam and Tyler describe each postulate and how it relates to the PAM-L2. In the end, a comparison between SLM and PAM emerges. Below are the SLM postulates from Flege 1995a.

1) *The mechanisms and processes used in learning the L1 sound system, including category formation, remain intact over the life span, and can be applied to L2 learning.*

[4] See discussions below on the nature of Russian "plain" consonants, which in actuality are velarized. Further below implications for this fact are explored.

2) *Language-specific aspects of speech sounds are specified in long-term memory representations called phonetic categories.*
3) *Phonetic categories established in childhood for L1 sounds evolve over the life span to reflect the properties of all L1 or L2 phones identified as a realization of each category.*
4) *Bilinguals strive to maintain contrast between L1 and L2 phonetic categories, which exist in a common phonological space.*

PAM agrees with P1, although disagrees in the objects of perception, namely that speakers extract invariants of articulatory gestures, instead of forming categories from acoustic cues. PAM disagrees with P2, stating that listeners do not require phonetic categories, and that they perceive gestural, phonetic or phonological information, depending on perceptual goals and context. Gestural, phonetic and phonological perception represent three levels of attentional focus, with the latter two most important for language. For P3, PAM agrees with the spirit, but again, not with the notion of phonetic categories. PAM agrees that over time listeners adapt to perception of different articulatory gestures and their phasing, and that when listeners assimilate a non-native speech sound, they do so based on the relation between its articulatory gestures and their phasing. For P4, PAM mostly agrees with the notion of phonetic categories existing in a common phonological space, but disagrees in that both phonetic and phonological categories of L1 and L2 can interact, which predicts that for example, if a speaker perceptually assimilates a phonological category in the L2 to an already existing L1 phonological category, but for which the two phones differ phonetically and perceptibly, eventually the phonological category might shift away from a single, shared one, to multiple different categories for the two phones (Best & Tyler 2007).

2.3.1. On learning in PAM

Different initial assimilation patters described above (TC, SC and CG) have been analyzed in Best and Tyler's (2007) work extending the PAM to L2 learning. Best and Tyler make several predictions for perceptual learning of non-native contrasts. Of particular interest are CG and SC assimilation types. Perceptual learning is predicted for CG types, for the more deviant member of the contrast, by first establishing a separate phonetic category, and later depending on its phonetic distance from the original L1 phonological category, a new phonological category. For SC types, Best and Tyler predict lower perceptual learning, as these contrasts are the most difficult. Again, if perceptual learning occurs, they hypothesize first the establishment of a new phonetic category. The frequency of the contrast was hypothesized to be a factor in perceptual learning of such difficult SC contrasts. If the contrast is frequent, or present in high frequency words, or lexical items necessary for everyday communication, then learning should be more probable. This could be relevant for this dissertation, as different plain/palatalized contrasts are much more frequent than others. For example, Kochetov (2002) reported relative frequencies of Russian plain and palatalized coronal and labial stops based on a corpus of Russian prepared by the same author containing approximately 33,000 words. Of the four relevant stops (plain coronals, palatalized coronal, plain labial, palatalized labial) the relative frequencies he reports for the relevant consonants are: /t/: 46%; /t^j/: 21%; /p/: 26%; and /p^j/: 7% (Kochetov 2002). These are relative frequencies, meaning that they all add up to 100%, so it does not indicate absolute frequencies. So, it is seen that palatalized voiceless labials are much less frequent than either plain consonants or palatalized coronals. For UC cases, the predictions are basically that if the two phones are close to each other, that is, if they are similar enough to the same set of L1 categories, discrimination and learning will be poor. However, if they are far apart, that is, each one is similar enough to different sets of L1 categories, then discrimination could be good and new phonological categories could be learned. For NA cases, it is not known how or whether speakers will perceptually learn the contrasts.

In this dissertation we test the PAM for initial exposure to an L2 for naïve listeners, as well as for learning, as described in the previous paragraph. Several perception experiments (to be described in Chapter 5-Chapter 8) were designed to test PAM for both naïve listeners and language learners. If phones assimilate to L1 categories as similar phones, learning is predicted to be worse than if phones do not assimilate as similar categories, as learners will be able to form *new* categories and production and perception will improve. These hypotheses will also be tested. In this section I have described some aspects of speech perception in general, and perception of non-native speech in particular from the perspective of PAM.

Chapter 3 Palatalization

In this chapter I introduce and describe several different phenomena that fall under the term “palatalization.” I describe different ways palatalization is manifested, as well as acoustic, articulatory and phonological characteristics of palatalization. I describe the kinds of palatalization that are present in Russian and English and differences between them. Lastly I review previous studies on perception of Russian palatalization and provide the context for the present dissertation.

3.1. General Introduction to Palatalization

The term “palatalization” has been applied to many phenomena all involving the raising or fronting of the tongue - usually blade or body - towards the general region of the hard palate (e.g. Bhat 1978; Ladefoged & Maddieson 1996). Palatalization can refer to both diachronic or synchronic phenomena, as well as phonological, articulatory, phonetic properties. Starting with diachronic palatalization, Bhat (1978) provides three different diachronic processes, tongue-fronting, tongue-raising, and spirantization. Bhat gives many examples of languages and how they differ in palatalization, and what type of palatalization occurs (fronting, raising, spirantization) and the relevant environments. Bhat argues that the three processes are distinct and that overall apicals usually undergo tongue raising, velars usually undergo tongue fronting and spirantization occurs as a separate process most often before glides. As described in Bhat’s 1978 work it is evident that the phonetic manifestations of palatalization are varied, from a secondary palatal articulation to strident frication. An example of diachronic palatalization that is seen in English occurs within words with /j/ initial suffixes: *press/pressure* and *habit/habitual*

(Zsiga 2000). In these cases, the alveolar consonants [s] and [t] presumably front to alvoelo-palatal /ʃ/ and /ʧ/. A fuller description of relevant types of palatalization seen in English is forthcoming later in this section.

Types of palatalization differ depending on whether a primary or secondary articulation is the articulatory gesture undergoing the palatalization (Bhat 1978, Bateman 2007). A secondary palatal articulation occurs with the addition of a palatal ("i"-like) articulation to a consonant, as in most Russian palatalized consonants, for example consonants /t/, /p/ as compared to /t^j/ and /p^j). The other type of palatalization (Bateman's *primary* palatalization type) occurs when the primary articulator moves towards a palatal articulation, as seen in the English example above. Although most cases of Russian palatalization involve secondary articulations, velar palatalized consonants are examples of primary palatalization, as the tongue body itself is active in the velar closure as well as fronting to the palatal region.

Examples of different kinds of palatalization are visible in dialects of Polish where the locative singular case of the lexical item [łapa] 'paw' is realized in five different Polish dialects as: [łap^je], [łapje], [łapçe], and [łapɕe], which spans the range between secondary palatalization ([p^j]) and strident frication ([pɕ]) (Kochetov 1998)[5]. As this example shows, one dialect employs secondary palatalization, the next dialect employs a full palatal offglide, and the latter two dialects go a step further in employing some version of a palatal fricative. The Polish dialects demonstrate how manifestations of palatalization can vary.

---

[5] Kochetov describes the difference between [łapçe], and [łapɕe] as between a non-strident palatal fricative in the former case, and a strident pre-palatal fricative. Since this is just an explanatory example of palatalization types we can treat these two as the same.

In this section I have introduced the term palatalization and have shown that there are different kinds of phenomena that can be labeled as "palatalization." In the following sections I further describe palatalization in Russian and how English differs.

3.2. Palatalization in Russian

Palatalization is a well-known property of Russian phonology. Most Russian consonants have non-palatalized and palatalized variants. The non-palatalized variants have been described in different contexts as being either plain or velarized. There is extensive evidence presented in the literature that the contrast in Russian is really a palatalized-velarized contrast as opposed to a palatalized-plain consonant contrast. For the purposes of this dissertation I use the term "plain" to refer to the non-palatalized variants of the palatalized-plain (or velarized) contrasts in Russian. However, we cannot ignore the fact that in many instances the consonants are not actually plain but velarized. However, it has also been found that especially in the context of back vowels, the velarization present in the non-palatalized variant is not always salient, and that especially for coronals, velarization is also not as strong as for other places of articulation (Kedrova et al. 2008, Bennett et al. to appear). From this point on, we usually refer to the non-palatalized consonants as either "plain" or "non-palatalized" but we do acknowledge that they are indeed velarized.

As an example of a plain-palatalized contrast, consider the laterals /l-l^j/, as in /luk/ 'onion', /l^juk/ 'hatch'. Languages that employ a palatalization contrast involving secondary articulation of a palatal gesture most commonly employ this contrast only in syllable initial position. Russian employs this contrast in both onset and coda syllable positions as well as in consonant clusters with some restrictions (Kochetov 2002). However, even though this contrast is not as common cross-linguistically in the coda position a number of other languages also share this feature with Russian (e.g. see Ní Chiosáin & Padgett (2012) for palatalization in Irish). As seen above, certain Polish dialects have secondary palatalization contrasts in onset environment; however Polish lacks a similar contrast syllable-finally. The restrictions occur in heterorganic place clusters where the first consonant would be p^j – that is, the palatalized labial cannot occur as coda before non-labial consonants (Kochetov 2002).

	Lab	**Lab-dent**	**Dental**	**Alv**	**Pal-alv**	**Pal**	**Vel**
Stops	p b p^j b^j		t d t^j d^j				k g (k^j)
Affricates				ts	ʧ		
Fricatives		f v f^j v^j	s z s^j z^j		ʃ ʒ		x
Nasals	m m^j		n n^j				
Rhotics				r r^j			
Laterals			l l^j				
Glide						j	

Table 3.1: Russian consonant table (adopted with minor changes from Hamilton 1980).

Table 3.1 presents the palatalization contrasts in Russian. It is also possible in Russian to have a three way contrast between plain consonants, palatalized consonants, and consonants + glide sequences (Ladefoged & Maddieson 1996). There are also relatively rare cases of palatalized consonants followed by a high front vowel and palatal glide (e.g. $C^{j}ij$). These sequences are infrequent and as such the $C^{j}j$ vs $C^{j}ij$ contrast is not extensive (Diehm 1998; Babel & Johnson 2007). Examples of each type of sequence are shown below. Note how each type of sequence is represented differently in Russian orthography.

Sequence type:	orthography:	example:	gloss:
C:	пот	**p**ot	‘sweat’
C^{j}:	петр	**p^{j}**otr	‘Peter’
Cj:	пьет	**p^{j}j**ot	‘drink -3sg’
Cij:	биология	b^{j}ijolo**g^{j}ij**a	‘biology’

The contrasts involving Cj and Cij sequences are possible only in syllable initial contexts, as there are no instances of C+j in coda position in Russian. We now turn to some phonetic, articulatory and acoustic properties of palatalization in Russian.

3.2.1. Articulatory properties of palatalization in Russian

In this section I describe the articulatory properties of palatalization in Russian. Russian palatalization has been described as a primary articulation of the tongue or lips along with a secondary articulation of tongue body raising or fronting towards the hard palate (Bolla 1981; Jones & Ward 1969; Hamilton 1980; Timberlake 2004; Ladefoged & Maddieson 1996). One important aspect of secondary palatalization is the timing or coordination between the gestures involved in palatalization, as (for coronals and labials, at least) this constitutes the main difference between palatalization involving a secondary articulation, and a sequence of consonant and glide[6].

The coordination between the primary and secondary gestures in palatalization is more simultaneous than comparable consonant + glide sequences. To start, consider a description of palatalized Russian consonants given by Jones and Ward, which shows how the palatalization gesture is timed with respect to the primary gesture. They state that a palatalized consonant must contain a:

"[...] slight 'glide', like a j-sound, and known as an 'off-glide.' and [...] Between a vowel and a following palatalized consonant there is often an 'on-glide', like a faint suggestion of an i-like vowel gliding from the vowel into the consonant." (Jones & Ward 1969: 82).

According to Jones and Ward, alveolar palatalized stops, /t^j/ and /d^j/, have a different configuration than labials due to the fact that both articulators that are responsible for the primary (tongue tip) and secondary (tongue body) articulations reside on the same organ – the tongue. In these cases, the blade of the tongue makes contact with the teeth, as the tongue body raises towards the palate. Jones and Ward (1969) in their text provide specific instruction for English speakers on how to accurately produce these sounds:

---

[6] For velars, however, palatalization will cause the primary articulation of the consonant to be shifted towards the palatal region more so than for the sequence of k+j.

“They must [...] make the j-element very short, trying to combine it as closely as possible with the preceding stop element.” (Jones & Ward 1969: 106).

For other alveolar consonants such as nasals and laterals, Jones and Ward give similar descriptions of the timing of the articulators as for the stops. For other places of articulation they give different descriptions, but their categorization of the timing of the gestures is roughly the same.

For labials, in the case of /p^j/, and /b^j/ there is a simultaneous articulation of p (or b) and j (Jones & Ward 1969) and “it must on no account be exaggerated to give the effect of p^j followed by the consonant j.” (p 93). It should be noted that Jones and Ward wrote their work with the purpose to serve as a pronunciation guide for English speakers learning Russian. It is not a rigorous phonetic or phonological analysis, but rather, a pedagogical tool. However, their descriptions do provide useful insights into the way Russian palatalized consonants should be pronounced, and specifically in the relative timing between the primary and secondary articulators. We have seen that Russian palatalized consonants are characterized by a relatively simultaneous tongue body raising gesture along with a primary constriction gesture, as compared to sequences of consonant and glide.

For more on Russian palatalization let us consider the work of Alexei Kochetov (e.g. Kochetov 1998, 2002, 2004). Kochetov’s (2004) study investigated the plain - palatalized contrast specifically in coda position. As was the focus for his 2002 dissertation, his 2004 study’s focus was to determine which factors affect the distributional asymmetries of palatalization in Russian and other languages. Palatalization in Russian and across languages in general is more frequent in onset position than in coda position. Palatalized consonants also undergo greater restrictions in their phonotactic distribution with other consonants. In the rarer cases (e.g. Russian and see See Ní Chiosáin & Padgett (2012) for Irish) that allow coda palatalization, coronals are usually more common or frequent than labials and velars (see previous chapter or Kochetov 2002 on relative frequency between labial and coronal palatalized consonants).

Kochetov performed EMMA articulatory studies (e.g. Kochetov 2002, 2004) to investigate onset and coda plain and palatalized Russian consonants in a variety of consonantal and vocalic contexts. Most importantly for our purposes are those environments where palatalized and plain consonants occur in both onset and coda in vocalic contexts (e.g. …CV#CV, CVC#V…). By means of these EMMA articulatory studies, Kochetov has found more detailed differences in the phasing between the gestures active in palatalized consonants in different contexts and of different places of articulation. Following is a summary of Kochetov's findings for the coordination of TB and consonantal (t or p) gestures for onset and coda palatalized labial and coronal stops.

The findings for labials are discussed first. Russian palatalized labials in onset position were found to coordinate the TB gesture more closely with the release of the LA gesture than with other landmarks such as onset or target of LA. Coda labials were found to coordinate the TB gesture more closely with the onset of LA, than with other landmarks. Thus, the results suggest a positive lag of TB and LA for onset labials, and a negative lag for coda labials. Kochetov found on average, less lag in coda position than onset position for labials in general.

As for the labials, both onset and coda coronals (stops and nasals) were found to coordinate the TB gesture more closely with the closure of the TT gesture than with another landmark such as onset or release of the closure gesture (Kochetov 2006a). Taken together, these results from Kochetov's work suggest that the TB gesture in palatalized labials is timed with slightly greater lag (positive in onset, and negative in coda) than that for palatalized coronals.

Kochetov also collected perception (consonant identification) data from Russian speakers and Japanese speakers (Kochetov 2002, 2004). Participants were asked to identify plain and palatalized labials and coronals in different contexts (coda, onset, consonantal and vocalic environments). One result was that Japanese listeners were worse than Russian listeners in identifying coda consonants, which is evidence for language-dependent factors in perception. Another result is that both groups of listeners were significantly more successful in identifying onset consonants than coda consonants. This is taken as evidence for language independent factors affecting perception of syllable-position differences, due to the observation that Russian employs all consonants in coda position, whereas Japanese employs almost none in coda position. These studies constitute further support for arguing that language-independently, onset position is a privileged position for perception of consonant contrasts (Kochetov 2004).

Kochetov provides an explanation based on perceptual reasons for the timing difference between labials and coronals. The explanation is that Russian palatalized coronal stops are often produced with an audible short period of frication that accompanies the palatalization. Palatalized labials lack this period of frication, and as such, for perceptual saliency, a relatively longer palatal offglide could be implemented for labials, over coronals.

In this subsection we have reviewed some important work investigating the timing of articulators involved in Russian palatalization and some perceptual consequences. We have seen results from certain studies that show differences in timing between TB and consonantal gestures for labials, coronals and onset and coda syllable positions.

3.2.2. Acoustic properties of palatalization

I now focus on the acoustic properties of palatalization, describing two of the more salient and important acoustic properties of palatalization in Russian – raised F2 and consonant release.

3.2.2.1. Raised second formant

Acoustically, palatalization is characterized by a high second formant (F2) (Halle 1959; Bolla 1981; Kochetov 2002). For stops, more specifically, which lack spectral information during the closure, the transitions between the consonant and surrounding segments carry the important changes in F2. Thus, for /t^j/ in onset or intervocalic position, F2 will fall from a high frequency into the following vowel (as long as it is not a high front vowel), or rise from a preceding vowel into the palatalized stop. Several studies have found average differences in F2 transition values between plain and palatalized stops.

Kochetov in his work analyzed not only articulatory, but also acoustic data in his studies (e.g. Kochetov 2002). His work found that the plain-palatalized contrast in Russian is signaled both in the final consonants themselves and also in the vowel preceding the consonants. He found that the palatalization of a following consonant is encoded mainly in the second half of the preceding vowel, and the main acoustic property of this palatalization is a raised F2. The results also demonstrate an observation important for the present dissertation. Since the F2 trajectory differed for vowels preceding plain and palatalized consonants, and since this difference was significant in the second half of the vowels, the acoustic properties are evidence that palatalization for post-vocalic consonants begins at least half way into the preceding vowel. This means that at least in coda environment, the tongue body is raising well before the closure of the consonant itself, as opposed to raising much closer to the articulation of the consonant. This does not mean that the target, or maximum constriction of the palatalization gesture is necessarily occurring before the stop, but rather that the onset of the gesture is beginning before the stop closure. In the present study, the timing of the palatalization is of interest, because it is argued that this is one of the main differences between Russian and English.

3.2.2.2. Stop releases

Previous studies have found that, in addition to F2, the spectra and duration of release bursts for palatalized stops are different to plain stops. Namely, palatalized stops on average have longer bursts with less diffuse and rising spectra than plain stops. This is especially true of palatalized coronals, and less so for palatalized labials (Kochetov 2002).

The release bursts of palatalized coronals, in addition to having rising spectra, can also exhibit considerable frication (Kochetov 2002). All of these characteristics serve to distinguish palatalized from plain stops. Of course, for laterals, nasals, rhotics and fricatives no such release bursts are present, thus only for the stops is this additional acoustic cue present.

3.3. Palatalization in English

This dissertation investigates the relationship between Russian and English in terms of palatalization. For this reason we must first introduce the extent to which palatalization exists in

English. We can distinguish the types of palatalization in English in two dimensions, the first is morphological boundary status (whether the palatalization spans morphologoical or word boundaries) and the second dimension is type of palatalization (primary, secondary or other, see below for a description).

3.3.1. Lexical Palatalization

English does not exhibit any type of palatalization like what we see in Russian. Our use of the term palatalization to describe certain phenomena in English does not entail any equivalency between the phenomena in Russian. With that in mind, what we are calling "lexical palatalization" in English will be described below, and importantly it does not occur at word or morpheme boundaries. Although this is not true palatalization it involves cases of an offglide following certain consonant contexts. This type of palatalization mostly occurs as a sequence of consonant followed by glide + high back vowel[7], so in a sense it isn't strictly a type of palatalization, however, for our purposes we will use the term palatalization to describe these sequences in English. The manifestation of this type of palatalization varies by dialect of English. In British, Canadian and some dialects of American English, several consonant places of articulation can elicit this kind of palatalization as seen in words like: "tune" [tjun], "beauty" [bjuti] and "cute" [kjut]. Most speakers of American English, however, have a more restricted distribution, exhibiting offglides only following labials and velars: "beauty" [bjuti], "cute" [kjut], "few" [fju]. Such American English speakers will lack the offglide following coronals in such environments, e.g. "tune" [tun], (Labov, 2006). Importantly, for our study, the English speaking

[7] However, it is possible that these cases are actually consonant+high front vowel sequences, and not true glides. We will discuss this possibility shortly (see e.g. Davis & Hammond 1995).

subjects were asked to read such words and it was verified that no speakers exhibited the offglide following coronals, only velars and labials.[8] It is important for our purposes, however, to point out that what we are calling lexical English palatalization and Russian palatalization are different in that in Russian palatalization involves a single phonological segment involving simultaneous coordination of articulatory gestures, whereas in English it involves multiple phonological segments involving less simultaneous coordination of articulatory gestures. Further discussion on these differences will appear below.

3.3.2. Post-lexical or phonological palatalization

In addition to those examples just described, English exhibits a kind of palatalization that occurs at morpheme boundaries. This type of palatalization is distinct in both the environment (lexical vs post-lexical) and the phonetic manifestation. Phonetically, this post-lexical type of palatalization occurs when a non-palatal consonant is produced in close proximity to a palatal glide at word or morpheme boundaries. When this happens, the following palatal glide will cause the preceding consonant to palatalize, most often exhibiting primary palatalization as described in previous section, that is, the primary articulation of the preceding consonant will undergo tongue body raising and/or fronting to a more palatal position. To exemplify, English speakers have been found to exhibit extensive post-lexical palatalization of sequences of coronal + glide or high vowel, as in “miss you” /miss ju/ -> [miʃju] or “get you” /gɛt ju/ -> [geʧu]. This phenomenon causes the alveolar fricatives or stops to sound like alveopalatal fricatives or

[8] The reader should also be aware that the instances of consonant + onglides in American English are also restricted not only to labial or velar consonantal contexts, but high back vowels as well. This should be kept in mind when discussing the relationship between English and Russian.

affricates (Zsiga 2000).

3.3.3. Distribution

As seen above, two different kinds of palatalization occur in English, the first, lexical kind, is just a sequence of a consonant followed by a glide (or high front vowel) in some lexical items. Crucially, this kind of palatalization can only ever occur in onset position. English does not exhibit any form of lexical palatalization in coda environment. However, in the coda position, English does have several vowels which can be analyzed as rising diphthongs (e.g. [ai] and [ei]) which can be compared to the onset lexical palatalization types described earlier. These syllable nuclei will be discussed in more detail below as well as possible assimilation categories for the PAM for coda consonants.

It is also true (more on this below) that the type of palatalization English exhibits is very different from the kind Russian exhibits. In terms of the post-lexical kind of palatalization, this type of palatalization only ever occurs at word or morpheme boundaries, however, it too also occurs at syllable onset position. We can never have this type of palatalization occurring in coda position or utterance finally. In this way too, this type of palatalization differs from Russian. Below we go into more detail on the differences between Russian and English palatalization types.

3.4. Differences between Russian and English Lexical Palatalization

As discussed above, Russian coordinates the secondary palatalization tongue body (TB, raising and fronting) gestures and primary articulation tongue tip, or labial (TT, LA) gestures simultaneously. In our comparison in this section we start by discussing how the Russian palatalization configuration compares to English lexical palatalization (sequences of consonant + [i] or [j])[9]. First, the most obvious difference is distributional – Russian allows for palatalized consonants to exist in both onset and coda positions, while English lexical palatalized sequences only occur as consonant clusters in onset position. However, if we consider the cases involving English rising diphthongs, the tongue body (TB) is moving towards a palatalized position towards the end of the syllable nucleus in words like "bait" and "fight".

Where Russian and English can be compared, English exhibits more sequential timing of such gestures in instances of English palatalization (Diehm 1998; Kochetov 2002; 2006a; 2009). The studies by Kochetov which were crucial in establishing timing relations between the gestures of Russian palatalized consonants of different places of articulation do not however, directly compare English and Russian timing properties. However, Diehm's (1998) study investigated the perception and production of onset palatalized consonants and consonant+glide sequences (p^ja vs pja) by Russians and English speakers and she collected acoustic data for both Russian and English speakers producing lexical palatalization examples in both languages. Diehm's work is important in establishing some facts about English lexical palatalization. She analyzed cases where English speakers produce instances of lexical palatalization as sequences of consonantal gesture followed by tongue body gesture. In using terminology of Articulatory Phonology, these sequences would be analyzed as a consonant cluster of competing gestural organization in the onset. That is, for a sequence like initial [bj] in "beauty" we can analyze the palatalization as a case of consonantal gestures LA (for [b]) and TB

---

[9] From now on, unless we are explicitly referencing the possibility of high vowel as onglide, we will refer to this configuration as consonant + glide, or CJ throughout this dissertation.

(for [j]) as being coupled in anti-phase mode with each other, but together in in-phase mode with the following vowel. The crucial thing here is that, being coupled in anti-phase mode with each other, they should exhibit greater gestural lag than similar Russian palatalized consonants, which should exhibit much closer gestural coupling.

Diehm's analysis showed that the timing between the palatalization TB gesture and the LA gesture for Russian [p^j] is more simultaneous than for the same gestures in native English speakers' productions of English /p+j/ tokens as heard in words like "beautiful." Although Kochetov's (2006a) findings show that the timing between the TB and LA gestures in /p^j/ is not strictly simultaneous in that the targets for both TB and LA gestures are not achieved exactly synchronously, the phasing between them is tighter than for English TB and LA sequences. It should also be noted that all the cases of lexical palatalization in English that Diehm studied are for labial + glide sequences, because these are the only cases in American English of stop + glide. The basic difference between Russian and English is that Russian coordinates the TB and consonantal gestures much more tightly than does English for consonant + glide sequences, which can be analyzed as clusters.

The work by other researchers described in this and previous sections together show that for stops, the gestures involved in Russian palatalization are more closely coordinated than those involved in English. In the next paragraphs I describe further differences between Russian and English consonants looking at post-lexical or morphological palatalization.

3.4.1. Morphological palatalization

We move on to describing the pattern of English palatalization seen at word boundaries. This type of palatalization was studied by Zsiga (2000). She investigated English and Russian consonant overlap at word boundaries. She studied different types of consonant clusters spanning word boundaries, but in addition she looked at instances of word final /s/ followed by word initial /j/. The premise of the study was to investigate consonantal overlap differences between English and Russian at word boundaries. Russian has been found to exhibit less gestural overlap than English at word boundaries, and so Zsiga set out to investigate to what extent s and j overlap in each Language, and compared the overlap of the two consonants with singleton /s/ and /j/ as well as the fricative /ʃ/ and in Russian the palatalized /s^j/. The hypothesis was that English exhibits greater gestural overlap at word boundaries than Russian and so the prediction would be that for sequences of /s/ + /j/ the fricative /s/ would remain closer to [s] throughout the sequence in Russian, whereas for English there would be a realization of the /s/ + /j/ sequence as somewhere between /s/ and /ʃ/ (Zsiga 2000).

Zsiga (2000) found that indeed there is greater gestural overlap in these environments in English than in Russian and that the gestural overlap in English /s+j/ sequences is different both from the Russian /s/ + /j/ sequences as well as the Russian palatalized fricative /s^j/. English /s+j/ sequences can be characterized with gestural overlap and blending, to result in a sequence of s changing into a post-alveolar fricative ʃ through the consonant sequence. The palatalization does not span the entire sequence, however, as it tends to increase towards the end of the fricative. Zsiga states:

"Thus Russian /s^j/ and English /s+j/ both exhibit overlap of the coronal and palatal gestures. They apparently differ in the timing of the two gestures, and in the care taken to keep the two simultaneous articulations separate." Zsiga (2000: 95)

The main point to take away from this work is that indeed, English exhibits what we can call primary palatalization at word boundaries, where word final consonants followed by word initial /j/ can undergo gestural overlap causing the articulator of the first consonant to be produced with greater tongue body raising or fronting. However, we can also see from this section that this type of palatalization is different from the lexical palatalization seen in Russian.

In this section I have outlined how Russian lexical palatalization differs from the different types of palatalization exhibited in English. We have seen that English exhibits certain types of lexical palatalization which are consonant sequences of consonant + /j/. In American English, the consonant will most always be a labial (e.g. “beauty”, “few”) and the vowel following the glide is always a high back vowel (/u/). The second type of palatalization seen is at word boundaries, where gestural overlap occurs between the word final consonant and word initial /j/. Both cases are different from Russian in the following ways. First, Russian palatalized consonants occur in different syllable context, whereas English palatalization only occurs in onset position for lexical palatalization, and spans word boundaries in phonological palatalization. Thus, the distribution of palatalization in the two languages is very different. Second, the actual articulation of the palatalization is different. For lexical palatalization in English, we see an anti-phase coordination of the two consonant gestures (consonant and /j/ gesture), while for Russian we see a much tighter coupling of the two gestures. For phonological palatalization in English, we see a different pattern, namely a blending of the consonantal and palatal gestures, whereas for Russian the articulation of the consonantal and palatal gestures are quite distinct (Zsiga 2000).

The purpose of this section has been to show that English lacks sequences that resemble

Russian palatalization. Although some sequences in English can involve similar gestures as Russian palatalization (e.g. consonant + glide), there is ultimately a difference in the timing or coordination of the gestures between the two languages.

3.5. Previous studies on non-native exposure to Russian palatalization

This section describes previous work on the perception and production of Russian palatalization by non-Russian speakers. This topic has been studied to varying degrees by Kochetov (2002), Diehm (1998), Babel and Johnson (2007), and in the pedagogical literature (e.g. Jones and Ward). Diehm (1998) performed a series of perception and production experiments with native Russians and native English speaking learners of Russian. The experiments tested speakers' perceptual and production abilities of certain Russian contrasts varying in degree of palatalization in onset CV syllables; CV - C^jV - C^jj - C^jij. The term "degree of palatalization" refers here to a theoretical continuum of palatalization of a given consonant or sequence, ranging from one extreme of no palatalization, to simultaneous palatalization (C^jV) to an offglide sequence C^jjV, and finally to the extreme of a sequence of palatalization, high vowel and glide (C^jijV). Thus, in Diehm's view, the term "degree of palatalization" represents the duration of the palatal(ized) portion of a given sequence.

In Diehm's study, all participants produced and perceived tokens of each of the four types of Russian sequences. They also produced English words containing types of English palatalization described earlier. The perception tasks consisted of forced choice identification tasks on the same tokens spoken by a native Russian speaker. The tasks involved the subjects circling on a sheet of paper the orthographic representation they best felt matched the token heard. The native English speaking participants were at an advanced level in learning Russian, and as such were able to perform identification using Russian orthography. The response variable for analysis was % correct identification for each token. A further analysis involved comparisons of the perception data to the production data.

Both production and perception results are now summarized. The author determined the accuracy of non-native Russian participants' productions by measuring F2 offset frequency and duration. Higher F2 values were found at the C-V transition for C^jjV and C^jijV than for C^jV. To explain the difference in F2 values observed between C^jV and sequences involving a full glide, she concluded that the tongue is in a higher position for the sequences involving a palatal glide than for sequences involving secondary palatalization. An alternative proposal is that the tongue begins to lower from an equally raised position earlier for C^jV than C^jjV. This would be consistent with a view that treats C^jV as having more simultaneous articulation of C and j than C^jjV, which would constitute a timing difference only.

To compare their Russian productions to their own native English palatalization productions, some English participants produced English tokens containing the kinds of palatalization seen in words like "beauty." Learners produced longer (later) palatal durations for English sequences than for Russian palatalized (C^jV) sequences, and the palatalization in their Russian sequences was in turn longer than for native Russians' productions of Russian palatalized (C^jV) sequences, thus the author concluded that learners had an "intermediate L2 phonetics." In terms of the different Russian sequences, learners generally did not produce C^jV, C^jjV and C^jijV differently, whereas native Russians did.

With regards to the perception data, Diehm found that learners successfully identified all three palatalization types more often than native Russian speakers identified palatalization + ij sequences (C^jij). That is, native Russians did not successfully perceive the difference between consonant + palatal glide sequences (C^jjV) and consonant + high vowel + palatal glide sequences (C^jijV), whereas English speaking learners of Russian were able to.

Taking production and perception data into account, she concluded therefore that learners hear a difference that they do not normally produce, namely C^jV and C^jjV, whereas native Russians do not normally hear a difference that they do produce, namely: C^jjV and C^jijV. Diehm concludes from this surprising result that native Russians perceive Russian "linguistically" while learners perceive "phonetically", and so learners were able to perceive the difference between C^jj and C^jij sequences better than native Russians. However, Diehm and others (Babel and Johnson 2007) acknowledge that this contrast has very limited functional load in Russian. C^jij is a relatively rare sequence and it constitutes a near merger with C^jj, which could affect native Russians' perception of these sequences. Also, given that only one identification task was carried out in Diehm's study, further research into the perception and production of Russian palatalization by non-Russian speakers is warranted.

Babel and Johnson (2007) conducted a perception study on a subset of these Russian sequences. Unlike Diehm's, their study involved a discrimination task and a rating task. The discrimination task was an AX same/different paradigm, and the rating task asked listeners to rate the perceptual similarity between tokens. The study found no statistically significant difference in how native Russians and English speakers discriminated the tokens, but they did find significant differences in rated perceptual similarity. English speakers rated the palatalized (C^jV), palatalized + glide (C^jjV) contrasts as more similar than native Russians. This study shows that although English speakers were able to discriminate different levels of Russian palatalization, they differed in how similar they considered them as compared to native Russians. They also demonstrated that discrimination for English speakers was better for faster response times, perhaps eliciting a more "phonetic" mode of processing. The two separate tasks were designed to illicit different modes of perception, the AX discrimination targeted psychoacoustic perception while the rating task targeted phonetic or linguistic perception. As such, the differences seen in participants' performance in the two tasks on the same stimuli indicated that Russians and English speakers differed in their linguistic perception and not in the psychoacoustic perception.

The studies reviewed above provide a starting point for the studies proposed in the present dissertation. This dissertation will examine the perception and production of Russian plain consonants, palatalized consonants and palatalized consonants + glide by both English speakers and Russian speakers. The present study also examines the contrast in coda position, unlike previous studies. Also unlike the previous studies, the present dissertation employs both discrimination and identification tasks with a rating component, as well as perception of synthetic stimuli, in order to examine the perception of the timing of articulatory gestures involved in palatalization. We now discuss the hypotheses for the current study, and following chapters introduce the methods and stimuli.

3.6. Importance of the Relation between English and Russian for the current study

In this section I describe the specific consonants and contexts to be investigated and their reasons for inclusion. Then in the next section I provide an overview of the hypotheses and further expand on these hypotheses of perception, production and learning. This study consists of several perception tasks and one production task. Before the consonants and hypotheses are presented in greater detail, some brief background on non-native speech and perception tasks is provided.

The perception tasks employed here are similar to traditional cross-language perception tasks, in that the perception of non-native *contrasts* are investigated. Pairs of consonants that contrast in one language (Russian) are to be presented to speakers of a language that lacks this contrast (English). One way in which an L2 contrast can differ from any in the L1 is if a phone in the L2 is non-existent in the L1. Such an example is the Farsi /G/-/g/ contrast which is absent due to the absence of uvular consonants in English. A different way in which a contrast can be non-native is when a language lacks both members of a contrast as phonemes. An example of this is the case of Japanese speakers' difficulty with English liquid /l/-/r/ contrast as Japanese differs from English in lacking a rhotic – lateral contrast.

Considering palatalization the timing of the two gestures themselves is non-native to English. This means that English lacks any segments which coordinate the tongue body gesture in a tightly coupled manner with a consonantal gesture as occurs in for Russian palatalized consonants. However, as seen above, English does have sequences of consonants preceding palatal glides, as in "beauty" and "few." We have seen that these sequences of English (labial) consonant + glide (Cj) are present at syllable onsets, and we have also seen that diphthongs can represent analogous sequential TB and consonantal gestures in syllable coda (e.g. [bejt or beɪt]). Russian allows palatalized consonants in coda position. Given these observations, and the relation between English and Russian, several predictions arise.

This study investigates a series of Russian contrasts which relate in different ways to the phonology of English. The primary focus of the study is an investigation into non-native perception and production of palatalization and in doing so, I examine consonants contrasting in degree of palatalization (plain, palatalized, Cj), differing in place of articulation (labial vs coronal) and syllable position (onset versus coda).

More specifically I examine voiceless stops /t/ and /p/ with varying degrees of palatalization. This will yield the following consonants and sequences; in the non-palatalized series we have /t/ /p/, in the palatalized series we have /t^j/ and /p^j/; and for the consonant + glide series we have /t/+/j/ and /p/+/j/.

In addition to degree of palatalization, we also examine place of articulation (coronal versus labial), and syllable context (onset versus coda). Ultimately, the questions this study will address are, whether non-native speakers perceive a synchronously coordinated TB palatalization gesture, and if so, how they perceive the timing of the coordination; whether and to what extent differences in such perception occur in different places of articulation, and syllabic contexts. These issues will then be compared to production of the same, focusing on the link between perception and production across groups of listeners varying in L2 experience.

3.6.1.1. Degree of Palatalization

In the onset, Russian consonants can occur in two degrees of palatalization to yield 3 way contrasts. This 3 way contrast is between plain (e.g. [t]), palatalized ([t^j]) and palatalized-j ([t^jj]) consonants, or sequences. This is really a difference in timing, but we can think of this as also a difference in degree, in that for consonant+glide sequences, more of the palatalization will extend following the initial consonant.

How English speakers perceive each of these types of consonants will depend on whether English speakers perceive a palatalization gesture, and to what extent the non-native timing affects speakers' perception[10]. Numerous studies have found differences between how English consonants are timed in the onset and coda (Browman and Goldstein 2000, Gick et al. 2006; and others). Syllable-initial consonants are timed in an in-phase relation (simultaneously) with the vowel while at the same time are timed in a sequential anti-phase mode with each other (Marin & Pouplier 2010) thus resulting in a competing phasing relationship in the onset between consonants being in in-phase mode with the vowel, but anti-phase with each other.

The effect of this configuration is that the vowel is phased with the mean of the centers of the activation intervals of each onset consonant gesture. This observation has been called the C-Center effect (Browman and Goldstein 1988; 2000). The situation for codas in English is different. In the coda position the leftmost consonant is phased in an anti-phase mode with the preceding vowel, while each remaining coda consonant is in-turn coordinated in an anti-phase mode with its preceding consonant. Since each coda consonant is phased sequentially with its neighbor and only the leftmost consonant has a phasing relationship to the vowel, we tend to see tighter gestural coordination in onset position in consonant clusters for English than in coda position.

The situation described above can apply to English consonant clusters (e.g. [bj] sequence

[10] By non-native timing we are referring to the timing of the TB gesture for palatalization and the consonantal gesture, we are not focused on the status of velarized consonants in this dissertation.

as in “beauty”). This differs markedly from the coordination of the gestures of a palatalized Russian consonant in coda position, whereby both consonantal gestures (TB, and C) are synchronous. Therefore, syllable final inter-gestural timing in Russian palatalized consonants can be said to differ to a greater extent from English than that in onset position. Across both syllable positions assimilation patterns are used to model discrimination patterns as in PAM, and if the assimilation pattern holds, then discrimination should be worse in this position than in the onset.

3.6.1.2. Place of articulation

Russian palatalized labials and coronals differ in at least two ways (other than place of articulation), first in timing between the articulators, and second with allophonic experience and relative similarity to English. Both of these differences could influence how English speakers perceive Russian labials and coronals.

In terms of timing, as already mentioned, several studies (by Kochetov and others) have found that in general, in onset, palatalized labials time the TB with the release of the LA, while coronals coordinate the TB with the target of TT. This means that articulators undergo greater lag in palatalized labials, and they should be perceived more like English multi-segmental p+j sequences, than their coronal counterparts.

On the second difference between labials and coronals, some English speakers will have instances of coronals+glide in their native English phonology. These sequences are mostly limited to high back vowel contexts as in “tune” and “assume”, constituting a certain degree of allophonic experience with sequences that could be considered phonetically similar to Russian palatalization. Speakers (all of the English participants in this study) who lack offglides following coronals are predicted to more readily assimilate palatalized labials to b+j sequences than palatalized coronals, since they would have much more allophonic experience with such English sequences involving labials over coronals.

However, in addition to the notion of allophonic experience, or "non-phonemic" experience of the type described in Burnham (1986), psychoacoustic salience can also play a role in perception. Here the inclusion of coronal stops in the present study is important. The studies on Russian palatalization have found that the plain/palatalized coronal stop contrasts are relatively acoustically salient compared to other contrasts, or at least compared to the palatalized – plain labial contrast (Kochetov 2002; Ní Chiosáin & Padgett 2012). Palatalization in general produces salient acoustic consequences, but to a greater degree for coronal stops. The coronal palatalized stops are often pronounced with a high energy, high frequency release burst which is not present in the plain counterparts nor in the palatalized labials. This release burst is more prominent in palatalized coronals than in consonants of other places of articulation (Kochetov 2002).

The palatalized coronal stops are therefore important for testing both allophonic experience, as well as acoustic saliency. In order to more fully test the factors of psychoacoustic saliency, timing of articulators, and allophonic experience, an additional manner of articulation is ideal.

3.7. Hypotheses: Perception of naïve group by syllable position

This first section deals with predictions and hypotheses for the perception studies for the naïve group (no Russian experience). I first provide predicted assimilation patterns for individual consonants. Then I provide assimilation patterns and their predicted discrimination rate (high/low) for the contrasts given the patterns for individual consonants. I lay out these predictions for the contrasts in 3 contexts: onset, coda +vowel, and coda + glide. As will be shown below, in some cases the Russian consonants will be predicted to assimilate to single English categories (e.g. /p/ or /t/) while in other cases the consonants are predicted to assimilate to multi-category English sequences (such as /pj/ or /tj/). PAM as described above models the assimilation of non-native speech sounds onto the native L1 phonetic and phonological categories. What is interesting about the phenomena described in this dissertation is that we expose the possibility of speakers assimilating speech sounds into multiple native phonological categories. If we take the principles of Articulatory Phonology and assume that the primitive phonological entities are articulatory gestures, and not segments, then the issue of a single Russian segment assimilating for an English speaker to multiple segments is moot. If we consider that the Russian segment is a constellation of at least two gestures (TB and LA for example) and the same is true for the English sequence of two segments (TB and LA) it becomes easy to describe this possibility. We expand on this further below.

3.7.1. Perception Onsets

Consistent with the PAM, the following predictions are based on articulatory differences between English and Russian. The following gestural scores (Figure 3.1 and Figure 3.2) represent English CjV syllables, (as in "beauty"), and Russian palatalized CV syllables and show how Russian palatalized stops differ from closest English sequences. These figures are schematic representations of sequences generated in the TADA system (Task Dynamic Application; Nam et al. 2004) (see Chapter 7 for more on TADA). For our purposes here, the TADA system allows the user to construct gestural scores from text inputs. In this case I constructed three figures, one representing the American English labial + /j/ sequence (as in tokens like “beauty”). This representation shows the sequential (or anti-phase) coordination of two articulatory gestures, the LA (labial) and TB (tongue body) gesture representing the glide /j/. The green highlighted boxes represent the activation trajectories of these two gestures, and as shown they do not overlap. We can assume for our purposes that the Russian consonant + glide sequences will have a similar representation. All other gestural activations can be ignored for our discussion.

The second and third figures represent Russian palatalized onset consonants for both labial and coronal place of articulation. The labial palatalized consonant is represented by activations of the LA (labial) and TB representing the palatalization gesture. In this case, in contrast to the sequential pattern in the first figure, we see that the activation trajectories overlap considerably. Finally, we see the gestural activation pattern for the onset palatalized coronal consonant. Recall from Kochetov’s analyses that onset coronal Russian palatalized consonants were found to have more tightly coordinated palatal and consonantal gestures than labials. This is represented in the third figure by much more overlap in the two relevant (green boxes) gestures.

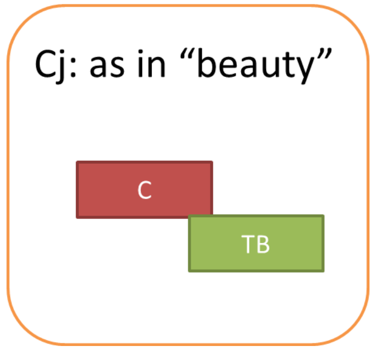

Figure 3.1: Gestural activation trajectories for English (or Russian) labial consonant + glide sequence (e.g. [bja]). The green highlighted boxes represent the two relevant (LA and TB) gestures and their relative timing.

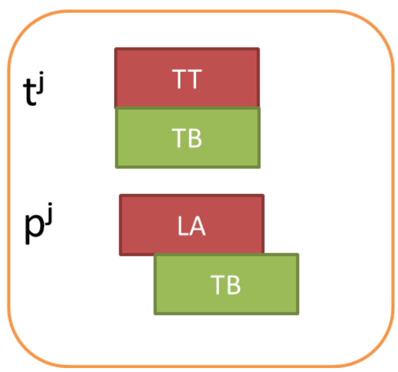

Figure 3.2: Gestural activation trajectories for Russian onset coronal (top) and labial palatalized consonant (e.g. [t^ja] and [[p^ja]).

For onset labials I predict the following assimilation patterns, each based on articulatory similarities and differences between English and Russian. The plain labial stop /p/ should assimilate as English /p/. Several possibilities exist for the palatalized labial /p^j/. The first possibility is that speakers will perceive a sequence of p+j, and assimilate this consonant to English multi-category /pj/. This is possible because Russian palatalized labial stops consist of labial and TB gestures, although timed more simultaneously, the same gestures involved in a sequence of p+j. This possibility should arise for listeners if they perceive a palatal gesture correctly, but do fail to incorporate the correct Russian timing. The second possibility is that speakers will perceive a variant of plain English /p/, or some non-ideal exemplar of English /p/. This possibility is predicted due to the timing difference between English and Russian, and should arise if the more simultaneous timing of the gestures hinders the possibility of assimilation to sequential multi categories (pj). An additional option is that speakers will fail to perceive this consonant in terms of any categories in L1. Such a case would result in an Uncategorized assimilation pattern where the speech sound is not perceived as an instance of any one L1 sound.

The question of what English listeners perceive if they do not assimilate /p^j/ to either /p/ or /pj/ arises. The experiments involving synthetic speech and the production experiments that follow the actual real speech perception experiments are designed to answer this question, which phrased differently asks how exactly English speakers are perceiving the timing of Russian palatalization.

For onset coronals, I predict the following assimilation patterns. The plain coronal stop /t/ will assimilate to English /t/. As with labials, several possibilities exist for the palatalized coronal /t^j/. The first possibility is that /t^j/ would assimilate to the multi-category English /tj/ sequence for the same reasons as for the labials, namely that same gestures, TT and TB compose /t^j/ as those involved for a hypothetical tj English sequence, although timed differently. The second possibility is for /t^j/ to assimilate to a non-ideal variant, or exemplar of English /t/. A third possibility exists, that /t^j/ would assimilate to the palato-alveolar affricate /ʧ/. Finally, it is possible that this consonant fail to assimilate to any L1 category. The tendency for palatal coronals to exhibit more simultaneous phasing of the TT and TB gestures than labials, as well as the presence of the affricate as well as the less common instances of /tj/ sequences than /pj/ sequences could favor listeners assimilating /t^j/ to a non-ideal variant of /t/, or /ʧ/, over /tj/.

Depending on the assimilation of particular consonants for each participant, I predict certain overall assimilation patterns and discrimination success for contrasts. Table 3.2 shows the predicted categorization types for each Russian consonant and Table 3.3 shows for each contrast, the possible assimilation patterns for individual consonants, the assimilation type for the contrast, and the predicted discrimination success rate.

Russian consonant	**Cat: EN/NA**
p	p
p^j	p
	pj
	U/NA
pj	pj
t	t
t^j	t
	tj
	ʧ
	U/NA
tj	tj

Table 3.2: Possible predicted assimilation types of individual onset consonants.

Rows 1-3 in Table 3.3 below show for each possible assimilation category of palatalized /p^j/ (from Table 1), the most likely discrimination pattern and rate. For instance, if /p^j/

assimilates to /pj/, then the contrast /p-p^j/ (row 2b) will have a TC assimilation pattern, and discrimination will be high. If /p^j/ assimilates to some variant of /p/, as in row 1a, a SC or CG discrimination pattern should arise and discrimination success rate will be lower. Given the assimilation patterns for the consonant /p^j/, we can also predict discrimination rates for the contrast /p^j-p^jj/ seen in row 2. If /p^j/ assimilates to /pj/, then this contrast will resemble a SC/CG pattern (row 2b), and discrimination will be worse than if /p^j/ assimilates to a variant of /p/ (2a), which will yield a TC pattern and higher discrimination.

Row	Russian contrast	En Categorization	Type	Discrimination rate
1	p-p^j	p-p	SC/CG	LOW
		p-pj	TC	HIGH
		p-NA	UC/NA	HIGH
2	p^j-p^{j}j	p-p	TC	HIGH
		pj-pj	SC/CG	LOW
		NA-pj	UC/NA	HIGH
3	p-p^{j}j	p-pj	TC	HIGH
4	t-t^j	t-t	SC/CG	LOW
		t-tj	TC	HIGH
		t-ʧ	TC	HIGH
		t-NA	UC/NA	HIGH
5	t^j-t^{j}j	t-tj	TC	HIGH
		tj-tj	SC/CG	LOW
		ʧ-tj	TC	HIGH
		NA-dj	UC/NA	HIGH
6	t-t^{j}j	t-tj	TC	HIGH

Table 3.3: Assimilation, discrimination patterns of different possibilities for contrasts

Similar predictions hold for coronals (rows 4-6). If /t^j/ assimilates to /tj/, a TC pattern should arise for /t-t^j/ and a SC/CG pattern should arise for /t^j-t^{j}j/. If /t^j/ assimilates to a variant of /t/, then we should see SC/CG for /t-t^j/ and TC for /t^j-t^{j}j/. If /t^j/ assimilates to the affricate (or some other category), then we should see closer to TC assimilation patterns for both contrast types.

I have provided several possible assimilation patterns for the palatalized consonants in syllable onset position. I suggested that for the labial stop /p^j/ the most probable assimilation pattern would be multicategory /pj/ sequence, while for the coronal stop /t^j/ an assimilation to an imperfect English /t/, or /ʧ/. We could attribute such a pattern to the observation that the TB gesture exhibits greater lag in labials than coronals, that is, the two gestures are phased more simultaneously in /t^j/ than in /p^j/ in Russian, thus, English speakers should perceive /p^j/ as /pj/ more often than /t^j/ as /tj/. In summary, according to PAM, we predict more SC/CG assimilations and worse discrimination for the contrasts: /t-t^j/ and /p^j-p^jj/ and more TC assimilations and better discrimination for the contrasts: /t^j-t^jj/ and /p-p^jj/.

3.7.2. Perception Codas

The following gestural scores (Figure 3.3) represent Russian palatalized consonants in the coda position of the syllable. Although there is no possible English variant of consonant+glide in this context since English has no palatalization at all in coda position, we provide a schematic representation of a possible English variant of glide (or diphthong) + consonant case as seen in sequences like (“bait” [beɪt]/[bejt] or “cape” [keɪp]/[kejp]). Figure 3.3 represents Russian palatalized coda consonants for both labial and coronal place of articulation. The labial and coronal palatalized consonants are represented by activations of the consonantal (LA or TT) and TB representing the palatalization gesture. In this case we see that the activation trajectories overlap in both cases.

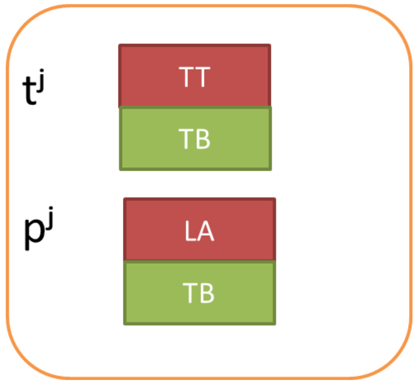

Figure 3.3: Gestural activation trajectories for Russian coda labial palatalized consonant (e.g. [ap^j]). The green highlighted boxes represent the two relevant (LA and TB) gestures and their relative timing. In this case the TB gesture is phased with the LA closure. Note: we cannot show any English figures in this section since English lacks any palatalization in this context.

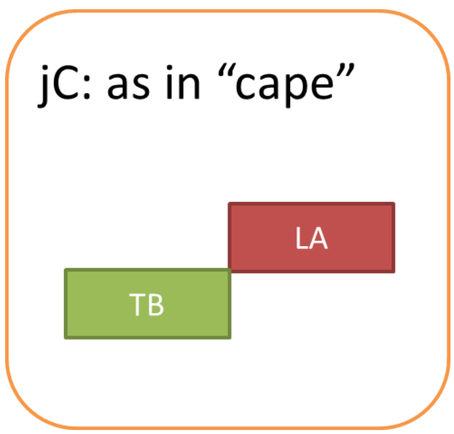

Figure 3.4: Gestural score representing glide (or diphthong) + consonant sequences as in English words like "cape" or "bait".

Two different coda contexts are investigated, word-final position preceding a vowel initial word (C#V), and word-final position preceding a glide-vowel sequence (C#j). In coda position, only one type of contrast exists, unlike onset, due to the lack of Cj sequences in this environment, which is why the additional C#jV sequences are included, to mirror the 3 way onset contrast.

In C#V (coda, pre-vocalic C) position, we expect Russian /p/ to assimilate to English /p/ even though these two consonants are not identical to each other due to the fact that Russian plain consonants are actually velarized and also involve additional phonetic differences like aspiration in English. We still expect the Russian non-palatalized variants to assimilate to English plain consonants. The palatalized consonant /p^j/ has the three following possibilities, English listeners could perceive it as a non-English Russian consonant, an imperfect variant of English /p/, a cross-word sequence of p#j, or they could assimilate it to a sequence of glide (or high vowel) + consonant. Similar predictions are maintained for the coronal stops, however, an additional possibility exists that listeners will assimilate /t^j/ to the affricate /ʧ/.

Like the onset consonants, the following predictions are based primarily on the phasing or coupling of the gestures in Russian and their relation to possible English syllables. English gestures are phased in a more sequential manner in general within the syllable, and in particular in the coda. In AP, the only relevant and established coupling relation in the coda is an anti-phase mode between vowel nucleus and the first coda consonant. The two gestures in Russian palatalized consonants exist in a roughly simultaneous, or in-phase, coupling mode. How do we expect the coupling mode of TB and TT gestures in coda /t^j/ to transfer to an English speaker's L1? I maintain that the answer lies in how English speakers perceive the timing between these gestures.

The possible assimilation categories are essentially the same as for onset consonants, however, it is predicted that palatalized consonants will assimilate less consistently, or less often to multicategory Cj sequences. This is due to the fact that English lacks any such sequences in coda position, unlike in onset position. An additional reason for this is due to the observation that syllable-finally the TB gestures tend to have reduced duration and magnitude for some Russian speakers (Kochetov 2002). In sum, if listeners perceive coda palatalized consonants as variants of plain English categories (t or p) then we expect to see more SC/CG patterns and less discrimination success than in the onset.

3.8. Hypotheses: Learning

The predictions given in previous sections applied for naive listeners with no prior exposure to Russian. However, PAM, and other theories of SLA such as the SLM provide predictions on which non-identical (to L1) consonants should have the highest probability of being learned adequately through time and exposure to the L2. The basic premise is that if a consonant assimilates consistently, or well, to an L1 category, and yet is not phonetically (in our case, articulatorily) identical to that category, it will not be readily learned. Conversely, if a consonant is not assimilated as well to any native L1 category, then this consonant should be more adequately learned. Russian provides an interesting test case for this theory. Specifically the different Russian palatalized consonants in onset and coda environments each pose different possibilities for assimilation, and as such, different predictions for learning success. More specifically, confusion matrices for each consonant for naive speakers are compared to those for more experienced learners. If the PAM (and SLM) predictions hold, we should see a correlation in confusion matrices. For example, if consonant /p^j/ is most consistently assimilated to English /pj/, while consonant /t^j/ is not consistently assimilated to a particular English category, then we might expect that learners improve their perception (and production: see next section) for contrasts involving /t^j/ over /p^j/. This would indicate that learners have learned /t^j/ more fully than /p^j/ as a result of the initial assimilation patterns.

3.9. Hypotheses: Production

The relation between speech perception and production is an increasingly prominent area of study but is still not well understood. The question of what exactly the objects of perception are has also been a matter of debate for some time. This dissertation is limited to exploring the relative ease or difficulty of perception over production in non-native speech. In this study I collect both perception and production data for all speakers for the same sequences and contrasts. One hypothesis to be tested is that perception ability should be as good, if not better than production ability. This is in line with previous findings that speakers develop perceptual abilities before production abilities, or that production errors result from perception errors (Flege 1995; Pisoni 1995). However, a few previous studies have found that some speakers are able to produce non-native sounds better than they are able to perceive them. Sheldon and Strange's (1982) for example, examined whether perception necessarily needs to precede production, specifically in language learning and in the context of teaching a L2. Both of these possible hypotheses will be tested. A more general hypothesis is that speakers will perceive and produce similarly. This would be the case if speakers produce the consonants the same way that they perceive them. This would happen, for example, if a speaker perceives /t^j/ as /tj/ and produces it the same. If a speaker categorizes the palatalized consonant as a sequence of stop + glide, then it is reasonable to assume they will also produce it likewise.

3.10. Summary of hypotheses

Given the above, we expect the following to follow from this dissertation. Onset palatalized consonants should be better perceived than coda palatalized consonants (ONS - COD). If naïve English speakers are assimilating palatalized labial consonants to sequential consonant + glide sequences, discrimination should be good for the plain-palatalized labial consonants. If, also, naïve English speakers are simply making use of salient psychoacoustic properties present in the release of the coronal palatalized labial, discrimination for the palatalized contrast should also be good. In essence, it is predicted that if two members of a contrast are assimilated to different categories (whether one of the assimilated categories is present in the L1 as a phonological category or not) this should yield better discrimination.

Chapter 4 Experiments

In this chapter I describe the three perception and production tasks employed in this dissertation. First, to test the PAM, both categorization and discrimination data were collected. Discrimination tasks were performed first to minimize any stimulus categorization effects on performance (Best et al 2007). Second, a forced choice identification task was administered as the categorization task. The third task was an AX discrimination task involving synthetic speech. The fourth task was a production repetition task.

The experiments were conducted in a single session for each participant, lasting between 60 and 90 minutes, depending on individual speed in completing the tasks. The participants performed the tasks while seated at either a Mac or PC running PRAAT MFC (Multiple Forced Choice) experiment software (Boersma & Weenink 2007). The auditory portions of the experiments were presented to the participant through Sennheiser headphones. The productions for the production task were recorded either with the same Marantz recorder, or a USB microphone directly connected to the PC or MAC at a 44khz sampling rate. The experiments all took place in the Phonetics Lab or sound booth at the linguistics department of Yale University.

Some remarks on the nature of the identification, or categorization, task are warranted. Certain difficulties arose in eliciting responses for the categorization of individual consonants by non-native speakers of Russian and naïve speakers with no knowledge of Russian due to orthographical differences between Russian and English. Naïve listeners have no knowledge of Russian or its orthography and as such cannot perform forced choice tasks using Russian orthography, nor can they write down Russian words and English orthography provides no way to accommodate palatalization. The closest way to represent palatalization in English is writing the consonant followed by "y" or "i" as is done in common English transliterations of Russian palatalized consonants (e.g. "Tanya/Tania" for Таня; "Soviet" for совет). It is impossible for English speakers to provide responses or transliterations of actual palatalization on the consonants themselves (unless of course they are familiar with the IPA).

For the forced choice options in the categorization task for naïve listeners the palatalization is represented somewhere other than the consonant itself, namely the surrounding environment through the addition of "y" or "i." Any such categorization task must then be performed with these considerations in mind. The difference between Russian and English is that Russian allows for the C^j- C^jj contrast to be encoded orthographically, whereas English does not. Below I describe a categorization paradigm that attempts to mitigate these issues.

The third perception task involving synthesized speech was carried out as a way to acquire more detailed information on listeners' perception of the timing of gestures. This task involves a series of tokens created by manipulating the relative timing between primary (TT or LA) and secondary palatalization (TB) gestures.

4.1. Materials and participants

In this section I describe in detail the materials and participants. The materials are Russian sequences of labial and coronal voiceless stops differing in their degree of palatalization and syllable context. The same materials are used for all tasks except the TADA perception task which uses synthesized stimuli. Below are more detailed descriptions of the materials and participants.

4.1.1. Materials

Twelve Russian sequences constituted the tokens for the experiments involving natural speech. The sequences differed by having a consonant of interest differing in three dimensions: place of articulation (labial, coronal), palatalization (plain, palatalized, consonant+glide) and syllable position (onset, coda). Thus, the sequences contained either a labial or coronal stop, either plain or palatalized or followed by a glide, in onset or coda position.

The sequences were two-word non-word tokens of the following CV structure: CVC(J)VC. The boundary between the two words varied as a function of the syllable position of the consonant of interest and was placed either after the first vowel, or after the second consonant depending on whether the consonant of interest (in bold and underlined) was onset or coda: (CV#<u>**C**</u>(J)VC and CV<u>**C**</u>#(J)VC). The vowels, which remained constant, were low back vowels /a/. For each of the two places of articulation (labial, coronal) and each of the two syllable contexts (onset, coda), three types of contrasts are investigated: plain-palatalized (e.g. p-p^j), plain-glide (p-p^jj), and palatalized-glide (p^j-p^jj). Table 3.1 contains the list of sequences included in this study in both Russian orthography and IPA. The sequences compose the contrasts mentioned above. A table with the contrasts themselves is provided in Chapter 5.

	Labials		Coronals	
CV#CVC onset	**IPA**	**RU**	**IPA**	**RU**
plain	pa **p**ap	па пап	ta **t**at	та тат
palatalized	pa **pʲ**ap	па пяп	ta **tʲ**at	та тят
glide	pa **pʲj**ap	па пьяп	ta **tʲj**at	та тьят
CVC#VC coda				
plain	pa**p** ap	пап ап	ta**t** at	тат ат
palatalized	pa**pʲ** ap	папь ап	ta**tʲ** at	тать ат
glide	pa**pʲ j**ap	папь яп	ta**tʲ j**at	тать ят

Table 4.1: Russian plain-palatalized contrasts by place and manner of articulation and the tokens in which consonants appear. The consonants and sequences of interest are underlined and bolded.

A female native speaker of Russian from Belarus recorded ten repetitions of each sequence by reading the items off a list into a Marantz recorder with head mounted microphone at a sampling rate of 44khz. The recordings were made in a sound booth at the Linguistics department at Yale University during one sitting.

4.1.2. Participants

A native Russian speaker recorded the original tokens. For the remaining experiments, three separate groups of speakers participated. All participants were recruited on the Yale University campus and the New Haven, Connecticut area and were paid for their time. The study was approved for the IRB.

The first English group consisted of native speakers of American English with no prior exposure or knowledge of Russian. This group consisted of 10 participants (F=8, M=2). The second English group is a group of native speakers of American English with at least one semester or equivalent of Russian knowledge. This group had 8 participants (F=4, M=4). The participants in this group varied as to their knowledge or experience in Russian which ranged between 2 years and 9 years of experience. The last group was the group of native Russian speakers which had 9 participants (F=6, M=3). We should first make comment on the nature and size of the three different groups. First, no group has more than 10 participants, this means that some of the results reported in this dissertation that do not reach statistical significance could do so with more participants. It would be even more beneficial as an idea for future work to further replicate some of these studies with larger groups of participants.

Second, the fact that one of the groups, the learner group, is less homogenous with respect to language knowledge or experience than the other two groups deserves mention. Since the learner groups consists of 8 participants who differ between each other in their experience in Russian presents another problem in that any comparisons between groups have to be taken with this caveat – the fact that the group itself is not homogenous and is variable. Again, for future study it would also be beneficial to further subdivide the group into smaller sub-groups based on other factors or by years of experience, but for our purposes we present analyses at the group level, but also at the sub-group level as well.

4.1.3. Materials for subsequent experiments

As stated above, a native Russian speaker recorded ten repetitions of each of the twelve stimuli. For each stimulus, five of the ten recorded repetitions were to be used as stimuli for the perception and productions tasks. I first listened to each of the stimuli and eliminated any stimuli deemed too acoustically unique, distinct or flawed in any way. I then examined formant transitions and durations of vowels, consonants and releases to further eliminate any tokens which were overly distinct from the other tokens of the same category. The final step was to construct the AXB stimuli (see Chapter 5) to make sure that in a discrimination task any pairs of tokens did not sound overly distinct from any other pairs of the same contrast. After the five physical tokens for each consonant were chosen, they were scaled for average intensity using Praat's scale intensity function to further eliminate any non-relevant differences in intensity between tokens. The following chapters each describe a perception or a production experiment, results and relevant discussions.

Chapter 5 Task 1: AXB discrimination

The first of three perception experiments is an AXB task, which is designed primarily to examine subjects' ability to discriminate between two stimuli, in this case, two speech tokens or phones. An AXB task itself contains trials involving three stimuli presented in sequence. The first and third stimuli represent two separate categories (A and B) and the medial stimulus (X) is a member of either the first or the third category. There are 4 possible orderings for any two A – B category pairings, namely: AAB, ABB, BAA, and BBA.

AXB tasks also inherently measure the ability to categorize since the listener must hear all three members of a trial and determine whether the medial member is more like the first or third, it is implied that the listener must initially categorize the phones themselves in order to perform this task (MacKain, Best and Strange 1981; Beddor and Gottfried 1995).

5.1. Methods

As stated in the previous chapter, five of the ten recorded Russian tokens (see Table 4.1 in Chapter 4) which were most acoustically similar to each other were chosen for the experiments and employed for all subsequent tasks. Acoustic similarity was determined by examining segment durations, formant frequencies and burst qualities.

The basic unit or trial type for this type of task is a contrast. Here the term contrast refers to the actual contrasting Russian sequences from Table 4.1. Table 5.1 below presents the 12 contrasts investigated.

Onset	**Simplified contrast type**	**Labials**	**Coronals**
plain-palatalized	CV-C^jV	pa **<u>p</u>**ap – pa **<u>p^j</u>**ap	ta **<u>t</u>**at – ta **<u>t^j</u>**at
plain – glide	CV-C^jjV	pa **<u>p</u>**ap – pa **<u>p^jj</u>**ap	ta **<u>t</u>**at – ta **<u>t^jj</u>**at
palatalized – glide	C^jV-C^jjV	pa **<u>p^j</u>**ap – pa **<u>p^jj</u>**ap	ta **<u>t^j</u>**at – ta **<u>t^jj</u>**at
Coda			
plain – palatalized	VC-VC^j	pa**<u>p</u>** ap – pa**<u>p^j</u>**ap	ta**<u>t</u>** at – ta**<u>t^j</u>**at
plain – glide	VC-VC^jj	pa**<u>p</u>** ap – pa**<u>p^j j</u>**ap	ta**<u>t</u>** at – ta**<u>t^j j</u>**at
palatalized – glide	VC^j-VC^jj	pa**<u>p^j</u>** ap – pa**<u>p^j j</u>**ap	ta**<u>t^j</u>** at – ta**<u>t^j j</u>**at

Table 5.1: Contrasts examined for the perception and production studies.

The table presents the contrasts by the three relevant conditions of the experiment, namely syllable position (onset, coda), palatalization (plain-palatalized, plain-consonant+glide, palatalized-consonant+glide), and place of articulation (labial, coronal). The description of the construction of the AXB stimuli is as follows.

For each contrast or AB pairing, (e.g. pa **p**ap – pa **p**^{j}ap) four stimulus orders (AAB, ABB, BBA, BAA) were available. Three distinct combinations of different physical stimuli were constructed for each of the four possible triad types. Each of these triads was repeated two times, thus for each contrast, a total of 24 AXB triad trials (4 orders x 3 pairings x 2 repetitions) was created. The entire AXB task therefore contains a total of 288 trials (24 triad trials x 12 contrasts) for each participant.

Each trial in the experiment consists of an AXB triad, where A and B represent different members of the contrast pair and X an exemplar of either A or B. In each AXB triad, the medial member (X) is always a different physical token than either the first or third members (A or B) to control for auditory factors influencing discrimination. In this way participants make decisions about phonetic and/or phonological identities of the phones, as opposed to psychoacoustic similarities between members of the triads.

The inter-stimulus interval (ISI) within each trial was 1 second, and the inter-trial interval between trials was also 1 second. The AXB triads were randomized into two blocks ensuring that each trial was presented once before any were repeated and that no two repetitions of the same trial appeared twice in a row.

The participants were instructed to listen to each trial and decide whether the medial member X most resembled the first (A) or the third (B) and to guess if unsure. Participants marked their choice by clicking one of two buttons on a computer screen marked either “first” or “third.” A response triggered the next trial. Every 50 trials the program automatically offered the participant a break. In addition to the response, reaction time (RT) was also logged for each trial.

5.2. Analysis

In this section I describe the analyses conducted for the AXB test. The design of the AXB task is meant to answer several questions. One question of interest is whether Russian knowledge (naïve, learners, native) is a significant predictor of discrimination ability. Another question is whether the factors associated with the stimuli are significant. The factors of interest are place of articulation (2 levels: labial, coronal), level of palatalization contrast (3 levels: plain-pal, plain-glide, pal-glide) and syllable context (2 levels: onset, coda). The remaining questions are whether groups differ with respect to the different factors. This question asks whether there are interactions between groups and the remaining factors. In order to answer these questions, several statistical tests were run. The tests were conducted on two measures, percent correct response (%correct) and reaction time (RT). The two analyses differ in their response variables. The RT response variable is the log transformation of the reaction time[11] in milliseconds (logRT). RT data often do not follow a normal distribution, thus the logarithmic transformation is a common one for data such as reaction time to alleviate this problem (Baayen 2008). Tests for both %correct and LogRT are run to gain a more complete picture of speakers' abilities in discriminating the contrasts. A lower % correct response for a particular contrast is interpreted as a more difficult contrast. Conversely, a higher RT is taken as evidence for a more difficult contrast. The reason for this is that a contrast that is more difficult to discriminate will take the participant longer to provide a response.

Below I summarize the main questions and in the next sections provide the statistical tests employed to answer them.

1) Are there differences between the groups (Group main effect)?
2) Are there differences between place over all groups (Place main effect)?
3) Are there differences between syllable position (syll main effect)?
4) Are there differences between contrast type (contrast main effect)?
5) Are there differences between onset and codas for different palatalized consonants and for different groups (interactions)?

[11] What this value actually represents is response time, not reaction time. Whenever we make reference to "reaction time" we are referring to response time.

5.2.1. Data processing

Several steps were taken before the perception data were ready to be submitted to the statistical analyses. The first step was to run visualization and summary statistics on the data. The data were summarized for % correct response and reaction time (RT). First, the RT data did not seem to follow a normal distribution for all subjects. A logarithmic transformation was performed (see above) which somewhat alleviated the distribution. Next, outliers with very long or very short reaction times were removed. Such outliers and influential points most likely resulted from erroneous button presses or other errors not relating to the stimuli at hand. Extremely large and small values were RTs of less than 0.05 seconds or greater than 15 seconds[12]. After such extreme data points were removed, other data points were removed if they exhibited a RT of more than 2.5 standard deviations away from the mean. Below the analyses are described in more detail.

5.2.2. Overall analysis for % correct response

This section describes the analysis undertaken to investigate all factors and all groups for % correct response. The data were analyzed using mixed effects logistic regression models with random effects for subject and experimental stimulus[13]. This and all subsequent statistical analyses were conducted the using the lme4 package in the R statistical language (see Baayen 2008). The response variable is proportion of correct responses (%correct) and the independent

---

[12] The reaction times for this particular experiment varied by subject, but values of between 2 and 6 seconds were considered valid.

[13] The random slopes for subject and stimulus were included as they both improved the model fit significantly.

variables are Group (native Russians, naïve English speakers, Learners), Contrast type (plain-palatalized, plain-glide, palatalized-glide), Syllable position (onset, coda) and Place of articulation (labials, coronals). Interactions for all the factors were also included in the model. A second test was carried out with logRT as the dependent variable on the same dependent variables.

5.2.2.1. Results: % correct response

The results of the overall linear mixed effects logistic regression model involving all of the factors is now provided. Table 5.1 summarizes the main trends in the data. The plots show the average percent correct response for the three groups.

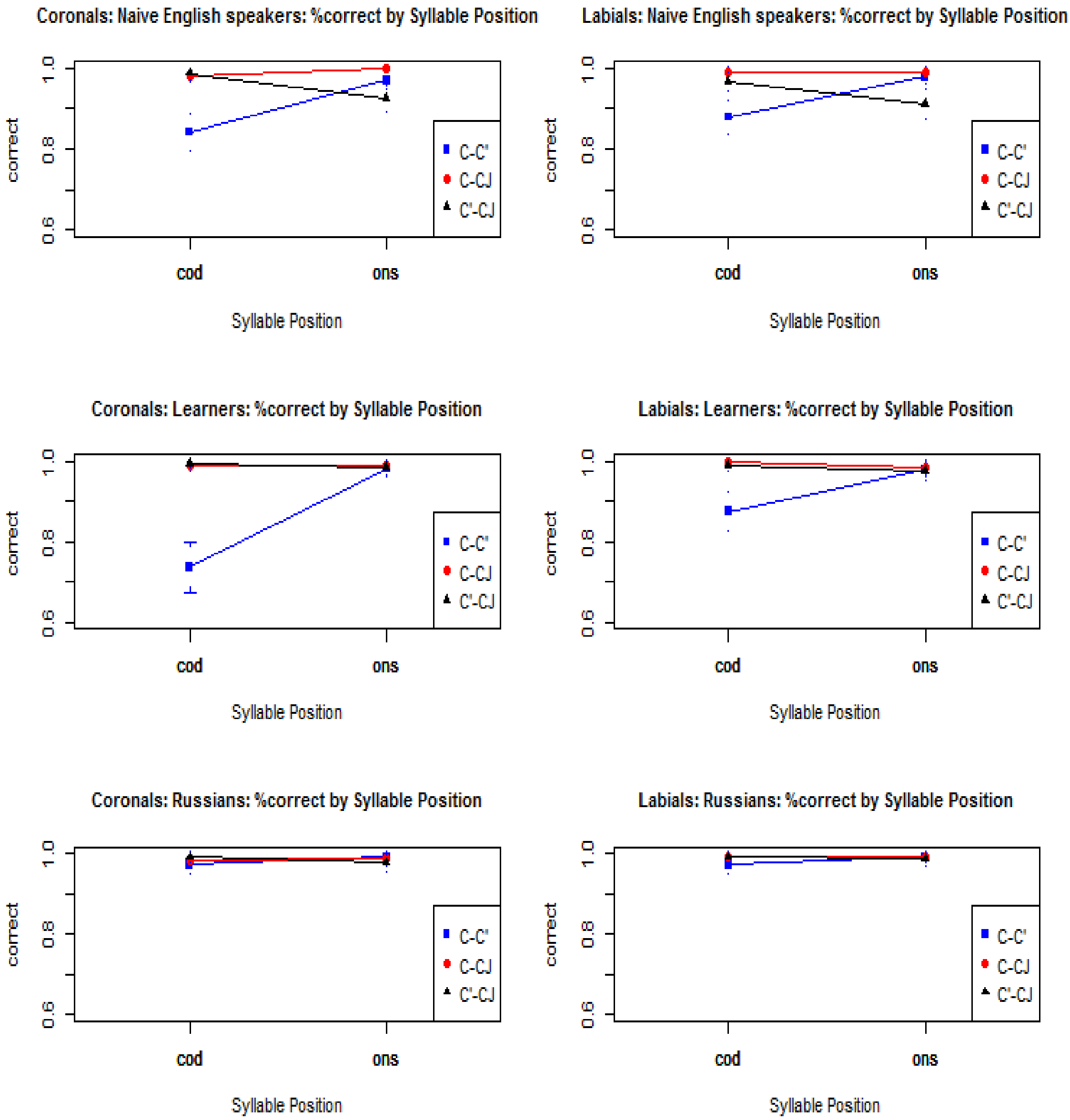

Figure 5.1: Average % correct by place of articulation. Left column: Coronals, Right column: Labials. Top row: naïve group, middle row: learners, bottom row: Russians.

Table 5.2 demonstrates possible effects of group, syllable position and place, as well as interactions between these factors. As such, the analysis described in the preceding section was carried out. As follows, significant effects are reported.

The analysis yielded significant main effects of Group (pr > z = 0.01), Syllable position ($p = 0.0002$), and Palatalization contrast type ($p = 0.0002$). No main effect of place was found. As for interactions, the only significant interaction was a syllable position by palatalization interaction ($p = 2.43e-05$). To determine if a better model could be fit, given that no other interactions nor the effect of Place were significant, several models were run removing the effect of Place as well as non-significant interactions. All of these models, as well as the more complex original, were compared using the anova() function in R for goodness of fit. The best fitting model did not contain the variable Place nor any interactions involving Place, however all other interactions were significant. For this reduced model (lacking Place and its interactions), all of the same effects and interactions, as well as additional interactions, were found to be significant.

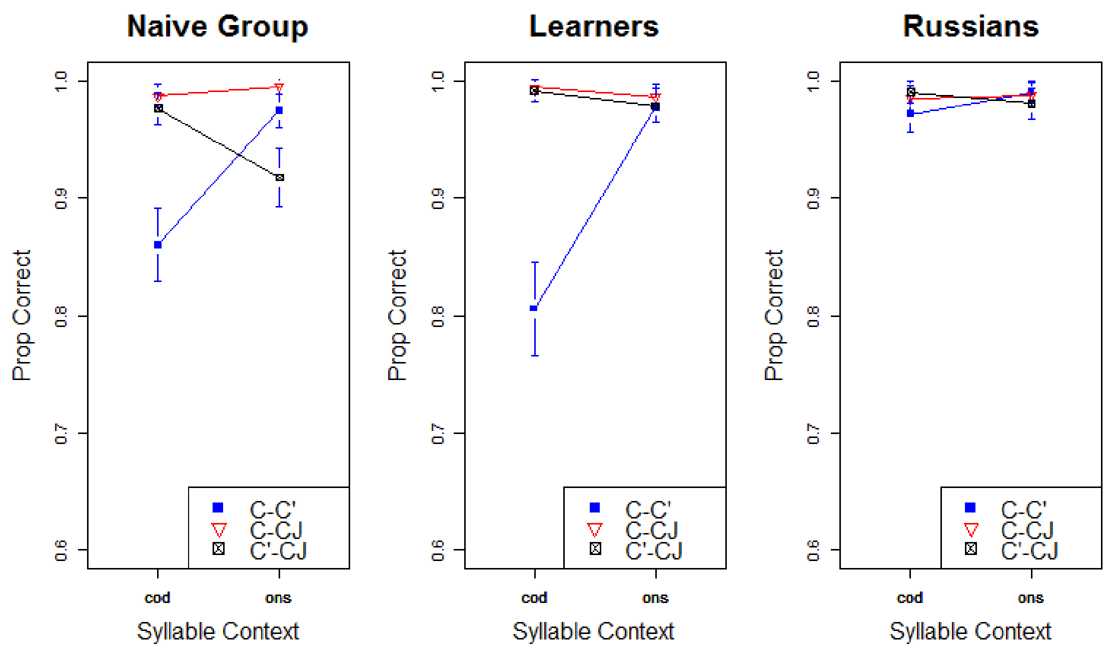

Figure 5.2: Average % correct response by contrast type for syllable positions and groups.

As in the first model, Group (pr > z = 0.001), Syllable position (p = 1.05e-07), and Palatalization contrast type (p = 1.72e-08) and the syllable position by palatalization interaction (p = 3.95e-10) were significant, and the significance level increased. Two additional interactions were significant; a group by palatalization type interaction (p = 0.003) and a group by syllable position by palatalization type interaction (p = 0.04). These significant effects can be interpreted as follows. First, the main effect of Group shows that the groups differed from each other in their success at discriminating the contrasts in general. Next, overall, onsets were better discriminated than onsets. Below I describe in greater detail more refined interpretations of these effects based on a closer examination of the coefficients and contrasts in the analysis. In our analysis, the baseline level for the factor Group is the naïve group, as such, the analysis compares the remaining levels: The learner group and the native Russian group against this baseline. So, we have information about whether the learner group and the native Russian group differ from the naïve group, but not whether the learners and Russians differ from each other. We do not encounter the same issue with the factor syllable position, because it has just two levels. In order to more closely examine whether Groups L and R differ from each other, and to examine the more complex interactions, a reveling of the factors was done, so that for example, a level other thank Group E was the baseline.

Below I present more informative results based on comparisons between the levels of the different factors. We start by examining differences between palatalization types, then syllable positions and then groups.

	Naïve		**Learners**		**Russians**	
Pal	coda	onset	coda	onset	coda	onset
C-C^j vs C^j-Cj	2.048 ***	-1.318 ***	3.55 ***			
C-C^j vs C-Cj	2.677 ***	1.84 *	3.95 ***			
C^j-Cj vs C-Cj	0.62	3.159 ***	0.4			

Table 5.2: Fixed effects coefficients for contrast type comparisons. Significant coefficients are indicated by asterisks indicating the level of significance in shaded cells14. A positive coefficient means that the contrast type on the right of the "vs" label in the "Pal" column has a significant increase in Log Odds of being discriminated correctly versus the contrast type on the left of the "vs" label for the particular Group and syllable position. For example, C^j-Cj contrasts in coda position for Group E, are discriminated significantly better than C-C^j contrasts.

[14] The significance levels are: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

I present more detailed results with regards to the factor palatalization contrast type (Pal), and any differences between onsets and codas and between different groups for these contrast types and syllable position. From Table 5.2 above, the results show that comparing plain-palatalized to palatalized-glide, (row 1); palatalized-glide contrasts are discriminated significantly better than plain-palatalized contrasts in coda position for both English groups. In onset position, the opposite pattern emerges for this comparison. The palatalized-glide contrasts are discriminated significantly worse in onset than plain-palatalized contrasts for the naïve group. Row 2 shows that plain-glide contrasts are discriminated significantly better than plain-palatalized contrasts in coda position for both English groups, as well as onset position for the naïve group. Finally, row 3 shows that plain-glide contrasts are discriminated significantly better than palatalized-glide contrasts only in onset position and only for the naïve group.

		Group		
Pal	Syllable differences	Naïve	Learners	Russians
$C\text{-}C^J$	Coda vs Onset	1.956 ***	2.547 ***	1.09
C-Cj	Coda vs Onset	1.11	-0.92	0.177
$C^J\text{-}Cj$	Coda vs Onset	-1.41 ***	-1.016	-0.72

Table 5.3: Fixed effects coefficients for Syllable comparisons. Asterisks indicate the level of significance in shaded cells. For each contrast type in the Pal column, a significant positive coefficient means that the syllable position “Onset” has a significant increase in Log Odds of discrimination performance versus the syllable position Coda, for each Group.

Now I present results focusing on syllable position differences. Table 5.3 above shows the results focusing on simple effects comparisons between syllable positions (Onset versus Coda) for each contrast type and group. For the contrast plain-palatalized, for both English groups, onsets were significantly better discriminated than codas. For the contrast palatalized-glide, the only significant result is that for the naïve group, onsets were discriminated significantly worse than codas. No other significant differences between syllable positions were found.

		Syllable position	
Pal	Group	coda	onset
$C\text{-}C^{j}$	Naïve - Learners	-0.588	0.002
	Naïve - Russians	1.827 **	0.96
	Learners - Russians	2.416 ***	0.959
C-Cj	Naïve - Learners	0.69	-1.35
	Naïve - Russians	-0.167	-1.1
	Learners - Russians	-0.86	0.247
$C^{j}\text{-}Cj$	Naïve - Learners	0.92	1.312*
	Naïve - Russians	0.87	1.561*
	Learners - Russians	-0.044	0.248

Table 5.4: Fixed effects coefficients for Group comparisons. Significant coefficients are indicated by asterisks indicating the level of significance in shaded cells. For each contrast type in the Pal column, a positive coefficient means that the group on the right of the “vs” label in the “Group” column has a significant increase in Log Odds of discriminating that contrast versus the group on the left of the “vs” label for the particular syllable position. For example, for C^{j}-Cj in coda position the native Russian group discriminated significantly than the naïve group.

The last group of results focuses on Group differences. Here the reader is referred to Table 5.4 above. The results are presented as comparisons between groups for each palatalization contrast type and syllable position. For the plain-palatalized contrasts in onset position, the native Russian group presented significantly better discrimination than both English groups. There were no significant differences between either English group for this contrast neither in coda position nor between any of the groups for this contrast in onset position. For the plain-glide contrast, no significant differences in discrimination were found between any of the groups in either syllable position. Finally, for the palatalized-glide contrast, significant differences were found only for syllable onset position between the naïve group versus the other two groups. Specifically, for this contrast type in onset position both the learner group and the native Russian group had significantly better discrimination than Group E.

5.2.2.2. Discussion: % correct response

The results for the mixed effects logistic regression for % correct response are discussed in further detail and with regards to the predictions outlined in earlier chapters. Overall, in terms of palatalization, the least problematic and best-discriminated contrast was the plain-glide contrast, as was predicted. The difference between plain (again, non-palatalized, or velarized) and glide consonants involves the presence or absence of a palatalization tongue body raising and fronting gesture. Also, Russian glide consonants are produced with considerably more lag with respect to the consonantal and palatal gestures than are palatalized consonants. All groups were able to discriminate these contrasts well. The two more problematic or difficult contrasts were the plain-palatalized contrast and the palatalized-glide contrast. However, there are significant interactions between all the factors tested, as was shown in the preceding section. Participants discriminated these contrasts better or worse depending on which group they belonged to and in some cases depending on whether the contrasts occurred in onset or coda position. The native Russians showed no significant differences in their discrimination abilities for any of the contrasts in any syllable position.

The remaining discussion will focus on the remaining two contrasts. For these two contrasts, the two native English groups showed some similarities to each other but also some differences. For the plain-palatalized contrasts, both English groups discriminated this contrast significantly worse than other contrasts in coda position. Thus both groups had more trouble discriminating a plain (or velarized) Russian consonant from a palatalized consonant in coda position, than either the plain-glide or palatalized-glide contrasts. The interesting pattern here is that for both English groups, palatalized consonants seemed to pose more difficulty in coda position than palatalized-glide sequences. Recall that the stimuli consisted of Russian consonants embedded within nonsense words, where the position of the consonant of interest varied with respect to syllable position. For the consonant-glide sequences, it was the case that the glide was onset to the following syllable, and it seems that the English groups analyzed the consonant as such. However, for coda palatalized consonants, the palatalization did not resyllabify to the onset of the following syllable, and remained as a relatively simultaneous gesture with the coda consonantal gesture. Thus, English listeners seem to have been more likely to misperceive coda palatalized consonants as being closer to plain consonants, than consonant glide sequences.

A different picture emerges when we consider onset contrasts. Here, it is no longer the plain-palatalized consonants that are the most difficult, as it was in coda position. In onset position, the only group that exhibited significant differences between discrimination of different contrast types was the naive group. This group exhibited significantly different discrimination performance between each of the three contrast types, where plain-glide were discriminated best, plain-palatalized less well and palatalized-glide contrasts the worst. This group also saw an improvement in perception performance between codas and onsets for plain-palatalized consonants, and a worsening in discrimination performance for the palatalized-glide contrasts. Since neither the learner group nor the native Russian group exhibited any significant differences between the contrast types in onset position, the naïve group was the only group which saw significantly lower discrimination performance in onset position for the palatalized-glide contrasts. This result is also in accordance with the predictions, since in onset position, English speakers might perceive palatalized consonants as sequences of consonant glide, and thus have trouble discriminating them from actual consonant-glide sequences.

5.2.3. Analysis: Reaction Time

When analyzing reaction time data for psycholinguistic experiments, a longer reaction time for a given stimulus is considered indicative of a greater processing time, or as indicating more difficult stimuli (Baayen and Milin, 2010). In this chapter we have seen results from reaction time data comparing the three contrast types to each other, syllable positions to each other, and groups to each other. When comparisons were made, a negative coefficient was indicative of a decrease in the logRT from one effect to another, while a positive coefficient had the opposite interpretation. A negative coefficient is thus associated with a faster response. Here I present the analysis of Reaction Time on the AXB data. Again, a mixed effects model using the same lmer() function in R was used. Since RT is a numerical variable, logistic regression was not used for this data. To begin with, all of the same independent variables were selected as for the previous analysis. Following Baayen and Milin (2010) I selected a full model with all interactions specified, as well as random intercepts for subject and stimulus, like the previous analysis. However, as described in Baayen (2008), it is possible for effects of stimulus order to affect reaction time. As such, I included another random effect of stimulus order (variable name = Trial) and by-subject correlations of Trial to model the possibility that subjects could either slow down or speed in their responses as the experiment progressed. As in the proportion correct analysis, a process of evaluating the model fit was performed. Models with increasingly complex random effects structures were tested against each other. The model with both random intercepts for stimulus and subject, as well as a by subject slope adjustment for Trial had the best model fit. Next, models were evaluated for their fixed effects. For this analysis, unlike for proportion correct, for which the factor of Place did not provide a significant increase in goodness of fit, a fully specified model involving all factors (Group, Syllable, Pal and Place) as well as their interactions provided the best model fit. This model was then run to determine the contribution of the individual factors and their interactions.

5.2.4. Results: Reaction Time

The model differed from that for the proportion correct response data in two ways. It included a by-subject correction for Trial, as well as the inclusion of the factor Place and its interactions. The increase in model fit between the model without the random effect of Trial by Subject and the model with the more complex term was significant (Pr > Chisq = 3.743e-06). This model incorporates an additional factor of Place (2 levels: labials and coronals) as well as the interaction between Place and the remaining factors, the multiplicative increase in degrees of freedom and subsequent comparisons renders the interpretation of the interactions from this single model difficult. As such, the data is partitioned into two parts, one corresponding to the coronal data and the other corresponding to the labial data. Therefore, two separate analyses are undertaken, one for coronal data and one for labial data. For each of the sets of data (coronals and labials) the results are presented by subsections referring to the analyses of interest, that is, the analyses referring to the factors Pal (palatalization contrast type), Group (differences between groups) and syllable position.

5.2.4.1. Results: coronals: Reaction time

The coronal data is presented first. The results are presented in a similar way to those for the %correct response data. That is, for the relevant factors of palatalization contrast type, syllable position and group separate comparisons are made and presented. Specifically I present more detailed results with regards to the factor Pal, and any differences between onsets and codas and between different groups for these contrast types and syllable position.

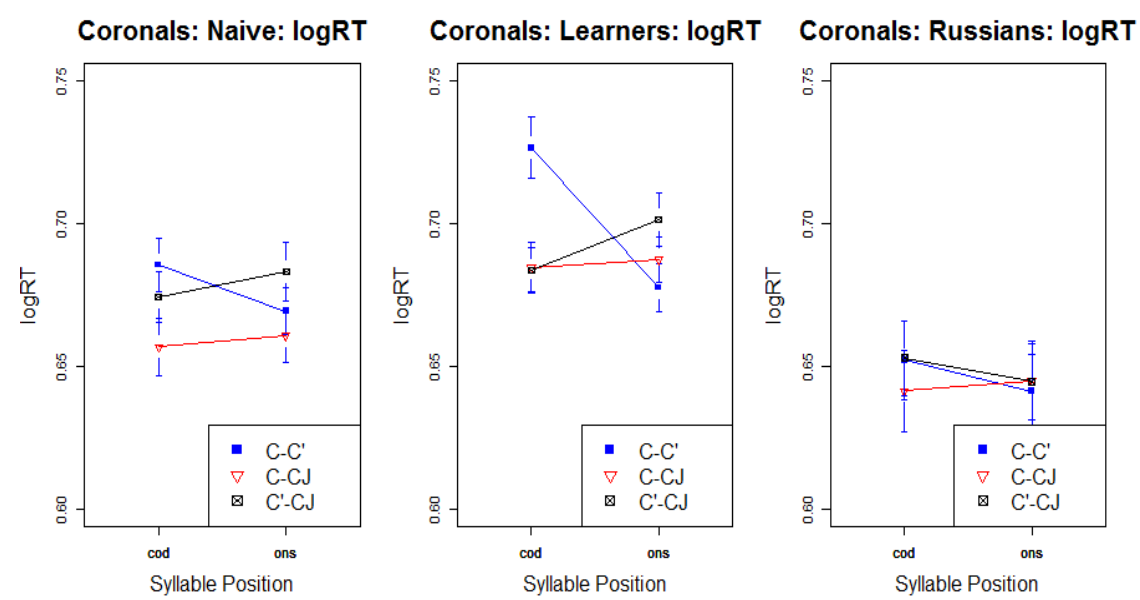

Figure 5.3: Average LogRT for coronals.

Pal	Naïve		Learners		Russians	
	coda	onset	coda	onset	coda	onset
$C\text{-}C^j$ vs $C^j\text{-}Cj$			-0.08 (0.0) ***	0.03(0.016)*		
$C\text{-}C^j$ vs $C\text{-}Cj$	-0.06 (0.0)***	-0.04 (0.0045)**	-0.08 (0.0) ***		-0.03(0.04)*	
$C^j\text{-}Cj$ vs $C\text{-}Cj$	-0.047 (0.003)**	-0.06 (0.0001) ***		-0.029 (0.056)*		

Table 5.5: Fixed effects coefficients for contrast type comparisons for logRT. Significant coefficients are indicated by asterisks indicating the level of significance in shaded cells. P-values are enclosed in brackets. A negative coefficient means that the contrast type on the right of the “vs” label in the “Pal” column has a significantly lower logRT than the contrast type on the left of the “vs” label for the particular Group and syllable position.

The first set of results focus on the differences in logRT between separate contrast types for each Group and syllable position. Table 5.5 shows that comparing plain-palatalized to palatalized-glide, (row 1); palatalized-glide contrasts in coda position have significantly lower logRT values only for the learner group In onset position, the opposite pattern emerges for this comparison. The palatalized-glide contrasts have significantly larger logRT values in onset than for plain-palatalized contrasts only for the learner group. Row 2 shows that plain-glide contrasts have significantly lower logRT values than plain-palatalized contrasts in coda position for both English groups. The same was true for this contrast in onset position for the naïve group, but no significant differences were found for these contrasts in onset position for the learner group. Finally, row 3 shows that plain-glide contrasts have significantly lower logRT values than palatalized-glide contrasts in onset position for both English groups, as well as coda position for the naïve group. The native Russian group saw significantly lower logRT rates only for coda plain-glide contrasts versus plain-palatalized in coda position.

			Group	
Pal	Syllable differences	Naïve	Learners	Russians
C-C^{j}	Coda vs Onset		-0.077 (0.0) ***	
C-Cj	Coda vs Onset			
C^{j}-Cj	Coda vs Onset	0.028 (0.047)*	0.04 (0.01)**	

Table 5.6: Fixed effects coefficients for Syllable comparisons for logRT. Asterisks indicate the level of significance in shaded cells. P-values are enclosed in brackets For each contrast type in the Pal column, a significant negative coefficient means that the syllable position “Onset” has a significant decrease in logRT performance versus the syllable position Coda, for each Group.

Table 5.6 shows results focusing on simple effects comparisons between syllable positions (Onset versus Coda) for each contrast type and group. For the plain-palatalized contrast in row 1, only for the learner group, onsets had significantly lower logRT values than codas. For the contrast palatalized-glide in row 3, the for both English groups, onsets had significantly lower logRT values than codas. No other significant differences between syllable positions were found.

Pal	Group	Syllable position: coda	onset
C-C^j	Naïve - Learners		
	Naïve - Russians	-0.087 (0.03)*	-0.077 (0.058)
	Learners - Russians	-0.144 (0.0008)***	
C-Cj	Naïve - Learners		
	Naïve - Russians		
	Learners - Russians	-0.08 (0.039)*	-0.09 (0.04)*
C^j-Cj	Naïve - Learners		
	Naïve - Russians		-0.1 (0.01)*
	Learners - Russians		-0.1 (0.01)*

Table 5.7: Fixed effects coefficients for Group comparisons for logRT. Significant coefficients are indicated by asterisks indicating the level of significance in shaded cells. P-values are enclosed in brackets For each contrast type in the Pal column, a positive coefficient means that the group on the right of the “vs” label in the “Group” column has a significant increase in Log Odds of discriminating that contrast versus the group on the left of the “vs” label for the particular syllable position. For example, for C^j-Cj in coda position the native Russian group discriminated had significantly lower logRT than the naïve group.

The last group of results focuses on Group differences. The results are presented as comparisons between groups for each palatalization contrast type and syllable position as seen in Table 5.7. For the plain-palatalization contrasts in coda position, the native Russian group presented significantly lower logRT values than both English groups. There were no significant differences between the English groups for this contrast in coda position. No significant differences were present in onset position for this contrast. For the plain-glide contrast, in coda and onset positions, the only significant differences were lower logRT values for the native Russians versus the learners. Finally, for the palatalized-glide contrast, significant differences were found only for syllable onset position between the naïve group and the native Russian group and between the learner group and the native Russian group. Specifically, for this contrast type in onset position the native Russian group had significantly lower logRT than both Groups English groups.

5.2.4.2. Results: Labials reaction time

In this section I present the labial data. As for the coronal data, the first set of results will focus on the differences in logRT between separate contrast types for each Group and syllable position.

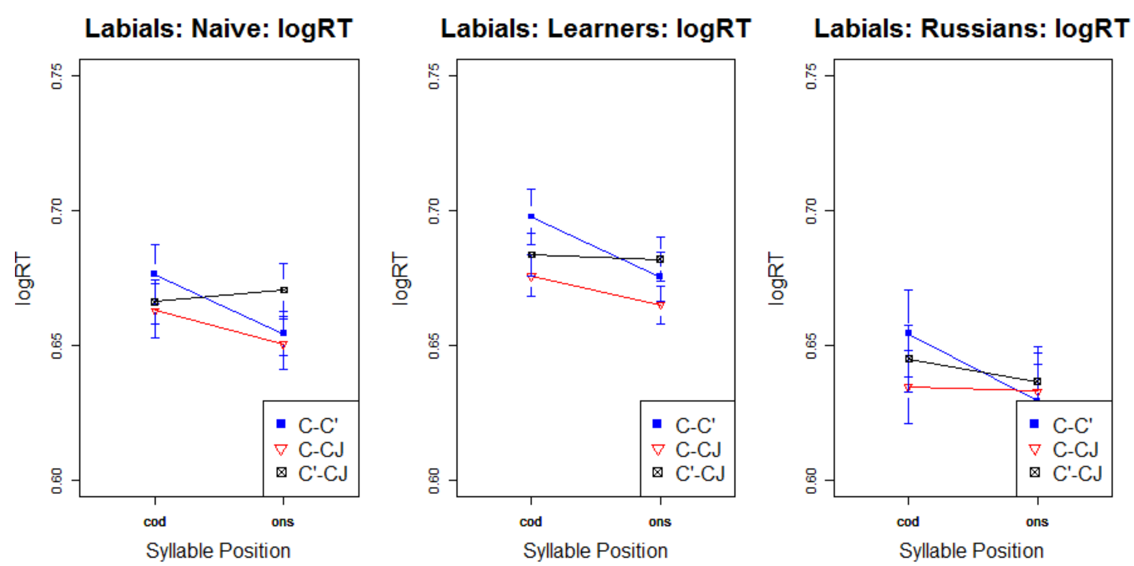

Figure 5.4: LogRT for labials for each group by syllable context.

Pal	Naïve		Learners		Russians	
	coda	onset	coda	onset	coda	onset
C-C^j vs C^j-Cj	-0.03 (0.01)*		-0.04 (0.005)**		-0.05 (0.0008)***	
C-C^j vs C-Cj	-0.03 (0.01)*	-0.02	-0.04 (0.001) ***	-0.02	-0.03 (0.03)*	
C^j-Cj vs C-Cj		-0.03 (0.03) *				

Table 5.8: Fixed effects coefficients for contrast type comparisons for logRT. Significant coefficients are indicated by asterisks indicating the level of significance in shaded cells. P-values are enclosed in brackets. A negative coefficient means that the contrast type on the right of the "vs" label in the "Pal" column has a significantly lower logRT than the contrast type on the left of the "vs" label for the particular Group and syllable position. For example in row 2, C-Cj contrasts in coda position for Group E have significantly lower logRT values than C-C^j contrasts.

Table 5.8 shows that between plain-palatalized to palatalized-glide, (row 1); palatalized-glide contrasts in coda position have significantly lower logRT values for all three groups. In onset position, the opposite trend emerges for this comparison, where palatalized-glide contrasts have larger logRT values in onset than for plain-palatalized contrasts for but this trend did not reach significance. Row 2 shows a similar pattern for the two English groups as in the coronal data. The plain-glide contrasts have significantly lower logRT values than plain-palatalized contrasts in coda position for both English groups. No significant differences were found for these contrasts in onset position. For this contrast in coda position the native Russian group exhibited the same result, a significantly lower logRT for plain-glide contrasts versus plain-palatalized. Finally, row 3 shows that plain-glide contrasts have significantly lower logRT values than palatalized-glide contrasts in onset position for the naïve group only. Unlike the coronal data, the learner group showed no significant difference in coda position between these two contrasts.

		Group		
Pal	Syllable differences	Naïve	Learners	Russians
C-C^{j}	Coda vs Onset			-0.03 (0.048)*
C-Cj	Coda vs Onset			
C^{j}-Cj	Coda vs Onset			

Table 5.9: Fixed effects coefficients for Syllable comparisons for logRT. Asterisks indicate the level of significance in shaded cells. P-values are enclosed in brackets. For each contrast type in the Pal column, a significant positive coefficient means that the syllable position “Onset” has a significant increase in LogRT versus the syllable position Coda, for each Group.

Now I present results focusing on syllable position differences for the labial data. Table 5.9 shows the results focusing on simple effects comparisons between syllable positions (Onset versus Coda) for each contrast type and group. The only significant result from for syllable position comparisons existed for the plain-palatalized contrast in row 1, for the native Russian group. Specifically, onsets had significantly lower logRT values than codas.

		Syllable position	
Pal	Group	coda	onset
C-C^J	Naïve - Learners		
	Naïve - Russians		-0.083 (0.066)
	Learners - Russians	-0.1 (0.03)*	-0.1 (0.028)*
C-Cj	Naïve - Learners		
	Naïve - Russians	-0.087 (0.052)	
	Learners - Russians	-0.1 (0.03)*	
C^J-Cj	Naïve - Learners		
	Naïve - Russians		-0.09 (0.04)*
	Learners - Russians	-0.08 (0.06)	-0.1 (0.03)*

Table 5.10: Fixed effects coefficients for Group comparisons for logRT. Significant coefficients are indicated by asterisks indicating the level of significance in shaded cells. P-values are enclosed in brackets. For each contrast type in the Pal column, a negative coefficient means that the group on the right of the “vs” label in the “Group” column has a significantly lower logRT value for that contrast versus the group on the left of the “vs” label for the particular syllable position. For example, for C^J-Cj in onset position Group R had significantly lower logRT values than either E or L.

The comparisons between groups for each palatalization contrast type and syllable position are presented in Table 5.10. For the plain-palatalization contrasts in coda position, the native Russian group presented significantly lower logRT values than the learner group. There were no significant differences between any of the groups for this contrast in coda position. In onset position for this contrast, the native Russian group also had a significantly lower logRT than the learner group. For the plain-glide contrast, only in coda position, the only significant differences were lower logRT values for the native Russian group versus the learner group.

Finally, for the palatalized-glide contrast, significant differences were found only for syllable onset position between each English group and the native Russian group. Specifically, for this contrast type in onset position the native Russian group had significantly lower logRT than both English groups.

5.2.4.3. Discussion RT

Here we highlight the main patterns from the reaction time data comparing contrasts, groups and syllable position, for each place of articulation. Table 5.11 below is a composite of Table 5.5 and Table 5.8 for coronal and labial data, showing for each Group, contrast pair and syllable position, whether the contrast on the right had a significant positive (P) or negative (N) increase in logRT versus the contrast on the left. First we discuss the overall patterns that hold between labials and coronals.

	Naïve				Learners				Russians			
Pal	Cor coda	Lab coda	Cor ons	lab ons	Cor coda	Lab coda	Cor ons	Lab ons	Cor coda	Lab coda	Cor ons	Lab ons
C-C^j vs C^j-Cj		N*			N***	N*	P*			N***		
C-C^j vs C-Cj	N***	N*	N**		N***	N***			N*	N*		
C^j-Cj vs C-Cj	N**		N***	N*			N*					

Table 5.11: Comparisons between contrast types for coronals and labials by group and syllable position. Directions (N = negative, P = positive) and significant effects are indicated by asterisks indicating the level of significance in shaded cells.

5.2.4.3.1. Commonalities: Coronals and labials in RT

First, the plain-glide contrast was discriminated faster than plain-palatalized for all three Groups in coda position. This result is predicted, as the plain-glide contrast is arguably the easier contrast for non-Russians (and Russians as well). Both plain consonants and consonant glide sequences exist in English. The plain-palatalized contrast is arguably the most difficult contrast in coda position, and thus, this comparison makes sense. For the learner group, the palatalized-glide contrast was discriminated faster than the plain-palatalized contrast in coda position. The plain-glide contrast was discriminated faster than the palatalized-glide contrast in onset position for the naïve group. This result is also predicted since the palatalized-glide contrasts were predicted to be more difficult in onset position than other consonants.

		Naïve		**Learners**		**Russians**	
Pal	Syllable differences	Cor	Lab	Cor	Lab	Cor	Lab
C-C^{J}	Coda vs Onset			NEG***			NEG*
C-Cj	Coda vs Onset						
C^{J}-Cj	Coda vs Onset	POS*		POS**			

Table 5.12: Comparisons between syllable position for each contrast type for coronals and labials by group. Directions (N = negative, P = positive) and significant effects are indicated by asterisks indicating the level of significance in shaded cells.

Table 5.12 presents results focusing on syllable position differences. No significant patterns hold between both labials and coronals.

		Syllable position			
Pal	Group	Cor coda	Lab coda	Cor onset	Lab onset
C-C^{J}	Naïve - Learners				
	Naïve - Russians	NEG*			
	Learners - Russians	NEG***	NEG*		NEG
C-Cj	Naïve - Learners				
	Naïve - Russians				
	Learners - Russians	NEG*	NEG*	NEG*	
C^{J}-Cj	Naïve - Learners				
	Naïve - Russians			NEG*	NEG*
	Learners - Russians			NEG*	NEG*

Table 5.13: Comparisons between syllable position for each contrast type for coronals and labials by group. Directions (N = negative, P = positive) and significant effects are indicated by asterisks indicating the level of significance in shaded cells.

The main patterns here have to do with differences between the native Russian group and the two English groups. No differences emerged between the two English groups. For the plain-palatalized and plain-glide contrasts, the native Russian group discriminated faster than the learners in coda position. The native Russian group discriminated the palatalized-glide contrast faster than either English group in onset position. These results suggest that overall; the learner group was slower than the native Russian group in discrimination. They also suggest that the palatalized-glide contrast again, is the more difficult contrast for English groups in onset position, as the Russians discriminated this contrast faster than either group.

By examining these tables we can surmise a series of major patterns that hold across places of articulation, but also some differences. One of the major patterns is evident when we look at the differences in logRT between plain-palatalized contrasts and palatalized-glide contrasts. The palatalized glide contrasts had significantly higher logRT values for both English groups for coronals in onset position. This result could be interpreted as palatalized-glide contrasts being more difficult than plain-palatalized contrasts in onset position.

5.2.4.3.2. Differences: Coronals and labials in RT

Now we turn to the differences that emerged in the RT data between coronals and labials. We focus first on palatalization contrast differences, and syllable position differences for each group. First, the palatalized-glide contrast was discriminated faster than the plain-palatalized contrast only for labials in coda position for Groups E and R. Staying with the same contrasts, the palatalized-glide contrast was discriminated slower than the plain-palatalized contrast only for the learner group for onset coronals. The plain-glide contrast was discriminated faster than plain-palatalized for the naïve group in coda position only for coronals. Finally, the plain-glide contrast was discriminated significantly faster than the palatalized-glide contrast in coda position only for the two English groups for coronals.

The Russian group also saw a significantly faster logRT for onsets for the plain-palatalized contrast in labial consonants over codas. If the plain-glide contrast is discriminated faster than the plain-palatalized contrast in coda position, then it makes sense that the plain-palatalized is also discriminated faster in onset position than in coda position.

Finally we consider some group differences between labials and coronals. Again, the main differences between labials and coronals here have to do with differences between Group R and the two English groups. No consistent differences emerged between the two English groups. For labials, the native Russian group discriminated plain-palatalized onsets faster than learners. For coronals, for coda plain-palatalized contrasts, the native Russian group discriminated faster than the naïve group. Also for coronals for plain-glide contrasts the native Russian group discriminated faster than the learner group in onset position. These results together suggest that overall; the two English groups were slower than the native Russian group in discrimination.

In sum, the RT data yield some general patterns. Overall, Russian speakers had the lowest average logRT values for all contrasts when compared to both naïve listeners and learners. There were no significant differences when comparing learners to naïve listeners in direct comparisons between the groups for contrast types. In terms of contrast types, the most difficult contrasts for the naïve group were onset palatalized-glide contrasts in both coronal and labial places of articulation, as well as plain-palatalized coda contrasts. The same is true for learners, coda plain-palatalized were discriminated more slowly than other contrasts. However, onset palatalized-glide contrasts were slower than other contrasts for the learner group for coronals only. The RT data also provide some similarities to the % correct response results, which will be discussed below.

5.3. General discussion

Both % correct and reaction time data were analyzed in this study. The statistical models did not benefit from the addition of a place of articulation factor when analyzing % correct response, but place became a significant factor in reaction time. The RT data present a more nuanced picture than % correct response. When comparing the results from logRT and %correct response data, it is assumed that contrasts that are discriminated more poorly at a significantly lower proportion of correct responses will also be associated with longer reaction times. This correspondence did hold for some contrasts, in that, contrasts which were discriminated at significantly lower rates tended to have higher reaction times. Below I discuss similarities and differences in the results for the two analyses.

5.3.1. Commonalities: RT and % correct response

Here we discuss the similarities between the two types of data. In the % correct response data, the main differences between contrast types occurred in both English groups. Similarly, the main differences between syllable positions occurred in these two groups. In terms of contrast types, naïve listeners had the most difficulty in discriminating plain-palatalized contrasts in coda position, and palatalized-glide contrasts in onset position. Learners had more difficulty in discriminating the plain-palatalized contrast in coda position as well. These main patterns in %correct discrimination were also observed in RT data. Naïve listeners also had more difficulty in discriminating plain-palatalized contrasts in coda position, as evidenced by longer logRT values. This group also had more difficulty discriminating palatalized-glide contrasts in onset position. Finally, learners had more difficulty in discriminating the plain-palatalized contrast in coda position as well. For these main patterns the reaction time and % correct data seem to correlate. The RT and % correct response data are related in that when we see a negative coefficient for % correct response data we see a positive coefficient for logRT data, indicating that, for example, a contrast that is discriminated correctly at a significantly lower rate will also be associated with a significantly higher logRT. Below we examine some differences between the two types of results.

5.3.2. Differences: RT and % correct response

Several differences between the RT and %correct response data emerge. The most common type of difference is a significant effect present in the logRT data that is not present in the % correct response. The most obvious differences occur for the learner group and the native Russian group. For the learners for onset contrasts, in the logRT data, the most difficult contrast was the palatalized-plain contrast. This same effect was not visible in the %correct response data. That is, the learner group showed significant differences in their logRT values in the expected direction while not showing significant differences in % correct response. A similar effect was visible for the Russian group for the comparison between plain-palatalized and plain-glide in the coda, where significant difference was observed for logRT but not for % correct response.

There also existed comparisons which were significant in terms of % correct response but not for logRT. An example is evident in the comparison of plain-palatalized and palatalized-glide contrasts in coda position. For % correct response for then naïve group, palatalized-glide contrasts were better discriminated than plain-palatalized contrasts. However, there was no difference between these two contrasts in logRT. So, it is also the case that some contrasts were discriminated worse, but showed no significant difference in logRT values.

5.4. Conclusions: AXB

The results from this experiment shed some light on the discrimination of Russian contrasts by three groups of listeners. The two types of data (% correct and RT) yield similarities and differences. The main findings are that both types of data provide evidence that native Russian speakers have the least amount of difficulty in perceiving Russian contrasts. This result is perhaps obvious. More interesting results are that the contrasts differ with respect to group as well as syllable position. Thus, some contrasts are more difficult than others in different positions. For the naïve group of English speakers, the most difficult contrasts were plain-palatalized contrasts in coda position and palatalized-glide contrasts in onset position. For the learner group a similar result was obtained when considering both the %correct and RT data, however, taking just % correct into account we see an improvement in learners' discrimination of the onset palatalized-glide contrasts over the naïve group. This difference was the most notable between the two groups. The results obtained here are compared to results from the remaining experiments in subsequent chapters.

Chapter 6 Task 2: Categorization

A categorization task was carried out following the discrimination tasks. The purpose of this task was to test how speakers categorize or perceptually assimilate the consonants. The categorization task tests very broadly how and to what extent speakers categorize Russian consonants to English L1 categories. In the case of native Russian speakers and learners, it tests whether they consistently categorize the consonants to the relevant Russian categories.

This type of information is relevant for the PAM model, as the PAM makes predictions on discrimination of non-native contrasts based on perceptual assimilation of the phones to native phonological categories.

6.1. Methods

For each of the 12 sequences, four of the same five recorded tokens from the AXB task were employed. The categorization task is a forced choice identification task where for each trial the token is presented along with three options on the screen. For the naïve group, each option was an English orthographic representation of a general pre-determined category. The first option is a representation containing palatalization preceding the primary constriction (e.g. “taytat”). The second option is one lacking any palatalization (e.g. “tatat”), and the third option involves palatalization following the primary constriction (e.g. “tatyat”). After a selection was made, participants rated on a scale of 1-5 how well they felt their selection matched what they heard. After a rating was selected, a new trial was presented.

An exemplar of each of the twelve sequences was presented a total of seven times. Three of the physical tokens were presented twice, while the fourth was presented only once (3 x 2 +1

= 7). This strategy was chosen to use as diverse as possible a number of individual recorded tokens while keeping the experiment at a manageable length. As in the AXB test, the tokens were presented in random order in blocks whereby each individual token appeared once before being repeated with the consideration of no two tokens appearing in doublets. The inter-trial interval was 1 second, and like the AXB test, a break was given after 50 trials. For 12 sequences there were 7 repetitions for a total of 84 trials per participant.

Overall, it can be considered that subjects decided whether a given token had a palatal TB gesture, and if so, whether this gesture occurred before or after the consonant. No option exists for a category involving more simultaneous palatalization, as is the case in Russian, as there is no way to represent this in English orthography. Identification or categorization of certain non-native phonemes is not always trivial as “Identification tests can assess this [categorization] only insofar as the experimenter can provide listeners with the appropriate labels for the categories.” (Beddor and Gottfried 1995; 225). In the present study only three labeled categories are provided, however the goodness rating essentially widens the category set (Beddor and Gottfried 1995). Thus, speakers could choose “tatyat” hearing something that sounds like “tatyat” with palatalization following the /t/, but could provide a lower rating, indicating that it is not an optimal choice. What the optimal choice would be is not easy to determine, again, given that English orthography does not allow for palatalization.

The groups of English speaking learners of Russian and native Russian speakers performed the task with one difference in labels of the categories presented. For the naïve English speakers, three English labels were presented for participants to choose from, while for the Russian learners and native Russians, four representations, representing four possible categories were presented. The categories were presented in Russian Cyrillic orthography and represented the following configurations: no palatalization (pa pap), palatalization (pa p^jap), glide (pa pjap) and pre-palatalization (paj pap). Each of these possibilities is represented through different orthographies. Thus, the Russian groups had the same 3 categories that naïve English speakers had to choose from, with an additional category representing phonemic secondary Russian palatalization.

Speakers' responses to these tasks were encoded to determine what perceptual assimilation pattern in PAM terms (SC, CG, TC) each contrast undergoes. A contrast will have Single Category (SC) assimilation if speakers choose the same orthographic representation with similar goodness ratings on a consistent basis for both members of the contrast. A contrast will have Category Goodness (CG) assimilation if both members of the contrast elicit the same English labels, but with consistently different goodness ratings for one member of the contrast. For example, for /p-p^j/ a CG assimilation will be encoded if a subject chooses "papap" for both /pa pap/ and /pa p^jap/ and the goodness rating for /pa p^jap/ is significantly different (lower) than that for /pa pap/. Lastly, if subjects consistently employ different responses for two members of a pair, a TC assimilation pattern will be encoded. A TC pattern will also be encoded if assimilations for one member of a pair are considerably less consistent than for the second member. For instance, if /pap^jap/ elicits a variety of responses ("paypap", "papap", and "papyap") while /papap/ elicits only one response consistently ("papap"), a TC assimilation will be encoded. For the present study it is predicted that relatively few TC assimilations will be encoded, thus the majority of assimilations will be either SC or CG.

6.2. Analysis

The patterns of assimilation of the Russian consonants to English categories (for the naïve group) and to the Russian categories (for learners and Russians) were compiled for each participant, into confusion matrices. In order to test PAM and to determine the assimilation pattern of a particular contrast, a "fit index" as done in previous work (Guion et al. 2000) was used. The fit index is a metric of goodness of fit of a particular response category to a speech token. As mentioned earlier, for each trial, the subject chooses from three (or four) options, (e.g. "taytat", "tatat", "tatyat") and then gives a goodness rating that ranges between 1-5. For each Russian consonant, the count of each type of response (pre-palatal, no-palatal, post-palatal) is multiplied by the mean goodness rating to determine the fit index. For instance, if a subject chooses for [tat^{j}at] the response "tatyat" 7 times with a mean goodness rating of 4, then the fit index for this response is 28 (7*4). The maximum fit index possible, 35 indicates a very good fit, and can be interpreted as perception of the phone as a very good instance of that particular category. Lower fit indices will indicate worse-exemplars of a particular category. Russian consonants with multiple fit indices associated with different categories of roughly equal value will indicate that the subject perceives that particular consonant as two possible English categories.

The fit indices are used to determine the assimilation pattern (SC, CG, TC) and are compared to the discrimination data to test PAM predictions that discrimination ability is related to perceptual categorization.

6.3. Results

The basic analysis for the categorization data is the confusion matrix. A confusion matrix is a table showing for each stimulus, or sequence, the various responses provided by the participant. In this results section, for each group I provide confusion matrices with each Russian consonant and the fit indices of the various categories.

6.3.1. Naïve listeners

In this section I provide the confusion matrices for the naïve group of participants. The confusion matrix in Table 6.1 represents the responses averaged over all participants in the naïve group.

Syllable	Consonant	C (no pal)	CJ (post pal)	JC (pre pal)
Onsets	pa	30.2	0	0
	p^ja	0.6	23.4	2.5
	p^jja	0	25.9	0.4
	ta	29.9	0	0
	t^ja	0.6	20.5	2
	t^jja	0.3	23.5	1.2
Codas	ap	24.7	0	0
	ap^j	8.4	4.4	9.4
	ap^jj	0	25.6	1.8
	at	21.7	0.8	0
	at^j	8	4.7	7.2
	at^jj	0.3	24.5	0.3

Table 6.1: Fit indices for assimilations of Russian consonants for the naïve group.

Table 6.1 represents the overall average fit indices (FI) over all participants in the naive group. Again, for each consonant in Column 1, the numbers represent the count of responses for a particular category (Assimilations columns C, CJ, JC) multiplied by the average goodness rating for that response. For example, for the plain labial stop in onset position (pa) the fit index is 30.2 for category C (no palatalization). Importantly, this consonant has a FI of 0 for both categories involving palatal representations (CJ and JC). This means that for this consonant, participants responded exclusively with the category C, meaning they clicked a box labeled "papap". Comparing this consonant to the palatalized labial onset consonant in the next row, we notice the FI for this consonant for the same category (C) is considerably lower at 0.6. This means that participants both responded less often and with a lower goodness rating with "papap" for the palatalized labial. The FI for the next category involving post palatalization (CJ or "papyap") is much higher at 23.4. However, this FI is lower than the FI of 30.2 for /p/ to English "papap", indicating that participants overall considered Russian /p/ a better fit to English /p/ than Russian /p^j/ to English /pj/.

The FI values in dark shaded cells represent modal assimilations, meaning a particular Russian consonant was assimilated to a particular category much better than other potential categories. This is true for ten of the twelve sequences. In onset position, all of the sequences (2 plain, 2 palatalized, 2 glide) were consistently assimilated to single categories. The plain consonants were assimilated to plain categories, and the palatalized and palatalized-glide sequences to Cj categories. Thus, in onset position, both palatalized and Cj sequences were consistently assimilated to the same categories, although to slightly differing degrees.

The two coda palatalized Russian consonants lack clear modal categories. This is evident when considering the rows in Table 6.1 corresponding to /ap^j/ and /at^j/. For both of these consonants, none of the three categories (C, CJ, JC) receive a clear dominant assimilation as is the case with other consonants. This is represented in lighter shaded cells. For these two consonants, the no palatalization (C) and pre-palatalization (JC) assimilations provided better fits than the post-palatalization assimilation (CJ).

Overall then, this table shows that on average, for naïve listeners, the consonants which proved most problematic in terms of assimilation to English categories were the coda palatalized consonants. Another result is the relatively similar assimilation patterns for onset palatalized and palatalized-glide sequences. These results will be compared to the discrimination data later in this chapter. Further discussion and comparisons between experiments will be presented in Chapter 9 and Chapter 10 as well.

6.3.2. Russian Learners

Syllable	Consonant	C (no pal)	C^j (Rus pal)	CJ (post pal)	JC (pre pal)
Onsets	**pa**	31	0	0	0
	p^ja	0	16	9	2
	p^jja	0	6	22	0
	ta	30	0	0	0
	t^ja	0	19	10	1
	t^jja	0	2	25	0
Codas	**ap**	26	5	0	0
	ap^j	8	16	0	4
	ap^jj	0	2	22	2
	at	19	11	0	0
	at^j	10	19	0	1
	at^jj	0	2	23	0

Table 6.2: Assimilations for each consonant for the learner group.

Table 6.2 represents the overall average fit indices over all participants in the learner group. Recall that the learners and native Russians perfromed assimilation tasks with four possible categories presented in Cyrrilic instead of three in English orthography. The additional category is one representing actual Russian palatalization, which as mentioned before, is not possible in English orthography. Thus, Table 6.2 has 4 columns of categories instead of 3 for Table 6.1. The new Russian palatalization category is represented in the second column, while the first, third and fourth are taken to be the same as for the naive group.

The two onset non-palatalized consonants, represented as “pa” and “ta” in the table have stable modal assimilations to a single category (C with no palatalization). The coda plain consonants also have modal assimilations but are less consistent than the onsets. The coda coronal is assimilated less well to the plain category than the labial, and also shows more assimilations to the Russian palatalized category, which is demonstrated by lighter shading in the cell. The fact that one of the non-palatalized consonants, the coda /t/ is being assimilated with a fit index of 11 to the palatalized category is something rather unexpected. It is possible that listeners are not considering the coda plain (actually velarized) consonants as good examples of a "plain" category because these consonants could be velarized and not actually plain.

Examining the palatalized onset consonants in the 2nd and 5th rows, we notice the FI for these consonants for the category (C) is zero. This means that learners perceived at least some palatalization during these trials. The FI for the category involving Russian palatalization is 16 for labials and 19 for coronals, and this FI is the highest for these consonants, indicating that learners in general preferred Russian palatalization as categories for these consonants to other categories. However, for both consonants, a sizeable FI for the category involving post-palatalization (CJ or "papyap") is observed (lighter highlighted cells). These patterns together indicate that for learners, subjects mostly assimilated these consonants to Russian palatalized categories, but exhibited some confusion as well in assimilating them to CJ categories. The Russian Cj sequences were consistently assimilated to CJ sequences.

In coda position, the labial consonant has high modal assimilations to the C category. The coronal has less consistent assimilations to the plain C category and higher assimilations to C^j. In this way, for the plain codas, labials and coronals differ. For the palatalized consonants the more common assimilation is to C^j with additional assimilations to C as indicated by the lighter shaded cells. In coda position the palatalized-glide sequences show consistent assimilations to the CJ (post-palatalization) category, as was the case for the naïve group. Overall for the learner group the consonants which proved most problematic in terms of assimilation to English categories were the palatalized consonants.

6.3.3. Native Russian group

Syllable	**Consonant**	**C (no pal)**	**C^j (Rus pal)**	**CJ (post pal)**	**JC (pre pal)**
Onsets	**pa**	34	0	0	0
	p^ja	0	31	1	0
	p^jja	0	2	30	0
	ta	34	0	0	0
	t^ja	0	32	0	0
	t^jja	0	1	31	0
Codas	**ap**	34	1	0	0
	ap^j	0	31	0	1
	ap^jj	0	0	33	0
	at	34	1	0	0
	at^j	0	34	0	0
	at^jj	0	2	29	0

Table 6.3: Assimilations for each consonant for the native Russian group.

Table 6.3 shows the same confusion matrix for the native Russian group. This group differs from both English groups in that for all consonants in both syllable positions and both place of articulations a clear modal assimilation category is observed. The assimilation patterns for the native Russian group are as expected, all plain consonants are consistently assimilated to plain categories; palatalized consonants are consistently assimilated to palatalized categories and Cj consonant sequences are consistently assimilated to the Cj categories.

6.4. Discussion

The results for each group show several patterns. First, onset consonants are generally associated with more consistent FI to modal categories than coda consonants for the two English groups. For Group E this is true for all onset consonants. For Group L, the palatalized onset consonants yield more confusion in that they take assimilations for both Russian palatalized and CJ categories. One difference between the tasks between the two English groups is that the learner group performed categorizations to actual Russian categories in Cyrillic, while naïve English listeners categorized to English orthographical representations. As such, Groups L and R had 4 options while Group E only three. If we consider the two English groups' responses to the onset palatalized consonants, it seems that Group E had a more consistent response pattern than Group L. Group E more consistently assimilated onset palatalized consonants to CJ than the learner group because this group had the additional option of assimilating these consonants to actual Russian palatalized consonants. Another difference between the two groups with regards to these two consonants is that Group E had a good FI for this category, meaning that participants were in general perceiving these consonants as good Cj categories, unlike Group L, which had lower FI for these consonants to this category.

6.5. Comparisons perception data: Discrimination and Categorization

In this section we report on the relation between the two perception experiments so far reported, discrimination and categorization. A central property of the PAM is a relation between a phone's perceived similarity – or assimilation – to the L1 and its predicted ease or difficulty in discrimination. Here we make comparisons between discrimination data from the AXB tests, and categorization data from the second experiment. First, for the three groups, the fit indices for each phone are compiled and compared with their average correct responses in discrimination. Then, subjects from each group are further divided into groups based on their assimilation patterns and discrimination results. For example, some subjects assimilated palatalized consonants to plain categories, while others assimilated them to sequential C+j categories. Analyses are run based on these divisions. The analyses are presented by group, starting with the naïve group.

6.5.1. AXB and CAT: Naïve group:

The Perceptual Assimilation Model described in previous sections has been posited to account for why certain phones are more difficult than others. In the following sections we extend the results from the individual experiments (categorization and discrimination) and unify them to arrive at a set of data for testing the PAM. In this section we focus on the naïve group.

The PAM was originally developed to account for non-native, naïve perception. The basic premise behind this model is that discrimination of two non-native phones should follow based on the assimilation of these two phones to native phonological categories. In this sub-section we compare the assimilation data from Chapter 6 to the discrimination data from Chapter 5. At first we will present a general comparison at the group level, then analyze the data more closely at the subject level, as individual participants have different abilities. Thus, the main purpose of this chapter is to examine, for each participant, whether discrimination can be predicted from categorization.

At the group level, we start by combining the discrimination data and categorization data with their predicted values based on the PAM introduced in the Predictions section. We see differences in discrimination rates for the two syllable positions and three contrasts. Coda plain-palatalized and onset palatalized-glide are the two most difficult.

The two tables below represent the onset and coda data. Each table shows for each palatalization contrast the predicted assimilation types according to the PAM, the actual assimilation types yielded from the categorization experiment, the predicted discrimination level and the actual percent discrimination yielded from the AXB experiment.

Russian contrast	Predicted Categorization	Predicted Type	Predicted Discrimination rate	Attested Type	Attested Discrimination
1) p-p^j	a) p-p	SC/CG	LOW	30/0.6	98%
	b) p-pj	TC	HIGH	30/23	
	c) p-NA	UC/NA	HIGH		
2) p^j-p^jj	d) p-p	TC	HIGH	0.6/0	91%
	e) pj-pj	SC/CG	LOW	23/26	
3) p-p^jj	f) p-pj	TC	HIGH	30/26	99%
4) t-t^j	g) t-t	SC/CG	LOW	30/0.6	97%
	h) t-tj	TC	HIGH	30/20.5	
	i) t-ʧ	TC	HIGH		
5) t^j-t^jj	j) t-tj	TC	HIGH	0.6/23.5	92%
	k) tj-tj	SC/CG	LOW	20.5/23.5	
	l) ʧ-tj	TC	HIGH		
6) t-t^jj	m) t-tj	TC	HIGH	30/23.5	100%

Table 6.4: Assimilation and discrimination comparisons at the group level for onsets for the naïve group.

Table 6.4 above shows the data for onsets for the naïve group. The first column shows the contrast type, the second column has all the most likely predicted assimilations to English categories for the respective phones for the contrast. For example, Row 1 shows for the contrast p-p^j, the three most likely predicted assimilations for these two phones. The penultimate column in this table shows the actual attested fit index for the particular contrast. Row 1)b) shows the most commonly attested assimilation type for this contrast. This contrast can be categorized to a single category or single category with category goodness (CG). The predicted discrimination rate for this categorization scheme is provided in column 4, in this example, LOW, meaning a single category or category goodness assimilation should yield a lower rate of discrimination. The final two columns show the actual attested categorization types and discrimination rates based on the Fit Index from the categorization experiment and proportion of correct response from the AXB experiment.

In the "Attested type" column we see two numbers. The first number represents the fit index for the first member of the contrast to the first member of the predicted English Categorization for that row. Similarly, the second number represents the fit index for the second member of the contrast to the second member of the predicted English categorization for that row. In the same example, in row b), the fit index for contrast member /p/ to predicted assimilation category /p/ was 30 (a high fit index), while the fit index of /p^j/ to /pj/ was also relatively high (but lower) at 23. This means that this contrast exhibited a two category (TC) assimilation pattern and as such, is predicted to have high discrimination rates. This is shown in the final column, where the actual discrimination rate for the entire group is shown. At 98% discrimination was good for this contrast. Compare this row to row a), the predicted possible SC/CG contrast type, which has a predicted lower discrimination rate, since category or category goodness contrast types should be discriminated less well than two category types. This assimilation type is not attested because for the second member of the contrast (p^j), the fit index to the plain category (p) was only 0.6. The actual discrimination rate of 98% follows from the actual most common attested type in row b), where both members of the contrast assimilated to different categories.

For the palatalized-glide contrast in row 2), the modal assimilation type is shown in the darkened row b). Both members of this contrast assimilated at relatively similar fit indices (23 and 26) to the same category /pj/. This type of assimilation, a SC or CG predicts a lower discrimination rate than the TC assimilation type. We see a discrimination rate of 91%. This discrimination rate was seen to be significantly lower than either of the other contrasts in the AXB discrimination chapter results. The plain-glide contrast saw a TC assimilation pattern with a 99% discrimination rate. For all three labial contrasts, the predictions hold. Meaning that clearly Two Category (TC) assimilation types are discriminated significantly better than CG or SC contrasts. The same patterns hold for the coronal contrasts.

Russian contrast	**Predicted Categorization**	**Type**		**Discrimination rate**	
1) p-p^j	a) p-p	SC/CG	24.7/8.4	LOW	88%
	b) p-pj	TC	24.7/4.4	HIGH	
	c) p-jp	TC	24.7/9.4	HIGH	
2) p^j-p^jj	d) p-pj	TC	8.4/25.6	HIGH	97%
	e) pj-pj	SC/CG	4.4/25.6	LOW	
3) p-p^jj	f) p-pj	TC	24.7/25.6	HIGH	99%
4) t-t^j	g) t-t	SC/CG	21.7/8	LOW	84%
	h) t-tj	TC	21.7/4.7	HIGH	
	i) t-jt	TC	21.7/7.2	HIGH	
5) t^j-t^jj	j) t-tj	TC	8/24.5	HIGH	99%
	k) tj-tj	SC/CG	4.7/24/5	LOW	
6) t-t^jj	l) t-tj	TC	21.7/24.5	HIGH	99%

Table 6.5: Assimilation and discrimination comparisons at the group level for codas.

Table 6.5 above shows the same comparisons for coda contrasts. Immediately noticeable are more shaded rows than there are for onset contrasts. This is indicative of some contrasts lacking a single modal assimilation type, unlike the onsets. The reason some coda contrasts lack a single modal assimilation type is due to the phones /t^j/ and /p^j/, the two palatalized consonants. These phones had more distributed assimilation patterns, assimilating relatively poorly to all three categories (C, CJ and JC). Looking first at the two plain-palatalized contrasts, the assimilation types with the highest fit index values are the SC/CG and TC assimilation types (darker shading). Depending on which of these assimilation types results, the discrimination can be very good or relatively low. Since these contrasts seem to have more varied and distributed assimilation types, the predicted discrimination should reflect this, and be more varied and distributed, which means, in effect, lower. The results for the AXB discrimination tests indeed showed significantly lower discrimination rates for these contrasts in coda position than the other two contrasts. These results are broadly in line with our predictions and with the PAM. First, we predicted that codas would cause more problems for the naïve group than onsets, and this is true. Second, it is the case that whenever we have instances of members of the contrast assimilated relatively well to different categories we have better discrimination.

Finally, the categorization results demonstrate that at least some of the time participants are perceiving the Russian palatalized consonants as sequences of glide (or high vowel) and consonant. In the coda, most often it is the tongue body which precedes the consonant in perception (assimilations to the JC category) in the onset it is usually perceived as tongue body following the consonant. We will discuss the reasons why Articulatory Phonology is used to explain these results later in the chapters related to the relation between perception and production.

6.5.2. AXB and CAT: Learner group:

We now move on to the comparison between the categorization and discrimination experiments for the learner group.

Russian contrast	Predicted Categorization	Predicted Type	Predicted Discrimination rate	Attested Type	Attested Discrimination
1) p-pʲ	a) p-p	SC/CG	LOW	31/0	98%
	b) p-pj	TC	HIGH	31/9	
	c) p- pʲ	UC/TC	HIGH	31/16	
	d) p-jp	TC	HIGH	31/2	
2) pʲ-pʲj	e) p-p	TC	HIGH	0/0	97%
	f) pj-pj	SC/CG	LOW	9/22	
	g) pʲ -pj	UC/TC	HIGH	16/22	
3) p-pʲj	h) p-pj	TC	HIGH	31/22	98%
4) t-tʲ	i) t-t	SC/CG	LOW	30/0	98%
	j) t-tj	TC	HIGH	30/10	
	k) t- tʲ	UC/TC	HIGH	30/19	
	l) t-jt	TC	HIGH	30/1	
5) tʲ-tʲj	m) t-tj	TC	HIGH	0/25	98%
	n) tj-tj	SC/CG	LOW	10/25	
	o) tʲ -tj	UC/TC	HIGH	19/25	
	p) ʧ-tj	TC	HIGH		
6) t-tʲj	q) t-tj	TC	HIGH	30/25	99%

Table 6.6: Assimilation and discrimination comparisons for the learner group for onsets.

Table 6.6 above shows the data for onsets. The layout for this table is the same as that for the naïve group. For example, Row 1 shows for the contrast p-p^j, the three most likely predicted assimilations for these two phones. The predicted discrimination rate for this categorization scheme is provided in column 4. The actual attested fit indices for each predicted categorization type are given in column 5, and finally the actual discrimination rate from the AXB discrimination data is given in column 6. For example, row 1)a) shows the possible predicted single category or single category with category goodness (p) assimilation pattern. In this example, the discrimination rate is given as LOW, meaning a single category or category goodness assimilation should yield a lower rate of discrimination.

For this contrast, row c) represents the actual attested type. In the "Attested type" column we see two numbers, the first number represents the fit index for the first member of the contrast to the first member of the predicted English Categorization for that row. Similarly, the second number represents the fit index for the second member of the contrast to the second member of the predicted English categorization for that row. In this example for row c), the fit index for contrast member /p/ to predicted assimilation category /p/ was 31 (a high fit index), while the fit index of /p^j/ to the actual palatalized category /p^j/ was also relatively high (but lower) at 16. Row b) also represented a less modal assimilation type, where the difference has to do with the assimilation of the palatalized consonant. Here, the palatalized consonant was assimilated with a fit index of 9 to the category /pj/ (sequential C+J). In either case, whether considering row b) or row c) this contrast exhibited a two category (TC) assimilation pattern and so is predicted to have high discrimination rates. The final column shows the actual discrimination rate for the learner group for this contrast. At 98% discrimination was good for this contrast, as predicted.

For the palatalized-glide contrast in row 2), the modal assimilation type is shown again in the darkened row c). Both members of this contrast assimilated at relatively high fit indices (16 and 22) to the same categories /p^j/ and /pj/. This type of assimilation, TC predicts a high discrimination rate, which we see at 97%.

Row 3) shows the data for the plain-glide contrast. Here, only one assimilation type is predicted, namely plain to /p/ and C+glide to /pj/. This contrast saw a TC assimilation pattern with a 98% discrimination rate. For all three labial contrasts, the predictions hold. Meaning that clearly Two Category (TC) assimilation types are discriminated well. The same patterns hold for the coronal contrasts.

Russian contrast	Categorization	Type		Discrimination rate	
1) p-p^j	a) p-p	SC/CG	26/8	LOW	87%
	b) p-pj	TC	26/0	HIGH	
	c) p-jp	UC/NA	26/4	HIGH	
	d) p- p^j		26/16	HIGH	
2) p^j-p^jj	e) p-pj	TC	8/22	HIGH	99%
	f) pj-pj	SC/CG	0/22	LOW	
	g) jp-	UC/NA	4/2	HIGH	
	h) p^j -pj	UC/TC	16/22	HIGH	
3) p-p^jj	i) p-pj	TC	26/22	HIGH	100%
4) t-t^j	j) t-t	SC/CG	19/10	LOW	74%
	k) t-tj	TC	19/0	HIGH	
	l) t-jt	TC	19/1	HIGH	
	m) t- t^j	UC/TC	19/19	HIGH	
Not Predicted:	n) t^j- t^j		11/19		
	o) t^j - t		11/10		
5) t^j-t^jj	p) t-tj	TC	10/23	HIGH	99%
	q) tj-tj	SC/CG	0/23	LOW	
	r) jt- t^j	TC	1/2	HIGH	
	s) t^j -tj	UC/NA	19/23	HIGH	
6) t-t^jj	t) t-tj	TC	19/23	HIGH	99%
not predicted	u) t^j -tj		11/23	HIGH	

Table 6.7: Assimilation and discrimination comparisons for the learner group for codas.

Table 6.7 above shows the same comparisons for coda contrasts. As for the naive group, there tend to be less clear modal assimilation types for codas than for onsets, however, the assimilation types do tend to be more consistent for the learner group. This is again indicative of some contrasts having less clear or less consistent modal assimilation types than onset consonants. As for the naive group, the two palatalized consonants (p^j and t^j) explain some of the less modal assimilation types. However, unlike for the naive group, the plain coronal consonant (t) is also responsible for part of the inconsistent assimilation for contrasts involving this consonant. The palatalized phones had more distributed assimilation patterns, assimilating relatively less consistently to a single category than the onsets.

We begin with the two plain-palatalized contrasts. For the labials, the assimilation types with the highest fit index values are the SC/CG and TC assimilation types (darker shading). The plain consonant /p/, assimilated at a relatively high fit index to the plain category, while the palatalized consonant assimilated relatively well to the palatalized category. The fit index for this assimilation was 16, while this consonant also assimilated relatively poorly (fit index of 8) to the plain category. Since there is a stronger assimilation for the palatalized labial to the Russian palatalized category, this contrast should exhibit an overall higher discrimination rate.

For the coronal plain-palatalized contrast, a slightly different pattern emerges. Here, the assimilation pattern for the palatalized consonant are similar to the labial consonant, assimilating relatively well to the Russian palatalized category, but also to the plain category. The difference between the labials and coronals lies in the assimilation of the plain consonant. This consonant assimilated relatively well to the plain category with a fit index of 19, but also assimilated less well to the Russian palatalized category with a fit index of 11. The plain labial did not assimilate to the Russian palatalized category to such an extent. The relatively high fit index of the plain coronal to the palatalized category means that this contrast exhibits more instances of a SC/CG assimilation type as opposed to just a TC assimilation type. SC and CG types, according to PAM, should yield lower discrimination than TC, which should yield the highest discrimination rate. Indeed, this coronal contrast has the lowest discrimination rate of all for this group (74%). This discrimination rate is also significantly lower than the plain-palatalized labial coda contrast (87%) and the difference, again, is due to the less consistent assimilation of the plain coronal consonant to the plain category. In this way, the PAM seems to hold.

Both palatalized-C+glide contrasts exhibited high discrimination rates (99% for both places of articulation). Again, the PAM is supported in that both contrasts exhibited either TC, or UC/NA assimilation types which are predicted to yield high discrimination rates. The reason these assimilation types are predicted to yield high discrimination rates is because in any case neither of the two members of the contrast assimilate to the same category. The basic premise of PAM is that whenever two members of a non-native contrast assimilate equally well to the same category, discrimination for that contrast is predicted to be poor. We see this pattern here, where as members of a contrast assimilate better to different categories, discrimination for that contrast is higher. Finally, the plain-C+glide contrasts also exhibited TC assimilation types and very high discrimination rates.

We return now to exploring why the “plain” coronal could have been assimilated relatively well to the palatalized category but not the plain labial. First, why should a plain (or non-palatalized) consonant assimilate at all to the palatalized category? We suggest that in coda position the contrast between non-palatalized and palatalized consonants in general is weaker and so non-native speakers will have difficulty with this contrast. Second, if we consider that the

Russian plain consonants are actually velarized consonants (as generally accepted in the literature), and if we also consider that velarization in Russian can be variable and impressionistically less salient, especially in the context of back vowels (e.g. see Kedrova et al. 2008; Bennett et al. to appear)[15] we can say that it's possible that the fact that what we are calling "plain" consonants are actually not good exemplars for the English plain category and therefore we see variability also in the assimilation of the velarized (non-palatalized) consonants. In terms of the difference between labials and coronals, given that there tends to be less salient velarization present for coronals over labials, we could also speculate that any increase in velarization for the labial consonant could be contributing to the perceptual difference between the plain labial and any English category involving palatal gestures. These are speculative reasons given for the question of why "plain" consonants would assimilate to non-plain categories and why labials should differ from coronals.[16]

6.5.3. AXB and CAT: Native Russian group:

For the native Russian group, a very different pattern emerges. Each consonant assimilates very well to only one category consistently, as predicted. In addition, near 100% discrimination rates are observed for each contrast as predicted.

---

[15] The consonants in this study were in the context of the back vowel /a/.

[16] The question of velarization is interesting and deserves further study. For example, the contrasts employed in this study could also be employed in a follow up study but instead of asking participants to force choice assimilate the consonant to a pre-determined category, the experimenter could ask the participants to provide a more granular judgement, perhaps on a scale, or perhaps asking the participant to write out what they hear. Alternatively, participants could be presented with plain (velarized) – palatalized contrasts and asked to select which two words they heard, with the options each representing plain, velarized and palatalized members (e.g. pa-pja or pwa-pja).

6.5.4. AXB and CAT: Individual subjects comparisons

Now that we have seen how the categorization data and discrimination data relate at the group level, we investigate this data at the subject level. Understanding the relationship between the two perception media requires understanding this relation at the subject level. Second, and relatedly, the learner group is less homogenous than either group since there is a range of Russian language knowledge in this group. By investigating the data at the subject level we can determine whether and to what extent different perception domains (and in later chapters, production) relate to each other by factoring in experience for this group. The data is split by place of articulation and syllable position and we start off with onset coronals.

6.5.4.1. Onsets

The figures in this section represent both AXB discrimination data, Categorization data and finally figures showing how these two experiments relate to each other at the subject level. During analysis at the subject level, individual subjects are coded via a numbering scheme, where subjects in the naïve group are represented with a leading letter "E" followed by a number. For example, subjects E1, E2 and so on are each individual subjects in the naïve group. The same convention applies for the learner group (L1, L2 and so on) and the native Russian group (R1, R2 and so on).

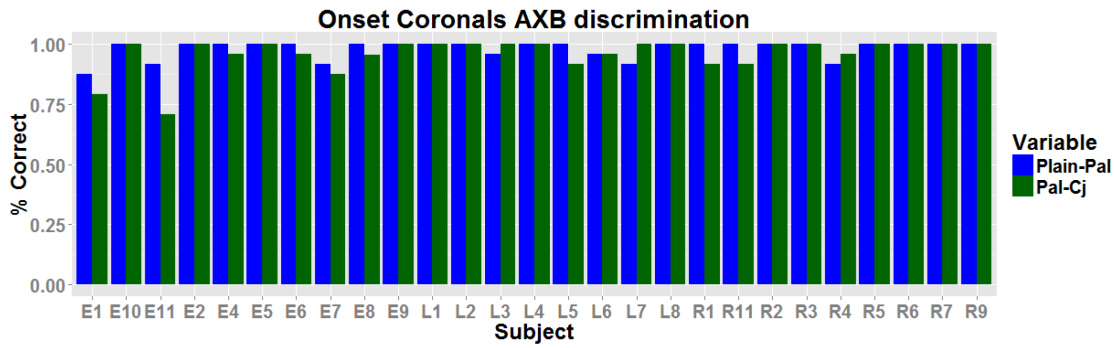

Figure 6.1: % correct response for AXB discrimination by subject for onset coronals for plain-palatalized and palatalized-C+j contrasts.

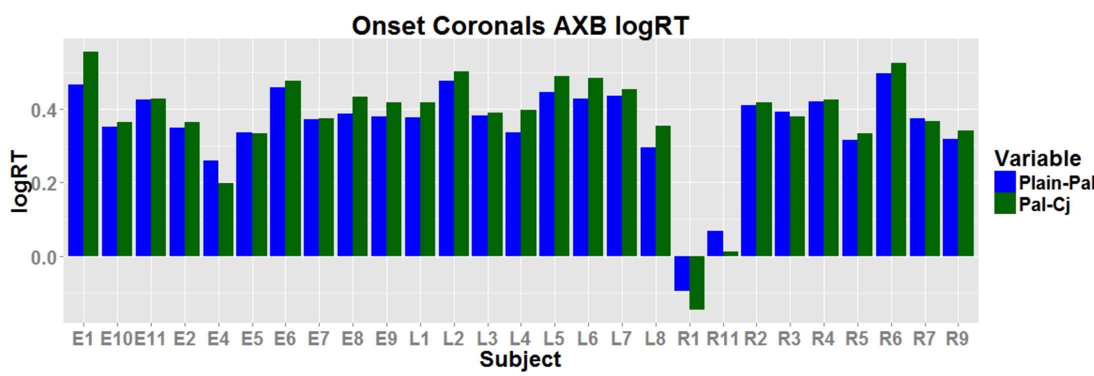

Figure 6.2: logRT for the discrimination data by subject for onset coronals for plain-palatalized and palatalized-C+j contrasts.

Figure 6.1 above shows the % correct response for each subject for two of the three contrast types for the AXB data (see previous chapter). The plain-Cj contrast type was left out because it was perceived the best by all groups and so we focus on the more challenging and interesting contrast types, those involving Russian palatalization. First, for the onset contrasts, we see that the Pal-CJ contrast was generally more difficult than the plain-pal contrast. Two subjects, E1 and E11 (from the naïve group) had % correct response near or below 75% for this contrast. These two subjects also had relatively high assimilation rates to the “CJ” category for the palatalized consonant. Other naïve English subjects, like E4 and E8 also had relatively high assimilation rates to this category; however these two subjects had relatively high discrimination rates. Figure 6.2, Figure 6.3, and Figure 6.4 show the logRT data for the same contrasts by subject, and assimilation data by subject for the onset palatalized coronal and consonant + glide sequences. A further comparison of these patterns is shown in Figure 6.5 below.

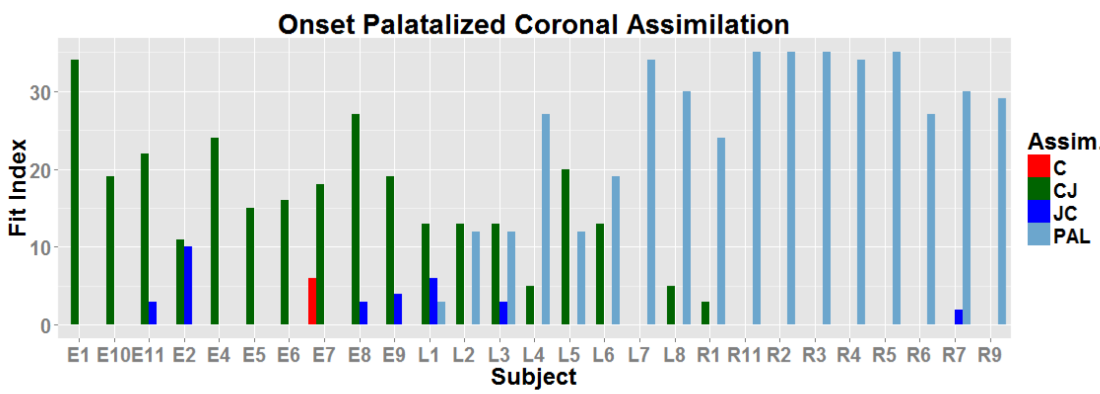

Figure 6.3: Assimilation fit index by subject for each of the assimilation categories for the onset palatalized coronal.

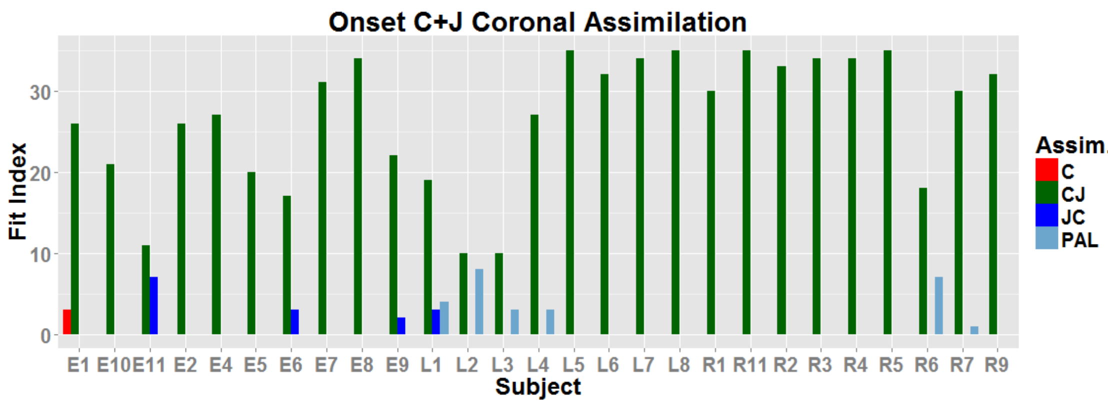

Figure 6.4: Assimilation fit index by subject for each of the assimilation categories for the onset consonant + glide sequence.

It is shown in Figure 6.5 that subjects E11 and E1 had relatively low discrimination rates (albeit still high) for the palatalized-CJ contrast (x-axis), as well as relatively high fit index values for assimilating the palatalized onset coronal to the CJ category (y-axis). We also see the position of E8 and E4 with higher discrimination rates for higher FI to assimilating the palatalized consonant to the CJ category. This means that these two subjects exhibited closer to a SC or CG assimilation pattern yet still discriminated the contrast well. The subject with the closest assimilation pattern to a Two Category (TC) assimilation pattern along with a very high discrimination rate is subject E2. This subject has a much lower FI for assimilating the palatalized consonant to the CJ category than the C+J sequence to the same category. In general, then, there is a weak relation between assimilation of the palatalized consonant to the CJ category and discrimination rate.

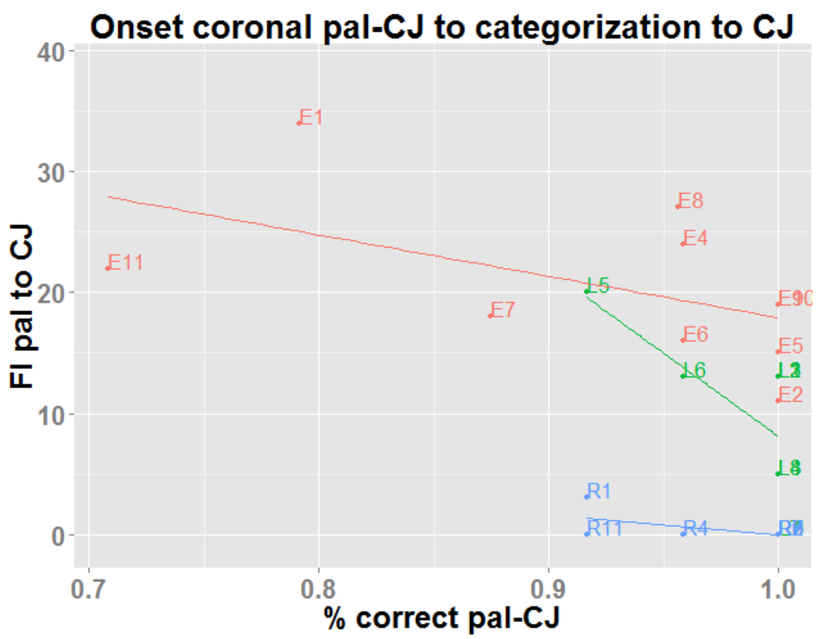

Figure 6.5: Comparison of assimilation fit of the onset coronal palatalized consonant to the consonant+glide category versus the % correct discrimination of the consonant+glide contrast.

This weak relationship is observed in the regression line for each group in Figure 6.5. The line represents the best fit of the relationship, for each group, for the proportion correct response to the palatalized-CJ contrast and the fit index for the assimilation of the palatalized consonant to the CJ category. We see a general negative slope for both English groups, indicating that, in general the better the subject chose to assimilate the palatalized onset coronal to the CJ category, the worse the perception was for the palatalized-CJ contrast. This is in line with the PAM model. If a subject perceives the palatalized consonant as relatively close to a multi-category consonant+glide sequence, discrimination between the palatalized and the consonant+glide contrasts would be worse. We see this, again for certain subjects (such as E11, and E1, and to a lesser extent, L5 and L6 for the learner group), but again, less clearly for other subjects. We move on now to examine onset labials.

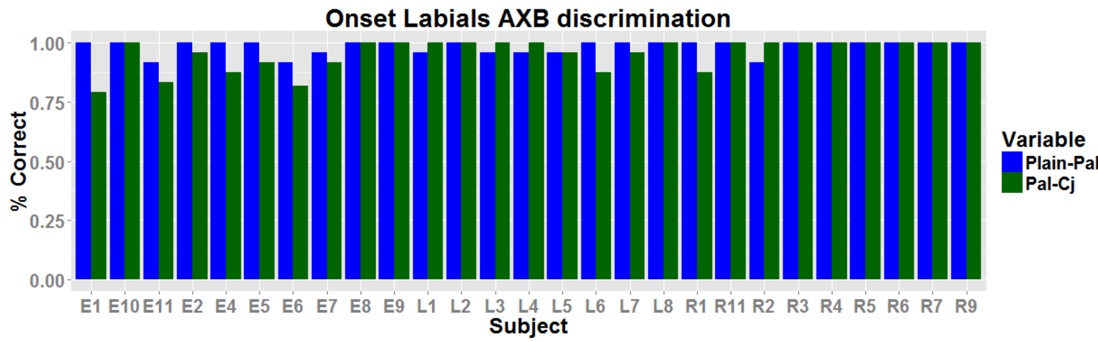

Figure 6.6: % correct response for AXB discrimination by subject for onset labials for plain-palatalized and palatalized-C+j contrasts.

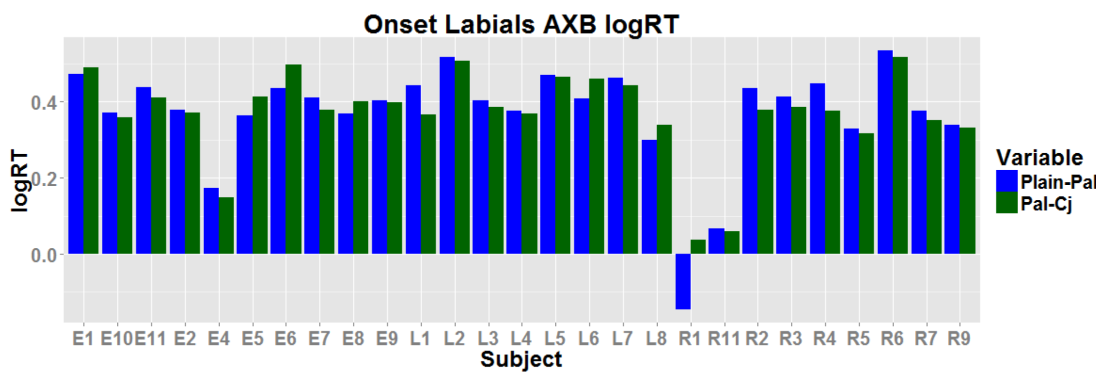

Figure 6.7: logRT for the discrimination data by subject for onset labials for plain-palatalized and palatalized-C+j contrasts.

We see % correct discrimination data plotted for labial onsets. We see again that the palatalized – consonant+glide contrast was more difficult than the plain-palatalized contrast. Three subjects, E1, E6 and E11 had % correct response near or below 75% for this contrast. These three subjects appear at the leftward end of Figure 6.10 below. E1 and E11 had similar values for the coronals as described above. Most subjects also had relatively high assimilation fit indices to the CJ category. There seems to be no pattern as to whether a naïve subject assimilates better the palatalized consonant to CJ and lower discrimination rates for palatalized – consonant+glide contrasts. Indeed, the regression line for the naïve group is positive, unlike that seen for onset coronals. In general we see relatively high assimilation fit index values for both the palatalized and consonant+glide sequences to the same category (C+j). We also see relatively high discrimination rates.

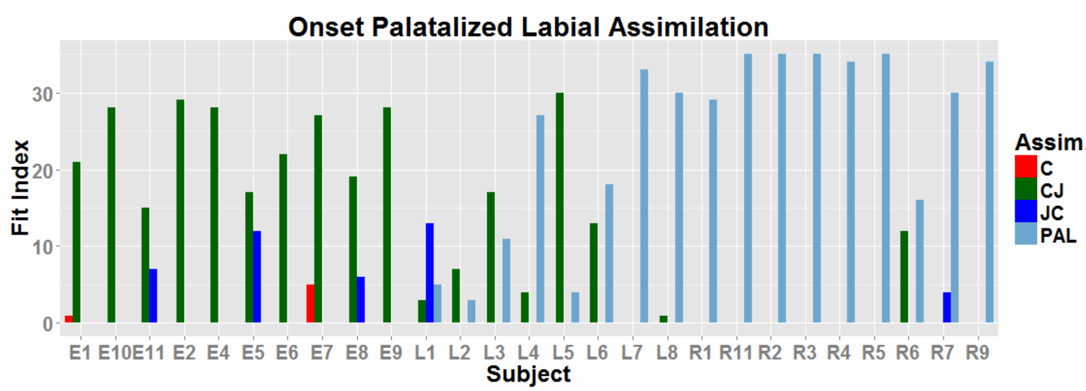

Figure 6.8: For labials, assimilation fit index by subject for each of the assimilation categories for the onset palatalized labial.

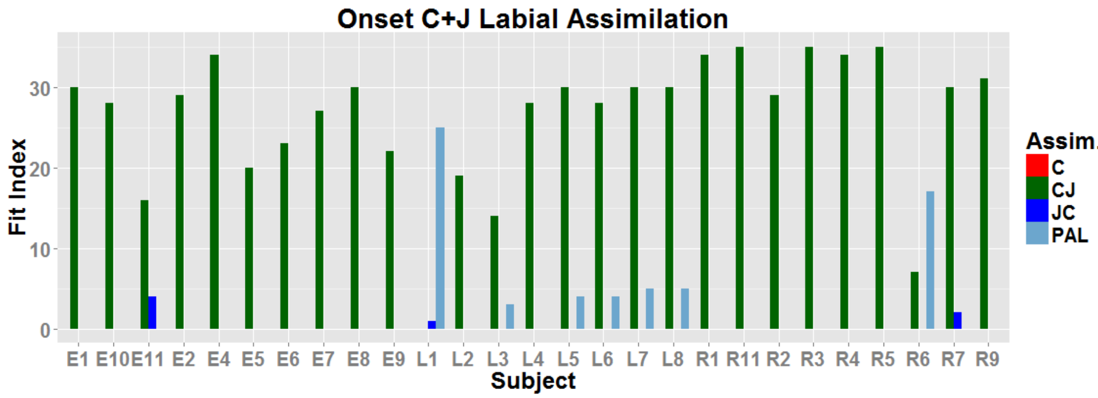

Figure 6.9: For labials, assimilation fit index by subject for each of the assimilation categories for the onset consonant + glide sequence.

Naïve subjects like E4 and E8 had relatively high assimilation rates to this category along with high discrimination rates. For the learner group, however, we see a negative slope in the regression line in Figure 6.10, similar to what we saw for coronals. We also see subject L6 having a relatively low assimilation FI for the palatalized consonant to the consonant+glide category as well as a relatively low proportion correct response for the contrast.

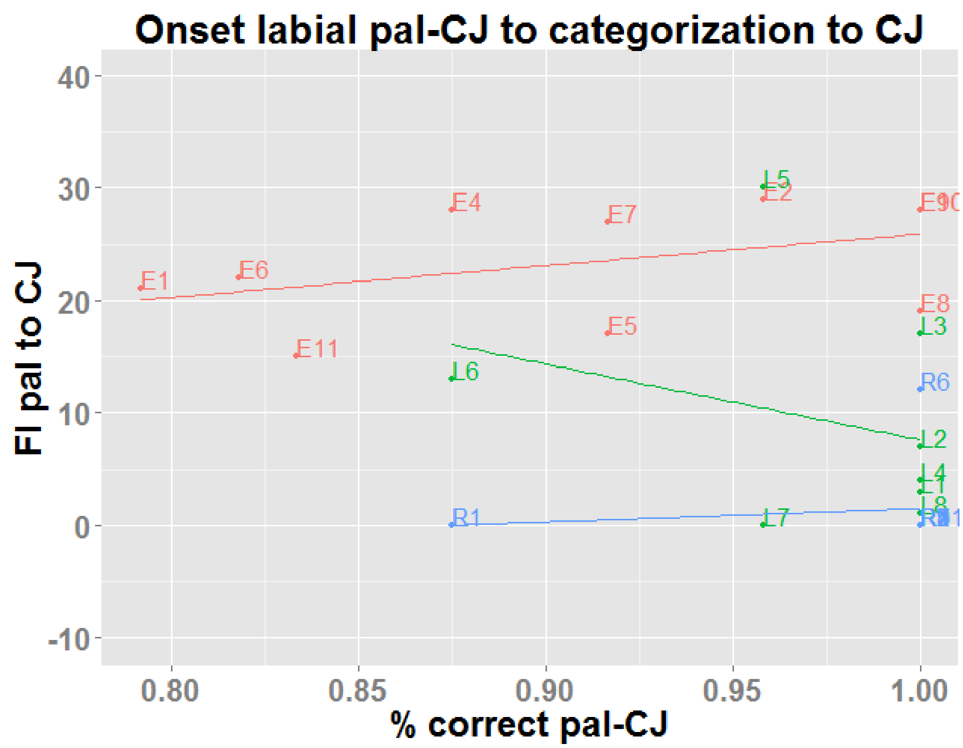

Figure 6.10: Comparison of assimilation fit of the onset labial palatalized consonant to the consonant+glide category versus the % correct discrimination of the consonant+glide contrast.

In sum, for onset consonants, three subjects consistently showed low discrimination rates for their group associated with assimilation of the palatalized consonant to the CJ category, approaching a TC assimilation pattern and thus exhibiting the expected behavior according to the PAM model. For other subjects, a similar pattern did not hold, mainly due to the high discrimination rate exhibited. The results for the comparisons at the subject level for onset consonants show that there is inter-subject variability, but some overall group trends which lend support to the PAM model. This latter observation is true more for coronals than labials, as the discrimination rate for labials for the naïve group was generally higher.

Overall for both labials and coronals in onset position, some subjects exhibited trends that support the PAM model, however, it must be stated that overall the discrimination rates across the board were relatively high, generally staying above 75%. There was also no reliable relationship at the subject level between discrimination rate and tendency to assimilate onset consonants to one category or another.

6.5.4.1.1. Learning Onsets

We now move on to a discussion on learning. The learner group consisted of subjects with varying degrees of Russian experience. We define Russian experience by the subject’s self-reported number of years of Russian exposure, whether formal education or other. Table 6.8 shows the subjects in the learner group sorted by experience in ascending order. The table shows for each subject the following information: their age; gender; years of Russian experience; the AXB discrimination % correct response for the coronal palatalized-glide contrast; the assimilation fit index of the palatalized consonant to the palatalized category; the assimilation fit index of the glide consonant to the glide category; the same values for the labial onset consonants.

We see right away that all members of this group have relatively high discrimination rates for onset coronals and labials for the palatalized – consonant+glide contrast.

Subject	Age	Gender	Years	AXB coronal pal-glide	t^ja to pal	t^jja to pal	AXB labial pal-glide	p^ja to pal	p^jja to pal
L1	25	F	2	1	3	4	1	5	25
L3	28	M	2.5	1	12	3	1	11	3
L8	49	M	4	1	30	0	1	30	5
L5	22	F	5	0.916667	12	0	0.958333	4	4
L6	23	F	5	0.958333	19	0	0.875	18	4
L7	19	M	5	1	34	0	0.958333	33	5
L4	28	M	8	1	27	3	1	27	0
L2	28	F	9	1	12	8	1	3	0

Table 6.8: Discrimination rates, assimilation patterns and years of Russian knowledge for the learner group for onsets.

There seems to be no strong relation between years of Russian experience and discrimination rate. There also seems to be no reliable relationship between assimilation to the palatalized category and years of Russian experience, although subject L1 has a low fit index for palatalized consonants assimilated to the palatalized category and even has high assimilation to this category for the labial C+j sequence. It seems, therefore, that all subjects in the learner group discriminate this contrast relatively well, but can struggle with assimilating the individual palatalized consonant to the proper category, regardless of years of experience.

6.5.4.2. Codas

We now move on to discussing the relation between AXB discrimination and categorization for coda consonants. We present similar figures showing both measures of AXB discrimination (% correct and logRT) and assimilation patterns for each subject. As for the onsets, we see considerable inter-subject variability in both AXB and categorization subject level results.

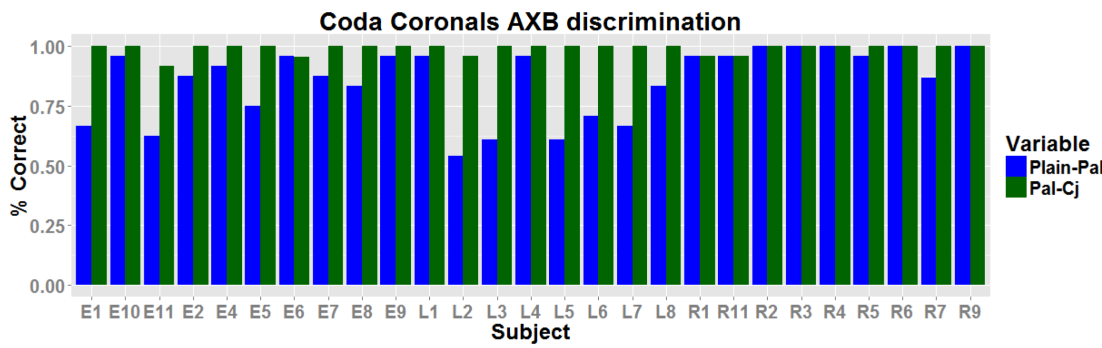

Figure 6.11: % correct response for AXB discrimination by subject for coda coronals for plain-palatalized and palatalized-C+j contrasts.

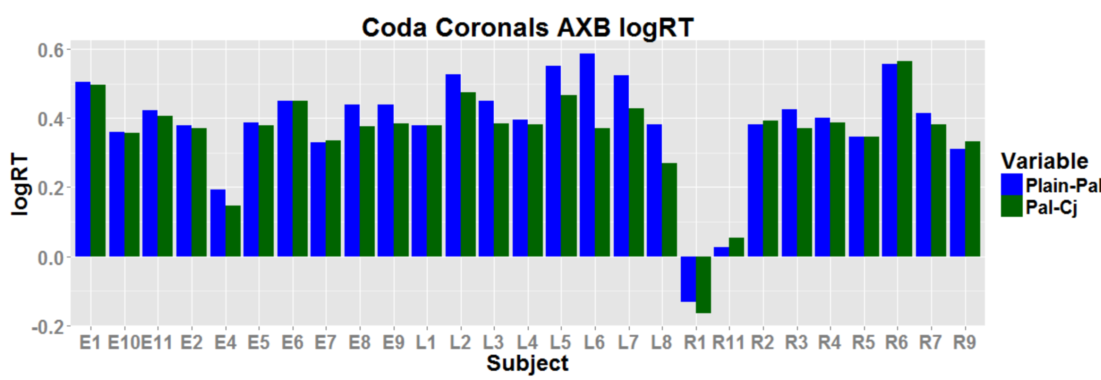

Figure 6.12: logRT for the discrimination data by subject for coda coronals for plain-palatalized and palatalized-C+j contrasts.

For coda coronals we see a range of % correct responses, especially for the plain-palatalized contrast (in blue) in Figure 6.11. This is true because the plain-palatalized contrast in coda position was more difficult for more subjects. We see that certain subjects, namely E1, E11, E5, L2, L3, L5, L6 and L7 performed at or below 75% correct response for the coda coronal plain-palatalized contrast. That is, 3 naïve English subjects and 5 learner subjects fell into this category.

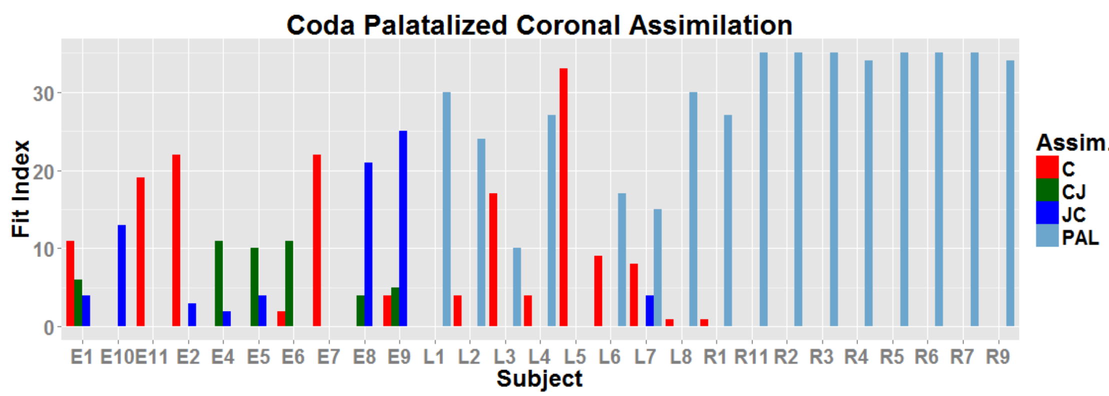

Figure 6.13: Assimilation fit index by subject for each of the assimilation categories for the coda palatalized coronal.

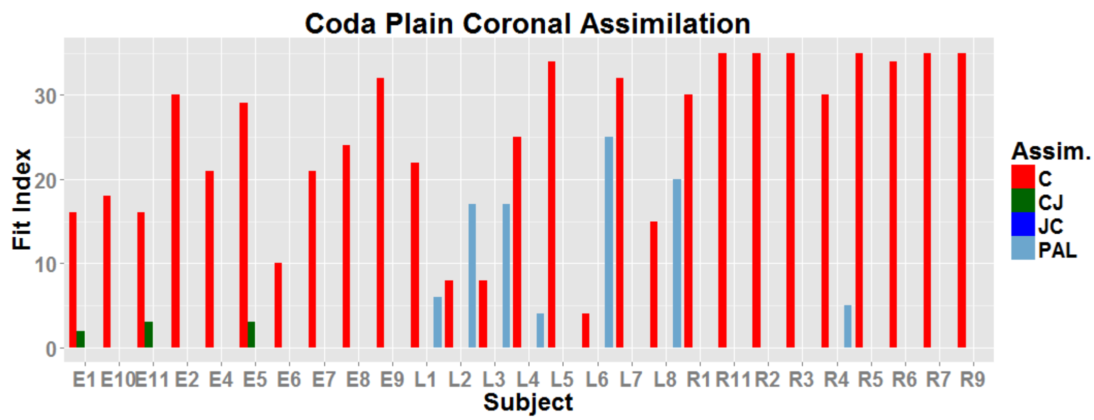

Figure 6.14: Assimilation fit index by subject for each of the assimilation categories for the coda plain coronal consonant.

Figure 6.15 below shows the proportion correct response for this contrast plotted against the fit index of the assimilation of the coda palatalized consonant to the plain category. We see the subjects listed above appear towards the left side of the plot, again demonstrating relatively lower discrimination rates for these subjects. It is interesting to note the diversity within each group. For instance, several learner subjects had lower discrimination rates for this contrast than several naïve subjects.

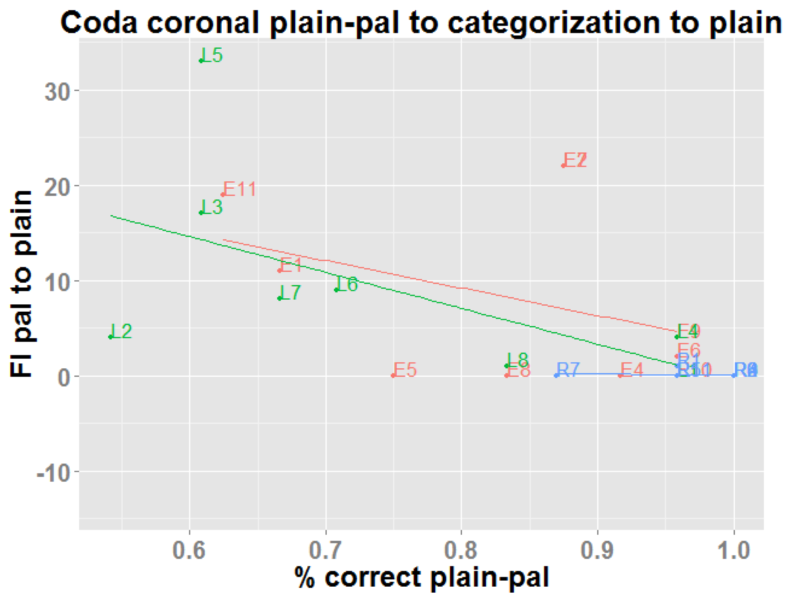

Figure 6.15: Comparison of assimilation fit of the coda coronal palatalized consonant to the plain category versus the % correct discrimination of the plain-palatalized contrast.

We also see that some of this subgroup of subjects, such as E11, L3 and L5, had relatively high assimilation rates for the palatalized consonant to the plain category. Of the subjects who had lower discrimination rates, a subset also have lower assimilation fit index values for the palatalized consonant to the plain category. Subject L2 had a lower assimilation FI value to the plain category than might have been expected. The reason a higher FI value for the palatalized consonant to the plain category could be expected for this subject is due to the lower discrimination rate for the plain-palatalized contrast. This same subject also had a higher assimilation FI of the palatalized consonant to the actual Russian palatalized category as shown by the light blue bar in Figure 6.13. Subject L2 had low discrimination rates for the plain-palatalized coda contrast while still having a relatively high fit index for the palatalized consonant to the palatalized category. This is unexpected because a high assimilation fit index for the palatalized consonant to the actual Russian palatalized category along with a low fit index for this consonant to the plain category should mean that this subject is able to identify the palatalized consonant relatively well, and according to PAM should be able to discriminate the contrast involving this consonant. So the question is why this subject who discriminates the palatalized-plain contrast relatively poorly still seems to assimilate the palatalized consonant to the appropriate category. When we consider the assimilation of the plain consonant in Figure 6.14 above it is clear that Subject L2 also has a relatively high FI for the plain consonant to the palatalized category. This result is unexpected in itself. It was not predicted that the plain consonant /p/ should be assimilated to anything other than a plain English category like /p/. However, considering that the consonants are not actually plain and are indeed velarized, this result might not be so surprising. In this case, for this subject assimilations occur more frequently to other categories such as /p^j/. The same is true for Subject L3.

Given these results, subject L2 has similar assimilation patterns for both the plain and palatalized consonants, which helps to explain the relatively low discrimination rate. The observation that this subject assimilated both plain and palatalized consonants relatively similarly goes to explain the relatively low discrimination rate of this contrast. However, an interesting result in itself is the observation that some subjects assimilate the plain (or velarized) consonant to the palatalized category. This is expanded upon in Section 6.5.4 below on learning.

We investigate one other assimilation type in detail for coda consonants, the assimilations of the palatalized consonant to the glide+consonant (JC) category. Figure 6.16 below plots the assimilation fit index for the palatalized consonant to the JC category against the proportion correct response for coda plain-palatalized contrasts.

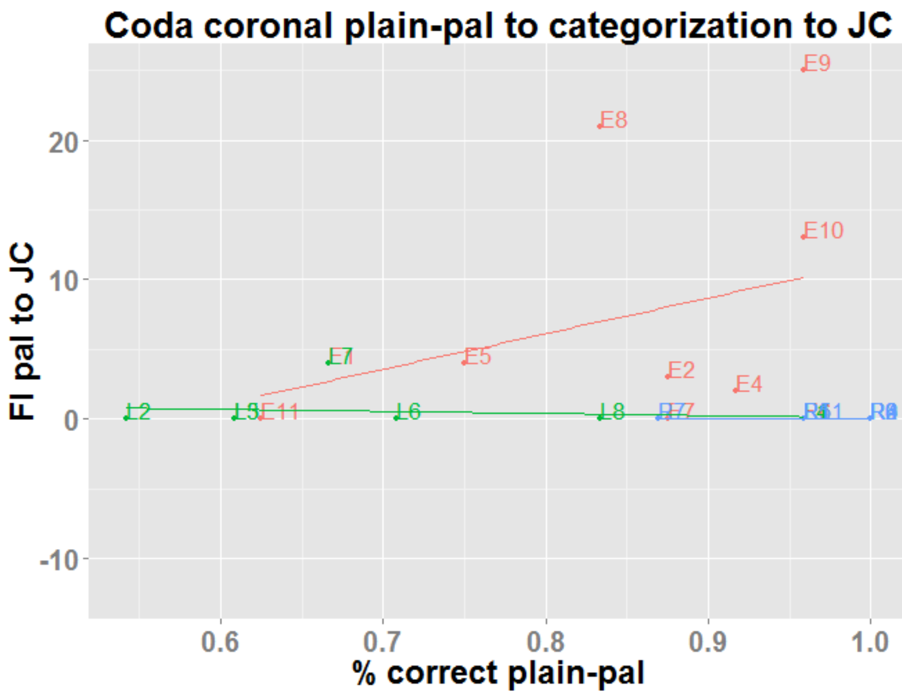

Figure 6.16: Comparison of assimilation fit of the coda coronal palatalized consonant to the glide+consonant category versus the % correct discrimination of the plain-palatalized contrast.

From Figure 6.16, we see that the few subjects who had relatively high assimilation fit indices for this consonant to the JC category (E8, E9, E10) also had relatively high discrimination rates for the plain-palatalized contrast. These subjects also did not have relatively high fit index assimilation values for this consonant to the plain category. The remaining subjects did not have consistent assimilations to the JC category. Considering all of this data we can place certain subjects into sub groups. The first group contains subjects that behave according to the PAM, in the second group those that exhibit results contrary to PAM.

The first group we can further subdivide on how they fall into this group. The first sub-group contains those subjects who perceived the plain-palatalized contrast relatively poorly and who had relatively high assimilation rates for the palatalized consonant to the plain category. This includes subjects E1, E11 for the naïve group and L5 for the learner group. The second group includes subjects who also discriminated the contrast relatively poorly but were able to assimilate the palatalized contrast to the Russian palatalized category but also assimilated the plain consonant to this same category, resulting in a SC or CG assimilation pattern but for somewhat unexpected reasons. This includes subjects L2, L3, and L6 for the learner group. Some subjects discriminated the contrast relatively well, and also had expected assimilation patters. Subjects who fell into this pattern were the three naïve subjects who assimilated the palatalized consonant to the JC category, this includes subjects E8, E9, 10. These three subjects also had high discrimination rates, as expected according to PAM. Subjects E2, E4, and E6 also had higher discrimination rates and assimilated the palatalized consonant to the CJ category thus also having TC assimilation patters. All of the native Russian speakers also fell into this sub-group which had TC assimilation patterns and relatively high discrimination rates.

The second sub group exhibited data that is not as expected according to the PAM. Subjects E5 and E7 both fall into this group for different reasons. Subject E5 had a relatively low discrimination rate but also exhibited a TC assimilation pattern. Subject E7 had a relatively high discrimination rate but exhibited closer to a SC assimilation pattern.

For the coda coronals, we saw a diversity in how subjects within and across groups assimilate and discriminate the contrasts. In general, as indicated by the regression lines in Figure 6.15 and Figure 6.16, a trend does occur which supports the PAM. If subjects assimilated the two members of the contrast to different categories, discrimination rates tend to be higher. However, this is not always the case, as evident by subjects like E5 and E7. It is clear that individual subjects differ in their perception of coda coronal consonants.

6.5.4.3. Coda Labials

We now move on to the labial consonants. Again we see inter-subject variability, especially for both English groups, while much less for the native Russian group. Again, the first two plots below show the discrimination data for each subject, the top panel being proportion correct response and the bottom panel logRT.

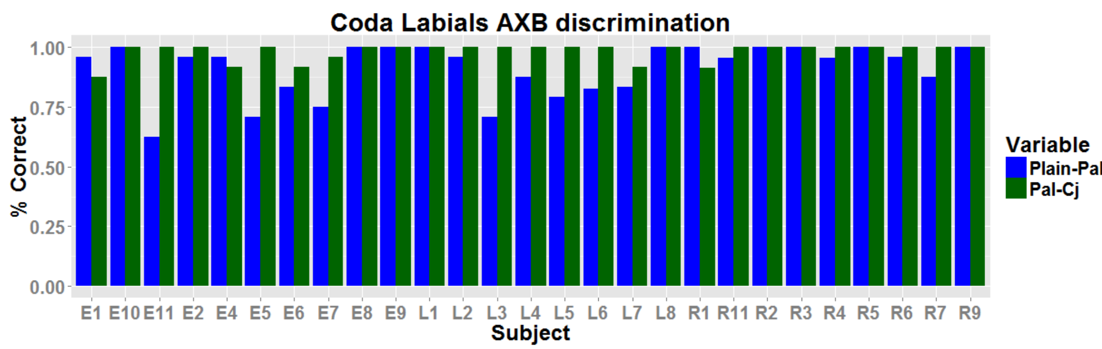

Figure 6.17: % correct response for AXB discrimination by subject for coda labials for plain-palatalized and palatalized-C+j contrasts.

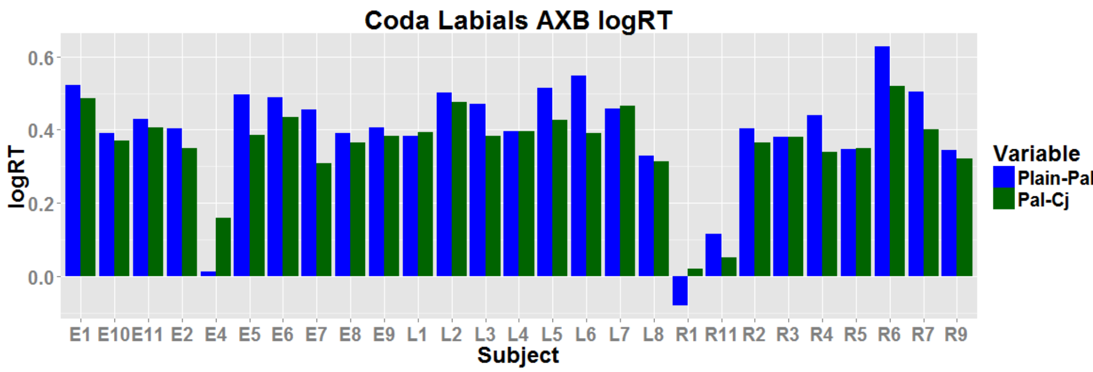

Figure 6.18: logRT for the discrimination data by subject for coda labials for plain-palatalized and palatalized-C+j contrasts.

We again see more variation for the plain-palatalized contrast (in blue) for Figure 6.17 and Figure 6.18. We see subjects, namely E5, E7, E11 and L3 that performed at or below 75% correct response for coda labial plain-palatalized contrasts. Below we see a plot of assimilation patters for the categorization data for the palatalized coda labial consonant (Figure 6.19) and the plain consonant (Figure 6.20). Figure 6.19 also shows the variety of assimilations between subjects, while Figure 6.20 shows more homogeneity given that this plot represents the assimilations for the plain labial, a consonant that is similar enough to an English consonant category that relatively high fit indices to the plain category are seen.

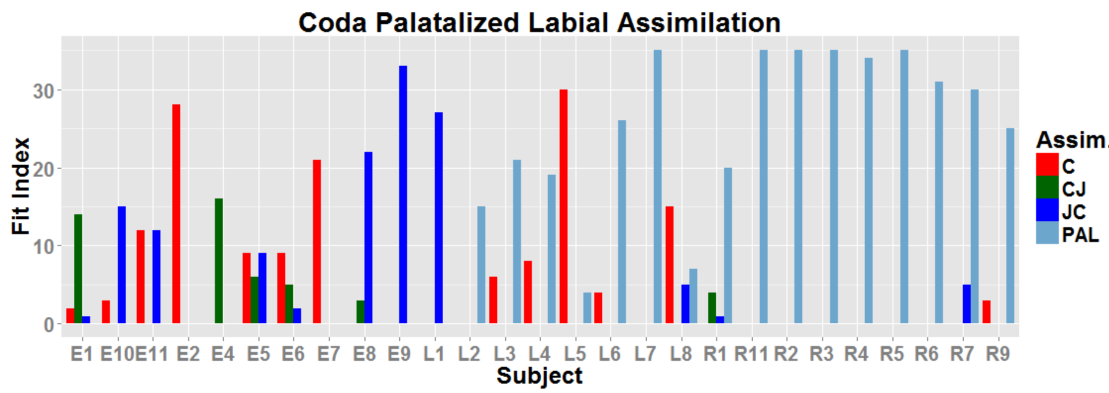

Figure 6.19: Assimilation fit index by subject for each of the assimilation categories for the coda palatalized labial.

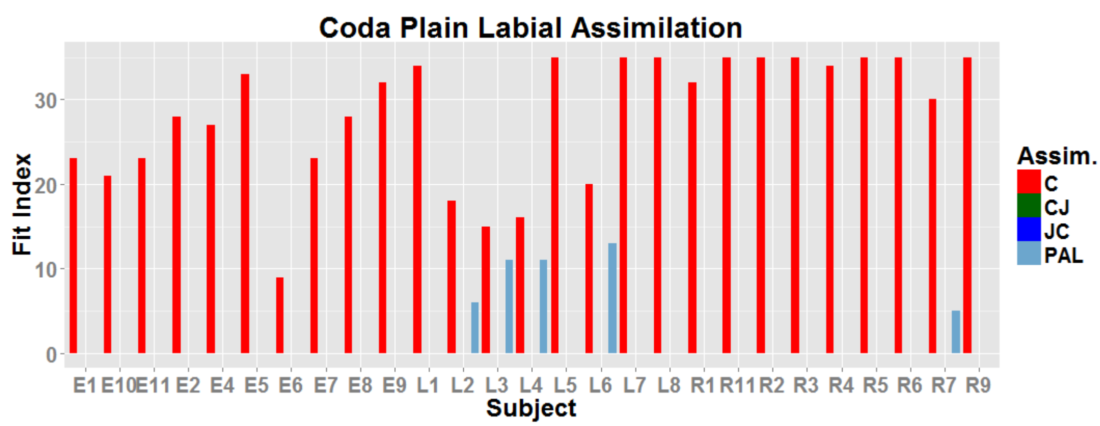

Figure 6.20: Assimilation fit index by subject for each of the assimilation categories for the coda plain labial consonant.

The plots above show that some subjects have a variety of assimilation types for the palatalized consonant (multiple colored bars) and others are more stable. Figure 6.21 below shows again, the fit index of the palatalized consonant categorized as plain plotted against the discrimination rate for the plain-palatalized contrast, as was shown for the coronals.

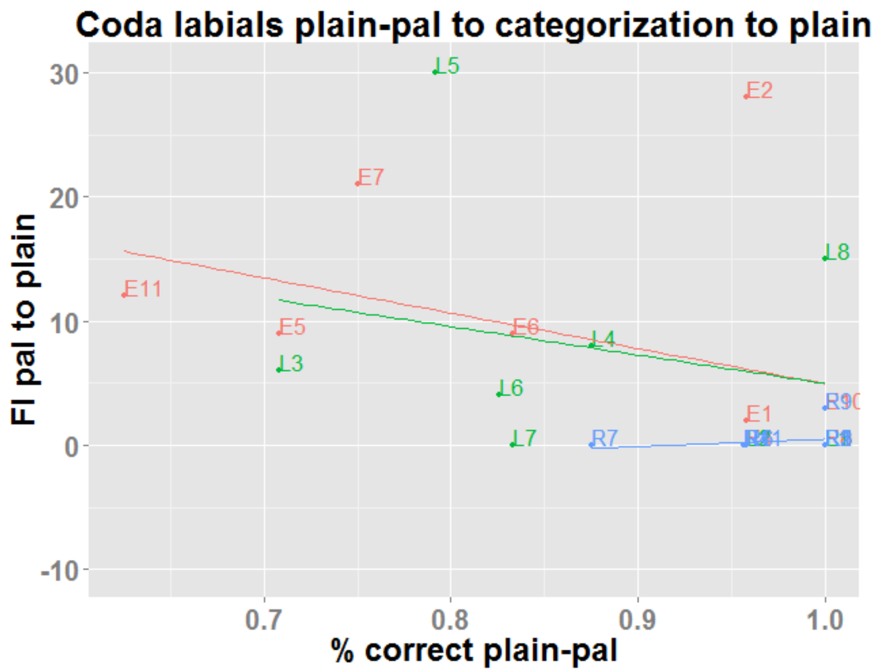

Figure 6.21: Comparison of assimilation fit of the coda labial palatalized consonant to the plain category versus the % correct discrimination of the plain-palatalized contrast.

We see the subjects listed above that had relatively low % correct response appear towards the left side of the plot. Again, there is diversity in the groups, for instance, subject L3 had relatively low discrimination rates, lower than several naïve subjects, like E1 and E2. Below we see a plot showing the fit index assimilation to the JC category. Here we see a difference in that for both the English groups the tendency is for a higher assimilation fit index to this category to correspond to a higher % correct response.

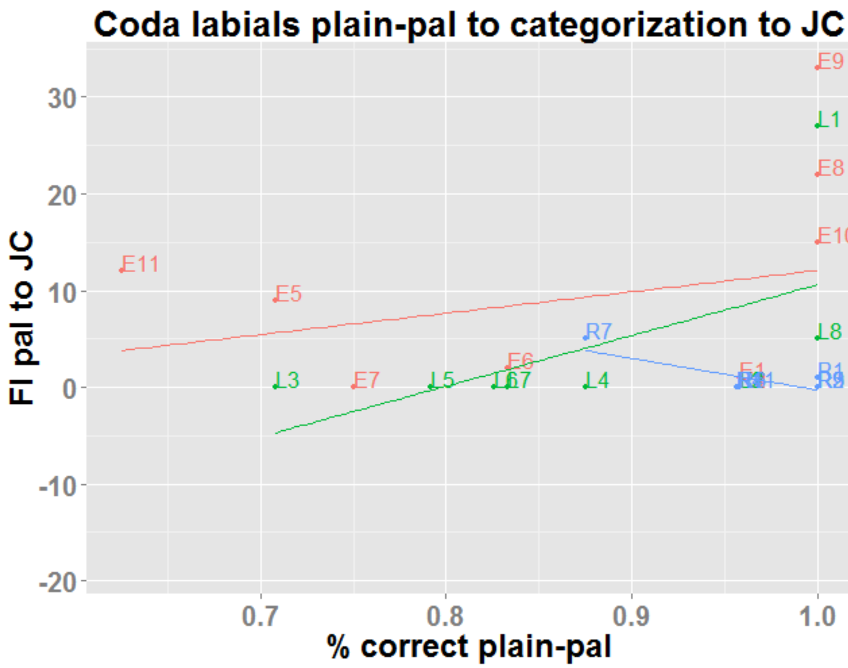

Figure 6.22: Comparison of assimilation fit of the coda labial palatalized consonant to the glide+consonant category versus the % correct discrimination of the plain-palatalized contrast.

Considering the above data in light of PAM, as we have done for the coronal consonants we can again subdivide the subjects into different sub-groups. As for the coronals the first group consists of subjects that behave according to the PAM; the second group those that exhibit results contrary to PAM.

The first group is further subdivided according to assimilation patters. The first sub-group contains subjects who perceived the plain-palatalized contrast relatively poorly and who had relatively high assimilation fit index values for the palatalized consonant to the plain category. This group includes subject E7 and to a lesser extent E5 and E11. The second group includes subjects who also discriminated the contrast relatively poorly but were able to assimilate the palatalized contrast to the Russian palatalized category but also assimilated the plain consonant to this same category, resulting in a SC or CG assimilation pattern but for somewhat unexpected reasons. Only L3 of the learner group fell into this category. Some subjects discriminated the contrast relatively well, and also had expected assimilation patters. A sub-group which fell in this pattern were the naïve subjects who assimilated the palatalized consonant to the JC category, this includes subjects E8, E9, 10. In addition, subject L1 from the learner group fell into this category. These four subjects also had high discrimination rates, as expected according to PAM. Subjects E1 and E4 also had higher discrimination rates and assimilated the palatalized consonant to the CJ category thus also having TC assimilation patters. All of the native Russian speakers also fell into this sub-group as well.

The second sub group exhibited data that is not as expected according to the PAM. Subjects E2, E5 and E11 fall into this group for different reasons. Subjects E5 and E11 had a relatively low discrimination rate but also exhibited assimilation patterns closer to Two Category assimilations. Subject E2 had a relatively high discrimination rate but exhibited closer to a SC assimilation pattern. The two naïve subjects E5 and E11 from the first subgroup who exhibited relatively low discrimination rates also assimilated this consonant to the JC category. These subjects exhibited relatively high FI assimilation values for the palatalized consonant to both the plain category and the JC category. In one case this constitutes a SC pattern, in the other a TC assimilation pattern. Either way there is inconsistency in how these subjects are assimilating the palatalized consonants and thus this is indicative of greater difficulty in perception for these subjects.

We expand upon this more in future sections when we discuss the relation between perception and production, but one of the most interesting results that emerges from considering the individual subjects is the observation that some subjects perceive the palatalized consonants as sequences of TB and consonantal gestures. The fact that a few (mostly) naïve subjects responded to the categorization task for the coda palatalized consonants as having heard sequences of glide + consonant is interesting because they seem to be assimilating these tokens to multi-category glide+consonant or rising diphthong+consonants (as in a word like “bate” [beɪt]). We have touched upon this earlier and we will come back to this again, but Articulatory Phonology is again well suited to describing this phenomena. What we are seeing is listeners attributing the gestures that constitute one category (or segment) in one language (Russian palatalized consonants) into multiple categories (or segments) in their L1. The reason Articulatory Phonology succeeds here is that we are considering gestures, or constellations of gestures, and not segments or strict categories.

6.5.4.3.1. Learning Codas

We now move on to a discussion on learning for coda consonants. Table 6.9 lists the subjects in the learner group sorted by experience in ascending order. In contrast to the onset consonants, subjects with varying degrees of Russian knowledge, from 2 to 9 years exhibit relatively low discrimination rates. This is evident as Subject L2 with 9 years of experience has a 0.54 proportion correct response discrimination rate for the coda coronal plain-palatalized contrast. This is actually the lowest proportion correct response, nearing chance. This subject also had relatively high assimilation fit index values for the plain coronal consonant to the palatalized category as described in previous sections. The fact that some learner subjects assimilated the plain consonant to the palatalized category was described earlier as unexplained and unexpected since this consonant is similar to the English category /p/ and therefore should have been well assimilated to this category. However, since these subjects have some Russian knowledge, and subject L2 having 9 years, a substantial amount of experience in the language, we can speculate that some learner subjects are over assimilating non-palatalized consonants to the non-native Russian palatalized category.

There seems to be no strong relation between years of Russian experience and discrimination rate. There also seems to be no reliable relationship between assimilation to the palatalized category and years of Russian experience, although subject L1 has a low fit index for palatalized consonants assimilated to the palatalized category and even has high assimilation to this category for the labial C+j sequence.

Subject	**Age**	**Gender**	**Years**	**AXB coda coronal plain-pal**	**at to pal**	**at^j to pal**	**AXB coda labial plain-pal**	**ap to pal**	**ap^j to pal**
L1	25	F	2	0.96	6	30	1	0	0
L3	28	M	2.5	0.61	17	10	0.71	11	21
L8	49	M	4	0.83	20	30	1	0	7
L5	22	F	5	0.61	0	0	0.79	0	4
L6	23	F	5	0.71	25	17	0.83	13	26
L7	19	M	5	0.67	0	15	0.83	0	35
L4	28	M	8	0.96	4	27	0.875	11	19
L2	28	F	9	0.54	17	24	0.96	6	15

Table 6.9: Discrimination rates, assimilation patterns and years of Russian knowledge for the learner group for codas.

6.6. Discussion: Comparisons

At the group level, the comparisons between the two experiments, AXB discrimination and Categorization seem to largely support the PAM. It is true that the closer the members of a contrast are to being assimilated as Single Category (SC) or Category Goodness (CG) the more likely it is for that contrast to have a lower discrimination rate. However, the reader should again be reminded that all of these contrasts were relatively easy for all groups, never reaching anything close to chance level. However, the significant results for the AXB task at the group level do show that for certain contrasts the groups did differ from each other significantly. Where they differed significantly they also exhibited less than perfect Two Category assimilation patterns in the categorization data, constituting a consistent finding in support of PAM.

Another interesting result discovered here is better observed by examining individual subjects. We saw that sub-groups of subjects behaved differently, and one of the interesting observations has to do with coda consonants. When subjects tended to perceive coda palatalized consonants as sequential glide+consonant they also discriminated the plain-palatalized contrasts better, further supporting PAM and also shedding light on one of the predictions outlined earlier in the dissertation. Some subjects would assimilate palatalized consonants to multi-category sequences. The reason we proposed this and are seeing these results is that if we take Articulatory Phonology seriously, we can claim that what listeners perceive are gestures, and they are reconstructing the gestures in a more nativelike timing manner. For some speakers, this involves sequential timing of the tongue body (TB) and consonantal gesture.

Chapter 7 Task 3: perception of timing: synthesized speech

An additional task involving synthesized speech tokens was carried out to test subjects' perception of the timing of the two gestures active in palatalized consonants. Previous studies have employed synthesized tokens in attempting to control for acoustic or phonetic features and cues while varying the relevant phonetic property along a continuum (Lisker and Abramson 1970; Gottfried and Beddor 1988; Beddor and Gottfried 1995; Underbakke et al. 1988). The goal of this experiment was to synthesize speech tokens varying in the timing between articulators to determine whether groups differing in language experience or background perceive different timed gestures differently. For example, we wish to test whether naïve English speakers perceive tokens that have greater lag in timing of tongue body gestures and consonantal gestures more similarly to tokens which have more simultaneous timing.

7.1. Methods

7.1.1. Tokens

For each of the two places of articulation studied in this dissertation, (labials and coronals) a synthetic continuum was constructed varying the timing of the primary and secondary gestures as a function of a "step" along this continuum. The continua were constructed in the TADA system (Task Dynamic Application; Nam et al. 2004), a task dynamics modeling suite that generates articulatory representations and subsequent acoustic output. In TADA, articulatory gestures and their relative timing, or phasing, can be modeled and manipulated. TADA is the modelling component of the Articulatory Phonology framework, and more specifically, incorporates the more general applications of Task Dynamics (Saltzman & Munhall 1989) which strives to reconcile the seeming contradictory ideas of higher level phonological constancy with phonetic and temporal variability.

The two continua were each constructed for both places of articulation for the consonants involving palatalization (/p^j/ and /t^j/). In Articulatory Phonology, gestures are modeled using tract variables. The tract variables involved in the tokens under investigation specified a primary lingual or labial constriction gesture: Tongue Tip Constriction Degree (TTCD) and Location (TTCL) for /t/; and Lip Aperture (LA) for /p/; as well as a glottal gesture for voicelessness (both /t/ and /p/). To represent the palatalization, a separate gesture with specifications for /j/ was superimposed on these tokens. The way this was done is described as follows. First, using the Gest function in TADA, a gestural score for an utterance involving the tract variables for the consonantal, glottal and tongue body gesture was generated. The specifications for the tongue body gesture were the same for both places of articulation, and the relevant tract variables for this gesture are Tongue Body Constriction Degree (TBCD) and Tongue Body Constriction Location (TBCL). Then, the entire continuum was constructed using the Gest function in TADA varying the relative temporal coordination between the tongue body and consonantal gestures. This process created tokens which were described above, consisting of coordinated consonantal and palatal gestures. The continua were then constructed by manipulating the beginning and end times of tract variables employed in the /j/ gesture for each token. Next, an acoustic output .wav file was extracted for each token.

Once the relevant palatalized tokens were created, the continua were constructed based on these palatalized items. For example, the continuum constructed based on the palatalized segment /t^j/ involves on one extreme, the TB gesture timed to considerably precede the primary

(TT) gesture, while on the other extreme, the TB gesture is timed to considerably follow the primary gesture. Between these two extremes, tokens varied incrementally based on the timing of the two gestures. Two continua were created, one for each of the palatalized consonants (t^j, p^j). For each of the two TADA continua (labials, coronals), a token along the continuum was designated as most closely resembling a naturally spoken Russian palatalized token.

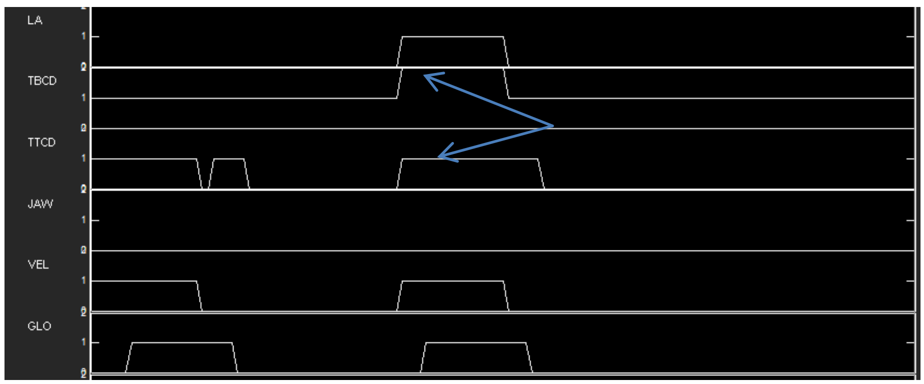

Figure 7.1: TADA token representing simultaneous coordination of TB glide gesture (TBCD) with voiceless tongue tip constriction gesture (TTCD).

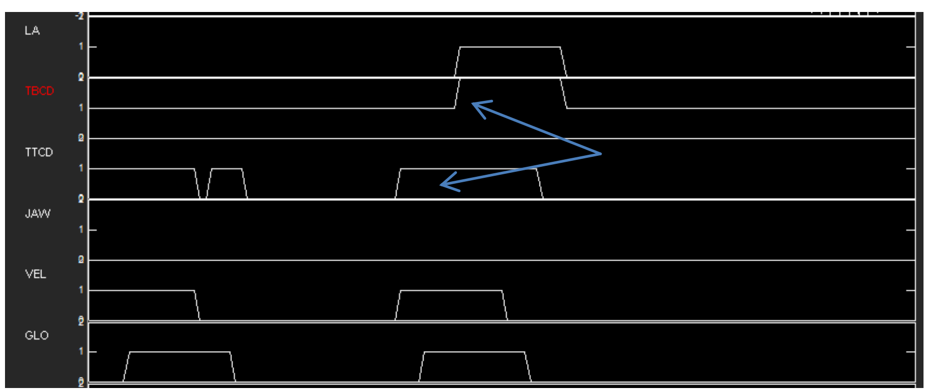

Figure 7.2 TADA token representing sequential coordination of TB glide gesture (TBCD) with voiceless tongue tip constriction gesture (TTCD). The TB gesture follows the TT constriction gesture.

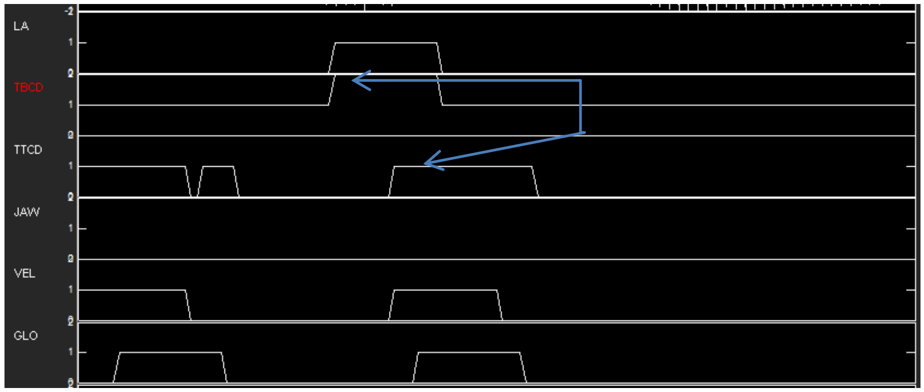

Figure 7.3: TADA token representing sequential coordination of TB glide gesture (TBCD) with voiceless tongue tip constriction gesture (TTCD). The TB gesture precedes the TT constriction gesture.

Figure 7.1 demonstrates the tract variables for this token. The tract variable TBCD (tongue body constriction degree) is phased in a simultaneous relation to the TTCD (tongue tip constriction degree) tract variable in the figure. Figure 7.2 shows the TADA token where the TB gesture considerably follows the TT gesture, and Figure 7.3 shows the TADA token where the TB gesture precedes the TT gesture. The figures above demonstrate examples from the labial continuum, but similar tokens were created for the labials.

The acoustic output of the TADA tokens was then extracted. The TADA acoustic tokens were then analyzed in the same way as the natural speech (see Chapter 8). That is, the F2 values were measured at the same locations as for the natural speech. For the vowel preceding the consonant, measurements were taken at vowel start, 90 ms prior to closure, 60 ms prior to closure, 30 ms prior to closure and at the VC transition. In the following vowel measurements were taken at the CV transition, 30 ms after release, 60 ms after release, 90 ms after release, at the vowel mid-point and at the vowel end. Likewise, the durations of V1, releases, and the consonantal closure were measured. Thus, for each TADA acoustic token and each natural token, the same measurements were obtained

The TADA token representing Russian palatalization (see Figure 7.1) was chosen by comparing the acoustic measurements of each TADA token in each continuum to the average of the values obtained from the natural speech. In both continua this token corresponded to the token where the TB and consonantal gestures were phased exactly synchronously.

The two continua of eleven tokens each varying in phasing of the relevant consonantal (p, t) and tongue body (j) gestures in equal steps were constructed by manipulating the starting and ending frame of the TB gesture in TADA. On one extreme the token phased the TB gesture to precede the TT or LA gesture to produce something like [jt] or [jp] (see Figure 7.3) while on the other extreme the token phased the TB target to substantially follow the release of TT/LA (e.g. [tj] or [pj]; see Figure 7.2). Each intervening token varied in its phasing in equal increments. There are five tokens in which the palatalization is timed to precede the primary constriction, one token in which timing is exactly synchronous, and five tokens in which the palatalization is timed to follow the primary constriction.

The acoustic stimuli that subjects hear were constructed as follows. The middle token representing Russian palatalization was paired with each of the members of the continuum, including itself, in AX trials. There are two possible AX pairings (XA and AX). For each of the two consonant continua (labials and coronals) there were 11 continuum steps for each AX ordering and each individual AX trial was repeated four times for a total of (11 pairs x 2 ordering x 2 consonants x 4 repetitions =) 176 AX trials per subject.

7.1.2. Procedure

The trials were presented in random order ensuring that no two identical trials appeared consecutively. The inter trial interval was 1 second and the inter-stimulus interval was 0.5 seconds. The task for the participants was to hear each AX trial and decided whether the two members were the same or different by clicking one of two buttons on the screen labeled “same” or “different”. Participants also provided a confidence rating on a 5 point scale of their confidence in the choice they made. Finally, the reaction time was also measured for each trial. The results from this task provide information on how speakers perceive the timing of the relevant gestures. This information is then used to correlate speakers’ perception to production with respect to timing. The results are also predicted to differ between groups of naive listeners, learners and native Russian speakers, where Russian speakers are predicted to have less sameresponses for AX trials pairing members of the continuum that differ in the phasing of the gestures than English speakers. The reason for this is due to the similar hypothesis for natural speech in that Russian speakers are predicted to synchronize the tongue body and consonantal gestures for palatalized consonants more synchronously than English speakers; the null hypothesis is that the same should hold for synthesized speech.

7.2. Analysis

The basic unit of measurement for this experiment is a binary variable where the two members of each AX stimulus were perceived as being either the same or different. A value of 1 indicates a “same” response, and a value of 0 indicates a “different” response. A function of same responses based on AX pairing was generated. Each AX pairing in the analysis below is called a “step”. The function of same response to step is presented in Figure 7.4 below. This curve is examined for steps where values diverge, and mixed effects models are run for each of these points of interest to determine whether group differences are significant. The most optimal mixed effects model accounts for both subject and stimulus as random effects, thus the differences reported take into account by-subject and stimulus idiosyncratic differences. Later in the chapter and in subsequent chapters, we more closely examine any relation between this data and other perception and production data both at a group level and for individual subjects.

In addition to proportion same response, reaction time and a confidence rating were also measured. The data will be interpreted as follows. First, if we see relative consistency in labeling for native Russians, and inconsistency in labeling for non-Russians, this will suggest that L2 speakers do not perceive the timing of the gestures as do native Russians (see Meister and Meister 2011 for discussion on inconsistencies in labeling for L2 learners). The TADA data will then be compared to the other perception data in this chapter at both a group and subject level

7.3. Results

Recall that the two continua each span 11 steps, 5 steps where the palatal gesture precedes the consonantal gesture, one step where coordination is simultaneous and 5 steps where the palatal gesture follows the consonantal gesture. We begin by examining the overall sameresponse rates for the coronal data and labial data.

7.3.1. Results: Proportion same response

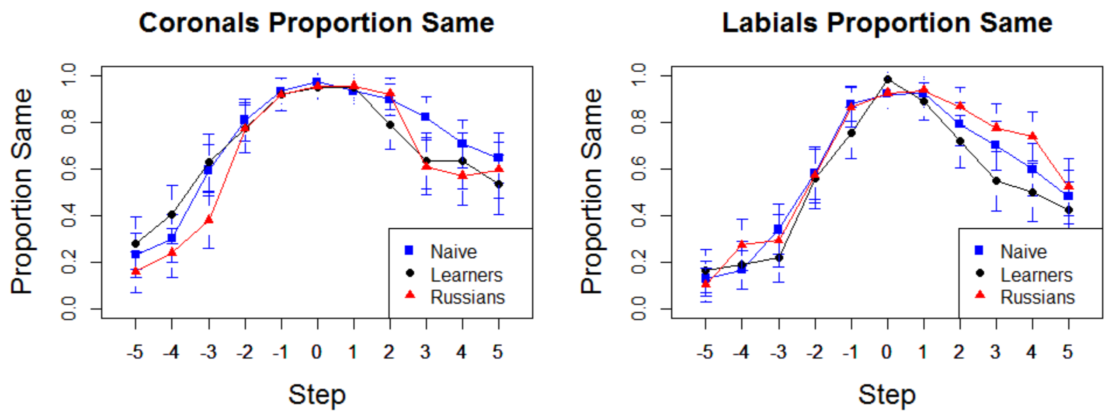

Figure 7.4: Proportion of same response values by group for coronals and labials for each step in the TADA continuum.

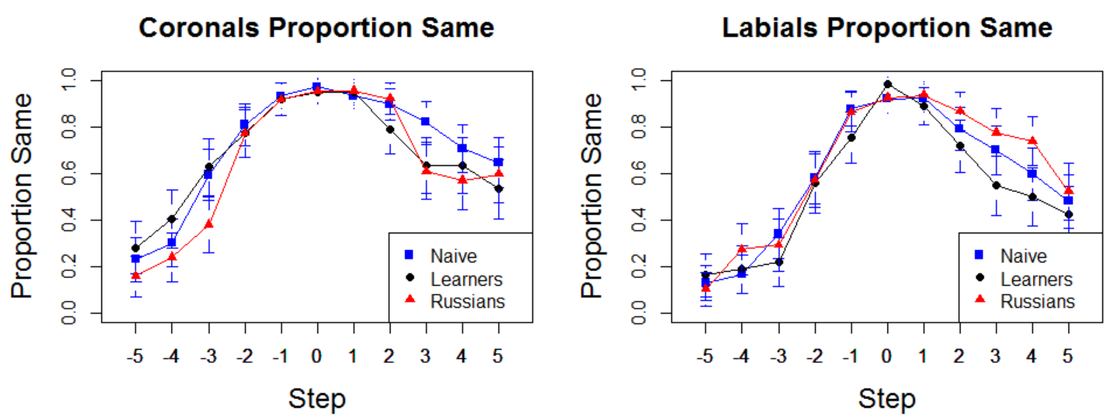

Figure 7.4 above shows the function of proportion of same response by step for each of the three groups for coronals on the left and labials on the right. Overall, for both places of articulation, the three groups show a similar trend. The curves are roughly parabolic in shape. The particular manner that the stimuli were constructed leads to this shape. For each AX pair, or step, one member was the token with simultaneous coordination of the consonantal and palatal gesture. The other member of the pair differed depending on the step. So, for instance, the step labeled 0 in the plot corresponds to the AX pairing where both members of the pair were the same token with simultaneous coordination of the two gestures. A very high proportion of same

responses is expected for this step, and this is exactly what is seen. High proportions of same responses are also expected at steps near 0, which we do see at steps -1 and +1 for all 3 groups.

For the coronal continuum, at point -2, all three groups diverge together, displaying a decrease in proportion same response. At points -3 and +2 the groups seem to diverge from each other. A significant difference between groups emerges at point -3 for the coronals. At this point the native Russian group has a significantly lower proportion same response the Learner group ($pr > z = 0.05$). The difference is not significant between the naïve group and either of the other two groups. This is the only significant difference that exists between groups for any of the steps for either continuum. A few trends are observed which do not reach significance, such as at point +2 the group of English learners of Russian shows a slightly lower proportion same response than the naïve group and native Russian group. There are other trends, but nothing reaches significance.

For the labials, a slightly different pattern emerges. The descriptions here are only trends, none of the differences for labials reached significance. All three groups have high proportion same response values at points 0, -1 and 1, as expected. For the negative points (-2 to -5), all three groups show similar values and show progressively lower proportion same response values. For the positive steps (2 to 5) we also see all three groups showing progressively lower proportion same response values. The three groups differ from each other slightly for points 2 to 5. In each case, the native Russian group has the highest proportion same response, followed by the naïve English group and followed last by the English learners.

7.3.2. Results: Reaction Time and Confidence Ratings

The next sub-section presents reaction time and confidence ratings information from the same experiment. The data is presented in the plots below, which demonstrate the average logRT for the three groups by step, and the average confidence rating for each group by step for coronals on the left and labials on the right.

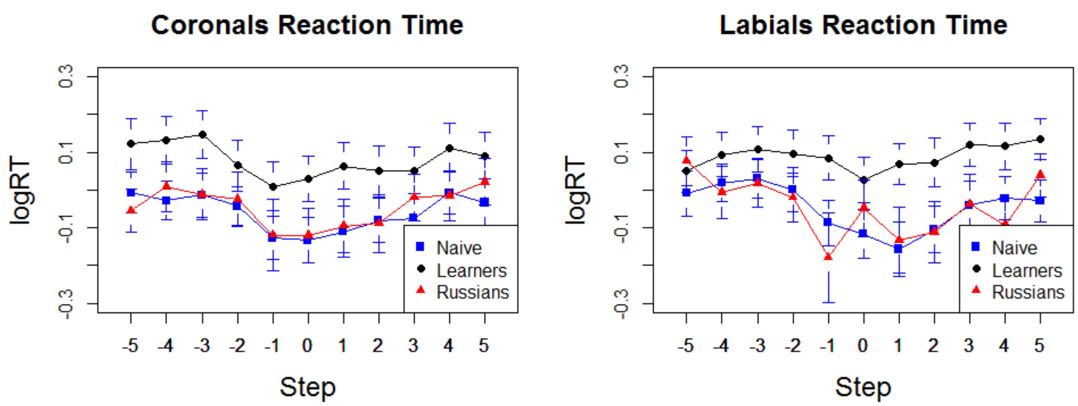

Figure 7.5: LogRT by group for coronals and labials for each step in the TADA continuum.

The plots in Figure 7.5 above show the mean logRT for the same steps. We notice immediately that the group of English learners of Russian seems to have an overall higher logRT across all steps than either naïve English speakers or native Russian speakers. All three groups, however follow a similar pattern for both coronals and labials, in that logRT values are lower at points near step 0, corresponding to those points with similar members of the AX pair. Points further away from point 0, where members of the pair were most distinct show higher logRT values than points near zero. This observation, which holds across groups is interesting because if we use reaction time as a proxy for how easy or difficult a particular stimuli is to discriminate, stimuli near point zero are easier than stimuli further away. That is, it is easier for subjects to agree that stimuli close to each other are the same, than it is for subjects to agree that stimuli further away are either the same or different, and this is true for all groups to similar degrees.

Also, interestingly, point -5 shows slightly lower logRT values than point 5. We could interpret this result in a similar way, that is, that point -5 is easier than point 5, meaning that subjects find tokens which have the TB gesture timed to precede the consonantal gesture more distinct from tokens with simultaneous phasing than they do tokens with the TB gesture phased to follow the consonantal gesture.

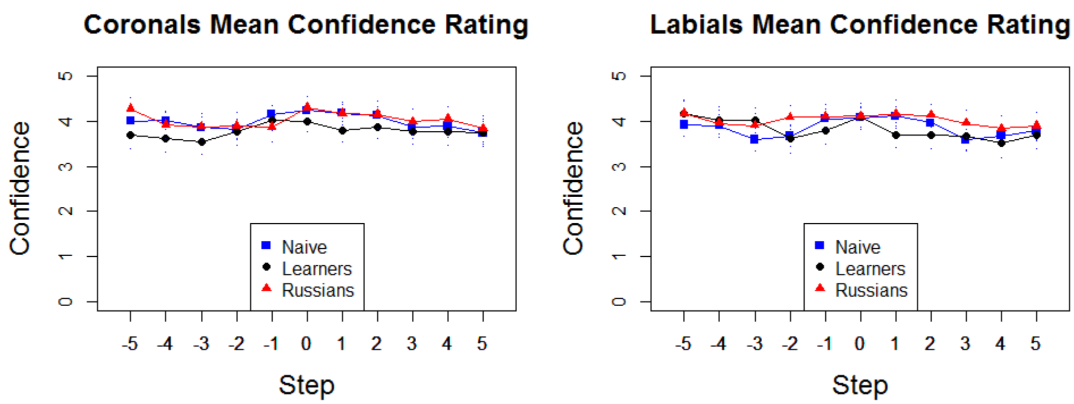

Figure 7.6: Mean confidence rating by group for coronals and labials for each step in the TADA continuum.

The plots in Figure 7.6 above show the mean confidence rating for the same steps. Each subject provided a confidence rating ranging from 1 (not confident) to 5 (very confident) in their selection of either a "same" or "different" response. Again, we see that subjects are most confident in their responses around steps 0, -1 and 1, as expected. Confidence levels also increase near the extremes to an extent. This means that, as expected, tokens that were more similar to each other near the center of the continuum were judged as being the same with greater confidence, and tokens which were more distinct from each other (near the extremes) were judged as being different with greater confidence as well. This is especially true for the negative end of the continuum, which also is in line with the reaction time data described above. The results at the group level are not as interesting as expected. For one, all three groups patterned together for most of the steps along the continua for both coronals and labials, the one exception being step -3 for the coronals. Even this significant result in the absence of any others could be a chance effect, and so it is difficult to attribute much to this difference. The most salient conclusion we can draw at the group level is that groups all patterned together and that there were differences between the steps. Taking the three measurements, proportion sameresponse, confidence ratings and reaction time, we can say that tokens closest to step 0 were perceived as easier than tokens further away. Also tokens closest to the ends of the continuum were also perceived as easier than tokens between zero and the ends. Given these results, we shift our focus to individual subject differences since the three groups behaved very similarly overall in their perception of synthesized speech.

7.3.3. Individual subject differences: Coronals

The results in the previous sections showed statistically significant differences between groups for proportion sameresponse for only 1 step (coronals -3). We now turn to investigating possible patterns amongst individual subjects. The upcoming sections also compare the results for the synthesized speech at the subject level to the previously presented categorization perception data. We do not present a separate section for comparisons at the group level due to the very similar perception patterns between groups for the synthesized data.

Any patterns or trends observed at the subject level for the synthesized speech perception data are also compared to the production experiment data in subsequent chapters. First we examine the coronal data.

Figure 7.7 shows the naïve group's proportion same response value for each step in the coronal continuum for each subject in the group. Immediately apparent is an increase in proportion same response towards the middle of the continuum for most subjects and a drop off towards the extremes, as was evident also in the overall group data. More subjects have greater proportion same responses towards the positive end of the continuum than the negative end, indicating that in general more subjects found items towards this end to be less distinct than when paired with negative steps. Again, this observation was seen when aggregating at the group level as well.

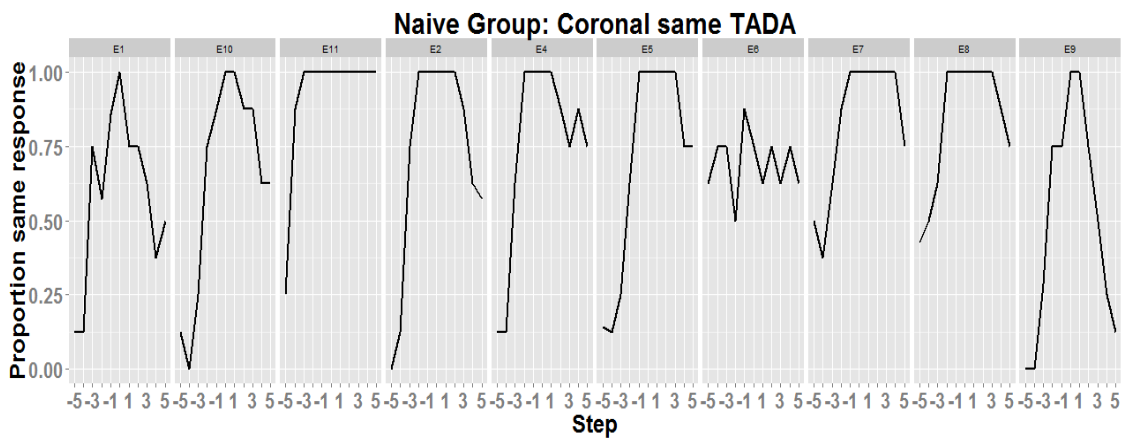

Figure 7.7: Proportion of same response values by subject for Group E for coronals for each step in the TADA continuum.

We can observe the following general trends from the plot in Figure 7.7. There are no subjects that tend to give mostly same responses for tokens in the negative range and only in the negative range. Providing same responses in the negative range means that the subject perceives the palatalization gesture preceding the consonantal gesture similarly to the palatalization gesture synchronous with the consonantal gesture.

However, some subjects do exhibit more distributed responses, providing relatively high proportion same responses in both negative and positive ranges. Subjects E6, E7, E8 and E11 exhibit this pattern. Another group of subjects exhibits higher rates of same responses in the positive range, namely E4 and E5. Lastly, subjects E1, E2, E9 and E10 exhibited only high proportions of same responses for tokens in the middle of the continuum.

In order to compare the synthesized speech perception to perception of natural speech we compare this data to the categorization data of palatalized consonants. We choose this modality and not the AXB discrimination results because the categorization task involved the subject hearing one consonant token and categorizing it to the best of their ability to one category. In this way we can test the hypothesis that a subject that has more same responses to tokens in the negative ranges for synthesized speech will exhibit a greater tendency to assimilate the palatalized consonant to the JC category in natural speech. The reason for this is due to the structure of the TADA continuum. The negative members of the continuum are composed of pairs of tokens, one token is always the token with simultaneous tongue body (palatalization) and consonantal (tongue tip, or labial) gesture coordination. The second token for the negative members is one where the tongue body gesture is coordinated to precede the consonantal gesture to yield a sequence of palatalization followed by consonant constriction similar to the acoustic sequence [jt] or [jp].

Beginning with coronal data for the naïve group, for the relation between the TADA results and categorization for onset coronals we see that most subjects exhibited relatively consistent assimilations to the CJ category. One subject, E2 had relatively distributed assimilations to both JC and CJ categories (fit index values of 11 and 10 respectively). For coda coronals, a series of subjects assimilated the palatalized consonant to JC categories, E8, E9 and E10. Subject E8 showed a distributed TADA pattern, while E9 and E10 showed high proportion same responses for the middle of the continuum. The subjects which mostly assimilated the palatalized coda coronals to the JC category did not show a preference for providing same responses at the negative end of the continuum, contrary to the prediction. This result further confirms that reached at the group level, namely that the TADA results do not seem to correlate meaningfully with natural speech. A similar situation holds when comparing the TADA data to the assimilations for the coda coronal categories.

Again we divided the subjects into subgroups based on their perception patterns. A group of subjects tended to assimilate the codas to JC categories (E8, E9, 10) but these subjects did not tend to prefer sameresponses in the negative end of the TADA continuum, as would be expected. The hypothesis was that subjects who perceived the coda palatalized consonants as sequences of glide+consonant would more likely also equate the synthesized tokens with palatal gesture preceding the consonantal gesture with those with more synchronous phasing of the two gestures. This did not hold in this case.

Another group of subjects assimilated the palatalized coda coronal mainly to the CJ category of consonant+palatalization (E4, E5, E6). Of this group, two subjects, E4, E5 exhibited more clear preferences for same responses for later members of the continuum. This means that these two subjects more often equated synthesized tokens with palatalization gestures following the consonantal gestures with tokens with synchronous phasing of the two gestures. For these two subjects at least, the categorization of natural speech and the perception of synthesized speech correspond. Overall, however, for coda coronals, different subjects exhibited different patterns in perception of synthesized speech, and these patterns did not always correspond as expected to perception of natural speech. For only two subjects did perception of natural speech and synthesized speech correspond closely.

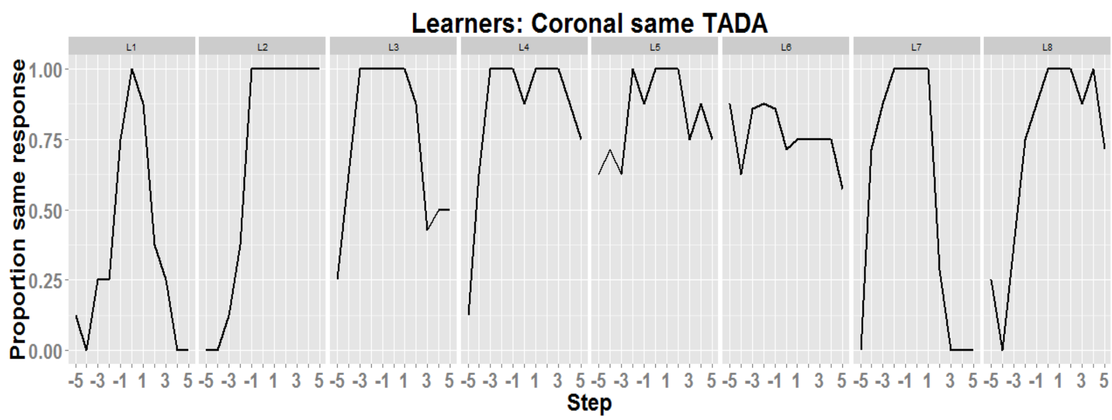

Figure 7.8: Proportion of same response values by subject for the learner group for coronals for each step in the TADA continuum.

We now move on to discuss the relation between TADA and categorization results for coronal onsets for the learner group. For natural speech it was predicted that the learners would exhibit greater assimilation values to the Russian palatalized category. In TADA, the prediction was that learners would exhibit higher preferences for same responses towards the middle region of the continuum. Again, the only difference exhibited at the group level occurred for step -3 on the continuum for the coronals. The difference was only significant between group R and Group L, that is, the learner group had a significantly higher same response value for step -3 for the coronals than the native Russian group. Looking at the data at the subject level TADA data in Figure 7.8 we do see some subjects who tended to have higher same responses in the negative ranges of the continuum, such as L3 and L7. Subjects L4, L5 and L6 had more distributed patterns, with relatively high sameresponse values across the continuum. Subject L1 had high sameresponses in the middle of the continuum and subjects L2 and L8 had responses clustered towards the positive end. We now proceed to see whether, at the subject level, these same subjects with differing synthetic speech perception patterns behave differently in categorization of the natural Russian palatalized coronals.

Subjects in this group also had differing categorization patterns. Subjects L2, L3, L5 and L6 exhibited a mix of assimilations of the coronal onset palatalized consonant to both the palatalized category as well as the CJ category. The remaining subjects showed modal assimilations to the palatalized category. What this means is that for the most part the Learner group assimilated the consonant to the palatalized category, as demonstrated in the group-level analysis in Chapter 8. The subset of subjects who also assimilated the consonant to the CJ category (L2, L3, L5, L6) did not all also tend to have sameresponse values in the positive range. Of this subset, one subject L3 exhibited a clearer preference for sameresponse values in the negative range while two had distributed patterns in both positive and negative ranges (L5, L6) and one (L2) exhibited mostly sameresponse values for positive members. Thus, as for the naïve group, there is not a clear pattern at the subject level as to a tendency for subjects to perceive the natural onset palatalized speech the same way as synthetic speech.

To further elucidate any subject-level tendencies we consider the perception patterns of the natural coda palatalized consonants compared to the synthesized speech tokens. First, no subjects in this group exhibited modal assimilations of the coronal coda palatalized consonant solely to the JC category. Subjects L3, L6 and L7 assimilated the palatalized coronal coda to both the plain category as well as the palatalized category. Two of these subjects, L3 and L7 had preferences for sameresponses at the negative end while the third L6 exhibited a preference for sameresponse values across the continuum. Subjects L1, L2 and L4 exhibited modal assimilations to the palatalized category and subject L5 to the plain category. For the coda consonants, the link between the TADA perception and the categorization of natural speech does not seem stronger than that for onsets for this group. The three subjects who assimilated the coda coronal to the palatalized category at least some of the time also had relatively high sameresponse values at the negative end of the continuum.

The three subjects who exhibited modal assimilations to the palatalization category for natural speech each exhibited different perception patterns for the synthesized speech. The subject who exhibited modal assimilations to the plain category exhibited distributed (both negative and positive) sameresponse rates for the synthesized speech. These results indicate that there is little association or pattern between the natural speech and synthesized speech for coda coronals for the learner group.

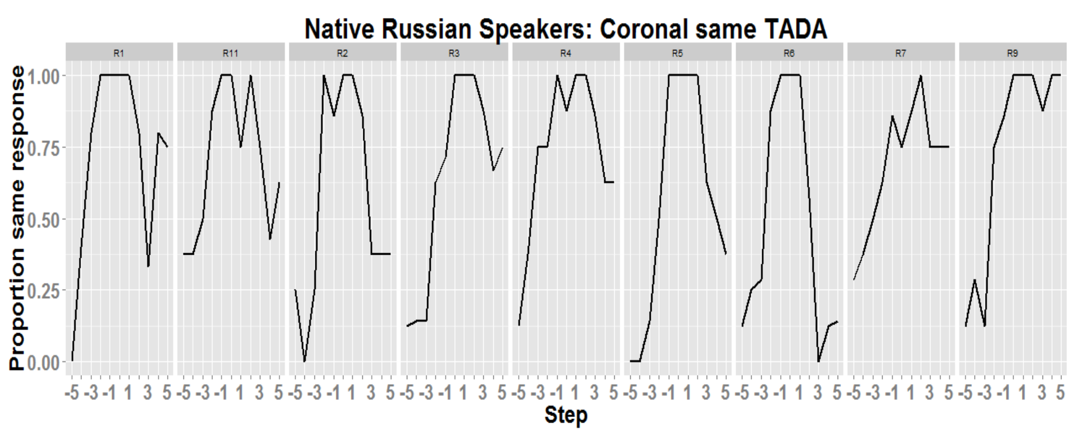

Figure 7.9: Proportion of same response values by subject for Group R for coronals for each step in the TADA continuum.

We will spend some time discussing the individual subject-level results for the native Russian group. There was also a variety of response patterns for this group for the synthesized TADA tokens seen in Figure 7.9. One subject tended to provide sameresponses in the negative range only (R1), while a subset provided sameresponses in both negative and positive ranges for a more distributed pattern (R4, R2, R11). Other subjects gave sameresponse patterns for the positive only (R3, R7, R9) while still others such as R5 and R6 provided responses centered In the middle of the continuum.

As seen in previous chapters, the majority of the time the native Russian group gave very high fit index values for the palatalized coronal onset and coda consonants to the palatalized category. This was not the case for either English group. As such, given the diversity in perception of the synthesized speech and the homogeneity of the perception in the natural speech for the coronals for the native Russian group, we can further say that there seems to be no strong link between the two modalities.

7.3.4. Individual subject differences: Labials

The results at the group level in previous sections for the labial consonants showed no significant group differences. We now turn to investigating possible patterns amongst individual subjects. Again, we will compare groups and subjects in subsequent chapters.

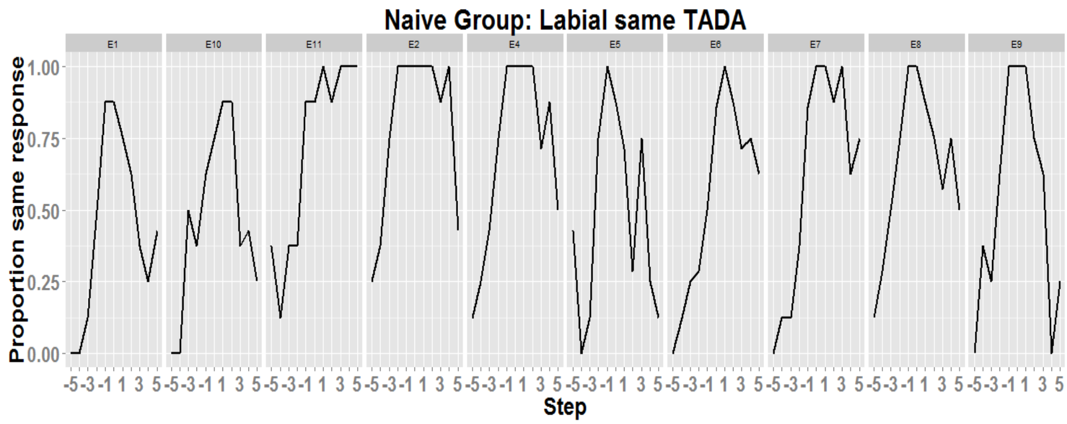

Figure 7.10: Proportion of same response values by subject for Group E for labials for each step in the TADA continuum.

We can observe the following general trends in Figure 7.10. There are no subjects that tend to give mostly same responses for tokens in the negative range, as happened for coronals. Again, this means that no subjects perceive the palatalization gesture preceding the consonantal gesture similarly to the palatalization gesture synchronous with the consonantal gesture.

Two subjects exhibited distributed responses, providing relatively high proportion same responses in both negative and positive ranges (Subjects E2 and E4). This is a different subset than the coronals (Subjects E6, E7, E8 and E11 exhibit this pattern for coronals). A larger group of subjects exhibits higher rates of same responses in the positive range, namely E5, E6, E7, E8 and E11. Lastly, three subjects, E1, E9 and E10 exhibited only high proportions of same responses for tokens in the middle of the continuum.

Now we compare this data to the categorization data. The same hypothesis as for the coronals was tested, in that a subject that has more same responses to tokens in the negative ranges will have more assimilations for the palatalized consonant to the JC category in natural speech and a subject that has more assimilations to the CJ category will have more same responses in the positive end. To summarize the categorization data for onset labials for the naïve group, all subjects exhibited a tendency to assimilate the onset palatalized labial consonant to the CJ category. A small subset consisting of subjects E5, E8 and E11 also showed some slightly lower tendencies to assimilate the consonant to the JC category, while still also exhibiting assimilations to the CJ category. The majority of subjects, therefore assimilated the consonant to the CJ category the majority of the time. Looking at the TADA data for the labials it is difficult to establish clear patterns between behavior of subjects in categorization of natural speech and perception of synthesized speech. Given that the majority of the subjects behaved similarly with categorization it is difficult to say how they would have differed in perception of synthesized speech had they differed more in categorization.

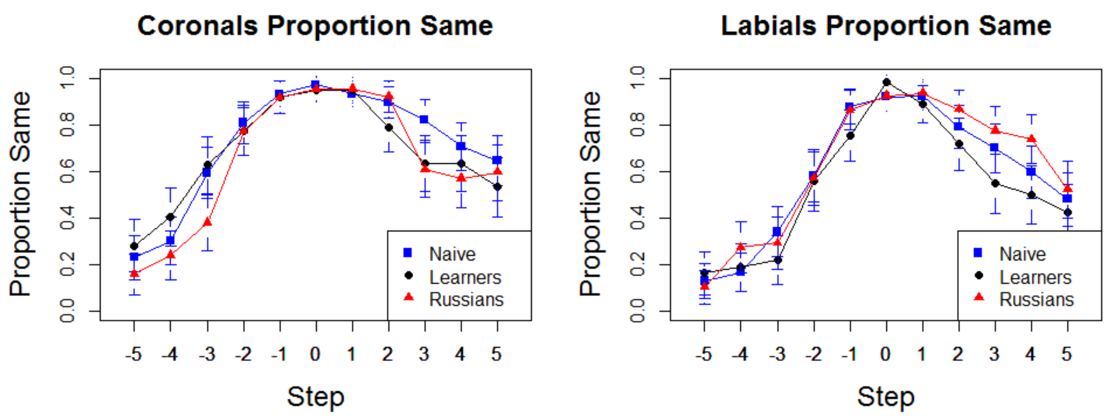

Figure 7.4 in 7.3.1 above showed a trend at the group level between labials and coronals for the synthesized speech perception, namely, that subjects across groups tended to attribute more same responses to coronal tokens in the negative range than labials. From the previous chapter, Table 6.1 shows the fit index values at the group level for Group E. We see a slight tendency as well for the labial palatalized consonant to be associated with a higher fit index assimilation to the CJ category (23.4) than the coronal palatalized onset to the CJ category (20.5), although both of these assimilations were the preferred assimilations for each. These observations at the group level suggest a similar trend between the synthesized speech perception and the categorization data, that labials were perceived better as sequences of consonant+glide than coronals. This observation should not be considered a firm conclusion, but just an observation present in both perception modalities.

For codas, we see the following comparisons between categorization and TADA for labials for subjects in the naïve group. Similarly to the coda coronals, there seem to be no real patterns at the subject level in comparing TADA perception behavior to categorization of natural speech. A group of subjects tended to assimilate the codas to JC categories (E8, E9, 10) but these subjects did not tend to prefer sameresponses in the negative end of the continuum, as no subjects preferred to give sameresponse values solely in the negative range. The hypothesis that subjects who perceived the coda palatalized consonants as sequences of glide+consonant would more likely also equate the synthesized tokens with palatal gesture preceding the consonantal gesture with those with more synchronous phasing of the two gestures did not hold, much like the coronal codas as well.

Another group of subjects assimilated the palatalized coda labial mainly to the CJ category of consonant+palatalization (E1 and E4). Neither of these two subjects exhibited clear preferences for same responses for later members of the continuum, unlike what would be expected if the synthesized speech perception were to mirror assimilation of natural speech. Subjects E5 and E6 had more distributed assimilation patterns for the natural speech, exhibiting assimilations to the plain, CJ and JC categories to varying degrees. These two subjects tended to give higher samresponse values for members of the TADA continuum in the positive range. Subject E11 also had more distributed assimilation of the palatalized labial coda to only two categories, C and JC, however this subject also showed greater tendency to use sameresponse labels in the positive end of the continuum. For this subject, assimilation of natural speech and perception of synthesized speech also did not correspond. The final subgroup consisted of subjects E2 and E7 which both assimilated the palatalized coda labial to the plain category in natural speech. These two subjects exhibited different strategies in perception of synthesized speech, with subject E2 having a more distributed sameresponse distribution, while subject E7 had more sameresponse values in the positive end of the continuum.

Overall, however, for coda labials, like coda coronals, differences existed between subjects in perception of synthesized speech as well as the relation between perception of natural speech and synthesized speech, and in general we cannot conclude that these differences correspond in patterns of relations between the two modalities, that is, there is no clear link between how a subjects perceived the natural speech and the synthesized speech.

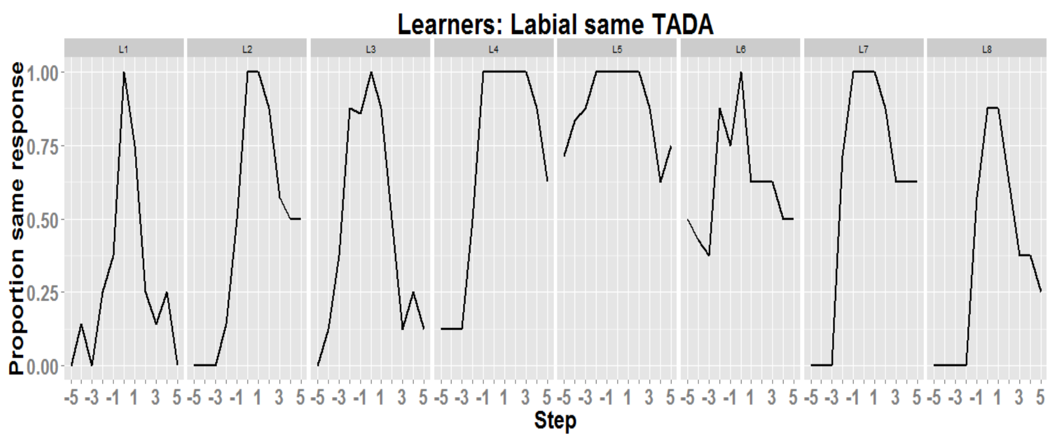

Figure 7.11: Proportion of same response values by subject for the learner group for labials for each step in the TADA continuum.

We move on now to describe the subject-level behavior of learners with respect to the synthesized TADA perception and its relation to natural speech for labial consonants. For the synthesized speech, the following patterns emerge for this group visible in Figure 7.11. A subgroup of subjects (L5, L6) exhibited distributed patterns of sameresponses in both negative and positive regions. Another subgroup (L2, L4, L7) had higher sameresponse values mostly in the positive region. A third and final subgroup (L1, L3, L8) had higher sameresponse values mostly in the middle of the continuum. Comparing this data to the natural speech we begin with the onset labials. Much like for the coronals, subjects in this group differed in assimilation patterns. Subjects L2, L3 and L6 exhibited a mix of assimilations of the coronal onset palatalized consonant to both the palatalized category as well as the CJ category, each of these subjects also exhibited similar patterns to the coronal onsets. Subject L1 showed a relatively low fit index value to the JC category. Subject L5 showed more consistent assimilations to the CJ category and the remaining subjects showed modal assimilations to the palatalized category. What this means is that subjects in the Learner group assimilated the consonant to the palatalized category as well as other categories. This differed from the coronal assimilation patterns which were overwhelmingly to the palatalized category.

The subject (L1) who assimilated the onset labial to the JC category did not all also tend to have sameresponse values in the negative range as no subjects did. This subject preferred the middle of the continuum. Further, subjects who assimilated the consonant to the same categories exhibited different patterns in perception of the synthesized speech. For example, subjects L4, L7 and L8 all assimilated the palatalized consonant to the palatalized category but did not all prefer sameresponse values solely in the middle of the continuum as would have been expected. It was the case as for the naïve group, there also seem to be no clear patterns at the subject level for perception of natural onset palatalized speech the same way as synthetic speech.

Finally we consider the perception patterns of the natural coda palatalized labials compared to the synthesized speech tokens. First, one subject (L1) in this group exhibited modal assimilations of the coronal coda palatalized consonant solely to the JC category, but again, no subjects including L1 had high sameresponse values to the negative end of the TADA continuum. Subjects L3 and L4 assimilated the palatalized coronal coda to both the plain category as well as the palatalization category. These subjects had sameresponse values in the middle or positive ends of the continuum. Further, subject L5 assimilated this consonant to the plain category while L2, L6 and L7 assimilated this consonant to the palatalization category. These subjects again exhibited varying patterns in the TADA perception. As for other syllable positions and place of articulation, no clear subject-level patterns emerge in the relation between natural and synthetic speech perception for this study and group.

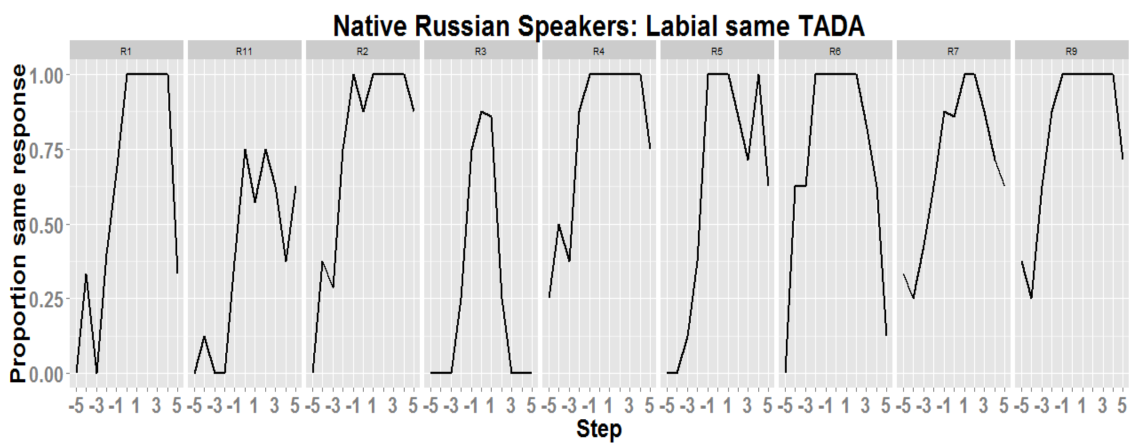

Figure 7.12: Proportion of same response values by subject for Group R for labials for each step in the TADA continuum.

We move on to describe the patterns of perception of labial synthesized and natural speech at the subject level for the native Russian speakers. There were no native Russian speaking subjects that tended to provide predominantly sameresponses in the negative range for labials. A subset of subjects provided responses which were distributed across both negative and positive ends (R6, R7, R9) while another subset provided sameresponse values mostly in the positive end of the continuum (R1, R2, R4, R5, R7, R9). Finally, one subject, R3, had a very pronounced preference for sameresponse values only in the middle of the continuum. The native Russian subjects overwhelmingly assimilated the palatalized onsets and codas to the palatalized category, again showing little relation between natural speech and synthesized speech perception.

7.4. Discussion

In this chapter, we have presented the results from the synthetic speech data experiment. In general, results were expected, but also unexpected in different ways. They were expected in that for all groups, the points closest to step 0 garnered the highest proportion of same response values. This was also found for most subjects when examining data at the subject level. This can be interpreted as follows. A high proportion same response, along with a high confidence rating and low reaction time can be interpreted as subjects being very confident that the two stimuli for that particular pair are the same. Since stimuli near step 0 were either the same (step 0) or very similar (steps -1 and 1), this makes sense. The interpretation here is that step 0 represents to Russian palatalized consonants and listeners generally were able to easily qualify them as being the same.

More interesting results appear as we move further from Step 0. As we move away from step 0 the stimuli paired become more distinct. As stated earlier, it is always true that for each step one member of the stimuli pair is the token with simultaneous coordination of the two gestures. As we move further away from step 0 in either direction, the other member of the stimulus pair will have more sequential timing of the two relevant gestures, thereby making it more distinct from its pair. We saw that groups tended to differ (although not reaching significant differences except in only one point) as steps moved away from zero.

The plots showing individual subject data demonstrated that there is indeed inter-subject variability for proportion same responses for various points along the continuum for both coronals and labials for all groups. We saw, for example, that two learner subjects (L5, L6) both had relatively high proportion same response for both labials and coronals when the simultaneous phased token was paired with tokens phasing the TB gesture to precede the consonantal. We could propose a hypothesis that for such subjects, we could see an increase in the actual categorization of real speech to the category representing palatalization preceding the consonant, if we think that subjects are perceiving the gestures and their relative phasing a certain way based on the data from the synthesized speech. For instance, we could assume that, since both these learner subjects (L5 and L6) perceived synthesized tokens roughly equivalent to /tajta/ and /tat^{j}a/ as being the same a relatively high proportion, we might also expect these same subjects to categorize Russian /t^j/ as /jt/ at least some of the time. However, these links were not upheld. That is, it was not the case that subjects who perceived the simultaneous token as being the same as the negatively phased token also perceived natural speech in the same way.

Since the results for this experiment were not as clear as others, it deserves a mention as to why this is. It is hypothesized that since this was not natural speech, and since we used the acoustic output from the standard TADA application without any additional enhancements, participants might simply have been employing psychoacoustic modes of perception (e.g. see

Babel &Johnson 2007 for further discussion on this topic). As a suggestion for future work the experimenter could first increase the size of the subject pool, and second employ better enhancement of the acoustic output to yield a more natural sounding acoustic output from the TADA continuum.

Chapter 8 Task 4: production

The final experiment in this study is a production task. This task took the form of either an auditory repetition task or a reading task depending on the subject's native language. The repetition task was chosen for English speakers due to naive speakers' inability to read Russian. For most participants, this task was presented in two blocks, one block occurring at the beginning of the session, before any perception tasks, and the second block occurring at the end of the session, after all perception tasks. This was originally done due to possible time constraints. it was deemed favorable to run at least part of the production task at the very beginning, in case participants were slower in completing the tasks and would decide not to continue after the perception tasks were completed. Only one subject actually declined to participate in the final production block due to time constraints. It is immediately evident as well that two production tasks separated by an hour of exposure to Russian could have some effect on the production performance of a particular subject. Since the original motivation for breaking up the production task was for time constraints, any effects of learning that would appear in production are not systematically tested in this dissertation.

8.1. Methods

As mentioned above, for the English speaking subjects this was a repetition task. The exact same tokens as used in the Categorization task were used in the repetition task. For convenience, the description of the tokens and the way they were used is reproduced here. An exemplar of each of the twelve sequences was presented a total of seven times. Three of the physical tokens were presented twice, while the fourth was presented only once ((3 x 2) +(1 x 1))

= 7). This strategy was chosen to use as diverse as possible a number of individual recorded tokens while keeping the experiment at a manageable length. As in the AXB and Categorization tests, the tokens were presented in random order in blocks whereby each individual token appeared once before being repeated with the consideration of no two tokens appearing in doublets. Unlike the Categorization task, the experiment itself was presented in two separate blocks as described above. The first block consisted of 1 repetition of each of 4 physical tokens for each consonant, for a total of 48 trials (4 tokens x 12 consonants x 1 repetition). The final block contained one repetition of each of 3 physically distinct tokens per consonant for 36 trials (3 tokens x 12 consonant x 1 repetition). It is helpful to remind the reader that this in total, then, for both blocks, for 12 sequences there were 7 repetitions for a total of 84 trials per participant, as in the Categorization task.

The inter-trial interval was 1 second. For each trial, subjects heard one of the Russian tokens followed by a 1-second pause and a beep. Subjects repeated the Russian sequence they heard only after the beep in order to reduce involvement of memory to psychoacoustic/phonetic properties of the speech sounds themselves. After repeating the token into a microphone, the subjects were instructed to click a button labeled “next” and a new trial was presented. This task aimed to test speakers’ accuracy in production of each of the Russian sequences.

8.2. Analysis

Participants’ productions of the Russian contrasts were tested with the primary aim to examine how production relates to how speakers perceive the contrasts. In this section, the acoustic analyses employed are described.

8.2.1. Data processing and measurements

Table 8.1 below repeats again the materials that the subjects were producing. It should be reiterated here that given that the consonants of interest needed to be surrounded by vowels, but that syllable position was also a factor to test, word boundaries are present surrounding these consonants. It is possible some of the results presented below could be affected by the word boundaries.

	Labials		Coronals	
CV#CVC onset	**IPA**	**RU**	**IPA**	**RU**
plain	pa **p**ap	па пап	ta **t̠**at	та тат
palatalized	pa **p̠ʲ**ap	па пяп	ta **t̠ʲ**at	та тят
glide	pa **p̠ʲj̠**ap	па пьяп	ta **t̠ʲj̠**at	та тьят
CVC#VC coda				
plain	pa**p** ap	пап ап	ta**t̠** at	тат ат
palatalized	pa**pʲ** ap	папь ап	ta**t̠ʲ** at	тать ат
glide	pa**p̠ʲ j̠**ap	папь яп	ta**t̠ʲ j̠**at	тать ят

Table 8.1. The sequences (same as for perception experiment) that subjects produced in the production task.

Each speech token for each subject was processed and measured in the following ways. For each utterance, 14 measurements were taken: 7 measurements during the first vowel preceding the consonant (V1) and 7 measurements during the second vowel following the consonant (V2). The measurements were taken at equal intervals from the beginning to the end of each vowel. Thus, the 7th measurement for V1 corresponded to the transition between V1 and C. The 1st measurement for V2 corresponded to the transition from C to V2 and the 7th measurement for V2 was the measurement taken at the end of V2. The measurements close to either side of the consonant of interest, namely the final measurement at V1 (V17) and the first measurement at V2 (V21), are particularly important as F2 is a reliable acoustic cue for palatalization of stops (Kochetov 2002). Figure 8.1 demonstrates this by showing the mean F2 values at the 15 measurements for the naïve Group for the three types of sequences.

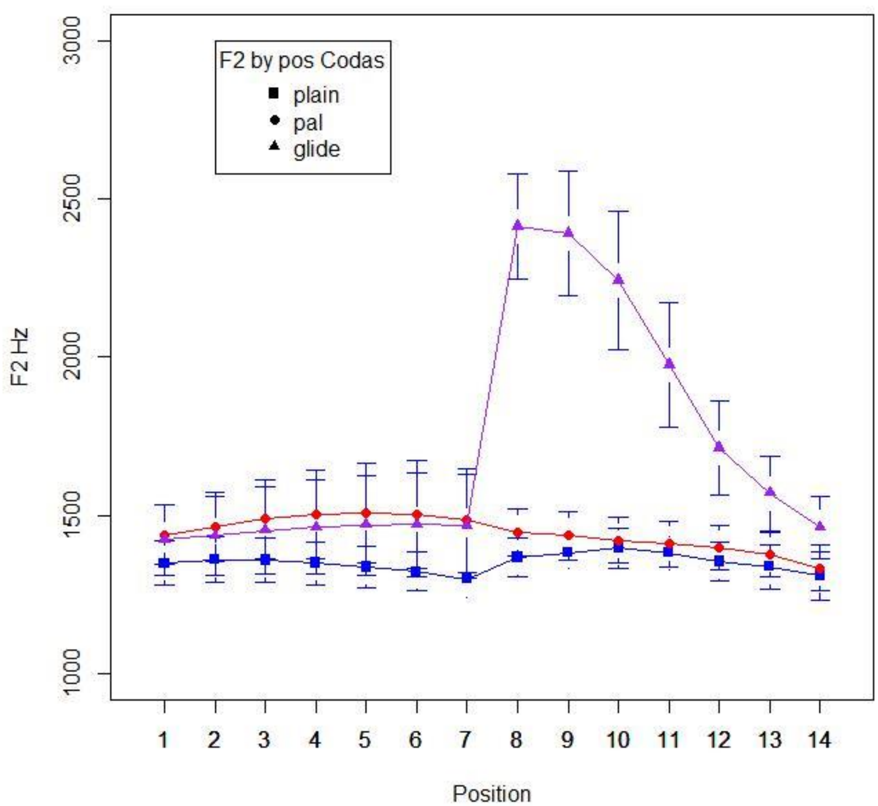

Figure 8.1: Mean F2 values for coda consonants by position for the naïve group.

The remaining measurements at set intervals during both vowels are important to more closely determine the relative timing of palatal and consonantal gestures. For example, higher F2 values at points further away from the transition would indicate TB raising at a greater lag or extent relative to the consonantal gesture than higher F2 values around the transitions and lower F2 values further from the transitions. The measurements also provide information about directionality of TB and C gestural timing. It is possible to determine whether the TB gesture is timed to precede, follow, or coincide with the C gesture. If we see higher F2 values at points preceding the consonant than at points following the consonant for particular tokens, we can interpret these results as exhibiting palatalization or TB raising preceding the consonant.

The analyses explained below are based on the predictions made about English speakers' possible strategies for incorporating a non-native gestural coordination (synchronous TB and TT gestures) such as palatalization in Russian. It is predicted that English speakers could time the TB gestures to precede or follow the C gesture more substantially than that which we see for Russian palatalized consonants. It was also predicted that naïve and less experienced learners could fail to produce palatalization altogether, especially in coda position. These data are also compared to the perception results in further sections.

8.2.2. Analyses for all groups and factors

The statistical analyses of the acoustic data are described as follows. Analyses were performed for two positions in the utterance, as well as 5 metrics derived from other positions. The first set of analyses takes the F2 values for two positions corresponding to the consonant-vowel transitions: VC transition and CV transition. The remaining analyses take differences, or subtractions, between different positions as dependent variables. These differences are: the difference between VC and CV transition; VC transition and preceding vowel point; VC transition and further preceding vowel point; CV transition and following vowel point; and CV transition and further following vowel point.

As in the statistical analyses in previous chapters, a process of evaluating the model fit was performed for each of the analyses. Models with increasingly complex random effects structures were tested against each other. The model with both random intercepts for token and subject had the best model fit. Next, models were evaluated for their fixed effects. For this analysis a fully specified model involving all factors (Group, Syllable, Palatalization and Place) as well as their interactions provided the best model fit.

As was the case for the reaction time data in Chapter 5, the increase in degrees of freedom and subsequent comparisons involved in the fully specified complex model would cause for a very difficult interpretation of a single model. As such, the data was partitioned into two sets, one corresponding to the coronal data and the other corresponding to the labial data. For each of the sets of data (coronals and labials) the results are presented by subsections referring to the analyses of interest, that is, the analyses referring to the factors Palatalization (palatalization contrast type), Group (differences between groups) and syllable position.[17]

To summarize, for each measurement position and each place of articulation (e.g. F2 at the CV transition for labials) a linear mixed effects model with subject and token as random effects and the value (e.g. mean F2 at CV in Herz) as response variable, and Group (E, L, R), Syllable (onset, coda) and Palatalization (plain, palatalized, glide) as independent factors was run. The interactions were also tested.

The above mentioned analyses are meant to test the value of F2 on the factors Place,

---

[17] This also means that we are not strictly comparing the results for the labials and the coronals against each other. The reader should keep this in mind.

Syllable position, Palatalization and their interactions at relevant locations in the sequences. These results would demonstrate how these factors affect the F2 near and further away from the consonants, and how different Groups differ. Below we provide the results for each position and place of articulation.

8.3. Results:

8.3.1. CV and VC transition: Coronals

First, the results from the analyses for the coronal data are presented for the two vowel transitions. We present the results for the VC transition first, and CV transition second.

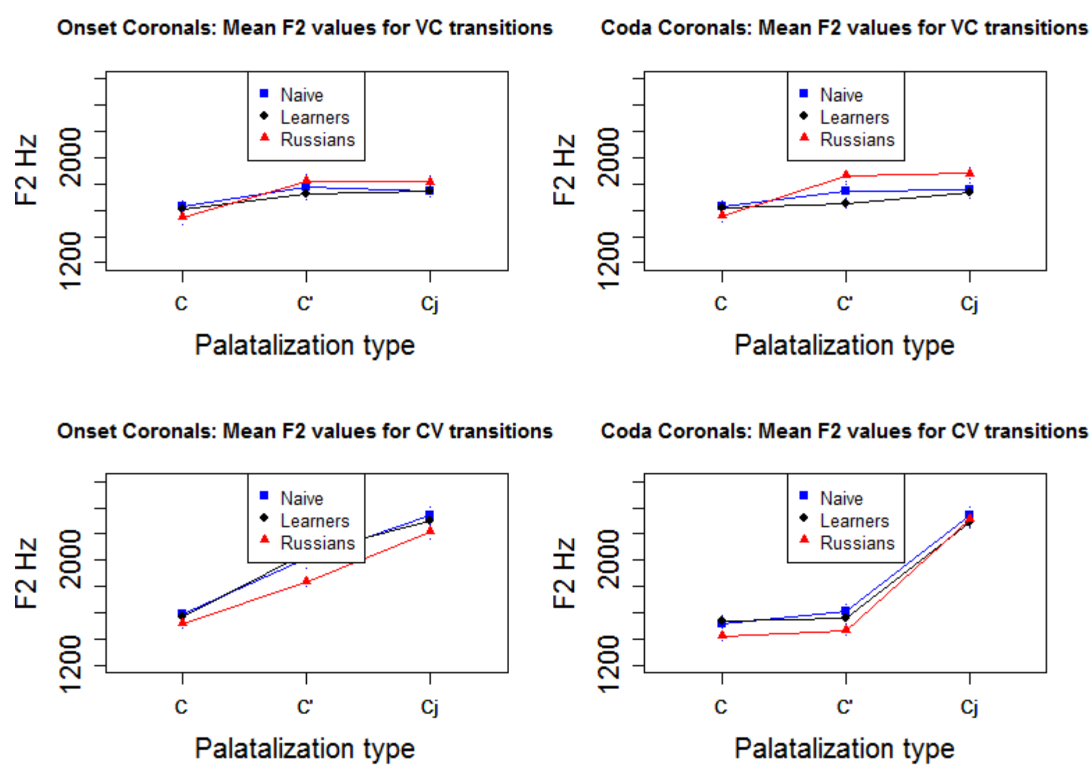

Figure 8.2: Mean F2 values for VC and CV transitions for coronal consonants.

Figure 8.2 shows average F2 values for the three palatalization types (x-axis) for each group (colored lines) for onset and coda coronals at both consonant transitions. As previously stated, the statistical analyses are conducted for each position. Thus, for the coronals, two analyses are presented here, one for the VC transition represented by the two upper plots, and another for the CV transition represented by the two lower plots.

8.3.1.1. VC Transition Coronals

The results are presented with the aid of tables analogous to those used in previous chapters, indicating significant coefficients and p-values.

	Naïve		**Learners**		**Russians**	
Pal	coda	onset	coda	onset	coda	onset
C vs C^j	112 (0)***	141 (0)***		115 (0)***	306 (0)***	280 (0)***
C vs Cj	127 (0)***	113 (0)***	118 (0)***	137 (0)***	324 (0)***	270 (0)***
C^j vs Cj			88 (0)***			

Table 8.2: Fixed effects coefficients for palatalization type comparisons for mean F2 at the VC transition for coronals. Significant coefficients are indicated by asterisks indicating the level of significance. P-values are enclosed in brackets. A positive coefficient means that the palatalization type on the right of the “vs” label in the “Pal” column has a significantly higher mean F2 than the contrast type on the left of the “vs” label for the particular Group and syllable position. For example in row 1, C^j contrasts in coda position for Group E have significantly higher F2 values than C contrasts in VC transitions.

Table 8.2 above shows the significant coefficients of the mixed models for each of the comparisons of the palatalization types for each group and syllable position for the VC transition for coronal consonants. Both groups E and R have significantly higher F2 values for both palatalized and consonant+glide consonants than for plain consonants. This is evident by the significantly positive coefficients for both of these groups in the first two rows, which compare the plain to the palatalized, and the glide consonant, respectively. The learner group has the same result only for onset consonants. For codas, the learner group has significantly higher F2 values for the consonant+glide sequence than for both plain and palatalized consonants.

		Group		
Pal	Syllable differences	Naïve	Learners	Russians
C	Coda vs Onset			
C^j	Coda vs Onset		72 (0.0002)***	-39 (0.027)*
Cj	Coda vs Onset			-65 (0.0002)***

Table 8.3: Fixed effects coefficients for Syllable comparisons for mean F2 at the VC transitions for coronals. Asterisks indicate the level of significance. P-values are enclosed in brackets. For each palatalization type in the Pal column, a significant positive coefficient means that the syllable position "Onset" has a significant increase in mean F2 versus the syllable position Coda, for each Group.

Table 8.3 above provides the significant coefficients of the mixed models for the comparisons between syllable positions for each group and palatalization type for the VC transition for coronal consonants. The learner group has significantly higher F2 values for the palatalized consonants in onset position over coda position. The native Russian group has significantly lower F2 values for palatalized and consonant+glide consonants in onset position over coda position. In this way, the native Russian group differs from the two English speaking groups.

		Syllable position	
Pal	Group	coda	onset
C	Naïve - Learners		
	Naïve - Russians		
	Learners - Russians		
C^j	Naïve - Learners		
	Naïve - Russians		
	Learners - Russians	213 (0.006)**	
Cj	Naïve - Learners		
	Naïve - Russians		
	Learners - Russians		

Table 8.4: Fixed effects coefficients for Group comparisons for mean F2 at VC transitions for coronals. Significant coefficients are indicated by asterisks indicating the level of significance. P-values are enclosed in brackets. For each palatalization type in the Pal column, a positive coefficient means that the group on the right of the “vs” label in the “Group” column has a significantly higher mean F2 value for that contrast versus the group on the left of the “vs” label for the particular syllable position. For example, for C^j in coda position the native Russian group had significantly higher F2 values than the learner group.

Now we present comparisons between groups for each consonant and syllable position for the VC transition for coronal consonants. In Table 8.4 above we see that the only significant difference between groups occurs between the learner group and the native Russian group for palatalized consonants in coda position. The native Russian group had significantly higher F2 values at the VC transition for coronals in this position than the learner group. The same trend occurs between the native Russian group and the naïve group for this consonant; and between the native Russian group and either English group for the consonant+glide sequences; however these differences did not reach significance.

8.3.1.2. CV Transition Coronals

	Naïve		Learners		Russians	
Pal	coda	onset	coda	onset	coda	onset
C vs C^j	99 (0)***	424 (0)***		488 (0)***		318 (0)***
C vs Cj	823 (0)***	745 (0)***	754 (0)***	733 (0)***	893 (0)***	698 (0)***
C^j vs Cj	724 (0)***	320 (0)***	726 (0)***	254 (0)***	849 (0)***	380 (0)***

Table 8.5: Fixed effects coefficients for palatalization type comparisons for mean F2 at the CV transition for coronals. Significant coefficients are indicated by asterisks indicating the level of significance. P-values are enclosed in brackets. A positive coefficient means that the palatalization type on the right of the "vs" label in the "Pal" column has a significantly higher mean F2 than the contrast type on the left of the "vs" label for the particular Group and syllable position. For example in row 1, C^j contrasts in coda position for Group E have significantly higher F2 values than C contrasts in CV transitions.

Table 8.5 above shows the significant coefficients of the mixed models for each of the comparisons of the palatalization types for each group and syllable position for the CV transition for coronal consonants. Group E has significant differences in mean F2 values between each of the three palatalization types in both syllable positions. Palatalized consonants have higher F2 values than plain consonants (row 1), and consonant+glide consonants have higher F2 values than plain, and palatalized consonants (rows 2 and 3). The learners and the native Russians have similar results, except that there is no significant difference between plain and palatalized consonants’ F2 values at CV transition for coda consonants.

Pal	Syllable differences	Group Naïve	Learners	Russians
C	Coda vs Onset	74 (0.002)**		97 (0.0001)***
C^J	Coda vs Onset	400 (0)***	494 (0)***	371 (0)***
Cj	Coda vs Onset			-97 (0.0001)***

Table 8.6: Fixed effects coefficients for Syllable comparisons for mean F2 at the CV transitions for coronals. Asterisks indicate the level of significance. P-values are enclosed in brackets. For each palatalization type in the Pal column, a significant positive coefficient means that the syllable position “Onset” has a significant increase in mean F2 versus the syllable position Coda, for each Group.

Table 8.6 above provides the significant coefficients of the mixed models for the comparisons between syllable positions for each group and palatalization type for the CV transition for coronal consonants. All three groups have significantly higher F2 values for the palatalized consonants in onset position over coda position. The naïve group and the native Russian group, but not the learners, have significantly higher F2 values for plain consonants in onset position over coda position. Finally, only the native Russian group has significantly lower F2 values for onsets over codas for consonant+glide sequences.

		Syllable position	
Pal	Group	coda	onset
C	Naïve - Learners		
	Naïve - Russians		
	Learners - Russians		
C^j	Naïve - Learners		
	Naïve - Russians	-142 (0.037)*	-170 (0.01)*
	Learners - Russians		-220 (0.002)**
Cj	Naïve - Learners		
	Naïve - Russians		
	Learners - Russians		

Table 8.7: Fixed effects coefficients for Group comparisons for mean F2 at CV transitions for coronals. Significant coefficients are indicated by asterisks indicating the level of significance. P-values are enclosed in brackets. For each palatalization type in the Pal column, a negative coefficient means that the group on the right of the "vs" label in the "Group" column has a significantly lower mean F2 value for that contrast versus the group on the left of the "vs" label for the particular syllable position. For example, for C^j in coda position the native Russian groupo had significantly lower F2 values than the naïve group.

We move on to comparisons between groups for each consonant and syllable position for the CV transition for coronal consonants. Table 8.7 above shows that the native Russian group has significantly lower F2 values at CV transition for palatalized coronals for both onsets and codas than the naïve group. the native Russian group also has lower F2 values at onset position for consonant+glide sequences than the learners. No other significant differences emerged between groups.

8.3.2. Results: CV and VC transition: Labials

We now present the results for the consonant transitions for labial consonants. Figure 8.3 shows the mean F2 values for these transitions for all groups, syllable positions and consonant types for labial consonants.

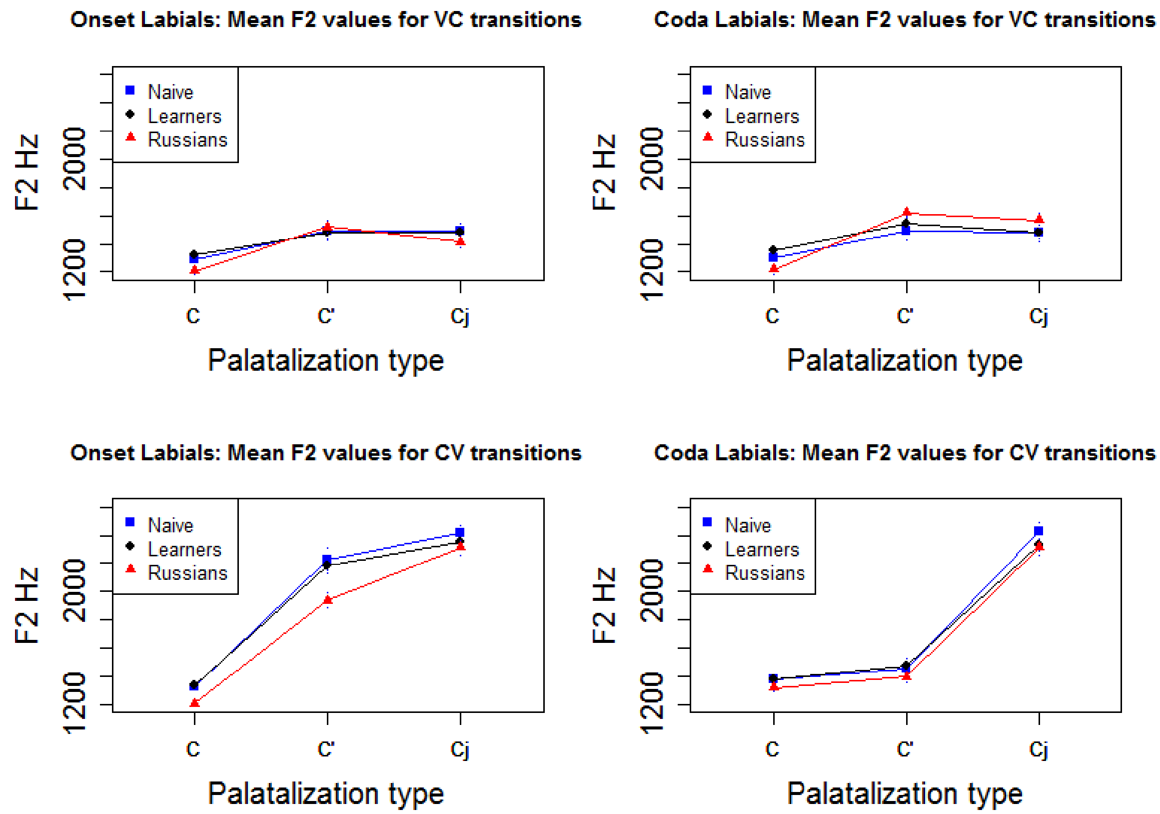

Figure 8.3: Mean F2 values for VC and CV transitions for labial consonants.

The results are presented first for VC transition and then for CV transition for labial consonants.

8.3.2.1. VC Transitions Labials

Pal	Naïve		Learners		Russians	
	coda	onset	coda	onset	coda	onset
C vs C^j	189 (0)***	196 (0)***	195 (0)***	158 (0)***	407 (0)***	310 (0)***
C vs Cj	174 (0)***	194 (0)***	133 (0)***	151 (0)***	349 (0)***	203 (0)***
C^j vs Cj			-61 (0.01)*		-58 (0.01)*	-107 (0)***

Table 8.8: Fixed effects coefficients for palatalization type comparisons for mean F2 at the VC transition for labials. Significant coefficients are indicated by asterisks indicating the level of significance. P-values are enclosed in brackets. A positive coefficient means that the palatalization type on the right of the "vs" label in the "Pal" column has a significantly higher mean F2 than the contrast type on the left of the "vs" label for the particular Group and syllable position. For example in row 1, C^j contrasts in coda position for the naïve group have significantly higher F2 values than C contrasts in VC transitions.

Table 8.8 above shows the comparisons between palatalization types for the VC transition for labial consonants. All three groups have significantly higher F2 values for both palatalized and consonant+glide consonants than for plain consonants. This is evident by the significantly positive coefficients for these groups in the first two rows, which compare the plain to the palatalized, and the glide consonant, respectively. The learner group and the native Russian group have significantly lower F2 values for consonant+glide CV transitions at coda position than palatalized consonants. The native Russian group also has this result for onsets as well.

		Group		
Pal	Syllable differences	Naïve	Learners	Russians
C	Coda vs Onset			
C^j	Coda vs Onset		-61 (0.01)*	-102 (0)***
Cj	Coda vs Onset			-151 (0)***

Table 8.9: Fixed effects coefficients for Syllable comparisons for mean F2 at the VC transitions for coronals. Asterisks indicate the level of significance in shaded cells. P-values are enclosed in brackets. For each palatalization type in the Pal column, a significant positive coefficient means that the syllable position "Onset" has a significant increase in mean F2 versus the syllable position Coda, for each Group.

Now we report on differences between syllable positions for VC transitions for labials. Table 8.9 shows that the learner group and the native Russian group have significantly lower F2 values for the palatalized consonants in onset position over coda position. The native Russian group also has this effect for consonant+glide sequences. No other significant differences were observed for labials here.

		Syllable position	
Pal	Group	coda	onset
C	Naïve - Learners		
	Naïve - Russians		
	Learners - Russians	-135 (0.03)*	
C^J	Naïve - Learners		
	Naïve - Russians	136 (0.02)*	
	Learners - Russians		
Cj	Naïve - Learners		
	Naïve - Russians		
	Learners - Russians		

Table 8.10: Fixed effects coefficients for Group comparisons for mean F2 at VC transitions for labials. Significant coefficients are indicated by asterisks indicating the level of significance. P-values are enclosed in brackets. For each palatalization type in the Pal column, a positive coefficient means that the group on the right of the "vs" label in the "Group" column has a significantly higher mean F2 value for that contrast versus the group on the left of the "vs" label for the particular syllable position. For example, for C^J in coda position the native Russian group had significantly higher F2 values than the naïve group.

We move on to comparisons between groups for each consonant type and syllable position for the VC transition for labial consonants. Table 8.10 above shows that the native Russian group has significantly lower F2 values at VC transition for plain coda labials than the learner group. The native Russian group has higher F2 values at coda position for palatalized consonants than the naïve group. No other significant differences emerged between groups.

8.3.2.2. CV Transitions Labials

Pal	Naïve		Learners		Russians	
	coda	onset	coda	onset	coda	onset
C vs C^j	80 (0.0014)**	906 (0)***	95 (0.0004)***	949 (0)***	76 (0.0028)**	729 (0)***
C vs Cj	1051 (0)***	1087 (0)***	948 (0)***	1012 (0)***	993 (0)***	1103 (0)***
C^j vs Cj	971 (0)***	180 (0)***	852 (0)***	163 (0)***	916 (0)***	373 (0)***

Table 8.11: Fixed effects coefficients for palatalization type comparisons for mean F2 at the CV transition for labials. Significant coefficients are indicated by asterisks indicating the level of significance. P-values are enclosed in brackets. A positive coefficient means that the palatalization type on the right of the "vs" label in the "Pal" column has a significantly higher mean F2 than the contrast type on the left of the "vs" label for the particular Group and syllable position. For example in row 1, C^j contrasts in coda position for the naïve group have significantly higher F2 values than C contrasts in CV transitions.

Table 8.11 above shows the comparisons between the palatalization types for each group and syllable position for the CV transition for labial consonants. All three groups have significant differences in mean F2 values between each of the three palatalization types in both syllable positions. This is immediately obvious since all cells are filled with significant coefficients. Palatalized consonants have higher F2 values than plain consonants (row 1), and consonant+glide consonants have higher F2 values than both plain, and palatalized consonants (rows 2 and 3).

		Group		
Pal	Syllable differences	Naïve	Learners	Russians
C	Coda vs Onset	-53 (0.03)*		-113 (0)***
C^j	Coda vs Onset	772 (0)***	712 (0)***	539 (0)***
Cj	Coda vs Onset			

Table 8.12: Fixed effects coefficients for Syllable comparisons for mean F2 at the CV transitions for labials. Asterisks indicate the level of significance. P-values are enclosed in brackets. For each palatalization type in the Pal column, a significant positive coefficient means that the syllable position “Onset” has a significant increase in mean F2 versus the syllable position Coda, for each Group.

We proceed to the differences between syllable positions for CV transitions for labials. Table 8.12 shows that the naïve group and the native Russian group have significantly lower F2 values for the plain consonants in onset position over coda position. All three groups have significantly higher F2 values for onsets over codas for palatalized consonants in CV position. No other significant differences were observed for labials here.

		Syllable position	
Pal	Group	coda	onset
C	Naïve - Learners		
	Naïve - Russians		
	Learners - Russians		-130 (0.04)*
C^J	Naïve - Learners		
	Naïve - Russians		-282 (0)***
	Learners - Russians		-250 (0.0001)***
Cj	Naïve - Learners		
	Naïve - Russians		
	Learners - Russians		

Table 8.13: Fixed effects coefficients for Group comparisons for mean F2 at CV transitions for labials. Significant coefficients are indicated by asterisks indicating the level of significance in shaded cells. P-values are enclosed in brackets. For each palatalization type in the Pal column, a negative coefficient means that the group on the right of the "vs" label in the "Group" column has a significantly lower mean F2 value for that contrast versus the group on the left of the "vs" label for the particular syllable position.

The last analysis is the comparisons between groups for each consonant and syllable position for the CV transition for labial consonants. Table 8.13 above shows that the native Russian group has significantly lower F2 values at CV transition for plain labials for onsets than the learner group. The native Russian group also has lower F2 values at onset position for palatalized consonants than either English group. No other significant differences emerged between groups.

8.3.3. Discussion Transitions

Separate analyses were performed for labials and coronals. The analyses presented above described, for each place of articulation, significant differences between the various factors of interest (palatalization type, syllable position, group). Here we present discussion of these results for the two consonant-vowel transitions.

8.3.3.1. Discussion: Coronals

We begin by discussing some main observations for VC transitions for coronals. We begin by discussing differences in syllable positions for the palatalized consonants. Naïve English speakers and Learners had higher F2 values at the VC transition for onsets over codas for palatalized coronals, but the Russian speakers had the opposite effect, signaling greater palatalization before the consonant for codas for Russian speakers, while both English groups exhibited greater palatalization at the VC transition for onsets. This effect could be interpreted as generally less palatalization for codas for the English groups.

Both English groups and Russian speakers had significantly higher F2 values for palatalized consonants at VC transitions for the onset palatalized consonants over onset plain consonants (the Learner group did not see such an effect for codas) but the coefficients (the extent of the increase), were much larger for the native Russians, as shown on Figure 8.2 in the top panel where both the Russian group and the naïve group increases between palatalized and plain consonants for the VC transition for both syllable positions.

One final pattern emerged for the coda coronal consonants at VC transition. This is evident in both Table 8.2 and Table 8.4. The Russian speakers had higher mean F2 values for palatalized coda consonants than the learner group, signaling an interaction. This result means that for plain consonants, and C+j sequences, Russian speakers and English groups did not differ in the F2 values at the VC transition for coda consonants, but they did differ in F2 values at this position for the Learner Group for palatalized consonants. This result is also evident by examining the differences between palatalization types for each group, where we saw that the learner group was the only group with no significant differences between plain and palatalized consonants in coda position. It seems, then, that Learners behaved less like native speakers at the VC transition than the naïve group, which exhibited altogether higher F2 values at this position. One possible explanation of this could have to do with naïve English speakers having a greater tendency to produce instances of palatalized coda consonants as sequences of glide+consonant which could help explain the group-level difference here.

For the CV transition, in terms of differences between syllable positions, all three groups had higher mean F2 values at the CV transition for onsets than for codas for the palatalized consonants. This is potentially due to onset consonants having much more pronounced CV transitions than coda consonants for palatalized consonants, thus explaining the higher F2 values at this position for all groups.

For onset coronals, all three groups also had progressively higher F2 values at the CV transition for all three consonant types. Consonant+glide sequences had the highest F2 values; palatalized consonants had lower F2 values, while plain consonants had the lowest values. This means that each group had higher F2 transitions from consonant to vowel for each type of palatalization, whether palatalized or glide. This is evident from the upward trend for the lower left panel on Figure 8.2. This means the three groups behaved similarly in the differences between consonant types, that is, all three groups had significantly different F2 values between the different consonant types.

Some context is added when we look specifically at group differences. For both onset and coda palatalized consonants we see that the naïve group had higher F2 values at the CV transition for than the Russian group. Since the naïve English group had higher F2 values for palatalized consonants at this transition than the Russian group this suggests this group exhibits more pronounced, or a higher TB gesture at this transition than the Russian group. For the Learners, this effect only happened for the onsets. The observation that both English groups had higher F2 values at the CV transition for palatalized consonants in onset position than the Russian group is interesting because it suggests a less pronounced difference between palatalized consonants, and consonant+glide sequences for the two English groups, while the Russian group seems to differentiate these two consonants more. This is in line with the predictions that English speakers would have less difference between palatalized consonants and consonant+glide sequences in production than Russian speakers. Although, all three groups had significant differences in mean F2 at the CV transition between these consonants, the differences were greater for the Russian group than for the English groups.

8.3.3.2. Discussion Labials

We begin by discussing differences between palatalization types. Both English groups as well as the Russian group had significantly higher F2 values for palatalized consonants at the VC transitions than plain consonants, but the coefficients (the extent of the increase), were much larger for the native Russians. This effect can be seen on Figure 8.3 comparing plain and palatalized consonants. We see the larger difference between these two consonants for the Russian group than the other two English groups. These effects go in line with the predictions that codas would be more problematic for English speakers.

In terms of differences between syllable positions for palatalized consonants, the native Russian speakers had lower F2 at this position for onsets over codas. The learner group also exhibited this effect. This, along with the effect described earlier (larger difference between palatalized and plain consonants for the Russian group) lend support to the idea that the Russian group differed in the implementation of palatalization. These two effects could suggest that the palatalization gesture begins earlier than for the English groups, at least for coda consonants. Further analyses will be undertaken to further shed light on this question.

Taking differences between groups explicitly, we saw significant differences between groups for plain consonants and palatalized consonants. For plain consonants, the Russian group had significantly higher F2 values than Learners for coda consonants at VC transition. The native Russian group also had significantly higher F2 values than naïve English speakers for coda palatalized consonants. This effect is related to the effects described above, where Russian speakers exhibited greater differences between plain and palatalized consonants at VC transition. This is evident also in the group differences.

One final interesting pattern that emerged with the VC transition data for labial consonants is that Russian speakers had lower F2 values for C+j sequences than for palatalized consonants for both onsets and codas. The Learner group had this effect only for codas. The Naïve group did not exhibit this effect, there were no differences between C+j sequences and palatalized sequences for this group. This effect suggests further support for a more lagged gesture for palatalized consonants for the naïve English group. This is an important result, highlighting differences between the two English groups as well as these two groups from the Russian group, for coda consonants (see upper right quadrant of Figure 8.3).

Moving on to the CV transition for labial consonants, we first discuss the differences between palatalization types. For all three groups and both syllable positions, significant differences in FV at CV transition existed between all three palatalization types. Thus, palatalized consonants had greater values than plain consonants, while C+j sequences had greater values than both plain and palatalized consonants. as was the case for coronals as well. This was mostly true for coronal consonants as well, except for the lack of a difference between plain and palatalized consonants for the learner group for codas (see previous section). This is significant because, it shows that all three groups, even the naïve English group, were able to differentiate the three consonant types at the CV transition. This is contrary to what happened at the VC transition, for both labials and coronals, where fewer differences between palatalized consonants and C+j sequences emerged. One interesting caveat to this result, however, is that, as for the coronal consonants, the differences between palatalized consonants and C+j sequences were smaller for the two English groups than they were for the Russian speaking group. This is seen in the difference between the Russian group and the two English groups in the lower panel of Figure 8.3.

This effect described above is also observed in the explicit group differences. The Russian group had significantly lower F2 values for palatalized consonants at the CV transition for onsets than either of the two English groups. There was no difference between groups for C+j sequences, indicating a difference in how both English groups produced palatalized onset labials from how native Russian speakers do (see lower left quadrant of plot). This is similar to what we saw with the coronals, where both English groups had higher F2 values at this position for palatalized onsets than the Russian group, suggesting a higher TB gesture at the CV transition for onset palatalized consonants for the two English groups than for the Russian group for both places of articulation.

8.4. Analyses: Differences between locations

A further set of tests was carried out to more specifically examine any differences between groups and the factors of syllable position and palatalization type with respect to the differences between F2 values at various points in the sequence (for example, between VC and CV transitions). The dependent variables for these tests are mean differences in F2 between various positions. These tests were done to more thoroughly determine differences in timing of any TB palatalization gestures for different groups and consonants. For example, if we wish to test whether there is a significant difference in the timing of a TB gesture between language groups or between different consonants then we could compare the mean difference in F2 between points 8 and 9 (CV transition and next point into the vowel). If this value is high for a particular group and/or consonant, then it means F2 has decreased between these two points, indicating a potentially sharp or steep F2 offset, like what is expected for Russian palatalization. If this value is low, it would firstly indicate a small (or no) change in F2 between these two points. However, the exact interpretation of this scenario would vary depending on the absolute F2 values at these points. If the particular participant did not produce any palatalization for a particular consonant (and concomitant F2 raising), then this value will be low because F2 did not raise at all throughout this sequence, or part of sequence. However, if the subject produced palatalization, and has timed the TB gesture to lag substantially with respect to the C gesture (as might be expected for English speakers), then a low (8 minus 9) value would indicate an F2 plateau, or steady state offset. Thus, with these values we can determine the relative timing of palatal gestures between groups, consonants and their interactions.

The first of such tests to be done was to test whether the groups exhibited significant differences between VC and CV transitions with respect to palatalization, place of articulation, and syllable position. This and all such tests were done using a mixed effects model with subject and token as random effects, mean F2 differences between VC and CV transitions (and for other tests, other points in the sequences) as a dependent variable, and the factors Group, Syllable (onset, coda), Place (labial, coronal), Palatalization (plain, palatalized, glide) and their interaction as fixed effects.

8.5. Results VC – CV

In this subsection we present the results for the mean difference in F2 between the VC transition and the CV transition (henceforth VC-CV). This metric is important because it provides more information on the timing of palatalization, and especially whether and to what extent palatalization is present before the consonant and after the consonant. For this metric, a higher value indicates a higher F2 before the consonant (at VC transition) than after the consonant (at CV transition). In essence, the higher the value, the more palatalization is present at VC transition as compared to CV transition. As above, we present the coronal results first, followed by the labial results.

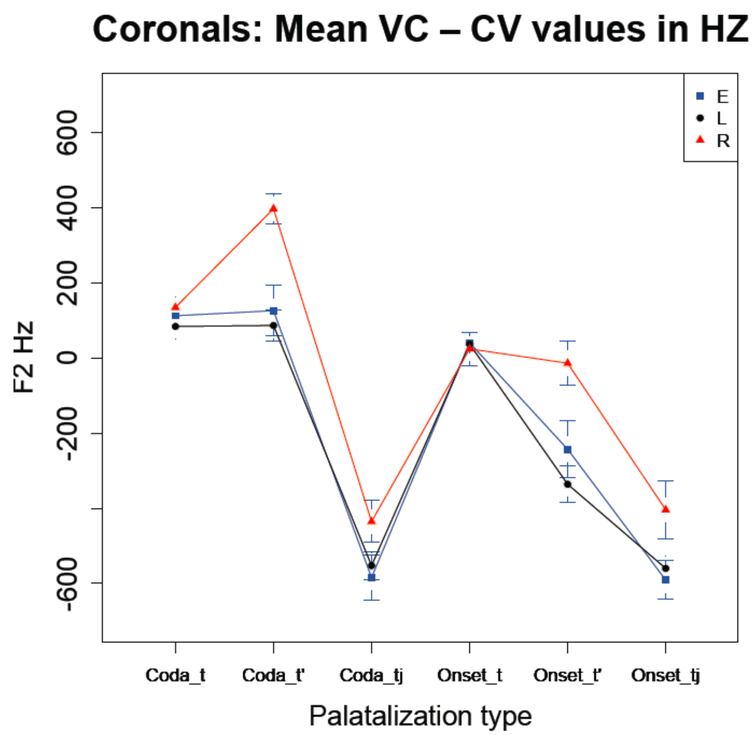

Figure 8.4: Mean differences in F2 between CV and VC transitions for coronal consonants.

8.5.1. Results: Coronals VC – CV

	Naïve		Learners		Russians	
Pal	coda	onset	coda	onset	coda	onset
C vs C^j		-283 (0)***		-372 (0)***	262 (0)***	
C vs Cj	-696 (0)***	-631 (0)***	-636 (0)***	-596 (0)***	-569 (0)***	-428 (0)***
C^j vs Cj	-710 (0)***	-347 (0)***	-638 (0)***	-223 (0)***	-832 (0)***	-389 (0)***

Table 8.14: Fixed effects coefficients for palatalization type comparisons for the mean subtraction in F2 between VC and CV transitions. Significant coefficients are indicated by asterisks indicating the level of significance. P-values are enclosed in brackets. A negative coefficient means that the palatalization type

on the right of the “vs” label in the “Pal” column has a significantly lower mean difference between VC and CV than the contrast type on the left of the “vs” label for the particular Group and syllable position. For example in row 1, C^j consonants in onset position for Group E have significantly lower VC-CV differences than C consonants.

Table 8.14 above shows the comparisons between palatalization types for VC-CV. First we highlight the similarities that exist between the three groups. All three groups, have similar values for the comparisons between plain consonant and consonant+glide sequences; and palatalized consonants and consonant+glide sequences. For both onsets and codas, all three groups had significantly lower mean differences in F2 between VC-CV for plain consonants than for palatalized consonants and C+j sequences. This is evident by the significant negative coefficients for these groups in the last two rows. For the comparisons between plain and palatalized consonants (first row), the two English groups have significantly lower VC-CV values for palatalized consonants than for plain consonants only for onsets. The native Russian group did not show this result for onsets. This group had significantly higher VC-CV for palatalized consonants over plain consonants in coda position, whereas the two English groups lacked this result.

		Group		
Pal	Syllable differences	Naïve	Learners	Russians
C	Coda vs Onset	-71 (0.01)*		-110 (0.0003)***
C^J	Coda vs Onset	-368 (0)***	-422 (0)***	-410 (0)***
Cj	Coda vs Onset			

Table 8.15: Fixed effects coefficients for Syllable comparisons for the mean subtraction in F2 between VC and CV transitions for coronals. Asterisks indicate the level of significance. P-values are enclosed in brackets. For each palatalization type in the Pal column, a significant positive coefficient means that the syllable position "Onset" has a significant increase in mean F2 versus the syllable position Coda, for each Group.

Here we report on differences between syllable positions for VC-CV transitions for coronals. Table 8.15 shows that all groups have significantly lower values for onsets over codas for the palatalized consonants. The naïve group and the native Russian group have significantly lower values for the plain consonants in onset position over coda position, whereas the learner group does not see this effect. No significant effects were seen between syllable positions for the C+glide sequences.

Pal	Group	Syllable position	
		coda	onset
C	Naïve - Learners		
	Naïve - Russians		
	Learners - Russians		
C^j	Naïve - Learners		
	Naïve - Russians	273 (0)***	231 (0)***
	Learners - Russians	310 (0)***	322 (0)***
Cj	Naïve - Learners		
	Naïve - Russians	151 (0.006)**	189 (0.0007)***
	Learners - Russians	117 (0.04)*	156 (0.008)**

Table 8.16: Fixed effects coefficients for Group comparisons for the mean subtraction in F2 between VC and CV transitions for coronals. Significant coefficients are indicated by asterisks indicating the level of significance in shaded cells. P-values are enclosed in brackets. For each palatalization type in the Pal column, a positive coefficient means that the group on the right of the "vs" label in the "Group" column has a significantly higher mean difference between the two transitions in F2 for that contrast versus the group on the left of the "vs" label for the particular syllable position.

This section presents differences between groups for VC-CV transitions for coronals. Table 8.16 shows that the only differences between groups occur for palatalized consonants and C+j sequences, the plain consonants exhibited no differences. For the two consonant types exhibiting differences between groups, the differences existed between both English groups and the Russian group, no differences existed between the two English groups. All of the coefficients are positive, meaning that for both the palatalized consonants and C+j sequences, the Russian group exhibited significantly larger values than either of the two English groups in both coda and onset position.

8.5.2. Results: Labials VC – CV

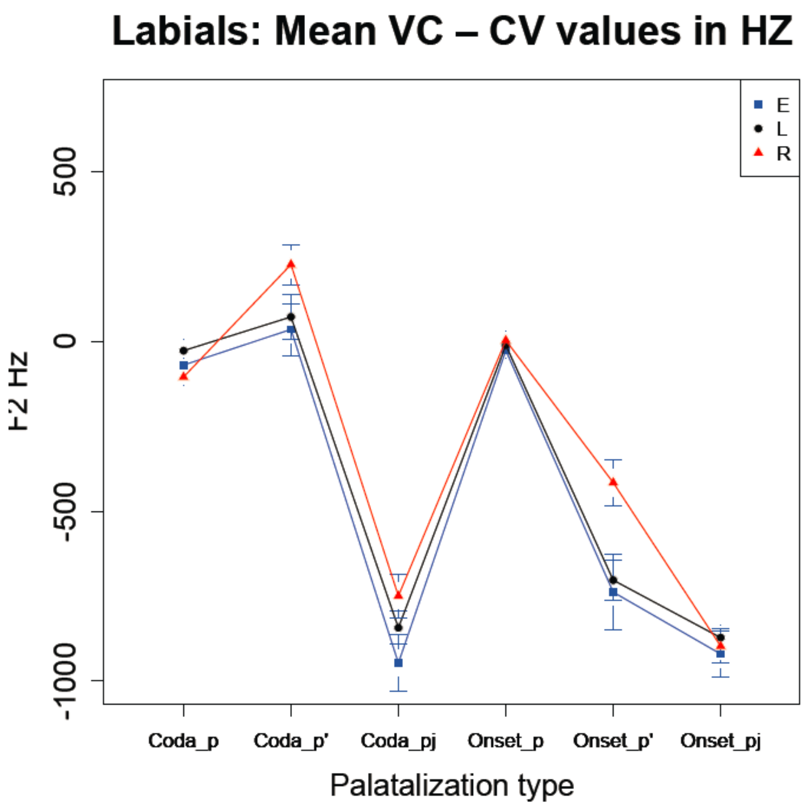

Figure 8.5: Mean differences in F2 between CV and VC transitions for labial consonants.

	Naïve		Learners		Russians	
Pal	coda	onset	coda	onset	coda	onset
C vs Cʲ	109 (0.002)**	-710 (0)***	99 (0.01)*	-691 (0)***	330 (0)***	-418 (0)***
C vs Cj	-876 (0)***	-892 (0)***	-815 (0)***	-860 (0)***	-644 (0)***	-899 (0)***
Cʲ vs Cj	-986 (0)***	-182 (0)***	-914 (0)***	-169 (0)***	-974 (0)***	-481 (0)***

Table 8.17: Fixed effects coefficients for palatalization type comparisons for mean F2 differences between VC and CV transitions for labials. Significant coefficients are indicated by asterisks indicating the level of

significance. P-values are enclosed in brackets. A positive coefficient means that the palatalization type on the right of the “vs” label in the “Pal” column has a significantly higher value than the contrast type on the left of the “vs” label for the particular Group and syllable position. For example in row 1, C^j contrasts in coda position for the naïve group have significantly higher values than C contrasts.

Table 8.17 above shows the comparisons between palatalization types for VC-CV for labial consonants. Again, we first highlight the similarities that exist between the three groups. For this analysis, all three groups have similar values for the comparisons between all three consonant types, as evidenced by the same significant coefficients and signs. As was the case for the coronals, for the comparisons between the C+j sequences and plain consonants; and the C+j sequences and palatalized consonants, for both onsets and codas, all three groups had significantly lower mean differences in F2 between VC-CV for these two comparisons. This is evident by the significantly negative coefficients for these groups in the last two rows. For the comparisons between plain and palatalized consonants (first row), all three groups have significantly lower VC-CV values for palatalized consonants than for plain consonants only for onsets. All three groups had significantly higher VC-CV for palatalized consonants over plain consonants in coda position.

		Group		
Pal	Syllable differences	Naïve	Learners	Russians
C	Coda vs Onset			107 (0.003)**
C^j	Coda vs Onset	-777 (0)***	-774 (0)***	-641 (0)***
Cj	Coda vs Onset			-147 (0)***

Table 8.18: Fixed effects coefficients for Syllable comparisons for the mean subtraction in F2 between VC and CV transitions. Asterisks indicate the level of significance. P-values are enclosed in brackets. For each palatalization type in the Pal column, a significant positive coefficient means that the syllable position "Onset" has a significant increase in value versus the syllable position Coda, for each Group.

Here we report on differences between syllable positions for VC-CV transitions for labials. Table 8.18 shows that all groups have significantly lower values for onsets over codas for the palatalized consonants. The native Russian group has significantly higher values for the plain consonants in onset position over coda position, whereas the two English groups do not see this effect. The native Russian group is also the only group to see significantly lower values for onsets over codas for the C+j sequences.

		Syllable position	
Pal	Group	coda	onset
C	Naïve - Learners		
	Naïve - Russians		
	Learners - Russians		
C^J	Naïve - Learners		
	Naïve - Russians	186 (0.004)**	322 (0)***
	Learners - Russians	154 (0.02)*	286 (0)***
Cj	Naïve - Learners		
	Naïve - Russians	197 (0)***	
	Learners - Russians		

Table 8.19: Fixed effects coefficients for Group comparisons for the mean subtraction in F2 between VC and CV transitions. Significant coefficients are indicated by asterisks indicating the level of significance. P-values are enclosed in brackets. For each palatalization type in the Pal column, a positive coefficient means that the group on the right of the "vs" label in the "Group" column has a significantly higher value for that contrast versus the group on the left of the "vs" label for the particular syllable position.

This section presents differences between groups for VC-CV transitions for labials. Table 8.19 shows that the only differences between groups occur for palatalized consonants and C+j sequences, the plain consonants exhibited no differences. For the two consonant types exhibiting differences between groups, the differences existed between both English groups and the Russian group, no differences existed between the two English groups. All of the significant coefficients are positive. For the palatalized consonants, the Russian group exhibited significantly larger values than either of the two English groups in both coda and onset position, while for the C+j sequences, the only significant difference occurred between the Russian group and the naïve English group for codas, where the Russian group had a significantly higher value.

8.6. Discussion: VC – CV

The VC-CV metric was devised to further determine the relative timing of the TB and consonantal gestures for the three palatalization types for the two groups for each syllable position. An analysis of average F2 values at the two transitions themselves was already presented. The metric constituting the difference between the transitions provides more information as to the timing of the gestures. A lower value indicates similar values at the two transitions. A value below zero indicates a higher CV transition, and possibly more lag for the tongue body gesture. A higher value indicates relatively higher values at VC transitions and less lag for the tongue body gesture. The discussions are presented first for the coronals followed by the labials.

8.6.1. Discussion: Coronals VC – CV

We first discuss the results for the coronal consonants. We begin by discussing the differences between palatalization type. The consonant glide sequences (Cj) all had significantly lower values than either plain or palatalized sequences for both onsets and codas for all groups. This result implies several things. First, as shown on Figure 8.4, the negative value for the C+j sequences means that, in general, these sequences had higher F2 values at CV transition than at VC transition. The same is true for the palatalized consonants for the two English groups in onset position. For the Russian group, a difference emerges, namely that for onsets, the difference between palatalized and plain consonants is not significant, whereas there was a significantly higher value for this group for coda position. This result can be interpreted as follows.

The positive coefficient for the difference between palatalized consonants and plain consonants in coda position for the Russian group means that only the Russian group exhibited significantly higher VC-CV values for palatalized consonants in coda position over plain consonants. The significantly higher VC-CV values for palatalized codas translate to higher F2 values at the VC position than at the CV position. Recall from the CV transition analysis, that Group R had lower F2 values than both groups in onset position, and higher F2 values than the learner group for coda position. In general, this means that the Russian group indeed had lower values at CV transition and higher values at VC transition, again, lending support to a more simultaneously timed TB gesture for the Russian group. Indeed, the Russian group exhibited significantly greater values than either of the two English groups for both palatalized and C+j sequences in both syllable positions.

It was also found that for palatalized consonants, onsets had significantly lower values than codas, for all groups. This is again indicative of a tongue body gesture with greater lag for onset palatalized consonants than coda palatalized consonants.

8.6.2. Discussion: Labials VC – CV

For the difference between VC and CV transition for labial consonants, we begin by discussing differences between palatalization types. The consonant glide sequences (Cj) all had significantly lower values than either plain or palatalized sequences for both onsets and codas for all groups, as was the case for the coronals. This means again that these sequences had higher F2 values at CV transition than at VC transition. The same is true for palatalized consonants for all the groups in onset position. For the three groups, for codas, the difference between palatalized and plain consonants has a significantly higher value. This was not the case for the coronals. These results can be interpreted as follows.

The positive coefficient for the difference between palatalized consonants and plain consonants in coda position for all groups means that, contrary to the coronals where only the Russian group exhibited significantly higher VC-CV values for palatalized consonants in coda position over plain consonants, all groups did so for labials. The significantly higher VC-CV values for palatalized codas translate to higher F2 values at the VC position than at the CV position. This is one way in which the coronals and labials differed, in that both English groups behaved more similarly to the Russian group than they did for the coronals.

It was also found that for palatalized consonants, onsets had significantly lower values than codas, for all groups. In terms of differences between groups, the Russian group exhibited significantly greater values than either of the two English groups for the palatalized consonants, as was the case for the coronals. While for the C+j sequence, the Russian group only had significantly higher values than the naïve group for codas.

In sum, the most important observation from labial consonants was a greater similarity in the performance of the three groups, meaning, English speakers behaved more like Russian speakers for this place of articulation than for coronal consonants.

8.7. Results: preceding vowel subtractions

We have seen already results for individual transition positions (CV and VC) as well as the differences between the two consonant transitions. Now we investigate further the differences between other positions in the vowels preceding and following the consonants. Four new metrics were calculated, two from the vowel preceding the consonant and two from the vowel following the consonant. The metrics are described as follows.

The two metrics from the preceding vowels are taken from subtractions of F2 at two different positions in the preceding vowel from the VC transition. Similarly, the two metrics from the following vowel are subtractions of F2 at two different positions in the following vowel from the CV transition.

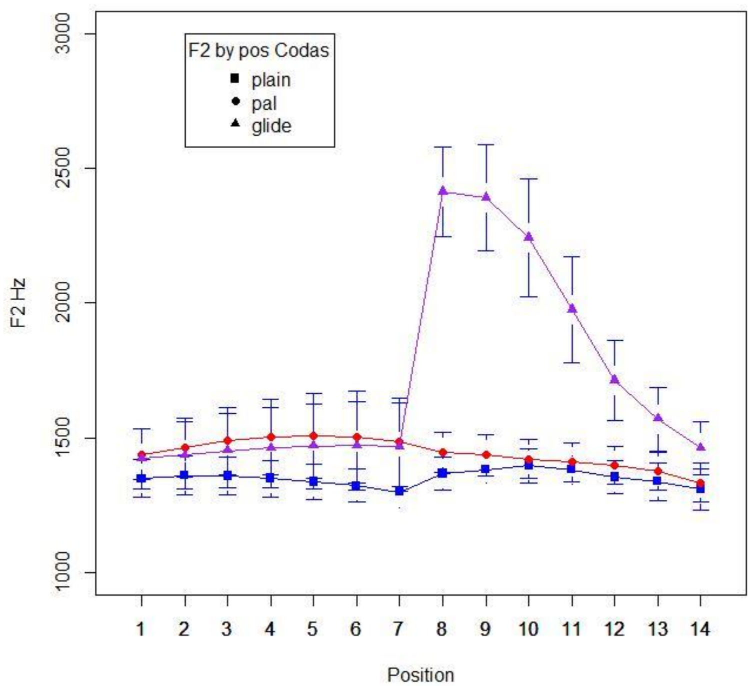

Figure 8.6: Mean F2 (Hz) by position for the plain, palatalized and consonant-glide sequence.

The plot above shows the positions where F2 measurements were taken. The position 7 corresponds to the VC transition, while position 8 corresponds to the CV transition. The subtractions occurred from the two positions labeled 5 and 6 in the preceding vowel and 9 and 10 in the following vowel. Thus, for the preceding vowel the first metric is the mean F2 at position 6 subtracted from the mean F2 at the VC transition (position 7). This metric will be known as VC-6. The second position is the mean F2 at position 5 subtracted from the VC transition, known as VC-5.

The first metric for the following vowel is the mean F2 at position 9 subtracted from the CV transition (position 8). The second metric is the mean F2 value at position 10 subtracted from the CV value. In general, the higher the value for these metrics, the greater the difference between the transition and the vowel points, indicating more abrupt raising or lowering of the tongue body. The data are presented in the same way as before, highlighting the comparisons between palatalization types, syllable positions and groups for each of the positions. We begin by describing the results from the metrics preceding the consonant.

8.7.1. Results: Coronals: preceding vowel subtractions

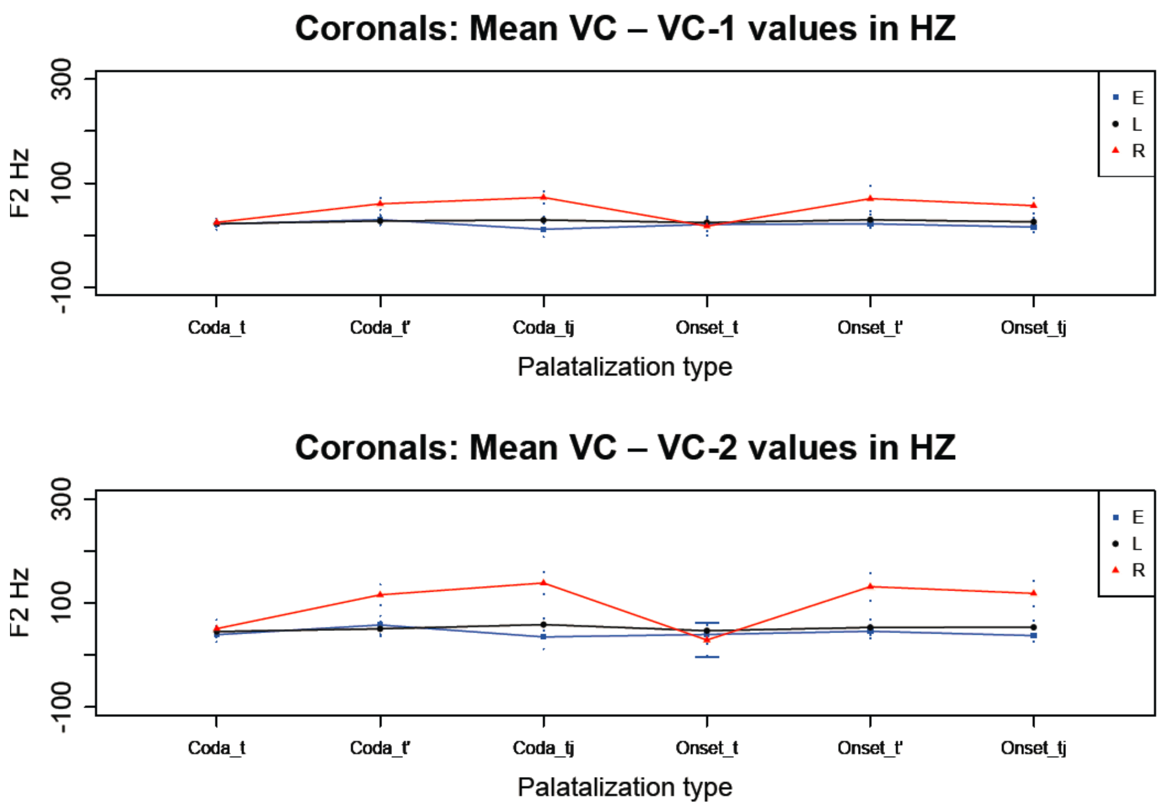

Figure 8.7: Mean differences in F2 between the VC transition and two points in the preceding vowel for coronals. The top panel represents the difference between the VC transition and the point closer to the VC transition, and the lower panel represents the difference between VC and the point further from the VC transition. The legend represent the groups: E: naïve group, L: learner group, R: native Russian group.

	Naïve		Learners		Russians	
VC - 6	**coda**	**onset**	**coda**	**onset**	**coda**	**onset**
C vs C^J					35 (0)***	52 (0)***
C vs Cj			-636 (0)***		47 (0)***	39 (0)***
C^J vs Cj	-18 (0.01)*					
VC - 5						
C vs C^J					65 (0)***	102 (0)***
C vs Cj					87 (0)***	89 (0)***
C^J vs Cj	-23 (0.03)*				22 (0.04)*	

Table 8.20: Fixed effects coefficients for palatalization type comparisons for mean F2 differences between VC transition and two preceding vowel points for coronals. The top panel represents the difference between the VC transition and the point closer to the VC transition, and the lower panel represents the difference between VC transition and the point further from the VC transition. Significant coefficients are indicated by asterisks indicating the level of significance. P-values are enclosed in brackets. A positive coefficient means that the palatalization type on the right of the “vs” label in the “Pal” column has a significantly higher value than the contrast type on the left of the “vs” label for the particular Group and syllable position.

We first present the data for the two positions preceding the consonant for the coronal consonants. Table 8.20 above shows the comparisons between palatalization types for the two subtractions (VC-6 and VC-5) for coronal consonants. The top panel represents the subtraction from VC transition to the closer vowel position (6), while the bottom panel represents the subtraction from VC to the further position (5).

The first thing to notice is that the Russian group exhibits more significant differences between the palatalization types than either English group. The coefficients for the Russian group are also all positive. The Russian group exhibited significantly greater difference in VC-6 for both palatalized and C+j consonants over plain consonants (first two rows) for both onsets and codas. This same effect held for VC-5 with the additional difference between the C+j and palatalized consonants in coda position (last row).

The only significant effects for the naïve English group are in differences between the C+j sequences and palatalized sequences in coda position, where C+j sequences had significantly lower values than palatalized consonants. This effect exists for both positions, VC-6 and VC-5. The English learner group had a significant effect between C+j and plain consonants. This effect is manifested as C+j having significantly lower values than plain consonants in coda position.

		Group		
VC - 16	**Syllable differences**	**Naïve**	**Learners**	**Russians**
C	Coda vs Onset			
C^J	Coda vs Onset			
Cj	Coda vs Onset			-15 (0.04)*
VC - 15	**Syllable differences**	**Naïve**	**Learners**	**Russians**
C	Coda vs Onset			
C^J	Coda vs Onset			
Cj	Coda vs Onset			

Table 8.21: Fixed effects coefficients for syllable position comparisons for mean F2 differences between VC transition and two points in the preceding vowel for coronals. The top panel represents the difference between the VC transition and the point closer to the VC transition, and the lower panel represents the difference between VC transition and the point further from the VC transition. Significant coefficients are indicated by asterisks indicating the level of significance. P-values are enclosed in brackets. A positive coefficient means that the syllable position on the right of the “vs” label in the “Syllable Differences”

column has a significantly higher value than the syllable position on the left of the “vs” label for the particular Group and palatalization type.

We move on to differences between syllable positions. As shown in Table 8.21, the only significant effect occurred between onsets and codas for the C+j sequences for the Russian group for the first subtraction. In this case, onsets had significantly lower values than codas.

		Syllable position	
VC - 16	**Group**	**coda**	**onset**
C	Naïve - Learners		
	Naïve - Russians		
	Learners - Russians		
C^j	Naïve - Learners		
	Naïve - Russians	30 (0.009)**	48 (0)***
	Learners - Russians	33 (0.009)**	40 (0.001)**
Cj	Naïve - Learners		
	Naïve - Russians	61 (0)***	41 (0.0005)***
	Learners - Russians	43 (0.0005)***	31 (0.01)*
VC - 15	**Group**	**coda**	**onset**
C	Naïve - Learners		
	Naïve - Russians		
	Learners - Russians		
C^j	Naïve - Learners		
	Naïve - Russians	58 (0.007)**	86 (0.0001)***
	Learners - Russians	65 (0.004)**	78 (0.0006)***
Cj	Naïve - Learners		
	Naïve - Russians	104 (0)***	82 (0.0002)***
	Learners - Russians	79 (0.0005)***	65 (0.004)**

Table 8.22: Fixed effects coefficients for group comparisons for mean F2 differences between the VC transition and two points in the preceding vowel for coronals. The top panel represents the difference between the VC transition and the point closer to the VC transition, and the lower panel represents the difference between VC transition and the point further from the VC transition. Significant coefficients are indicated by asterisks indicating the level of significance.

The next analyses involved differences between groups show in Table 8.22. Here, for palatalized consonants, and C+j sequences, but not for plain consonants, the Russian group differed significantly from both English groups. For both of these palatalization types, for both syllable positions, Russian speakers had significantly higher values than either of the two English groups. The English groups did not differ from each other. All of the coefficients for the differences between the Russian group and the two English groups were positive, meaning that for these two consonants and syllable positions, the Russian group had significantly higher VC-6 and VC-5 values.

8.7.2. Results: Labials preceding vowel subtractions

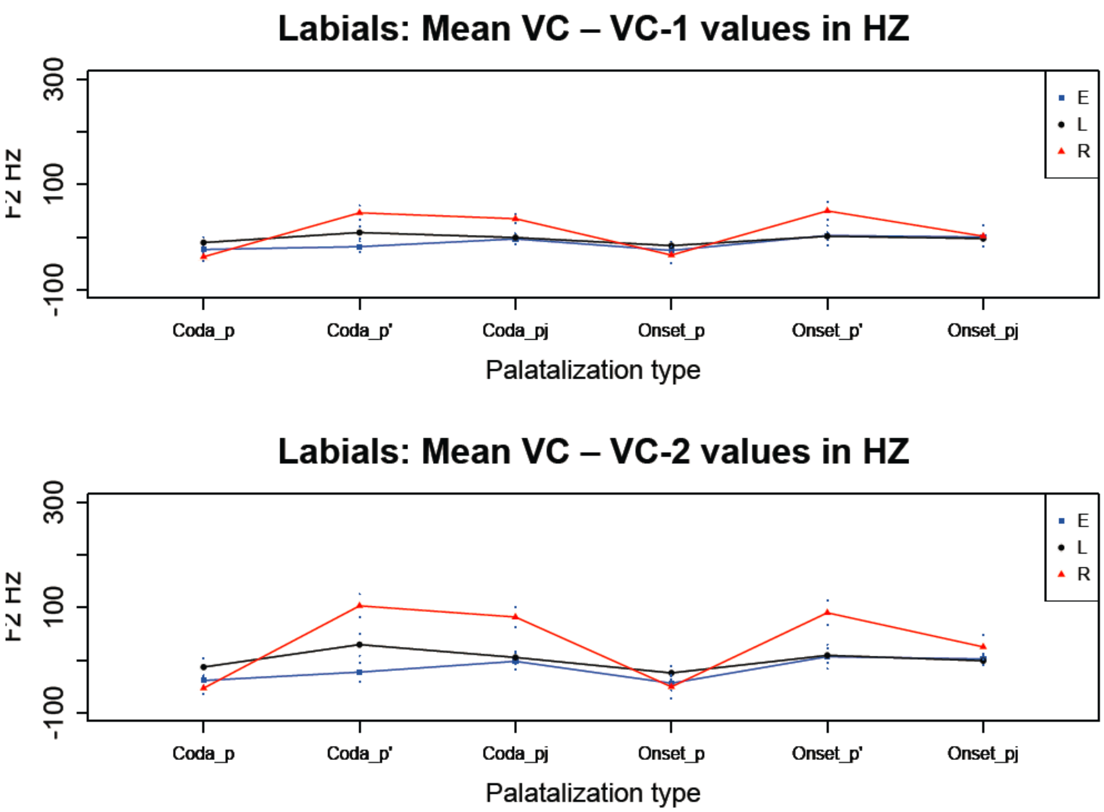

Figure 8.8: Mean differences in F2 between the VC transition and two points in the preceding vowel for labials. The top panel represents the difference between the VC transition and the point closer to the VC transition, and the lower panel represents the difference between VC and the point further from the VC transition. The legend represent the groups: E: naïve group, L: learner group, R: native Russian group.

	Naïve		Learners		Russians	
VC - 16	**coda**	**onset**	**coda**	**onset**	**coda**	**onset**
C vs C^j		28 (0.0001)***	19 (0.01)*	17 (0.02)*	83 (0)***	84 (0)***
C vs Cj	19 (0.007)**	25 (0.0007)***			72 (0)***	35 (0)***
C^j vs Cj	14 (0.047)*					-48 (0)***
VC - 15						
C vs C^j		50 (0)***	42 (0.0004)***	33 (0.005)**	156 (0)***	140 (0)***
C vs Cj	36 (0.001)**	45 (0)***			135 (0)***	75 (0)***
C^j vs Cj			-24 (0.04)*			-64 (0)***

Table 8.23: Fixed effects coefficients for palatalization type comparisons for mean F2 differences between the VC transition and two preceding vowel points for labials. The top panel represents the difference between the VC transition and the point closer to the VC transition, and the lower panel represents the difference between VC transition and the point further from the VC transition. Significant coefficients are indicated by asterisks indicating the level of significance. P-values are enclosed in brackets. A positive coefficient means that the palatalization type on the right of the "vs" label in the "Pal" column has a significantly higher value than the contrast type on the left of the "vs" label for the particular Group and syllable position.

The first thing to notice from Table 8.23 above is that the results seem less organized as the results for coronals. There are more significant differences between the palatalization types for labials than there are for coronals. For the naïve group, for codas, we see significantly higher values for C+j sequences than either plain or palatalized sequences for VC-6 position, whereas this effect disappears with the difference between Cj and palatalized consonants for VC-5 position. For the learner group, a different pattern for codas emerges, where palatalized consonants have significantly higher values than plain consonants, whereas this did not arise for the naïve group. Also, for the VC-5 position, C+j sequences had significantly lower values than palatalized consonants. The native Russian group had significantly higher values for both palatalized and C+j sequences over plain consonants in for codas in for both positions. No differences between C+j and palatalized consonants were found for codas for this group.

As for the onsets, the naïve group has significantly higher values for both palatalized and C+j sequences over plain consonants for both positions. The Russian group also has these same effects. The Russian group has an additional difference between the C+j sequences and palatalized consonants which either English group lack. This group has significantly lower C+j value sfor sequences over palatalized sequences. Finally, the learner group only exhibits significantly higher values for palatalized consonants over plain consonants.

		Group		
VC - 16	**Syllable differences**	**Naïve**	**Learners**	**Russians**
C	Coda vs Onset			
C^j	Coda vs Onset	21 (0.004)**		
Cj	Coda vs Onset			-33 (0)***
VC - 15	**Syllable differences**	**Naïve**	**Learners**	**Russians**
C	Coda vs Onset			
C^j	Coda vs Onset	29 (0.008)**		
Cj	Coda vs Onset			-56 (0)***

Table 8.24: Fixed effects coefficients for syllable position comparisons for mean F2 differences between VC transition and two points in the preceding vowel for labials. The top panel represents the difference

between the VC transition and the point closer to the VC transition, and the lower panel represents the difference between VC transition and the point further from the VC transition. Significant coefficients are indicated by asterisks indicating the level of significance. P-values are enclosed in brackets. A positive coefficient means that the syllable position on the right of the “vs” label in the “Syllable Differences” column has a significantly higher value than the syllable position on the left of the “vs” label for the particular Group and palatalization type.

We move on to differences between syllable positions. Table 8.24 shows the only significant effects occurred for the naïve group and the native Russian group. For the naïve group, onsets had significantly higher values than codas for palatalized consonants. The native Russian group had significantly lower values for onsets than codas for the C+j sequences. These effects hold for both positions (VC-6 and VC-5).

		Syllable position	
VC - 16	**Group**	**coda**	**onset**
C	Naïve - Learners		
	Naïve - Russians		
	Learners - Russians	-27 (0.01)*	
C^J	Naïve - Learners	27 (0.01)*	
	Naïve - Russians	64 (0)***	47 (0)***
	Learners - Russians	37 (0)***	48 (0)***
Cj	Naïve - Learners		
	Naïve - Russians	38 (0)***	
	Learners - Russians	35 (0.001)**	
VC - 15	**Group**	**coda**	**onset**
C	Naïve - Learners		
	Naïve - Russians		
	Learners - Russians	-40 (0.02)*	
C^J	Naïve - Learners	52 (0.002)**	
	Naïve - Russians	126 (0)***	84 (0)***
	Learners - Russians	73 (0)***	81 (0)***
Cj	Naïve - Learners		
	Naïve - Russians	84 (0)***	
	Learners - Russians	76 (0)***	

Table 8.25: Fixed effects coefficients for group comparisons for mean F2 differences between the VC transition and two points in the preceding vowel for labials. The top panel represents the difference between the VC transition and the point closer to the VC transition, and the lower panel represents the difference between VC transition and the point further from the VC transition. Significant coefficients are indicated by asterisks indicating the level of significance.

We now present results for differences between groups. Here, again, as for the coronals, for palatalized consonants, and C+j sequences, the Russian group differed significantly from both English groups. For both of these palatalization types, for both syllable positions, Russian speakers had significantly higher values than either of the two English groups. All of the coefficients between the Russian group and the two English groups were positive, meaning that for these two consonants and syllable positions, the Russian group had significantly higher VC-6 and VC-6 values. Unlike for the coronals, here this difference also existed between the two English groups, but only for palatalized consonants in coda position, where the learner group had significantly higher values than the naïve group.

Two effects also exist for coda consonants, but not for onsets. For C+j sequences in both positions, the Russian group has significantly higher values than either English groups (last two rows for each position). The final effect is with the plain consonants, where Russian speakers had significantly lower values than learners.

8.8. Discussion: Preceding vowel subtractions

In this section vowel points from the vowel preceding the consonant were incorporated into the analysis. Two metrics involving two different points were used. The metrics are the difference between the VC transition and each of the two vowel points. For these metrics, a higher value means a larger difference and a higher F2 value at the VC transition. A low value means less difference between the two points. A negative value means a lower VC transition than the point on its preceding vowel. This information, combined with the other metrics discussed above, will provide more information on timing of tongue body gestures.

8.8.1. Discussion: Coronals: preceding vowel subtractions

First, we discuss the pre-consonantal metrics for coronals. One of the main patterns that emerged for both metrics (VC-6 and VC-5) is that the Russian group had significantly higher values for both palatalized consonants and C+j sequences than plain consonants. This is important because we see significant differences between the native Russian group and the remaining English groups for positions preceding the consonant for both palatalization types.

Moving on to syllable position differences, the only significant effect involved a difference between onsets and codas for the C+j sequences for Russian speakers for the subtraction involving the position closest to the transition. This difference was barely significant.

The largest differences occurred between groups. Both of the English groups differed significantly from the Russian group for all comparisons for the palatalized consonants and C+j sequences. Again, if we recall the VC transition analysis (Table 8.4 in section 8.3.1.1), the only difference between the groups occurred between the native Russian group and the learners, where the native Russian group had significantly higher F2 values at the VC transition for coronals in coda position. The significant differences between the groups for the vowel midpoint subtraction metrics suggest the difference here stems from a difference between the vowel points themselves. This means that Russian speakers had significantly lower F2 vowel values than either English group, to result in a significant VC transition – vowel point subtraction.

8.8.2. Discussion: Labials preceding vowel subtractions

We now discuss the pre-consonantal metrics for labials consonants. One of the main patterns that emerged for both metrics (VC-6 and VC-5) and both syllable positions, is that the Russian group had significantly higher values for both palatalized consonants and C+j sequences over plain consonants, as was the case for the coronal consonants. The naïve English group saw this effect also, though only for onsets. In this way, this group differed from the Russian group. For codas, the naïve group only saw differences between C+j sequences and plain and palatalized consonants, but not between palatalized consonants and plain consonants.

The learner group also differed from both groups. This group only had significantly higher values for palatalized consonants over plain consonants for both syllable positions and both vowel points. In coda position at least, the three groups differed more from each other than they did for coronal consonants. In fact, for codas, the three groups differed significantly from each other for palatalized consonants in both vowel points. The Russian group had the highest value, the learner group had lower values than the Russian group, and the naïve English group had the lowest value. Recall from the VC transition analysis, that the naïve group had significantly lower values than the native Russian group.

We can see, therefore that the three groups differ from each other for coda consonants for values preceding the consonant. In this case, the Learners behaved somewhere between the naïve English group and the Russian group, exhibiting an effect of learning. They did not behave exactly like naïve English speakers, nor did they resemble native Russians as a group, unlike for coronal coda consonants. However, also like coronal consonants, we still see significant differences between the native Russian group and the remaining English groups for positions preceding the consonant for palatalized consonants and C+j sequences for codas.

For onset consonants, the Russian group had significantly higher values than either English group for palatalized consonants, like for onsets. There were no other differences, whereas for coronals both palatalized and C+j sequences exhibited differences between the

Russian groups and English groups.

Moving on to syllable position differences, the only significant effect involved a difference between onsets and codas for the C+j for Russian speakers for both vowel points. There was also a significant difference between onsets and codas for palatalized consonants for the naïve group. This is also evident in the group differences described above. This difference was barely significant.

8.9. Results: Following vowel subtractions

This section presents the results from the two metrics for the vowel positions following the consonant. The two metrics for the following vowels are taken from subtractions of F2 at two different positions in the following vowel from the CV transition.

Recall that the first metric for the following vowel is the mean F2 at position 9 subtracted from the CV transition and the second metric is the mean F2 value at position 10 subtracted from the CV transition value. The data are presented in the same way as before, highlighting the comparisons between palatalization types, syllable positions and groups for each of the positions. We begin by describing the results from the metrics preceding the consonant.

8.9.1. Results: Coronals: Following vowel subtractions

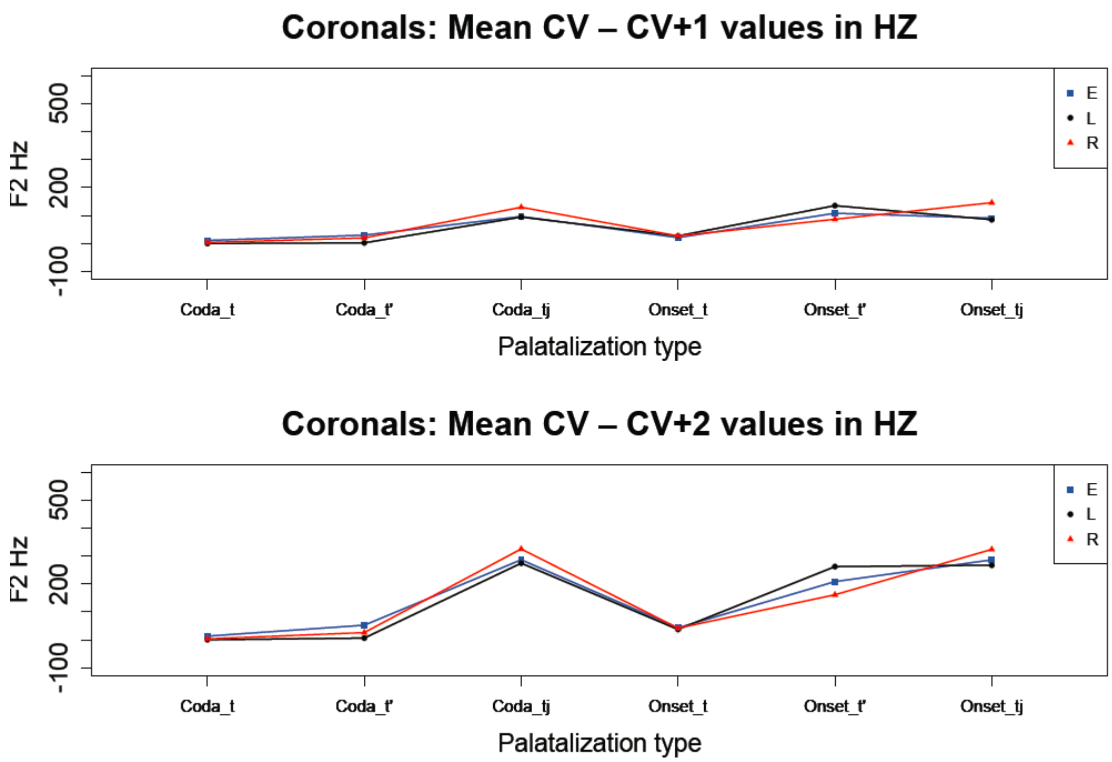

Figure 8.9: Mean differences in F2 between the CV transition and two points in the following vowel for coronals. The top panel represents the difference between the CV transition and the point closer to the CV transition, and the lower panel represents the difference between the CV transition and the point further from the CV transition. The legend represent the groups: E: naïve group, L: learner group, R: native Russian group.

CV – V1	Naïve coda	Naïve onset	Learners coda	Learners onset	Russians coda	Russians onset
C vs C^j		87 (0)***		108 (0)***		59 (0)***
C vs Cj	85 (0)***	69 (0)***	94 (0)***	58 (0)***	125 (0)***	118(0)***
C^j vs Cj	66 (0)***		92 (0)***	-50 (0.0001)***	109 (0)***	59 (0)***
CV – V2						
C vs C^j	39 (0.04)*	165 (0)***		224 (0)***		119 (0)***
C vs Cj	273 (0)***	243 (0)***	273 (0)***	229 (0)***	321 (0)***	281 (0)***
C^j vs Cj	234 (0)***	77 (0)***	268 (0)***		299 (0)***	162 (0)***

Table 8.26: Fixed effects coefficients for palatalization type comparisons for mean F2 differences between CV transition and two following vowel points for coronals. The top panel represents the difference between the CV transition and the vowel point closer to the CV transition, and the lower panel represents the difference between CV transition and the point further from the CV transition. Significant coefficients are indicated by asterisks indicating the level of significance. P-values are enclosed in brackets. A positive coefficient means that the palatalization type on the right of the "vs" label in the "Pal" column has a significantly higher value than the contrast type on the left of the "vs" label for the particular Group and syllable position.

Table 8.26 above shows the comparisons between palatalization types for the two subtractions (CV-V1 and CV-V2) for coronal consonants. The top panel represents the subtraction from CV transition to the closer vowel position (V1), while the bottom panel represents the subtraction from CV to the further position (V2).

For coda consonants, all three groups had similar significant effects for the first position

CV-V1. All three had significantly higher values for C+j sequences than either plain or palatalized sequences. For the position CV-V2, the learner group and the native Russian group showed the same effects for codas, whereas the naïve group shared those effects but had the additional effect of a higher value for palatalized consonants over plain consonants.

For onsets, the native Russian group had significant effects between all three comparisons for both positions. This group had significantly higher values for palatalized consonants over plain consonants, as well as significantly higher values for C+j sequences over both plain and palatalized consonants. The naïve group had significantly higher values for palatalized and C+j consonants over plain consonants for the first position, with the same effects as the Russian speakers (all three comparisons were significant) for the second position. The learner group also had significantly higher values for palatalized consonants and C+j sequences over plain consonants for both positions. The learner group differed also in that for this group, C+j sequences had significantly lower values than palatalized consonants only for the earlier position (CV-V1). This effect was not seen for either of the other groups.

		Group		
CV – V1	**Syllable differences**	**Naïve**	**Learners**	**Russians**
C	Coda vs Onset		27 (0.04)*	23 (0.05)*
C^j	Coda vs Onset	78 (0)***	133 (0)***	66 (0)***
Cj	Coda vs Onset	35 (0.04)*		
CV – V2	**Syllable differences**	**Naïve**	**Learners**	**Russians**
C	Coda vs Onset			38 (0.05)*
C^j	Coda vs Onset	155 (0)***	255 (0)***	135 (0)***
Cj	Coda vs Onset			

Table 8.27: Fixed effects coefficients for syllable position comparisons for mean F2 differences between CV transition and two points in the following vowel for coronals. The top panel represents the difference between the CV transition and the point closer to the CV transition, and the lower panel represents the difference between CV transition and the point further from the CV transition. Significant coefficients are indicated by asterisks indicating the level of significance. P-values are enclosed in brackets. A positive coefficient means that the syllable position on the right of the "vs" label in the "Syllable Differences" column has a significantly higher value than the syllable position on the left of the "vs" label for the particular Group and palatalization type.

We move on to differences between syllable positions. Table 8.27 shows for the naïve group, onsets had significantly higher values than codas for palatalized consonants and C+j sequences for the first position (top panel), but only for palatalized consonants for the second position. The learner group and the native Russian group both had significantly higher values for onsets over codas for both plain and palatalized sequences for the first position. The native Russian group had the same effects for the second position, whereas for the learner group this effect remained only for the palatalized consonants.

		Syllable position	
CV – V1	Group	coda	onset
C	Naïve - Learners		
	Naïve - Russians		
	Learners - Russians		
C^J	Naïve - Learners		
	Naïve - Russians		
	Learners - Russians		-48 (0.007)**
Cj	Naïve - Learners		
	Naïve - Russians	35 (0.04)*	57 (0.008)**
	Learners - Russians	35 (0.05)*	60 (0.0008)***
CV – V2	**Group**	**coda**	**onset**
C	Naïve - Learners		
	Naïve - Russians		
	Learners - Russians		
C^J	Naïve - Learners		
	Naïve - Russians		
	Learners - Russians		-100 (0)***
Cj	Naïve - Learners		
	Naïve - Russians		
	Learners - Russians		

Table 8.28: Fixed effects coefficients for group comparisons for mean F2 differences between the CV transition and two points in the following vowel for coronals. The top panel represents the difference between the CV transition and the point closer to the CV transition, and the lower panel represents the difference between CV transition and the point further from the CV transition. Significant coefficients are indicated by asterisks indicating the level of significance.

We now present results for differences between groups. Here, for the first position (CV-V1) for C+j sequences, the Russian group differed significantly from both English groups. For this palatalization type, for both syllable positions, Russian speakers had significantly higher values than either of the two English groups. The only other significant effect existed for the palatalized consonants in onset position. Here, the Russian group had significantly lower values than the learner group. No other differences between groups existed.

8.9.2. Results: Labials following vowel subtractions

This section presents the results for the labial consonants for subtractions of CV transitions and following vowel positions. Again, the results are presented in the same way as for preceding sections.

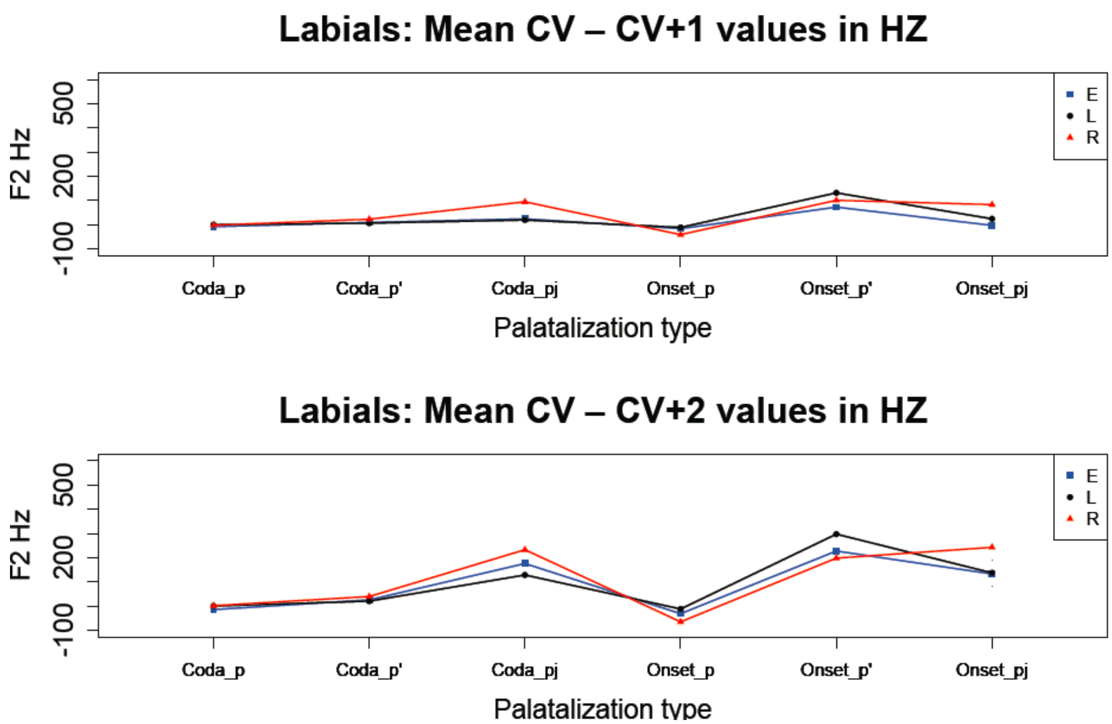

Figure 8.10: Mean differences in F2 between the CV transition and two points in the following vowel for labials. The top panel represents the difference between the CV transition and the point closer to the CV transition, and the lower panel represents the difference between the CV transition and the point further from the CV transition. The legend represent the groups: E: naïve group, L: learner group, R: native Russian group.

	E		L		R	
CV – V1	**coda**	**onset**	**coda**	**onset**	**coda**	**onset**
C vs C^j		88 (0)***		143 (0)		141 (0)***
C vs Cj	31 (0.02)*			35 (0.01)*	96 (0)***	124 (0)***
C^j vs Cj		-76 (0)***		-107 (0)***	71 (0)***	
CV – V2						
C vs C^j		259 (0)***		309 (0)***		263 (0)***
C vs Cj	190 (0)***	162 (0)***	128 (0)***	150 (0)***	231 (0)***	308 (0)***
C^j vs Cj	150 (0)***	-96 (0)***	108 (0)***	-159 (0)***	193 (0)***	44 (0.04)*

Table 8.29: Fixed effects coefficients for palatalization type comparisons for mean F2 differences between CV transition and two following vowel points for labials. The top panel represents the difference between the CV transition and the vowel point closer to the CV transition, and the lower panel represents the difference between CV transition and the point further from the CV transition. Significant coefficients are indicated by asterisks indicating the level of significance. P-values are enclosed in brackets. A positive coefficient means that the palatalization type on the right of the "vs" label in the "Pal" column has a significantly higher value than the contrast type on the left of the "vs" label for the particular Group and syllable position.

Table 8.29 has comparisons between palatalization types for the subtractions (CV-V1 and CV-V2) for labial consonants. The top panel represents the subtraction from VC transition to the closer vowel position (V1), while the bottom panel represents the subtraction from VC to the further position (V2).

For coda consonants for the first position (CV-V1), the naïve group and the native Russian group both had significantly higher values for C+J sequences over plain consonants. The native Russian group had significantly higher values for C+j sequences than for palatalized consonants for this position as well. The learner group did not show any significant effects for codas in this position. For the second position, (CV-V2) all three groups had similar significant effects for the first position for codas. All three had significantly higher values for C+j sequences than either plain or palatalized sequences.

For onsets, for the first position, all three groups had significantly higher values for palatalized consonants over plain consonants. For the difference between C+j sequences and plain consonants, Groups L and R had significant effects for the first position, where C+j sequences had significantly higher values. For the difference between C+j sequences and palatalized consonants only the two English groups showed significant effects, where C+j sequences had significantly lower values.

For onsets for the second position, all three groups had similar effects for the differences between the palatalized consonants and plain consonants, and C+j sequences and palatalized consonants. Palatalized consonants had significantly higher values than plain consonants, and C+j sequences had significantly higher values than palatalized consonants. Both English groups had significantly lower values for C+j sequences over palatalized consonants, while the Russian group had the opposite effect, with C+j sequences having significantly higher values than palatalized consonants.

		Group		
CV – V1	**Syllable differences**	**Naïve**	**Learners**	**Russians**
C	Coda vs Onset			-38 (0.007)**
C^J	Coda vs Onset	63 (0)***	125 (0)***	78 (0)***
Cj	Coda vs Onset	-27 (0.04)*		
CV – V2	**Syllable differences**	**Naïve**	**Learners**	**Russians**
C	Coda vs Onset			-66 (0.002)**
C^J	Coda vs Onset	202 (0)***	277 (0)***	158 (0)***
Cj	Coda vs Onset	-44 (0.03)*		

Table 8.30: Fixed effects coefficients for syllable position comparisons for mean F2 differences between CV transition and two points in the following vowel for labials. The top panel represents the difference between the CV transition and the point closer to the CV transition, and the lower panel represents the difference between CV transition and the point further from the CV transition. Significant coefficients are indicated by asterisks indicating the level of significance. P-values are enclosed in brackets. A positive coefficient means that the syllable position on the right of the “vs” label in the “Syllable Differences” column has a significantly higher value than the syllable position on the left of the “vs” label for the particular Group and palatalization type.

We move on to differences between syllable positions. Table 8.30 shows that for palatalized consonants, for all three groups, and for both positions onsets had significantly higher values than codas. The native Russian group also had an effect for plain consonants for both positions, in that plain onsets had significantly lower values than plain codas. The naïve group had an effect for C+j sequences in both positions, where onsets had significantly lower values than codas.

		Syllable position	
CV – V1	Group	coda	onset
C	Naïve - Learners		
	Naïve - Russians		
	Learners - Russians		
C^J	Naïve - Learners		60 (0.003)**
	Naïve - Russians		
	Learners - Russians		
Cj	Naïve - Learners		
	Naïve - Russians	71 (0.0004)***	87 (0)***
	Learners - Russians	74 (0.0004)***	58 (0.005)**
CV – V2	Group	coda	onset
C	Naïve - Learners		
	Naïve - Russians		
	Learners - Russians		
C^J	Naïve - Learners		73 (0.03)*
	Naïve - Russians		
	Learners - Russians		-99 (0.004)**
Cj	Naïve - Learners		
	Naïve - Russians		115 (0.0006)***
	Learners - Russians	104 (0.002)**	104 (0.002)**

Table 8.31: Fixed effects coefficients for group comparisons for mean F2 differences between the CV transition and two points in the following vowel for labials. The top panel represents the difference between the CV transition and the point closer to the CV transition, and the lower panel represents the difference between CV transition and the point further from the CV transition. Significant coefficients are indicated by asterisks indicating the level of significance.

The last section presents group differences. For the palatalized consonants, for both positions (CV-V1, CV-V2) for syllable onsets, the learner group had significantly higher values than the naïve group. No effects existed for this consonant in coda position. For the second position (CV-3), the native Russian group had significantly lower values than the learner group.

For the C+j sequences, for the first position, the native Russian group differed from the two English groups, where the native Russian group had significantly higher values than either English group for both syllable positions. The native Russian group also had significantly higher values for this sequence than both English groups for the second position (CV-3) for onsets, whereas this difference was only significant for coda consonants between the native Russian group and the learner group, not the naïve group.

8.10. Discussion: Following vowel subtractions

In this section vowel points from the vowel following the consonant were incorporated into the analysis. Two metrics involving the difference between the CV transition and each of the two vowel points were used. For these metrics, a higher value means a larger difference and also a higher F2 value at the CV transition than at the vowel point. A low value means less difference between the two points. A negative value means a lower CV transition than the point in the following vowel. This information will provide more information on timing of tongue body gestures.

In this section we present discussion of the results for the transition to following vowel subtraction metrics. We begin by discussing coronal consonants and end with the labial consonants.

8.10.1. Discussion: Coronals following vowel subtractions

Here we discuss the results for coronal consonants. For syllable codas, all three groups behaved similarly for the first position, having higher values for C+j sequences than either of the other two consonants. This is indicative of the main difference of the C+j sequences at CV transition and beyond in coda position[18]. More interesting differences for CV-vowel point subtractions occurred for consonants in the onset position. For the first subtraction, that is, CV subtracted from the closer vowel point, (CV-V1) the difference between palatalized consonants and C+j sequences is significant for learners but not for naïve listeners, meaning learners have a significantly greater difference in CV and vowel point for palatalized consonants than for glides while naïve learners do not.

Recall for the CV transition values, the two English groups had higher values in onset position than the native Russian group for palatalized consonants. No differences existed for onset C+j between the groups. All groups had significant differences between palatalization types within groups. For the first subtraction, that is, CV subtracted from the closer vowel point, (CV-2) the difference between C’ and CJ is significant for learners but not for naïve listeners, meaning learners have a significantly greater difference in CV and vowel point for palatalized consonants than for glides while naïve learners do not. At first glance, this could seem like a possible learning effect. However, the Russian group has the opposite effect, meaning the

[18] There was a difference in how speakers produced C+j sequences in this position versus the palatalized and plain consonants. Many times, syllabification of the /j/ to the following vowel (across the word boundary) occurred, thus yielding higher CV transition values and thus higher differences between the CV transition and following vowel points for these sequences.

difference is greater for C+j than palatalized consonants. This effect for the Russian group is not due to a higher F2 at the CV transition (see Table 8.10 for differences on the actual CV transition). This effect is clearly seen in Figure 8.8 where the Russian group has higher values for C+j sequences than other consonants, while the learner group has lower values.

More specifically, in terms of group differences, the Russian group also differs significantly from the other two English groups for this consonant (C+j) at onset and coda positions, as also seen on the plot. However, for the palatalized consonant, the only group differences were between the native Russian group and the learner group, where the native Russian group had lower values than the learner group for onsets for both vowel subtractions. Thus, the two English groups seem to behave more closely with each other than either does with the Russian group, for C+j and palatalized consonants.

8.10.2. Discussion: Labials: following vowel subtractions

Now we move on to the discussion for labial consonants. Starting with syllable codas, unlike for coronals, where all three groups behaved similarly for the first position, only Groups E and R had differences between C+j and plain consonants and C+j and plain and palatalized consonants, respectively. This is also indicative of the main difference the C+j sequences at CV transition and beyond in coda position. There was a difference, but not as pronounced as for the coronals, in how speakers produced C+j sequences in this position versus the palatalized and plain consonants. For the second position subtraction, for codas, all three groups had differences between C+j and the two other consonants, as was the case for the coronals.

Again, as for coronals, more interesting differences for CV-vowel point subtractions occur for consonants in the onset position. Consider the difference between C+j and palatalized consonants. For coronals, The learner group and the native Russian groups differed from each other in the direction of this difference. For the native Russian group, C+j sequences had higher values in both subtractions, while the learner group had lower values for C+j at the first subtraction, meaning Learners had significantly lower values for C+j than palatalized consonants, while Russian speakers had the opposite effect. For labials, the same pattern emerges for Group L, but it extends to the second subtraction (further into the vowel), and it is also present for Group E, unlike for the coronals.

Both English groups behaved more similarly to each other in the case of labials for the difference between C+j sequences and palatalized consonants in onset position, and the Russian group behaved differently from both groups. When looking specifically at Group differences, English groups differed from the Russian group for C+j sequences in onset position, where Russian speakers had significantly higher values. This was also the case for codas for the first subtraction, and codas for the second subtraction only between the native Russian group and the Learner group.

In terms of group differences for palatalized consonants, indeed less differences emerge. In the first and second subtractions, the learner group only had higher values than the naïve group. Recall, again, from the CV transition analysis that no differences existed between the two English groups here, thus the difference we see in this present analysis must be due to the vowel. That is, a greater CV – vowel point difference for the learner group than for the naïve group for palatalized consonants seems to be as a result of higher TB gesture activity further away from the CV transition for the naïve group than the learner group, possibly evidence of a learning effect for the learner group. This effect was not seen in the coronal consonants.

Chapter 9 Comparisons between experiments

This chapter introduces additional analyses and comparisons between the perception and production experiments presented above. The goal of this chapter is to compare the results between production and perception and to present discussion on the factor of learning on perception and production, in pursuit of two of the main goals introduced in the first chapter. The results from different perception experiments are compared with production data to more closely determine the relation between perception and production. This chapter is organized as follows; Section 9.1 is an analysis relating perception to production at the group level, and Section 9.2 compares perception and production at the subject level.

9.1. Comparisons perception and production

The relationship between perception and production is one of the main focuses of this dissertation. The goal was to investigate how speakers having different knowledge or experience in Russian perceive and produce the Russian palatalized consonants in different syllable positions and at different places of articulation. Another aim was to determine for each group, the relative proficiency and ability for perception versus production.

To more closely reveal the link between perception and production the results from the individual experiments are compared. This section introduces some analyses performed to arrive at these goals. The first analysis is a comparison between the categorization data and the production data at the group level. Since the categorization experiment in Chapter 6 and the production experiment in Chapter 8 both used the same number and identity of tokens a comparison between them is the most straightforward. Recall that each English participant heard the same 12 sequences a total of 7 times for both the categorization task and the production task. In one task the subjects answered questions to test their perception, while in the other they repeated the tokens into a microphone to test production.

The second analysis is also a comparison of the categorical and production data but at the subject level, to elucidate any within-group patterns. Finally, we will examine the relation between average standard deviation of F2 as a metric of variability in production to the performance in perception of both discrimination and categorization. This last analysis is undertaken to test one of the hypotheses laid out in 0– namely that decreased proficiency in perception is correlated with increased variability in production.

9.1.1. Categorization and Production comparison

Recall from Chapter 8, for the average difference between the VC and CV transitions, the native Russian group had significantly higher values than both English groups for both coronals and labials for both coda and onset palatalized consonants indicating a combination of a higher F2 value on average at the VC transition and also a lower F2 value at the CV transition. We will now compare this main result to the perception data.

For perception and specifically for assimilation, the native Russian group was much more consistent in assimilations of Russian palatalized consonants to the Russian palatalized category in coda and onset positions than either English group[19] thus showing a more difficult assimilation strategy for palatalized consonants, and also coda consonants for both English

[19] The learner group had the option to assimilate the consonant to the Russian palatalized category, the naïve group did not.

groups. The production results at the group level did not demonstrate great differences in the average F2 value at the either VC, CV transitions nor was there a great difference in the average difference in F2 between the two transitions for the two English groups.

In perception, the two English groups did differ more from each other, however. While the English learner group had the chance to assimilate consonants to a Russian palatalization category, and the naïve group did not we saw in Chapter 6 that the naïve group had in general more variable assimilations, especially for coda consonants. For coda consonants the naïve group assimilated the palatalized consonants almost as well to the JC category as the plain C category, and much less to the CJ category. Thus, we should see whether a similar difference or propensity for production of JC consonants exists in coda position.

In production, neither of the English groups differed from each other in VC, CV or VC-CV production metrics, indicating that at least at the group level, the relation between perception and production between the two groups was not the same. The two English groups were more similar to each other in production and more distinct from each other in perception. That is, it seems that perception can improve faster than production for groups overall, especially in coda position. The overall conclusion for the group level comparison between perception and production for the English groups therefore is that the groups differed more from each other in perception than production, and that as such, both groups differed to a similar degree from the native Russian group, while all three groups differed from each other more in perception than production. In order to more fully understand the relationship between perception and production in non-native speech we will examine the comparisons between perception and production results at the subject level.

9.2. Subject level comparisons:

We now move on to comparing the production of palatalized consonants to perception of palatalized consonants for individual subjects. The purpose of this section is not simply to examine each subject individually but to determine and expose any possible sub-groups within each larger group based on perception and/or production strategies. That is, we wish to determine whether individual subjects exhibit specific patterns in their production and perception and if so, to determine what these patterns are and whether the patterns hold across both perception and production. This endeavor will be similar to that already done for the comparisons between perception results.

In order to accomplish these goals we will compare production of not only palatalized consonants but plain consonants. For each place of articulation (coronals and labials) we will investigate the relation between two variables, a production variable and a perception variable. For the first group of analyses the production variable is always a metric of the difference between F2 at either the VC or the CV consonantal transition between the palatalized consonant and the pain consonant. For example, for the onset coronals, the production metric would be the mean difference in F2 at the VC transition between the palatalized onset coronal and the plain onset coronal. Another analysis will involve a production metric for the difference in F2 between palatalized and plain consonants at the CV transition. This production metric gives us an indication of the difference in the tongue body raising at both the VC transition into the consonant and the CV transition out of the consonant between the palatalized and plain consonants. In a way we are normalizing the effect of palatalization in F2 for each subject by subtracting the F2 values at the transitions for the plain consonants from the values for the palatalized consonants. Thus, a higher value for this metric will indicate that the transition (CV or VC) was produced with higher F2 for the palatalized consonant than for the plain consonant, and thus the palatalized consonant is associated with a higher tongue body raising gesture at that transition than the plain consonant.

Each syllable position and place of articulation combination will have two production metrics, one for the VC transition and one for the CV transition. For each of the syllable positions and place of articulation combinations we will compare perception and production with two metrics.

The perception metric is the fit index of the assimilation of the palatalized consonant to a particular perceptual category (the same variables analyzed in previous chapters). For example, we could be interested in comparing how production of the palatalized onset coronal relates to perception of this same consonant, in which case we would compare the difference in F2 metric described above to the assimilation fit index of the palatalized onset coronal to the plain perceptual category, for example. Since naïve subjects had three perceptual categories to choose from (plain, consonant+glide, and glide+consonant) we will have separate comparisons for the different categories. We will briefly go over each of these analyses for the naïve group and then discuss the overall observations at a higher level. For the learner group we will provide a separate discussion.

Using the difference between palatalized and plain consonants as a production metric allows us to easily see the effect of palatalization in production. It also allows for predictions in the relation between perception and production. For example, if a subject exhibits no difference in mean F2 value at the CV transition between the plain and palatalized consonants, we will be able to see whether the prediction that this subject also exhibit a certain pattern in perception will hold, namely that if no difference in production exists between plain and palatalized consonants, no difference in perception should also arise. Given this example we can formulate similar predictions across all of our comparisons. We can have as working hypotheses the following predictions for perception and production for each comparison. The predictions are formulated as relationships between the mean F2 difference value for production (positive or negative, higher or lower) and the assimilation of a palatalized consonant to a particular category for perception.

	Pal to plain	**Pal to JC**	**Pal to CJ**	**Pal to Pal**
VC transition	Negative	Positive	Negative	Positive (if both)
CV transition	Negative	Negative	Positive	Positive (if both)
Both	Negative	Negative	Negative	Positive

Table 9.1: Predictions for the relationship between perception and production for the naive group of subjects.

We formulate predictions for three perceptual assimilation patterns, palatalization to the plain category, to the consonant+glide category and to the glide+consonant category. Recall that these predictions and this sub-section relate only to the naïve group so far. Table 9.1 above demonstrates the general predictions for the association between assimilation of the palatalized consonant to each of these particular categories (perception) and the magnitude and direction of difference between F2 at each of the two consonant-vowel transitions between palatalized and plain consonants. For example, we predict that a high fit index of assimilation of the palatalized consonant to the plain category should be negatively associated with a high difference in F2 between the palatalized consonant and the plain consonant for both the VC and

CV transitions (first column). The reason for this is that a high fit index assimilation of the palatalized consonant to the plain category indicates that the subject perceived the palatalized consonant as a good exemplar of a plain consonant and as such under this hypothesis would produce the same consonant in a similar way – as a plain consonant. The second column shows the predictions for production for a high fit index of the palatalized consonant to the glide+consonant category. A high fit index to this category means that the subject perceived the palatalized consonant as a sequence of glide+consonant. This is predicted to also coincide with a high production metric for the VC transition but not necessarily for the CV transition. The reason for this is that a larger difference in F2 at the VC transition between palatalized and plain consonants indicates that the subject exhibited a higher degree of tongue body raising at the VC transition for the palatalized consonant and therefore, if perception and production are aligned, should also perceive the same consonants as a sequence of glide+consonant. The third prediction states that if a subject has high fit index assimilation values for the palatalized consonant to the consonant+glide category, this should coincide with a larger difference in F2 between the palatalized and plain consonants at the CV transition. This prediction states that if a subject perceives the palatalized consonant as a sequence of consonant+glide, then the subject should also produce the consonant in a similar way, with a higher F2 and thus greater tongue body raising , after the consonant (at the CV transition). We divide the analysis as we have previously done, by place of articulation and syllable position.

One further prediction of this dissertation was that greater variability in production should correspond to lower perception ability. To test this we also examined the standard deviation for each subject's F2 productions for each syllable position and place of articulation. The metric examined is the difference between average standard deviation at each consonantal transition (VC and CV) between palatalized and plain consonants. A higher value for this metric indicates a higher average standard deviation for the palatalized consonant. We predict that subjects who have higher values in standard deviation difference will also tend to have lower perception values. The perception values investigated here are the assimilation values as well as discrimination values. For each place of articulation, we begin with discussion on the mean F2 differences, and follow up with brief discussions on standard deviation.

9.2.1. Onset Coronals

For each of the analyses below, we first describe mean F2 differences, and we then analyze briefly the differences in standard deviation, or variability. As we have done throughout this dissertation, discussions will occur separately for each place of articulation and syllable position. We begin the discussion with onset coronals. We approach this discussion by first identifying sub-groups of subjects within each group that exhibit similar patterns in perception. For many of the plots in the following sections we again display information about individual subjects, where for subjects belonging to a particular group we have a letter followed by a number. For example, subjects E1, E2 and so on are each individual subjects in the naïve group. The same convention applies for the learner group (L1, L2 and so on) and the native Russian group (R1, R2 and so on).

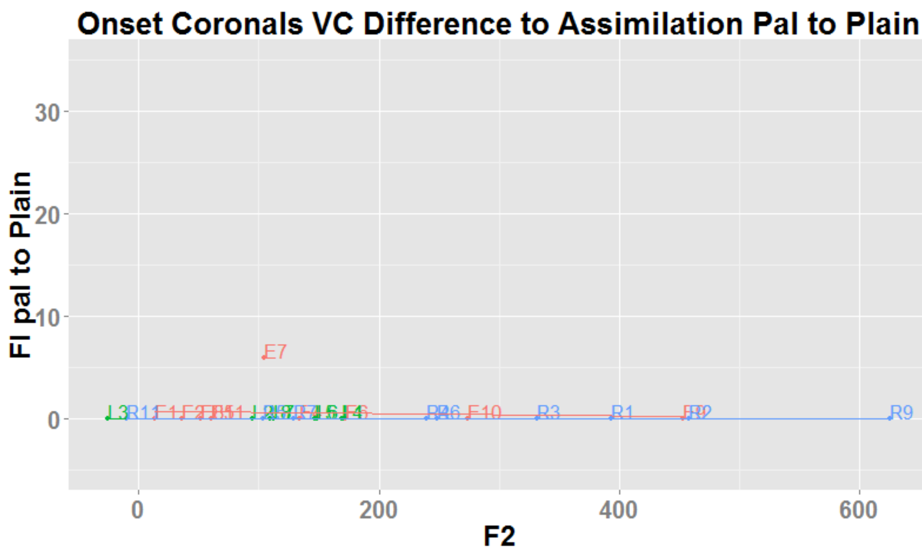

Figure 9.1: Onset coronals comparison of the fit index assimilations of the palatalized consonant to the plain category and the mean difference in F2 between palatalized consonant and plain consonant at the VC transition by subject.

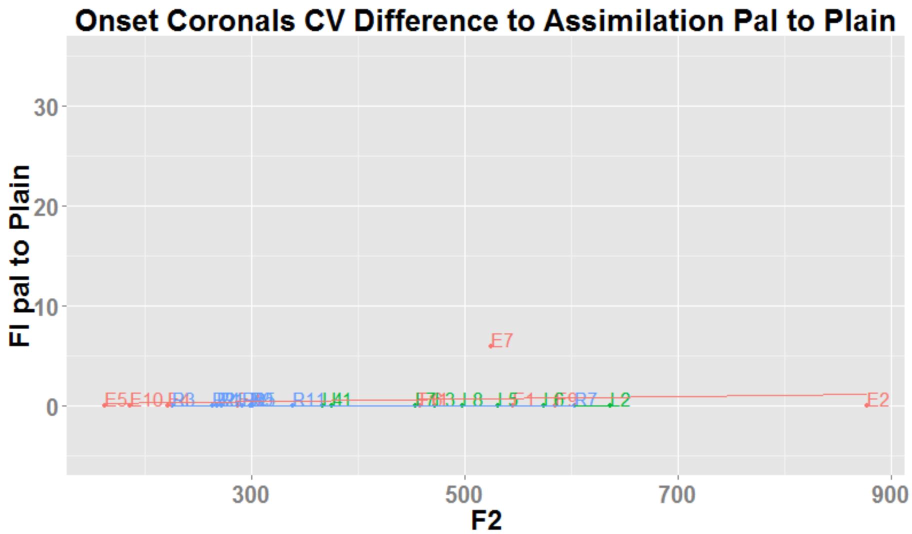

Figure 9.2: Onset coronals comparison of the fit index assimilations of the palatalized consonant to the plain category and the mean difference in F2 between palatalized consonant and plain consonant at the CV transition by subject.

Figure 9.1 and Figure 9.2 above demonstrate the relationship between F2 difference between the palatalized consonant and the plain consonant at the VC (top) and CV (bottom) transitions against the fit index assimilation of the palatalized onset coronal to the plain category. Notably, only subject E7 had any noticeable assimilation tendency for the palatalized consonant to the plain category. As such, the remaining subjects are clustered around zero.

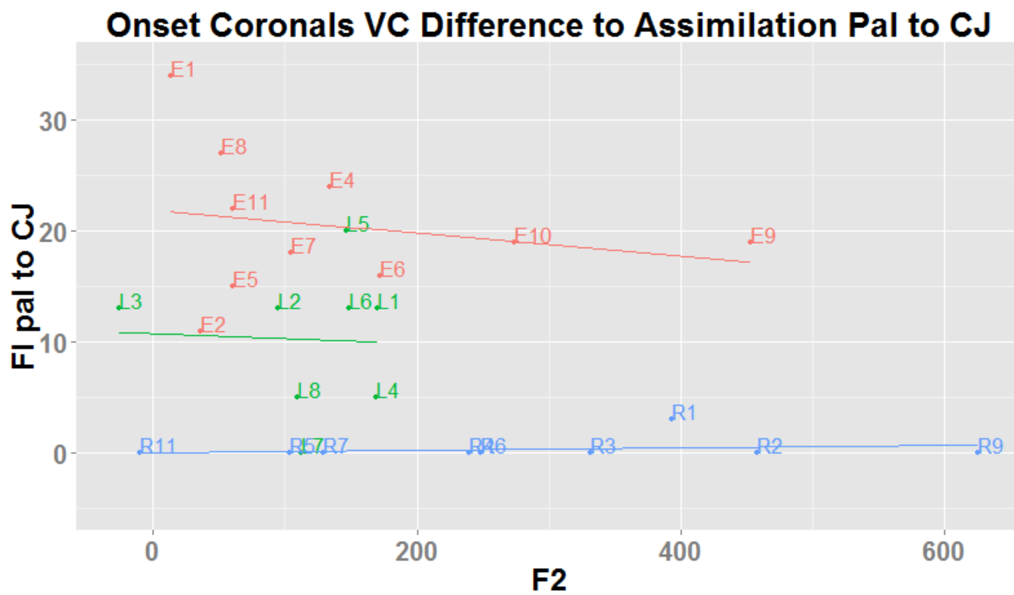

Figure 9.3: Onset coronals comparison of the fit index assimilations of the palatalized consonant to the consonant+glide category and the mean difference in F2 between palatalized consonant and plain consonant at the VC transition by subject.

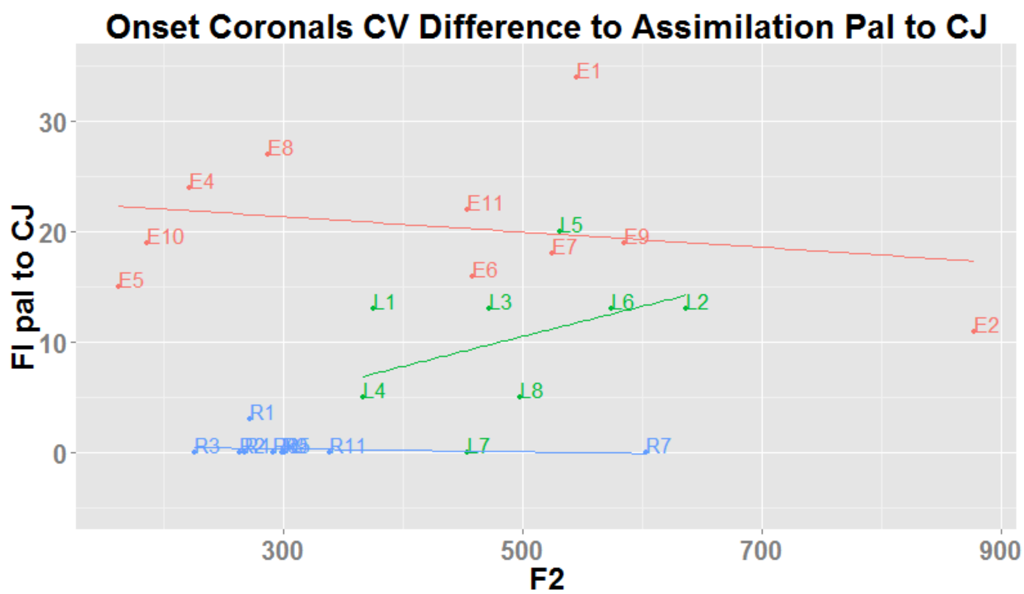

Figure 9.4: Onset coronals comparison of the fit index assimilations of the palatalized consonant to the consonant+glide category and the mean difference in F2 between palatalized consonant and plain consonant at the CV transition by subject.

Figure 9.3 and Figure 9.4 demonstrate the relationship between F2 difference between the palatalized consonant and the plain consonant at the VC and CV transitions against the fit index assimilation of the palatalized onset coronal to the consonant+glide category. For the VC transition, we see a slight negative association[20] for both English groups where the higher the fit index value for the palatalized consonant to the consonant+glide category, the lower the mean difference in F2 would be between the palatalized and plain consonants at the VC transition.

[20] When we discuss “associations” in this data they are not statistical analyses. The relatively small number of subjects unfortunately was not enough to run rigorous statistical analysis at the subject level. For future work again a more rigorous analysis of this data is welcome.

For the CV transition, we see a similar trend for the naïve group, where a higher fit index value to the consonant+glide category is associated with a lower difference in F2. We see the opposite trend for the learner group. Also of interest is that subject E9 had high F2 differences for both the VC and CV transitions, indicating a higher TB gesture for palatalized onset coronals of greater duration or magnitude that begins sooner and ends later and is also timed more simultaneously. The trend seen for the CV transition does not necessarily fit our predictions. We predicted a positive association between a high difference in F2 between palatalized and plain consonants and an assimilation fit index to the CJ category.

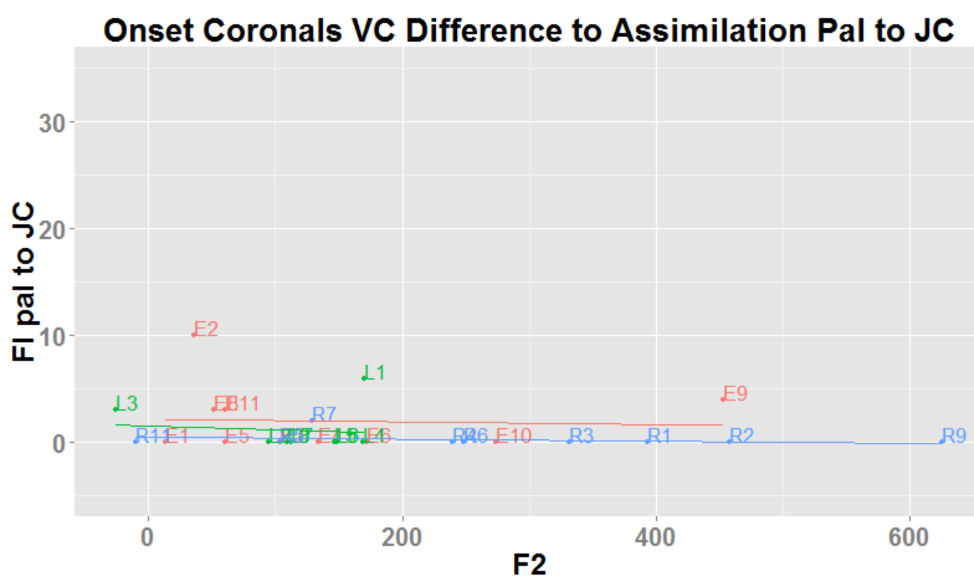

Figure 9.5: Onset coronals comparison of the fit index assimilations of the palatalized consonant to the glide+consonant category and the mean difference in F2 between palatalized consonant and plain consonant at the VC transition by subject.

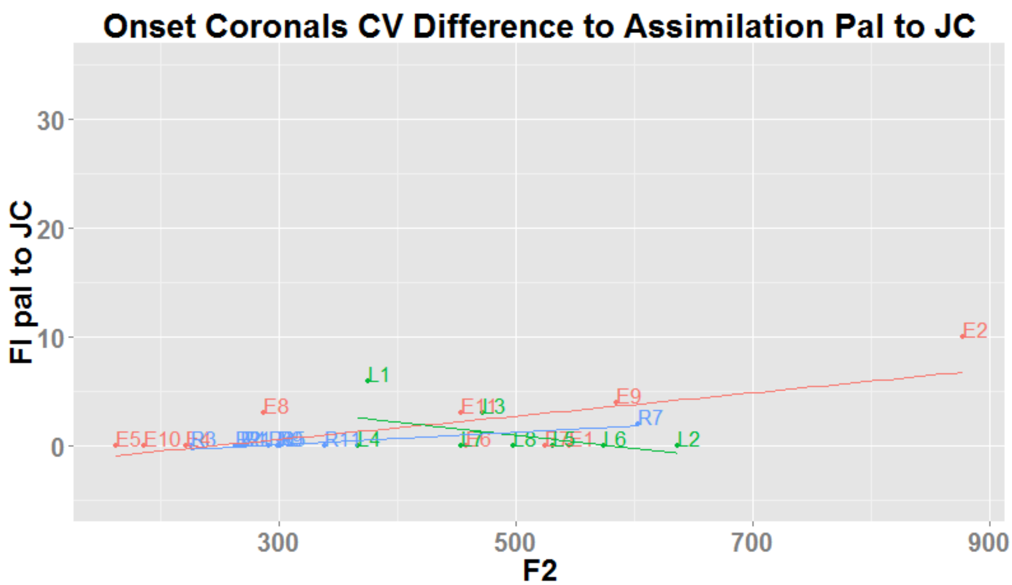

Figure 9.6: Onset coronals comparison of the fit index assimilations of the palatalized consonant to the glide+consonant category and the mean difference in F2 between palatalized consonant and plain consonant at the CV transition by subject.

Figure 9.5 and Figure 9.6 above demonstrate the relationship between F2 difference between the palatalized consonant and the plain consonant at the VC and CV transitions against the fit index assimilation of the palatalized onset coronal to the glide+consonant category. Only one subject, E2, had any noticeable assimilation of the palatalized onset coronal to the glide+consonant category. This subject had a very large difference in F2 between palatalized and plain consonants at the CV transition. This does not necessarily follow our predictions, since a tendency to assimilate the consonant to the glide+consonant category is predicted to be associated with a high increase in F2 at the VC transition for the palatalized consonants.

9.2.2. Coda Coronals

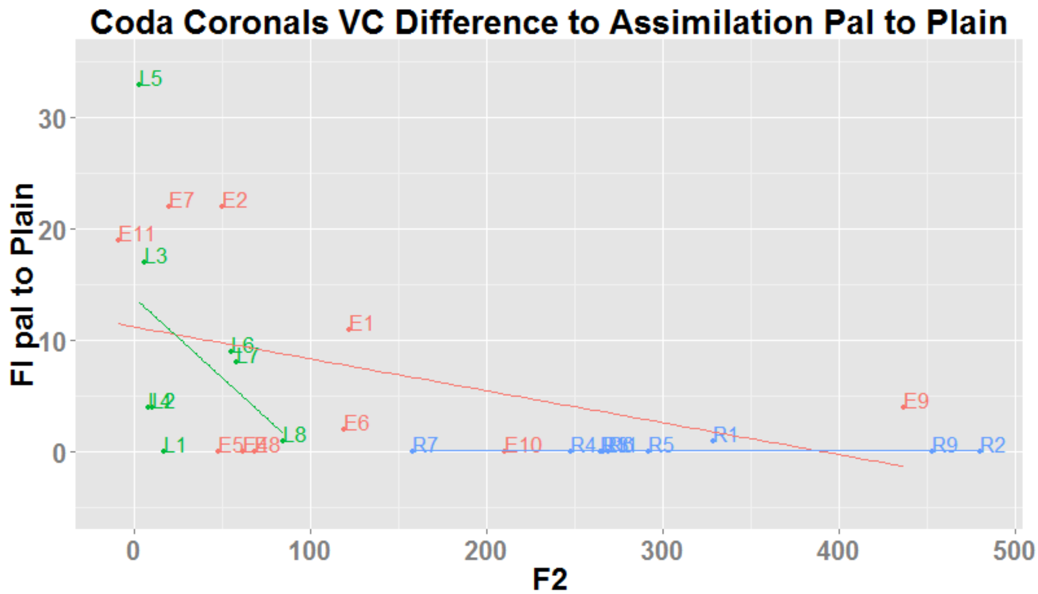

Figure 9.7: Coda coronals comparison of the fit index assimilations of the palatalized consonant to the plain category and the mean difference in F2 between palatalized consonant and plain consonant at the VC transition by subject.

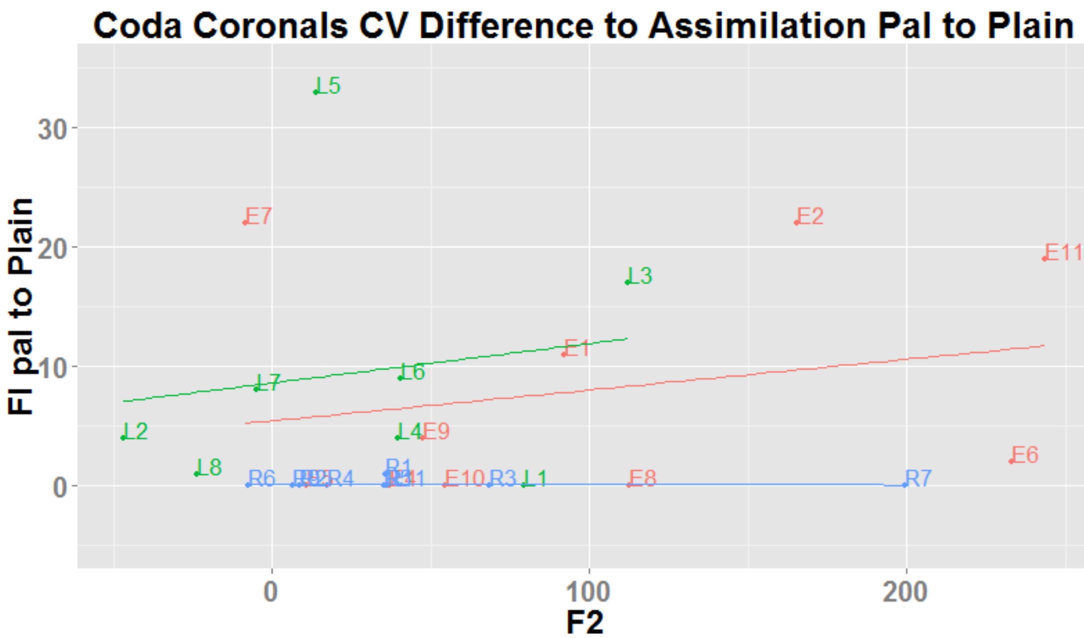

Figure 9.8: Coda coronals comparison of the fit index assimilations of the palatalized consonant to the plain category and the mean difference in F2 between palatalized consonant and plain consonant at the CV transition by subject.

Figure 9.7 and Figure 9.8 above show the relationship between the mean F2 difference between the palatalized consonant and the plain consonant at the VC and CV transitions against the fit index assimilation of the palatalized coda coronal to the plain category. Several subjects (including E7, who also showed this pattern it for the onsets) demonstrated assimilation tendencies for the palatalized consonant to the plain category. Subjects E11, E2 and E7 fall into this category. So do subjects L3 and L5 from the learner group. We will further investigate the learner group in more detail later in this chapter. In production these subjects also all exhibited relatively low difference between mean F2 at the VC transition between the palatalized and plain consonants. This fits in line with our prediction. This means that subjects who produced the plain and palatalized consonants similarly tended to perceive the coda coronal palatalized consonant as a plain consonant.

We do see, however that not all of these same subjects who had a low difference in F2 at the VC transition and high FI to the plain category in perception demonstrated a low F2 difference for the CV transition. A few of these subjects, (E2, E11, L3) had differences greater than 100 Hz at the CV transition. In general though, there is a slightly positive relationship between the assimilation of the palatalized consonant to the plain category and a difference in F2 between the palatalized and plain consonants at the CV transition for coda coronals.

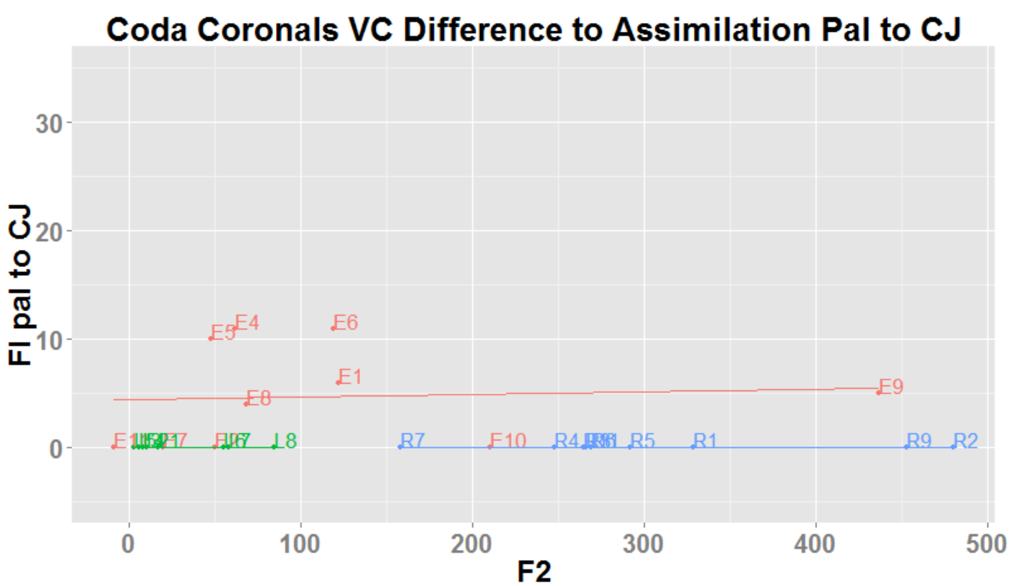

Figure 9.9: Coda coronals comparison of the fit index assimilations of the palatalized consonant to the consonant+glide category and the mean difference in F2 between palatalized consonant and plain consonant at the VC transition by subject.

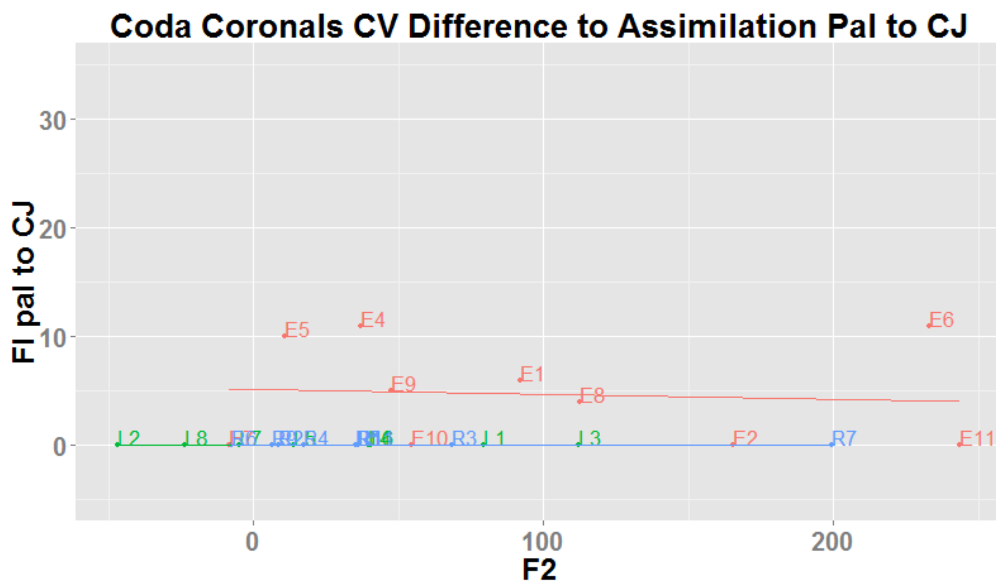

Figure 9.10: Coda coronals comparison of the fit index assimilations of the palatalized consonant to the consonant+glide category and the mean difference in F2 between palatalized consonant and plain consonant at the CV transition by subject.

Figure 9.9 and Figure 9.10 show the relationship between F2 difference between the palatalized consonant and the plain consonant at the VC and CV transitions against the fit index assimilation of the palatalized coda coronal to the consonant+glide category. There are only a few subjects from Group E who had any FI values of 10 or above for the coda coronal consonant to the consonant+glide category in perception. These four subjects all demonstrated relatively low mean F2 differences between plain and palatalized coda coronal at the VC transition. All subjects had relatively low difference between palatalized and plain consonants in production at the CV transition for coda coronals.

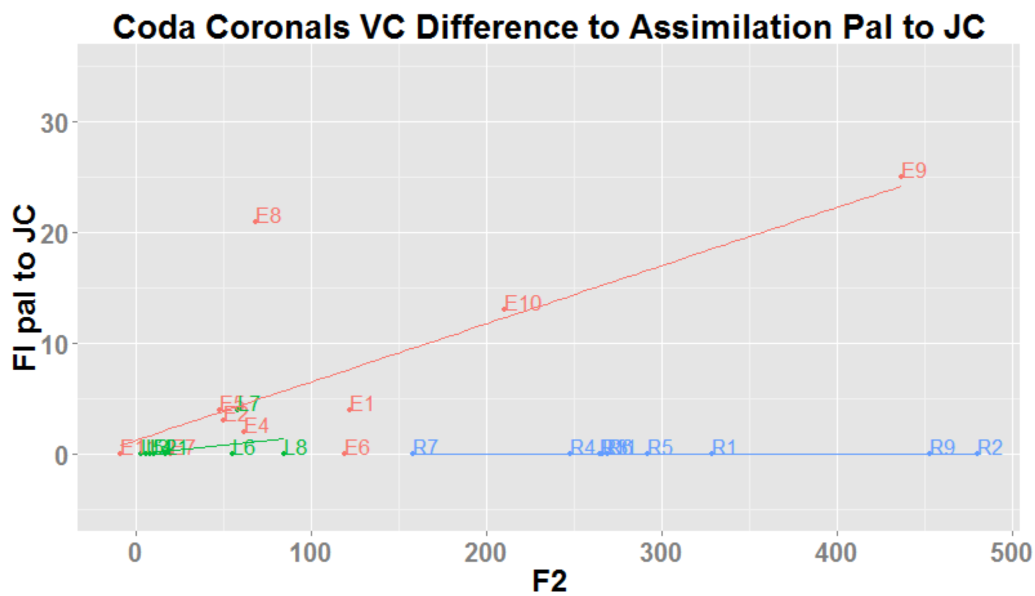

Figure 9.11: Coda coronals comparison of the fit index assimilations of the palatalized consonant to the glid+consonant category and the mean difference in F2 between palatalized consonant and plain consonant at the VC transition by subject.

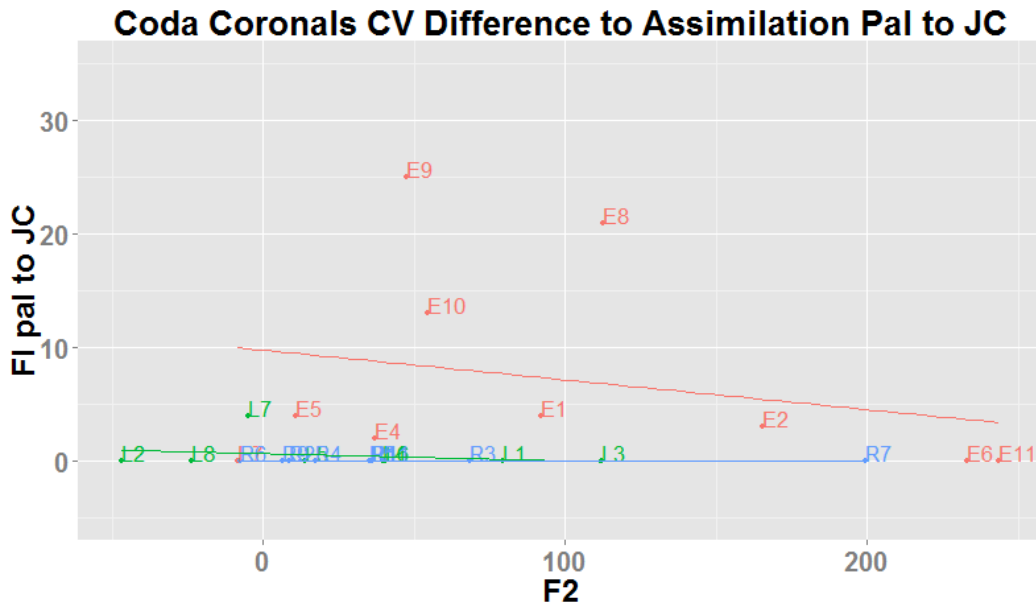

Figure 9.12: Coda coronals comparison of the fit index assimilations of the palatalized consonant to the glid+consonant category and the mean difference in F2 between palatalized consonant and plain consonant at the CV transition by subject.

For the relation between production and perception for the assimilation of the palatalized coda coronal to the glide+consonant category we see a relatively strong positive relationship between the fit index value and the difference in mean F2 between the palatalized and plain consonant at the VC transition for the naïve group (Figure 9.11). This means that there seems to be a positive relationship between perceiving this consonant as a sequence of glide+consonant and producing a higher F2 at the VC transition for palatalized consonants than for plain consonants. Subjects E8, E9 and E10 had relatively high fit index values for the coda coronal to the JC category. Subjects E9 and E10 had mean differences in F2 between palatalized and plain VC transitions of greater than 200 Hz, with E9 having a value of greater than 400 Hz. Subject, E8, however did not fit the trend and had a low difference in production. Therefore, for coda coronals, and for the relation between perception and production when subjects perceived the coda palatalized consonant as a sequence of glide+consonant, most subjects (except for Subject E8) performed in a manner that fit the predictions.

At this point we take a short detour to go into more depth on two specific subjects from the naïve group to illustrate how subjects are varying between each other and to also add commentary on perception of timing, and production of timing. Subject E1 had a tendency to assimilate the palatalized onset coronal consonant to the CJ (consonant+glide) category in perception, while subject E9 had a strong tendency to assimilate the palatalized coda coronal consonant to the JC category (glide+consonant). In terms of perception strategies, E1 and E9 differ in that we can say that E1 tends to perceive the palatalized consonant as a sequence of consonant+glide, while E9 tends to perceive (at least in the coda) the palatalized consonant as a sequence of glide+consonant. In examining their productions more carefully, we see that they each employ different production strategies as well.

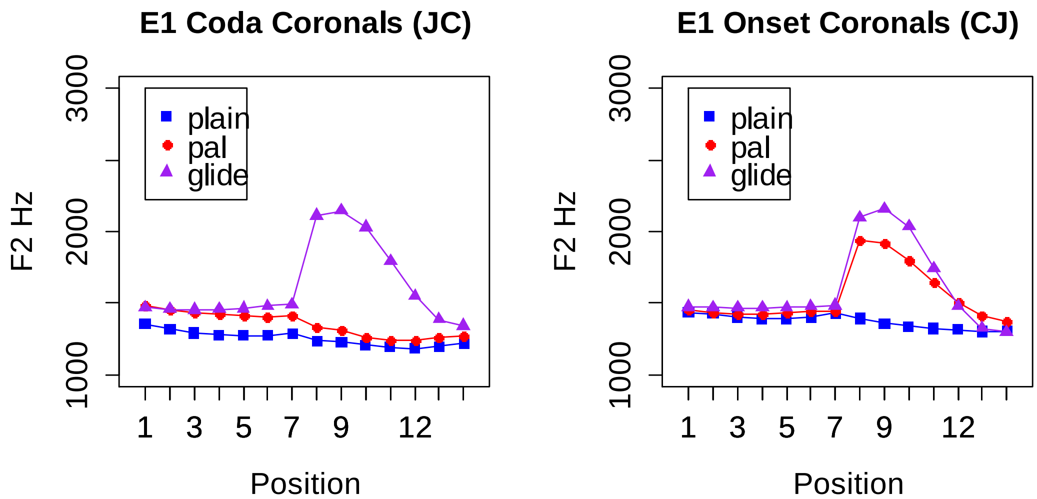

Figure 9.13. For subject E1 who tended to perceive the palatal gesture following the consonant, the plot on the left shows the mean F2 values at each of the points in the entire sequence (preceding vowel, transitions, following vowel) for the coda coronal, the plot on the right shows the same for the onset coronal.

Figure 9.13 shows two plots generated for Subject E1. On the left we see this subject's mean F2 measurements for each of the three sequence types (plain, palatalized, palatalized+glide) for the coda position, on the right the same for the onset. Figure 9.14 shows the same measurements for Subject E9. What is striking is that mean F2 values are much higher for Subject E9 for positions preceding the consonant (Position 7 = Consonant) for the palatalized productions (red circles). This is not the case for Subject E1. What this shows is that Subject E9 not only perceived the TB gesture as preceding the consonantal gesture, but produced the tokens in the same way. In this way, at least some subjects are perceiving and producing timing of gestures in the same way.[21]

[21] Of course, it is possible that these two subjects also heard different things, or assimilated each of these sequences to different L1 categories. For example, perhaps E9 actually assimilated the

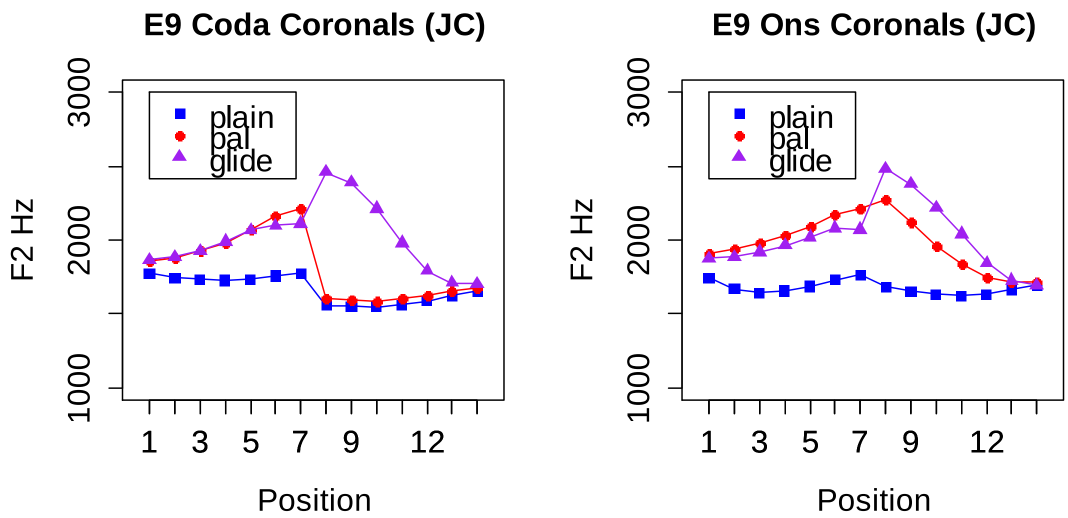

Figure 9.14. For subject E9, who tended to perceive a palatal gesture preceding the consonant, the plot on the left shows the mean F2 values at each of the points in the entire sequence (preceding vowel, transitions, following vowel) for the coda coronal, the plot on the right shows the same for the onset coronal.

9.2.3. Coronals Standard Deviation

---

palatalized codas to a sequence of vowel dipthong + consonant, and not glide + consonant and the perhaps they consciously produced them in the same way. Again, further study is warranted to tease out some of these possibilities.

NAÏVE GROUP ONSET CORONALS						
Subject	SD diff pal and plain (VC)	SD diff pal and plain (CV)	Perception plain-pal (axb)	Perception pal-glide (axb)	Perception FI Pal to Plain	Perception FI Pal to CJ
E1	34	8	0.88	0.79	0	34
E10	73	84	1	1	0	19
E11	54	200	0.92	0.71	0	22
E2	71	105	1	1	0	11
E4	21	111	1	0.96	0	24
E5	51	141	1	1	0	15
E6	22	78	1	0.96	0	16
E7	-85	58	0.92	0.88	6	18
E8	-11	117	1	0.96	0	27
E9	-175	-29	1	1	0	19

Table 9.2: The table above presents for each subject in the naïve group for onset coronals the mean difference in standard deviation between palatalized and plain consonants at the two transitions, and for perception the percent correct discrimination from the AXB tests for both plain-palatalized and palatalized-glide contrasts, as well as the fit index assimilation of the palatalized consonant to the plain category and the consonant+glide category.

We now discuss the variability in production for the naïve subjects for coronals and how this variability relates to perception. For onset coronals, we see a lot of variation as seen in Table 9.2. For example, subject E11 had the lowest percent correct discrimination of the palatalized-consonant+glide contrast (fourth column in Table 9.2), and also had by far the highest difference in standard deviation at the CV transition. This means that this subject had a much higher standard deviation for the CV transition for palatalized consonants than for plain consonants and also had a lower discrimination rate for palatalized-glide contrast. However, subject E1 also had a relatively low discrimination rate for this contrast but had a negligible difference in standard deviation.

NAÏVE GROUP CODA CORONALS						
Subject	**SD diff pal and plain (VC)**	**SD diff pal and plain (CV)**	**Perception plain-pal (axb)**	**Perception pal-glide (axb)**	**Perception FI Pal to Plain**	**Perception FI Pal to JC**
E1	22	-19	0.67	1	11	4
E10	43	10	0.96	1	0	13
E11	-2	294	0.63	0.92	19	0
E2	36	225	0.88	1	22	3
E4	-2	21	0.92	1	0	2
E5	80	44	0.75	1	0	4
E6	-8	132	0.96	0.96	2	0
E7	-80	-8	0.88	1	22	0
E8	3	99	0.83	1	0	21
E9	-47	18	0.96	1	4	25

Table 9.3: The table above presents for each subject in the naïve group for coda coronals the mean difference in standard deviation between palatalized and plain consonants at the two transitions, and for perception the percent correct discrimination from the AXB tests for both plain-palatalized and palatalized-glide contrasts, as well as the fit index assimilation of the palatalized consonant to the plain category and the glide+consonant category.

In general we see that most of the naïve subjects (group average of 87) had high difference in standard deviation values, indicating a higher degree of variation in production for palatalized consonants, which seems in line with the prediction, as most subjects exhibited less than perfect discrimination of at least the palatalized-glide contrast. For the coda coronals, we focus on different metrics, namely the VC transition (see Table 9.3). In general the difference in standard deviation was lower than for onset coronals (5 versus 87). Only one subject, E5 had relatively high standard deviation differences (80). This subject also exhibited relatively low discrimination percent correct (0.75) for the plain-palatalized contrast. However, subject E11 exhibited the lowest percent correct discrimination value (0.63) and did not have a high average standard deviation difference. It is worth noting that Subject E11 also exhibited a relatively high difference in F2 between palatalized and plain consonants at the VC transition (see previous discussion), thus we still do not see consistent associations between poor perception and production metrics for all subjects.

9.2.4. Onset Labials

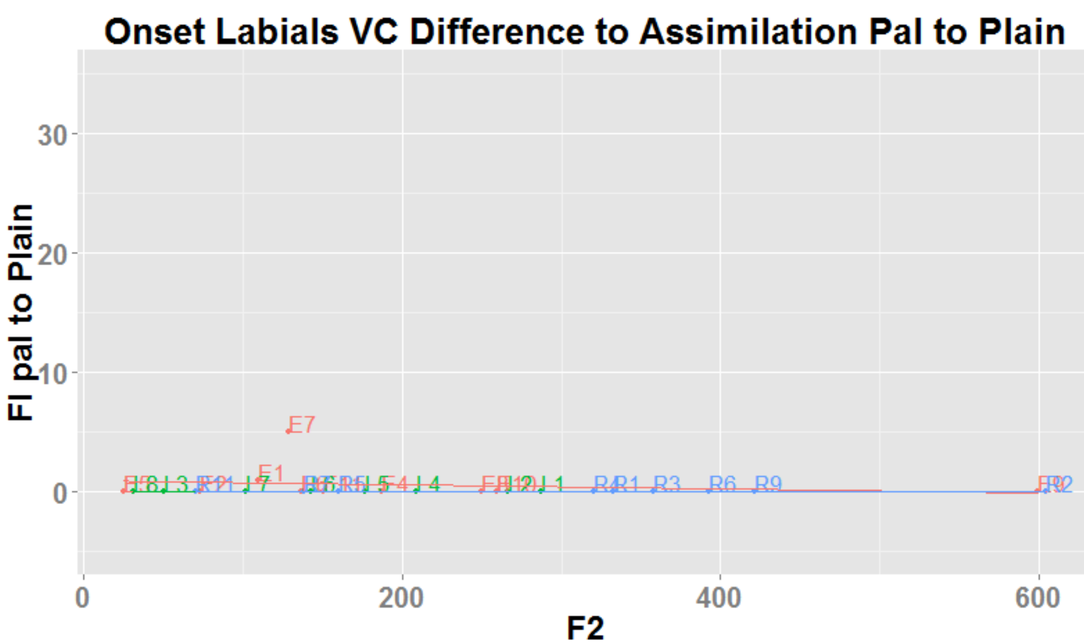

Figure 9.15: Onset labials comparison of the fit index assimilations of the palatalized consonant to the plain category and the mean difference in F2 between palatalized consonant and plain consonant at the VC transition by subject.

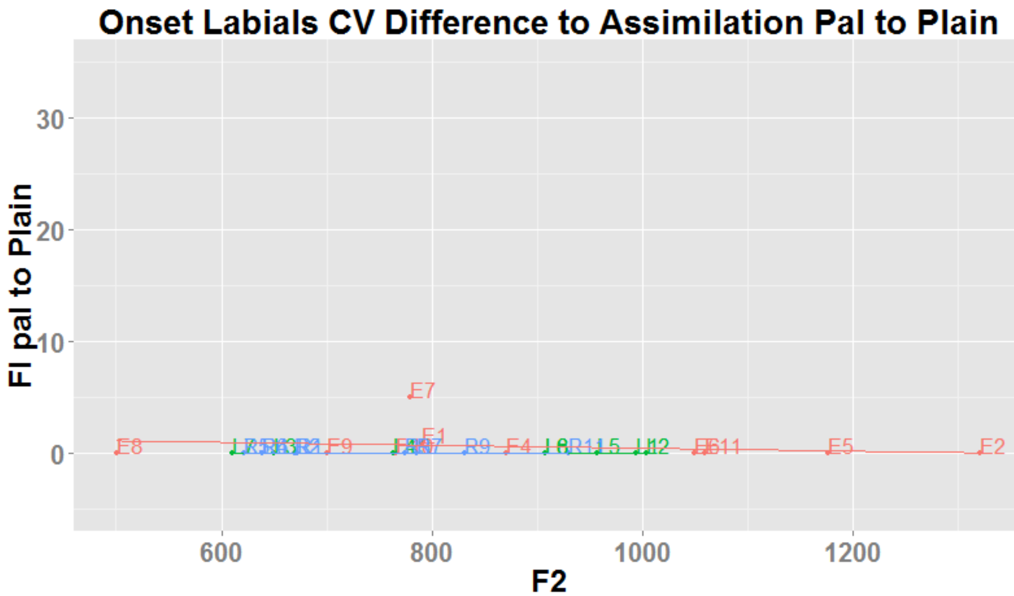

Figure 9.16: Onset labials comparison of the fit index assimilations of the palatalized consonant to the plain category and the mean difference in F2 between palatalized consonant and plain consonant at the CV transition by subject.

We now move on to discuss the same perception and production relations for labials as we have just done for coronals. Figure 9.15 and Figure 9.16 above show the relationship between F2 difference between the palatalized consonant and the plain consonant at the VC and CV transitions and the fit index assimilation of the palatalized onset labial to the plain category for the labials. As for the coronals, only subject E7 had any noticeable assimilation tendency in perception for the palatalized consonant to the plain category. Also of interest, subject E9 demonstrated high F2 differences for both VC and CV transitions, indicating a higher TB gesture for palatalized onset coronals of greater duration or magnitude that begins sooner and ends later and is also timed more simultaneously.

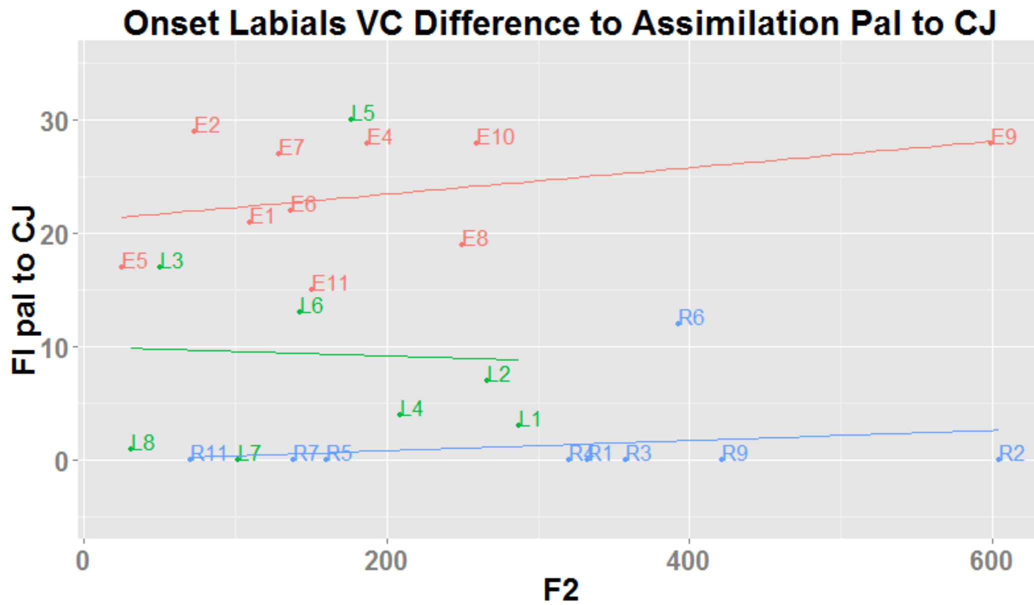

Figure 9.17: Onset labials comparison of the fit index assimilations of the palatalized consonant to the consonant+glide category and the mean difference in F2 between palatalized consonant and plain consonant at the VC transition by subject.

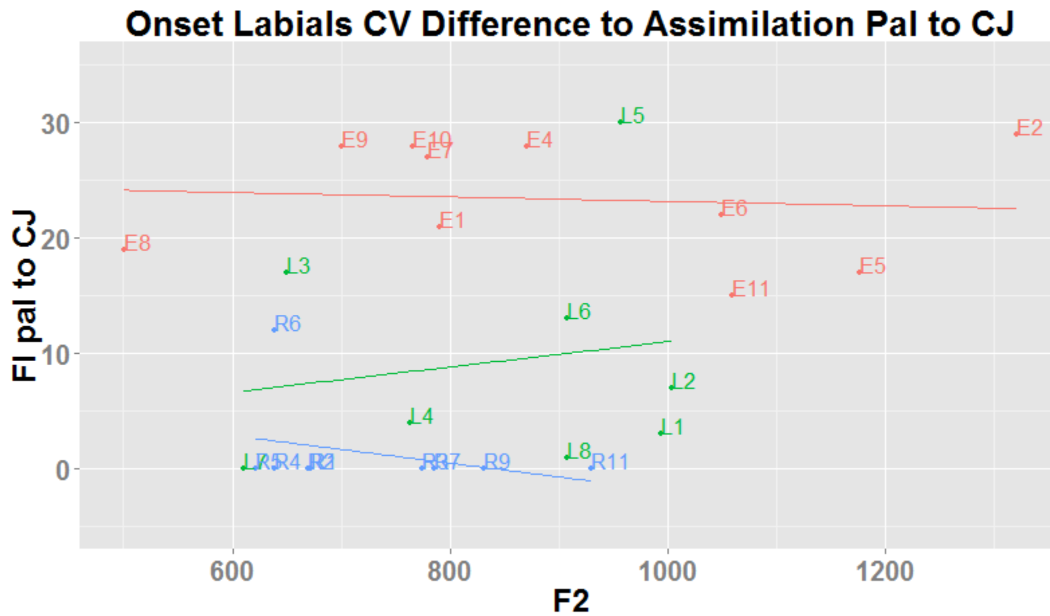

Figure 9.18: Onset labials comparison of the fit index assimilations of the palatalized consonant to the consonant+glide category and the mean difference in F2 between palatalized consonant and plain consonant at the CV transition by subject.

We now examine the relationship between the F2 difference between the palatalized consonant and the plain consonant at the VC and CV transitions against the fit index assimilation of the palatalized onset labial to the consonant+glide category. For the VC transition we see a slight positive (opposite direction from coronals) association for the naïve group and seemingly no relationship between the two variables for the learner group, where the higher the fit index value for the palatalized consonant to the consonant+glide category, the higher the mean difference in F2 would be between the palatalized and plain consonants at the VC transition. This is observed in Figure 9.17. This means that in general if a subject had a higher difference in F2 at the VC transition for palatalized onset labials than plain, they also tended to prefer assimilations to the consonant+glide category.

For the CV transition in Figure 9.18, we see a similar trend for the naïve group as we did for the coronals, where a higher fit index value to the consonant+glide category is associated with a lower difference in F2. We see the opposite trend for the learner group in this case, as was the case for the coronals. This means that in general, our prediction does not hold for the relationship between perception to the consonant+glide category and production of the palatalized consonant.

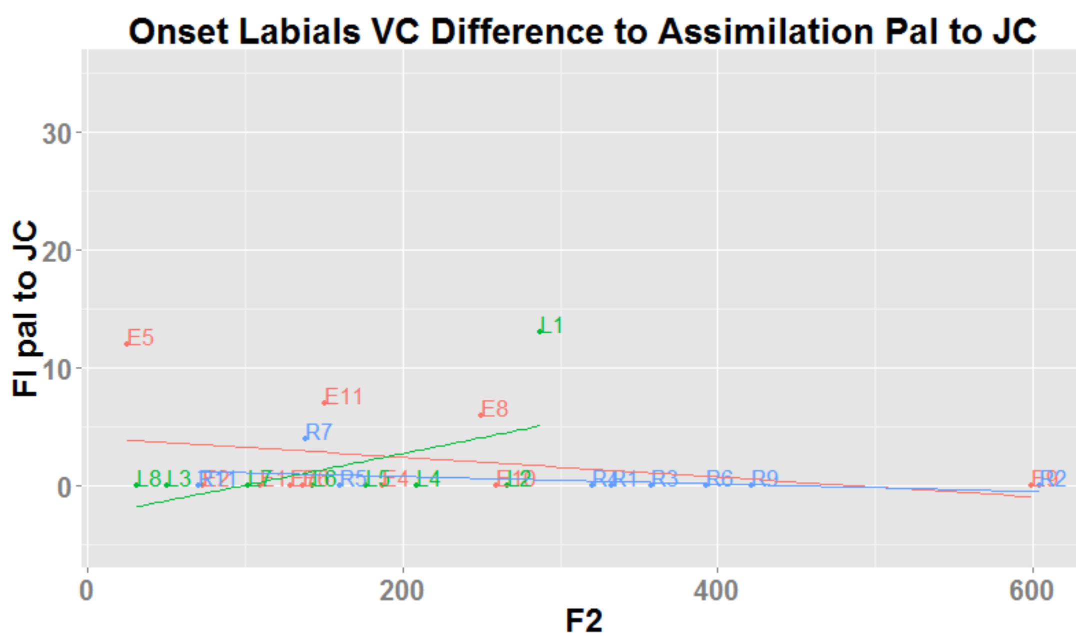

Figure 9.19: Onset labials comparison of the fit index assimilations of the palatalized consonant to the glide+consonant category and the mean difference in F2 between palatalized consonant and plain consonant at the VC transition by subject.

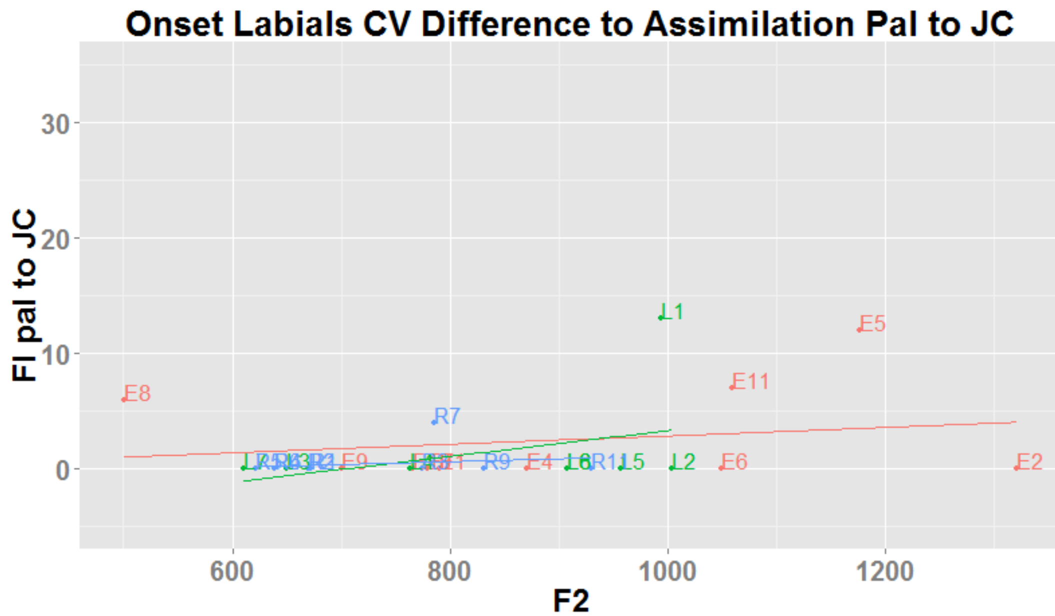

Figure 9.20: Onset labials comparison of the fit index assimilations of the palatalized consonant to the glide+consonant category and the mean difference in F2 between palatalized consonant and plain consonant at the CV transition by subject.

Figure 9.19 and Figure 9.20 above show the relationship between F2 difference between the palatalized consonant and the plain consonant at the VC and CV transitions against the fit index assimilation of the palatalized onset labial to the glide+consonant category. Subjects E5 and L1 had the most noticeable assimilation to this category for onset labials. Both of these subjects had high difference in F2 between the palatalized and plain consonants at the CV transition, but they differed in production at the VC transition.

9.2.5. Coda Labials

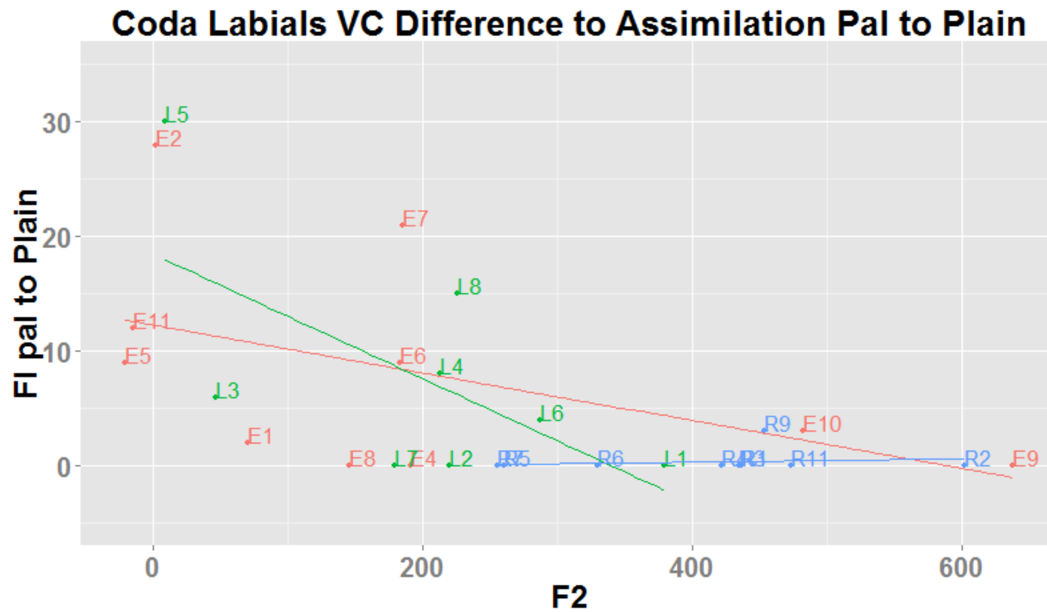

Figure 9.21: Coda labials comparison of the fit index assimilations of the palatalized consonant to the plain category and the mean difference in F2 between palatalized consonant and plain consonant at the VC transition by subject.

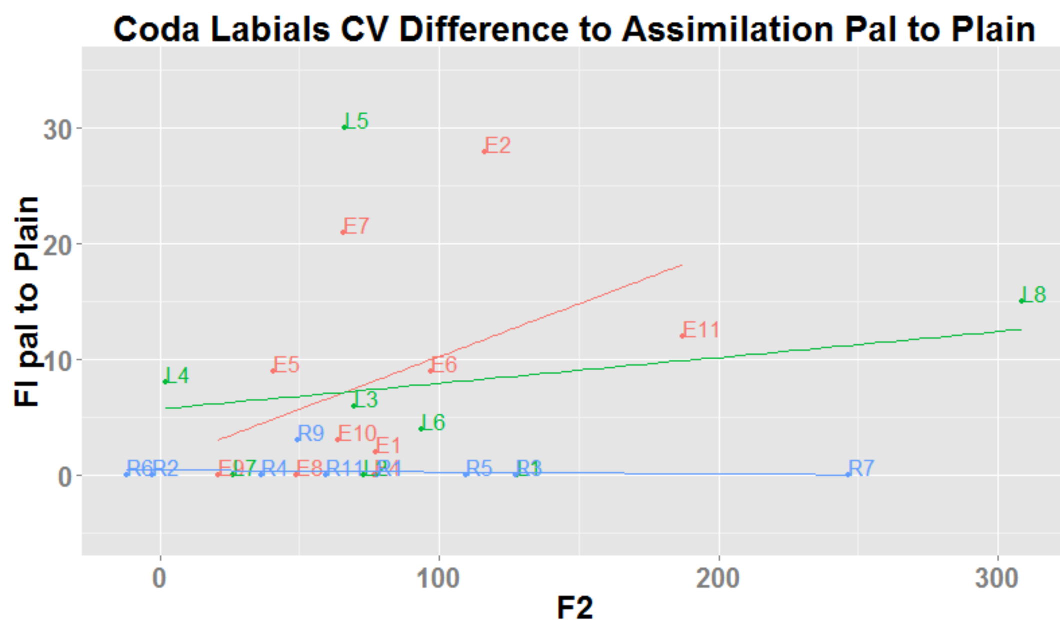

Figure 9.22: Coda labials comparison of the fit index assimilations of the palatalized consonant to the plain category and the mean difference in F2 between palatalized consonant and plain consonant at the Cv transition by subject.

We now move on to the last of the four comparison types, that for coda labials. Above we see the relationship between the mean F2 difference between the palatalized consonant and the plain consonant at the VC and CV transitions against the fit index assimilation of the palatalized coda labial to the plain category. There is a negative relationship between difference in F2 between palatalized and plain consonants at the VC transition and assimilation fit index to the coda labial to the plain category. That is, the higher the fit index to the plain category, the lower the tendency for a difference between palatalized and plain consonants. This is an expected result, as this indicates that at the VC transition the subjects who are exhibiting less distinction between plain and palatalized consonants are also perceiving these consonants most like plain consonants.

Several subjects (including E2 and E7, who also exhibited this tendency for coronals) had assimilation tendencies for the palatalized consonant to the plain category. L5 also falls into this category for both labials and coronals. These subjects also all exhibited relatively low difference between mean F2 at the VC transition between the palatalized and plain consonants.

We do see, however that not all of these same subjects who had a low difference in F2 at the VC transition and high FI to the plain category in perception as having also low F2 difference for the CV transition. A few of these subjects, (E2, E11 who also had it for coronals) showed differences greater than 100 Hz. This is still a relatively small difference. Subject L8 exhibited a difference of greater than 300 Hz. In general though, there is a slightly positive relationship between the assimilation of the palatalized consonant to the plain category and a difference in F2 between the palatalized and plain consonants at the CV transition for coda coronals. This trend was the same for the coronals as well.

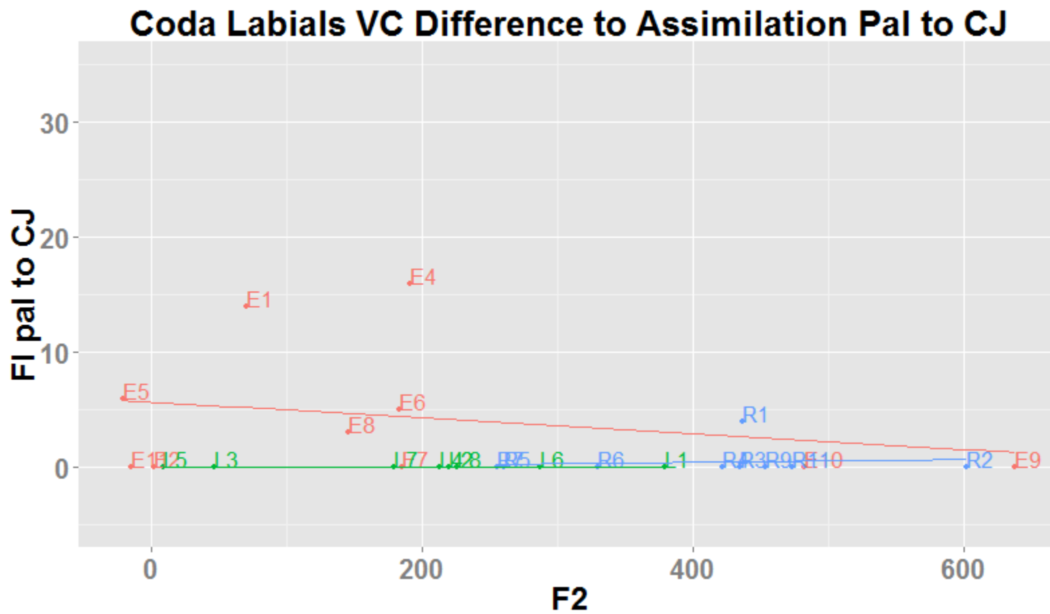

Figure 9.23: Coda labials comparison of the fit index assimilations of the palatalized consonant to the consonant+glide category and the mean difference in F2 between palatalized consonant and plain consonant at the VC transition by subject.

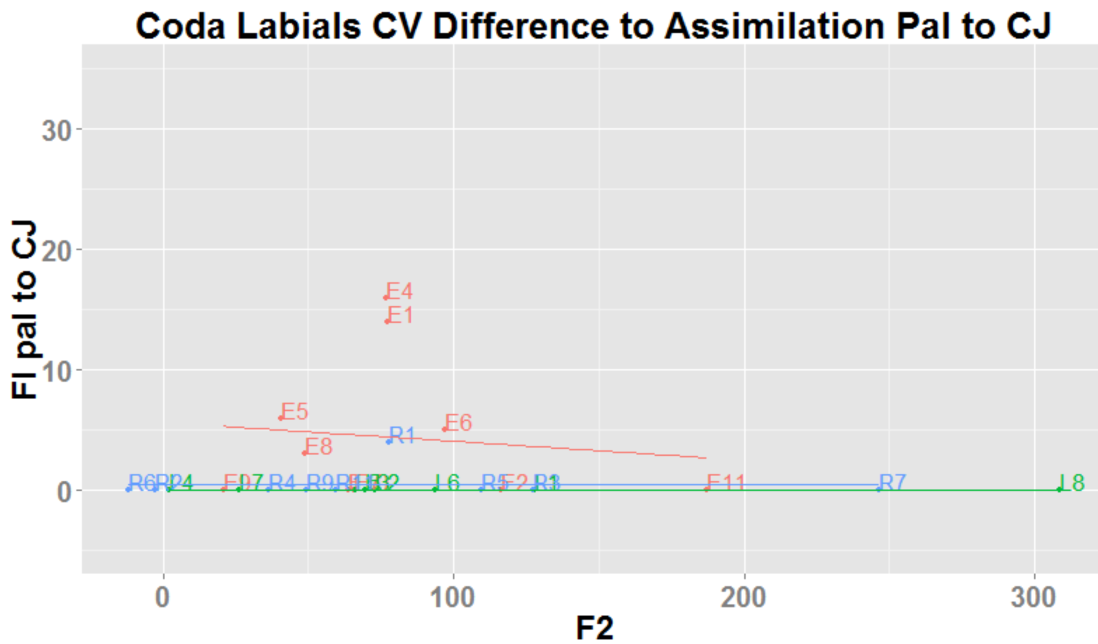

Figure 9.24: Coda labials comparison of the fit index assimilations of the palatalized consonant to the consonant+glide category and the mean difference in F2 between palatalized consonant and plain consonant at the CV transition by subject.

We move on to examine the relationship between the F2 difference between the palatalized consonant and the plain consonant at the VC and CV transitions against the fit index assimilation of the palatalized coda labial to the consonant+glide category. There are only a few subjects from the naïve group who had any FI values of 10 or above for the coda labial consonant to the consonant+glide category as was the case for the coronals. These subjects all had relatively low mean F2 differences between plain and palatalized coda coronal at the VC transition as seen in Figure 9.23. All subjects had a relatively low difference between palatalized and plain consonants in production at the CV transition for coda labials much like for the coronals as seen in Figure 9.24. We cannot conclude enough from this data to assert whether the prediction holds for this specific relationship.

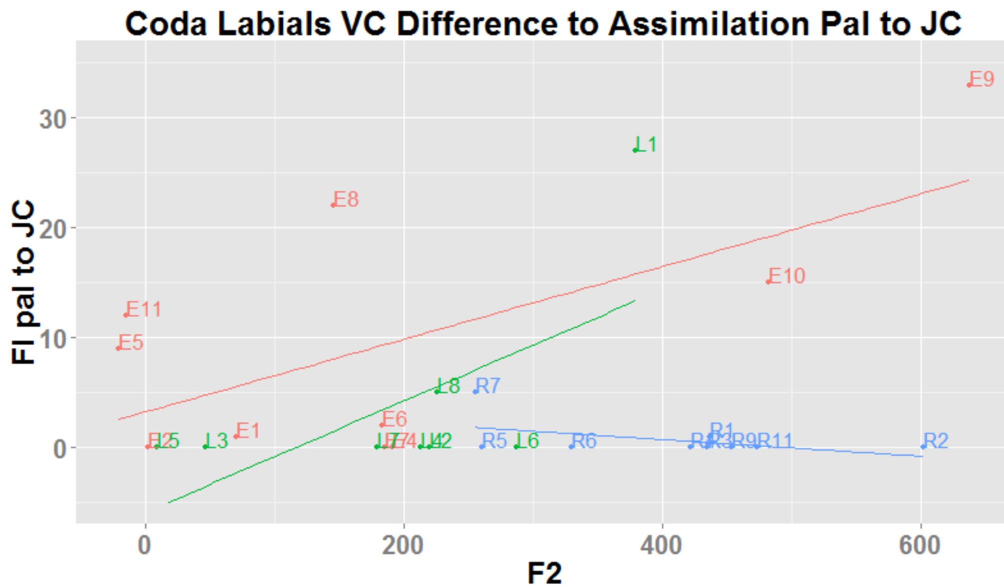

Figure 9.25: Coda labials comparison of the fit index assimilations of the palatalized consonant to the glide+consonant category and the mean difference in F2 between palatalized consonant and plain consonant at the VC transition by subject.

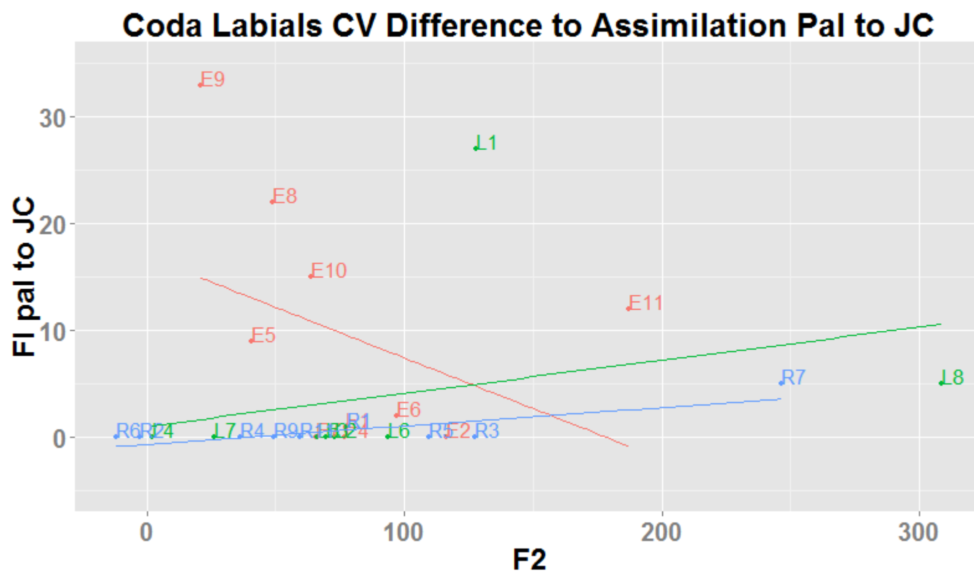

Figure 9.26: Coda labials comparison of the fit index assimilations of the palatalized consonant to the glide+consonant category and the mean difference in F2 between palatalized consonant and plain consonant at the CV transition by subject.

Figure 9.25 and Figure 9.26 above show the relationship between the F2 difference between the palatalized consonant and the plain consonant at the VC and CV transitions against the fit index assimilation of the palatalized coda labial to the glide+consonant category. For the relation between production and perception for the assimilation of the palatalized coda labial to the glide+consonant category we again see a relatively strong positive relationship between the fit index value and the difference in mean F2 between the palatalized and plain consonant at the VC transition for the naïve group. We saw this same trend for the coda coronal as well. We also see a similar trend for the learner group which is due to the behavior of one subject, L1. This subject had a high fit index of the palatalized coda labial to the glide+consonant category but did not show the same pattern in perception for the coda coronal.

In general, for most subjects the same interpretation that was attributed for the coronals can be provided here, namely that there is a trend for a positive relationship between perceiving this consonant as a sequence of glide+consonant and producing a higher F2 at the VC transition for palatalized consonants than for plain consonants. Subjects E8, E9 and E10 had relatively high fit index values for the coda coronal to the JC category as they did for the coronal consonants. Subjects E9 and E10 had mean differences in F2 between palatalized and plain VC transitions of greater than 400 Hz, with E9 having a value of greater than 600. Subject, E8, however did not fit the trend and had a low difference in production.

9.2.6. Labials Standard Deviation

NAÏVE GROUP ONSET LABIALS						
Subject	**SD diff pal and plain (VC)**	**SD diff pal and plain (CV)**	**Perception plain-pal (axb)**	**Perception pal-glide (axb)**	**Perception FI Pal to Plain**	**Perception FI Pal to CJ**
E1	6	110	1	0.79	1	21
E10	87	113	1	1	0	28
E11	136	68	0.92	0.83	0	15
E2	-18	4	1	0.96	0	29
E4	40	104	1	0.88	0	28
E5	115	271	1	0.92	0	17
E6	97	122	0.92	0.82	0	22
E7	73	220	0.96	0.92	5	27
E8	229	241	1	1	0	19
E9	24	67	1	1	0	28

Table 9.4: The table above presents for each subject in the naïve group for onset labials the mean difference in standard deviation between palatalized and plain consonants at the two transitions, and for perception the percent correct discrimination from the AXB tests for both plain-palatalized and palatalized-glide contrasts, as well as the fit index assimilation of the palatalized consonant to the plain category and the consonant+glide category.

NAÏVE GROUP CODA LABIALS						
Subject	**SD diff pal and plain (VC)**	**SD diff pal and plain (CV)**	**Perception plain-pal (axb)**	**Perception pal-glide (axb)**	**Perception FI Pal to Plain**	**Perception FI Pal to JC**
E1	-13	-9	0.96	0.88	2	1
E10	72	-5	1	1	3	15
E11	39	360	0.63	1	12	12
E2	74	235	0.96	1	28	0
E4	72	18	0.96	0.92	0	0
E5	126	104	0.71	1	9	9
E6	225	-14	0.83	0.92	9	2
E7	122	-2	0.75	0.96	21	0
E8	125	23	1	1	0	22
E9	44	25	1	1	0	33

Table 9.5: The table above presents for each subject in the naïve group for coda labials the mean difference in standard deviation between palatalized and plain consonants at the two transitions, and for perception the percent correct discrimination from the AXB tests for both plain-palatalized and palatalized-glide contrasts, as well as the fit index assimilation of the palatalized consonant to the plain category and the glide+consonant category.

For onset labials, as for coronals we also see variation in the relation between standard deviation differences and perception. As for the onset coronals the majority of subjects in the naïve group exhibited a high average difference in standard deviation in F2 between the palatalized and plain consonant CV transitions (132). But we don't see all of these subjects performing equally well or poorly in perception. For example, subjects E1, E6 and E11 had the lowest percent correct discrimination of the palatalized-consonant+glide contrasts and exhibited various ranges of mean standard deviation differences (6, 97 and 136 respectively). In contrast, subjects E5 and E8 exhibited the highest values for mean standard deviation difference and exhibited relatively high percent discrimination, with Subject E8 exhibiting 100% success in discrimination.

For coda labials we see in general a much higher difference in standard deviation than for the coda coronals for the VC transition (89 versus 5 for coronals). Subjects such as E5, E6, E7 and E8 exhibited high average difference in standard deviation. A subset of these subjects exhibited less than perfect discrimination of the plain-palatalized contrast. However, as for the onset labials, subject E8 did not have trouble discriminating this contrast. In short, there is not a coherent relationship between mean F2 differences, standard deviation and perception.

9.3. Learners:

In this chapter we have seen some results from all groups but we have not yet examined the relationship between experience in Russian and perception and production. For the subjects in the learner group, we examine similar comparisons as the naïve subjects. However, the learner group has an additional dimension of experience with the target language that naïve subjects lack. Although all subjects in this group had at least some experience in Russian, subjects exhibited different levels of experience. Russian experience ranged from 2 to 9 years as shown in the table below. As for the naïve group we will compare production and perception. In addition we also examine differences in either perception or production accuracy based on level of experience. In Table 9.6 below, each of the learner subjects are sorted by experience in increasing order.

Subjects	Years Russian Experience	Age
L1	2	25
L3	2.5	28
L8	4	49
L5, L6, L7	5	22, 23, 29
L4	8	28
L2	9	28

Table 9.6: Learner subjects sorted by years of experience in Russian. The age of the subjects is also reported in the third column.

The data in Table 9.6 confirms that subjects vary in years of Russian experience. Subject L1 reported one year of Russian experience, subject L3 reported 2.5 years, Subject L8 reported 4 years, Subjects L5, L6 and L7 reported 5 years, Subject L4 reported 8 years and Subject L2 reported the most experience with 9 years. We expect subjects with higher years of experience to behave in perception and production more closely to native Russian speakers, that is, we expect them to exhibit higher assimilation fit index values for the palatalized consonants to the palatalized category in perception. We also expect subjects with more experience to produce palatalized consonants more similarly to native Russian speakers. The way we expect production to be more similar to native Russians is that we expect speakers with more experience to produce palatalized consonants with higher F2 values at the VC transition to indicate tongue body raising earlier in the sequence and thus more simultaneous phasing between the consonantal and tongue body gestures.

In order to test this prediction we examine for each place of articulation and syllable position the differences in F2 between plain and palatalized consonants for production compared to the assimilation fit index of palatalized consonants to various perceptual categories. This is exactly what was done in the previous section except in this case we focus also on years of experience as an additional factor.

9.3.1. Onset Coronals

We begin by examining the onset coronals. As for the previous section, we divide the discussion by perception assimilation strategy, first examining assimilation to the plain category, then the consonant+glide category, followed by the glide+consonant category and finally the palatalized category. As seen in the plots in the previous section, no learner subjects assimilated the palatalized onset coronal consonant to the plain category.

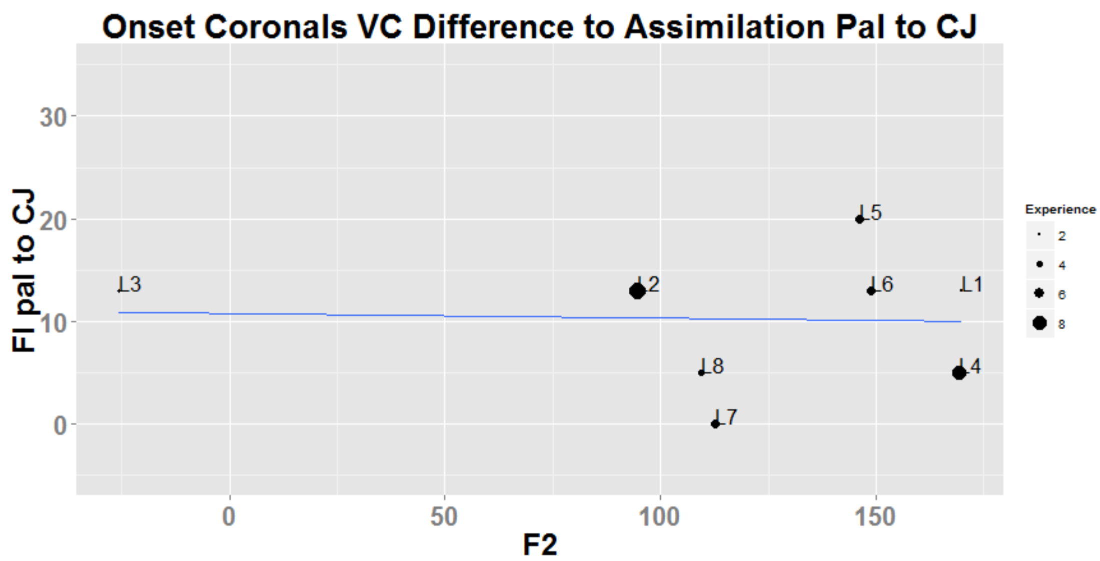

Figure 9.27: Onset coronals comparison of the fit index assimilations of the palatalized consonant to the consonant+glide category and the mean difference in F2 between palatalized consonant and plain consonant at the VC transition for the learner subjects. The size of the dots represent the years of experience for each subject.

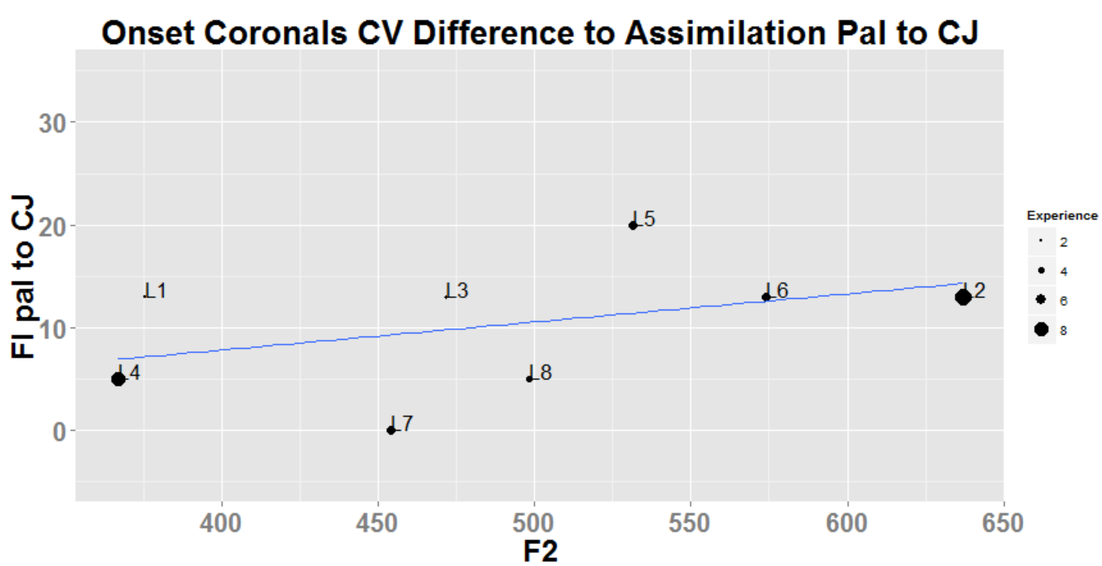

Figure 9.28: Onset coronals comparison of the fit index assimilations of the palatalized consonant to the consonant+glide category and the mean difference in F2 between palatalized consonant and plain consonant at the CV transition for the learner subjects. The size of the dots represent the years of experience for each subject.

Figure 9.27 and Figure 9.28 above show the relationship between difference in F2 and the assimilation of the onset coronal consonant to the consonant+glide category. We see that there is a slightly positive relationship only for the CV transition, that is, if there is a greater difference between the palatalized and plain consonant in the CV transition we would expect a higher fit index assimilation to the consonant+glide category. However, there seems not to be any relationship with experience. We see subject L2 and L4 with the most experience at opposite ends of the production scale, while the assimilation to the consonant+glide category ranges from 0 to 20.

According to the data presented in the previous sections, assimilation to the glide+consonant was also not significant for this group so we will skip this section for the onset coronal. We move on to examining the assimilation of the palatalized consonant to the palatalized category.

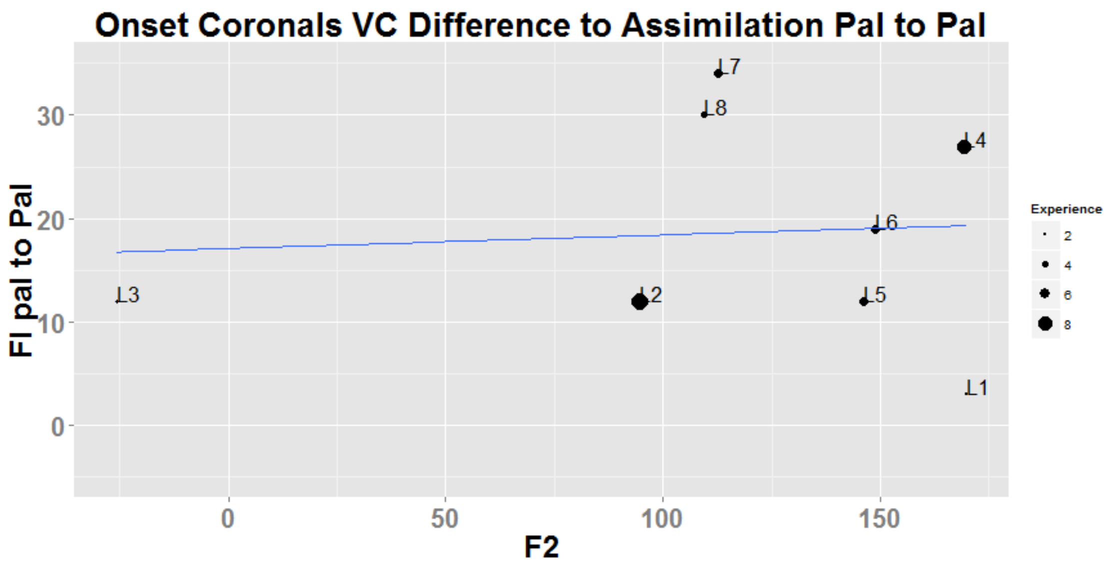

Figure 9.29: Onset coronals comparison of the fit index assimilations of the palatalized consonant to the Russian palatalization category and the mean difference in F2 between palatalized consonant and plain consonant at the VC transition for the learner subjects. The size of the dots represent the years of experience for each subject.

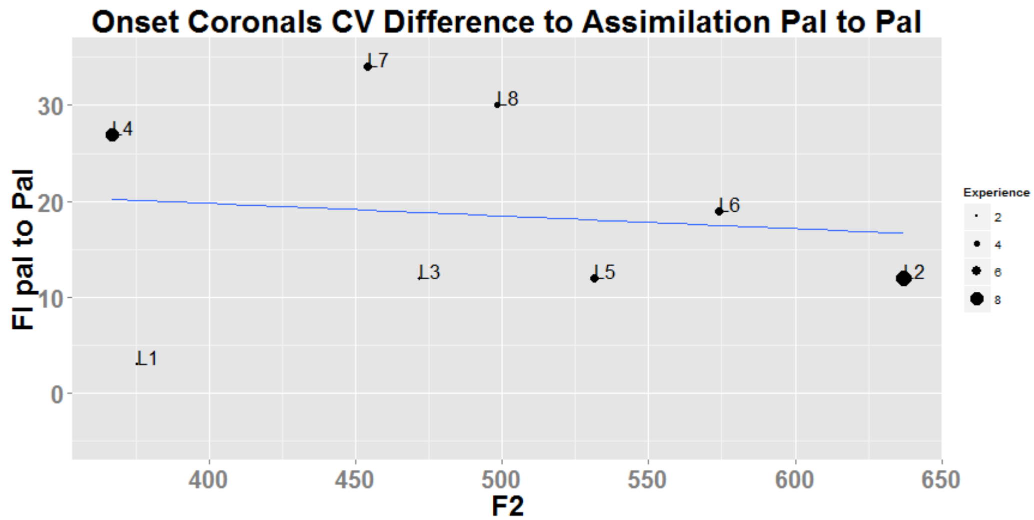

Figure 9.30: Onset coronals comparison of the fit index assimilations of the palatalized consonant to the Russian palatalization category and the mean difference in F2 between palatalized consonant and plain consonant at the CV transition for the learner subjects. The size of the dots represent the years of experience for each subject.

The plots above show the assimilation of the onset coronal consonant to the palatalized category against the VC and CV transition differences between palatalized and plain consonants. We see that there is a slight positive relationship between perception and production for the VC transition and a slight negative relationship for the CV transition. In terms of experience for the VC transition we see most subjects (from most to least experience) clustering at 100 Hz or greater for the difference between F2 between palatalized and plain consonants. Again, it seems that most learner subjects regardless of experience exhibited differences in production between palatalized and plain consonants at the VC transition. We see that there is more variability in the difference in production for the CV transition, however all subjects exhibit a relatively high difference. It seems therefore, that there is no discernible relationship between experience and perception of the palatalized onset coronals to the palatalized category, or between experience and production of the onset palatalized coronal.

9.3.2. Coda Coronals:

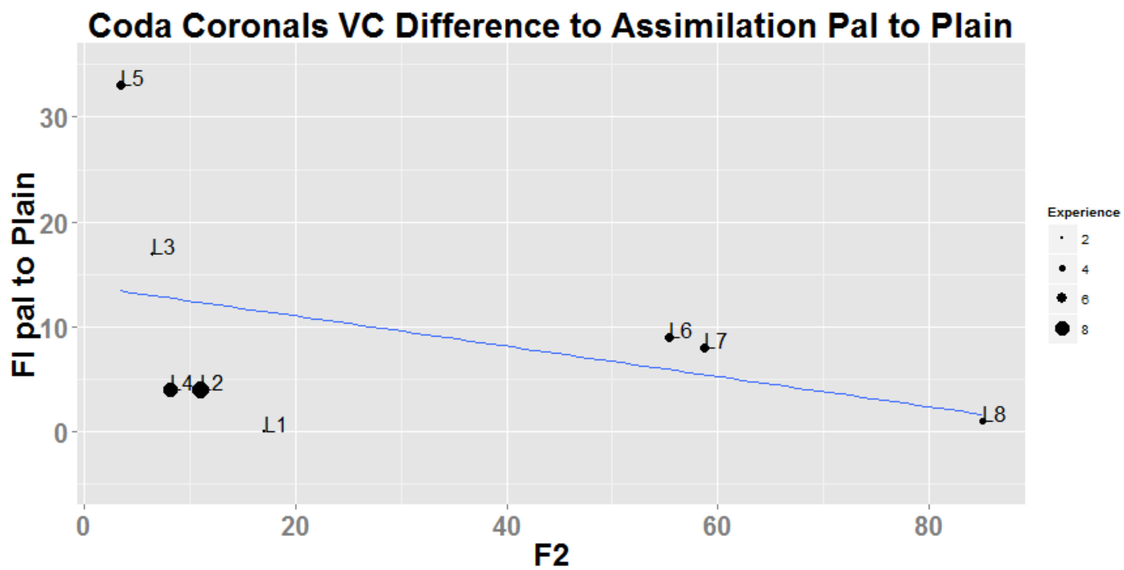

Figure 9.31: Coda coronals comparison of the fit index assimilations of the palatalized consonant to the plain category and the mean difference in F2 between palatalized consonant and plain consonant at the VC transition for the learner subjects. The size of the dots represent the years of experience for each subject.

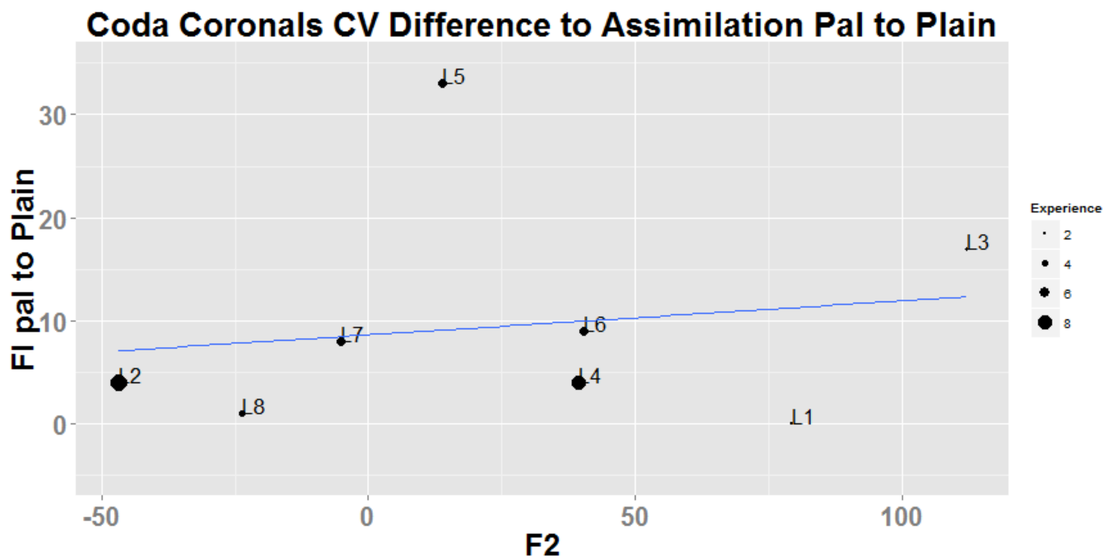

Figure 9.32: Coda coronals comparison of the fit index assimilations of the palatalized consonant to the plain category and the mean difference in F2 between palatalized consonant and plain consonant at the CV transition for the learner subjects. The size of the dots represent the years of experience for each subject.

We now move on to discussing the coda coronals. As seen previously and unlike for onset coronals, for the coda coronals we did see some learner subjects assimilate the palatalized coda to the plain category. Subject L5 with 5 years of experience exhibited high assimilation to the plain category along with low difference in VC (but also CV). Other subjects with high experience such as L2 and L4 had very low assimilation fit index values to the plain category. These two subjects both had low differences in production at the VC transition between plain and palatalized coda coronals but varied in their production at the CV transition, which subject L4 having a higher value than L2.

As seen above, no subjects assimilated the coda coronal consonants to the consonant+glide category. Also, only L7 had a minor fit index assimilation value to the glide+consonant category so we move on to examine the relationship between production and perception to the assimilation patterns to the palatalization category.

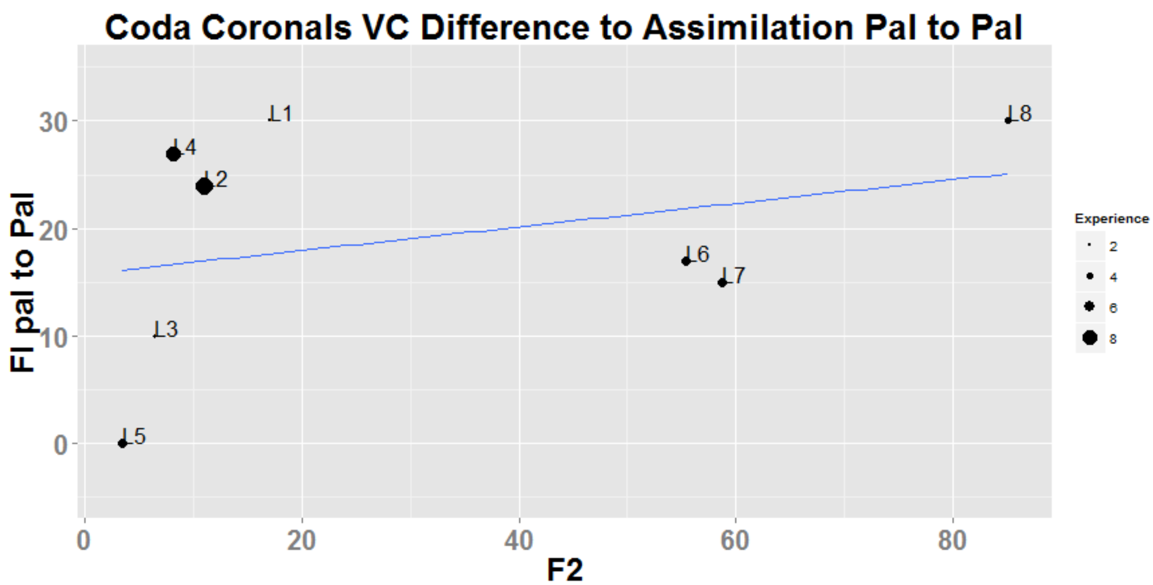

Figure 9.33: Coda coronals comparison of the fit index assimilations of the palatalized consonant to the Russian palatalization category and the mean difference in F2 between palatalized consonant and plain consonant at the VC transition for the learner subjects. The size of the dots represent the years of experience for each subject.

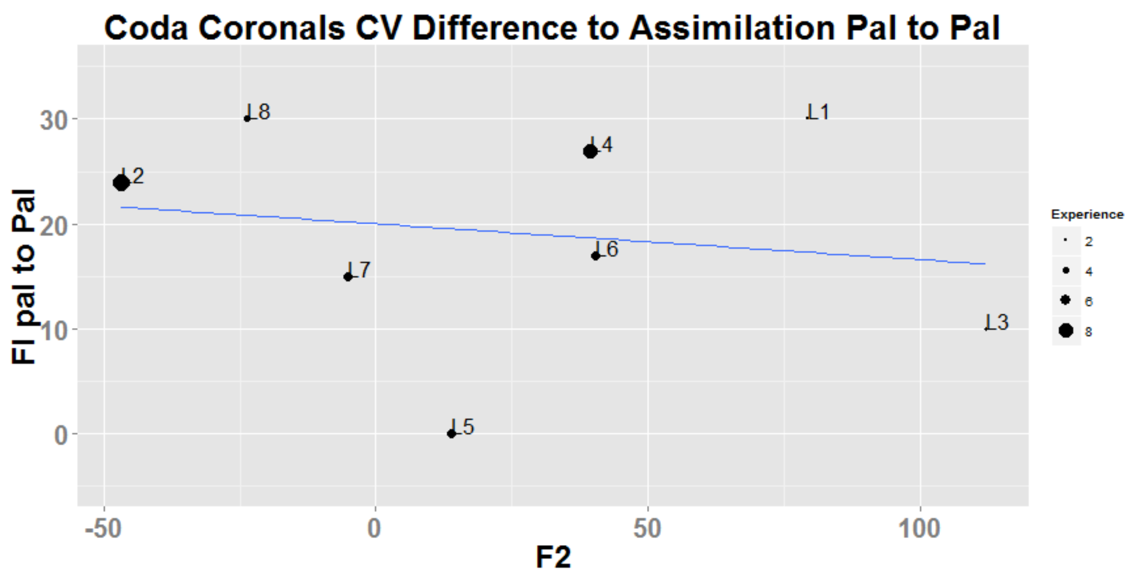

Figure 9.34: Coda coronals comparison of the fit index assimilations of the palatalized consonant to the Russian palatalization category and the mean difference in F2 between palatalized consonant and plain consonant at the CV transition for the learner subjects. The size of the dots represent the years of experience for each subject.

As seen in Figure 9.33 there is a slight positive relationship between the assimilation fit index value to the palatalized category and a difference in F2 between the palatalized and plain consonant at the VC transition. However, there seems not to be a clear relationship with level of experience and perception and production of the coda palatalized coronal. For example, subjects L2 and L4 with high experience exhibited relatively low difference in F2 between the palatalized and plain consonants in the VC transition. The two subjects differed slightly in the CV transition as seen Figure 9.34, however the magnitudes of the difference are small.

These observations tell us that, although these two, more experienced subjects accurately perceived the palatalized coda coronal as the correct, Russian palatalization category, their production of the coda coronal consonants was not accurate. We know this because of the higher difference in F2 between the palatalized and plain consonants for the native Russian speakers as compared to the learner group. As such, it seems that for this group of subjects, experience did not have a strong role in production. As far as perception, subjects from all experience levels, as long as they had some Russian experience, were able to assimilate the coda palatalized consonant to the palatalized category to some degree. As such, experience in Russian increases the ability to assimilate the coda coronal to the palatalized category, but not necessarily the ability to produce it accurately.

9.3.3. Coronals Standard Deviation

LEARNERS ONSET CORONALS							
Subject	**SD diff pal and plain (VC)**	**SD diff pal and plain (CV)**	**Perception plain-pal (axb)**	**Perception pal-glide (axb)**	**Perception FI Pal to Plain**	**Perception FI Pal to CJ**	**Perc FI Pal to Pal**
L1	83	157	1.00	1.00	0	13	3
L2	-27	47	1.00	1.00	0	13	12
L3	16	47	0.96	1.00	0	13	12
L4	13	6	1.00	1.00	0	5	27
L5	-116	-72	1.00	0.92	0	20	12
L6	41	-5	0.96	0.96	0	13	19
L7	69	199	0.92	1.00	0	0	34
L8	-24	131	1.00	1.00	0	5	30

Table 9.7: The table above presents for each subject in the learner group for onset coronals the mean difference in standard deviation between palatalized and plain consonants at the two transitions, and for perception the percent correct discrimination from the AXB tests for both plain-palatalized and palatalized-glide contrasts, as well as the fit index assimilation of the palatalized consonant to the plain category, the consonant+glide category and the palatalized category.

LEARNERS CODA CORONALS							
Subject	**SD diff pal and plain (VC)**	**SD diff pal and plain (CV)**	**Perception plain-pal (axb)**	**Perception pal-glide (axb)**	**Perception FI Pal to Plain**	**Perception FI Pal to JC**	**Perc FI Pal to Pal**
L1	-10	40	0.96	1.00	0	0	30
L2	18	27	0.54	0.96	4	0	24
L3	11	118	0.61	1.00	17	0	10
L4	-8	58	0.96	1.00	4	0	27
L5	-30	-13	0.61	1.00	33	0	0
L6	8	0	0.71	1.00	9	0	17
L7	11	-20	0.67	1.00	8	4	15
L8	-20	-60	0.83	1.00	1	0	30

Table 9.8: The table above presents for each subject in the learner group for coda coronals the mean difference in standard deviation between palatalized and plain consonants at the two transitions, and for perception the percent correct discrimination from the AXB tests for both plain-palatalized and palatalized-glide contrasts, as well as the fit index assimilation of the palatalized consonant to the plain category, the glide+consonant category and the palatalized category.

We now discuss in slightly more detail the differences in variability in production for the learner group. For onset coronals, we see some variation and in general we see fewer subjects in this group with high average standard deviation differences than for the naïve group. Most subjects in this group had relatively high percent correct discrimination. We can, however, look at the perception assimilation data and specifically the tendency for each subject to assimilate the palatalized consonant to the palatalized (correct) category. However, by comparing standard deviation to the assimilation of the consonant to the palatalized category we do not see a clear pattern. For example, subject L7 exhibited the highest difference in standard deviation (199) but actually exhibited the highest fit index of assimilation of all subjects (34). In short, this association does not hold for all subjects. For the coda coronals, none of the learner subjects exhibited high difference in standard deviation for the VC transition. The highest value shows was 18, for subject L2.

9.3.4. Onset Labials

As for the coronals we begin with the onsets. Like the coronals, no subjects assimilated the onset labials to the plain category in perception, nor was there any significant assimilation to the glide+consonant category. We proceed to the patterns of assimilation to the consonant+glide category.

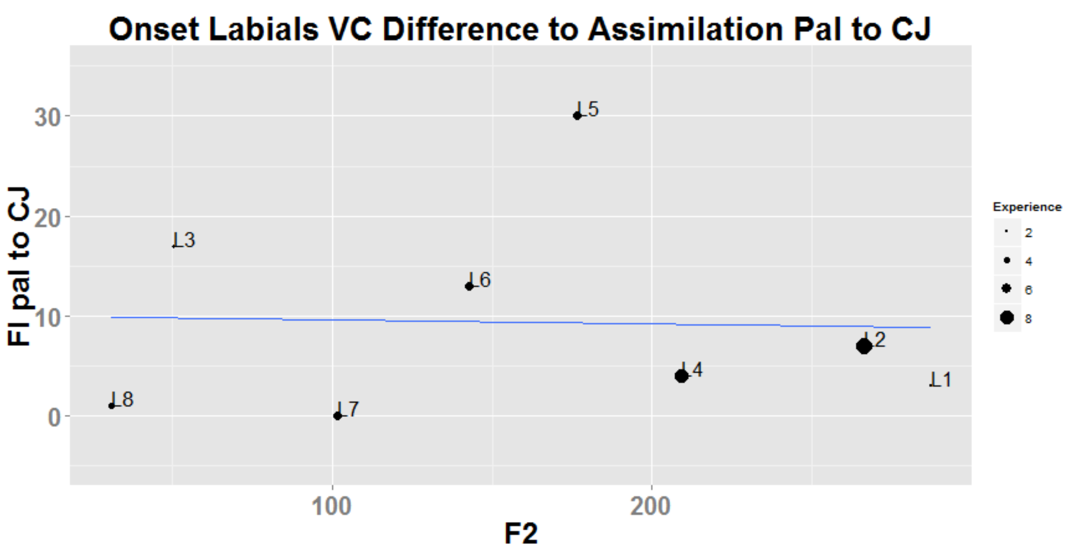

Figure 9.35: Onset labials comparison of the fit index assimilations of the palatalized consonant to the consonant+glide category and the mean difference in F2 between palatalized consonant and plain consonant at the VC transition for the learner subjects. The size of the dots represent the years of experience for each subject.

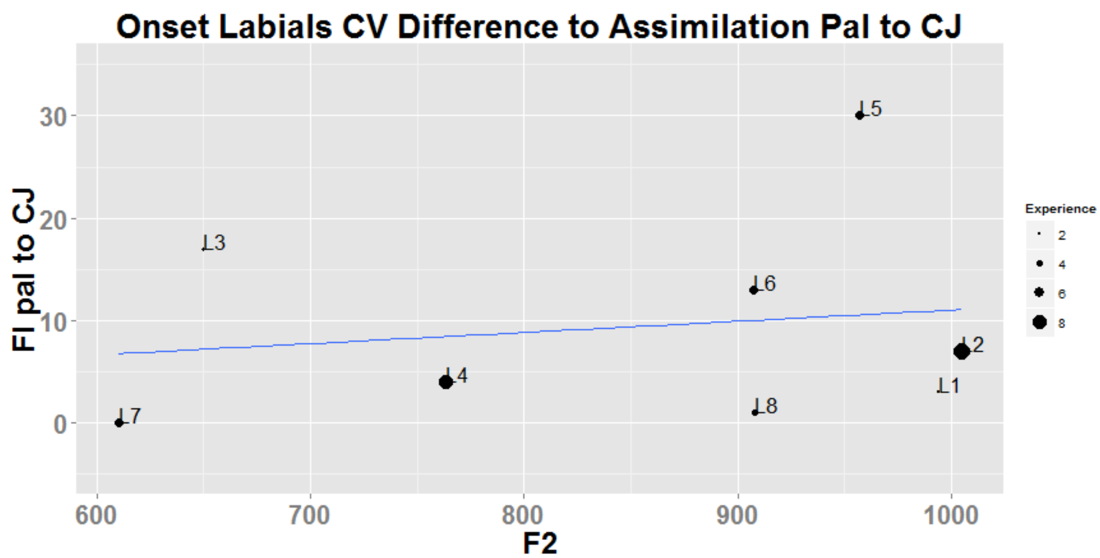

Figure 9.36: Onset labials comparison of the fit index assimilations of the palatalized consonant to the consonant+glide category and the mean difference in F2 between palatalized consonant and plain consonant at the CV transition for the learner subjects. The size of the dots represent the years of experience for each subject.

Comparing the relationship between perception and production of the onset coronals to the onset labials for the assimilation of the palatalized consonant we see a similar trend, in that there is a slightly negative association between fit index of perception of the palatalized consonant to the consonant+glide category and F2 difference between palatalized and plain consonants at the VC transition (Figure 9.35), and a slightly positive association at the CV transition (Figure 9.36). In terms of experience, as for the coronals there seems to be no relationship with experience. We see subject L2 and L4 with the most experience at different regions of the production scale, while the assimilation to the consonant+glide category ranges from 0 to 30 across subjects. Before moving on to examining the patterns of assimilation to the palatalized category we say a few words about assimilation to the glide+consonant category. We move on to examining the assimilation of the onset labial to the palatalized category.

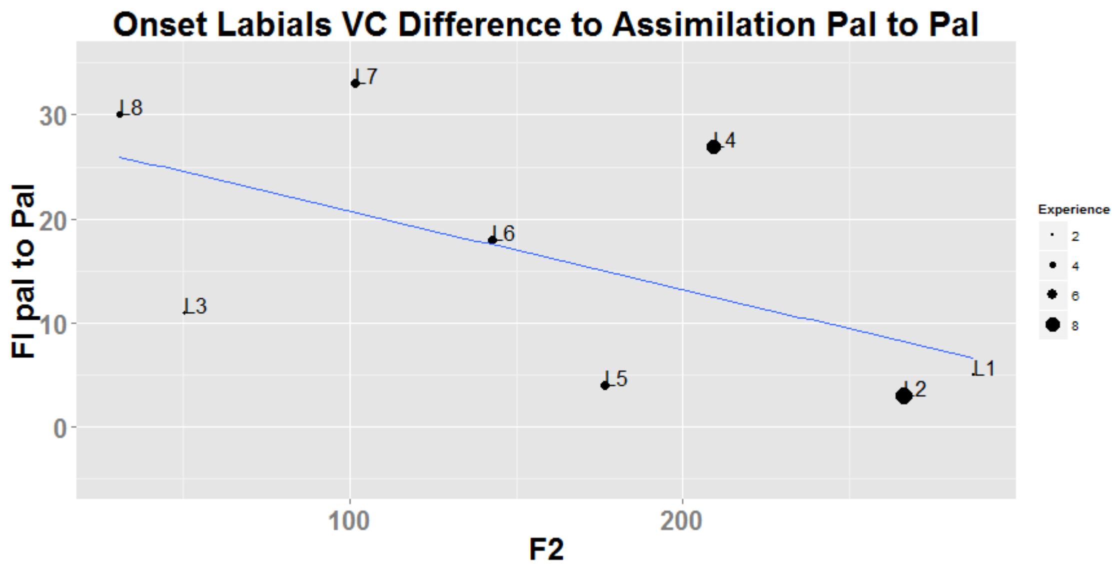

Figure 9.37: Onset labials comparison of the fit index assimilations of the palatalized consonant to the Russian palatalization category and the mean difference in F2 between palatalized consonant and plain consonant at the VC transition for the learner subjects. The size of the dots represent the years of experience for each subject.

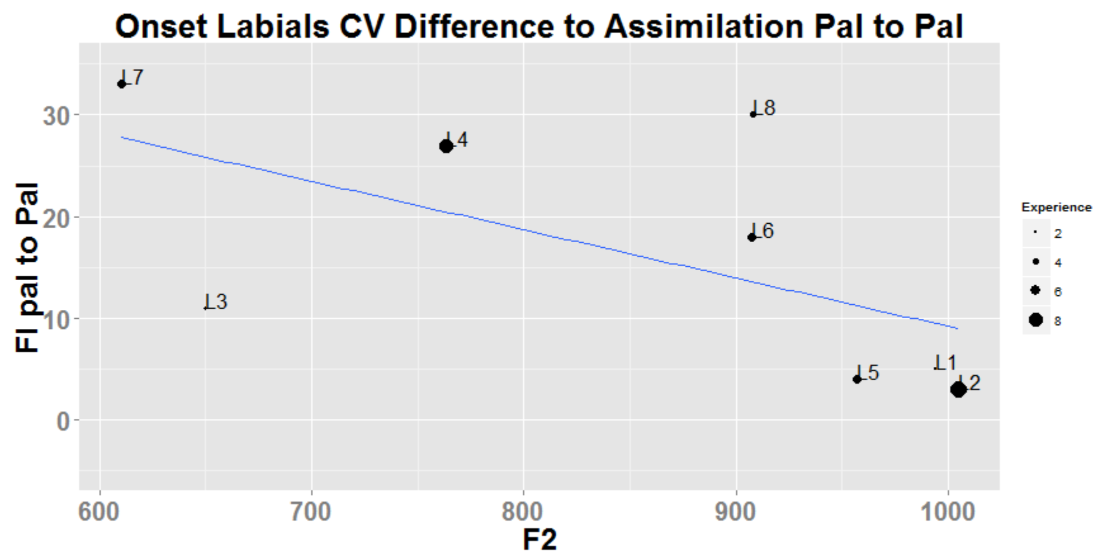

Figure 9.38: Onset labials comparison of the fit index assimilations of the palatalized consonant to the Russian palatalization category and the mean difference in F2 between palatalized consonant and plain consonant at the CV transition for the learner subjects. The size of the dots represent the years of experience for each subject.

Figure 9.37 and Figure 9.38 above show a slightly different pattern for the relationship between perception and production for the onset labials as for the coronals. Here we see more clear negative relationships for the differences between F2 at both transitions and the assimilation fit index of the consonant to the palatalized category. This means that as we see a higher difference in mean F2 between the production of the onset palatalized labial at both transitions over the production of the plain labial, so we see a lower tendency for that subject to assimilate the consonant to the palatalized consonant. This result is unexpected if production and perception were closely linked. That is, we would expect subjects that were learners of Russian who exhibited greater difference in production between palatalized and plain consonants to be able to assimilate the palatalized consonant to the accurate palatalized category, however this is not observed in this data. We also see that several subjects with higher years of experience exhibit lower assimilation fit index values to the palatalized consonant (for example L2). This is also not expected, since greater experience is predicted to be associated with behavior closer to the native Russian group.

9.3.5. Coda labials

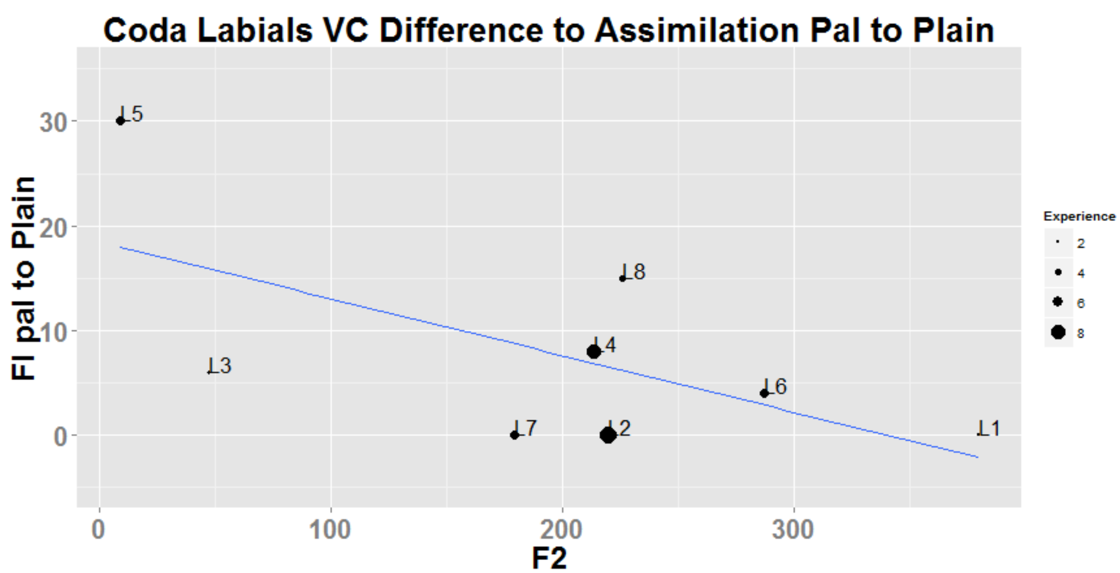

Figure 9.39: Coda labials comparison of the fit index assimilations of the palatalized consonant to the plain category and the mean difference in F2 between palatalized consonant and plain consonant at the VC transition for the learner subjects. The size of the dots represent the years of experience for each subject.

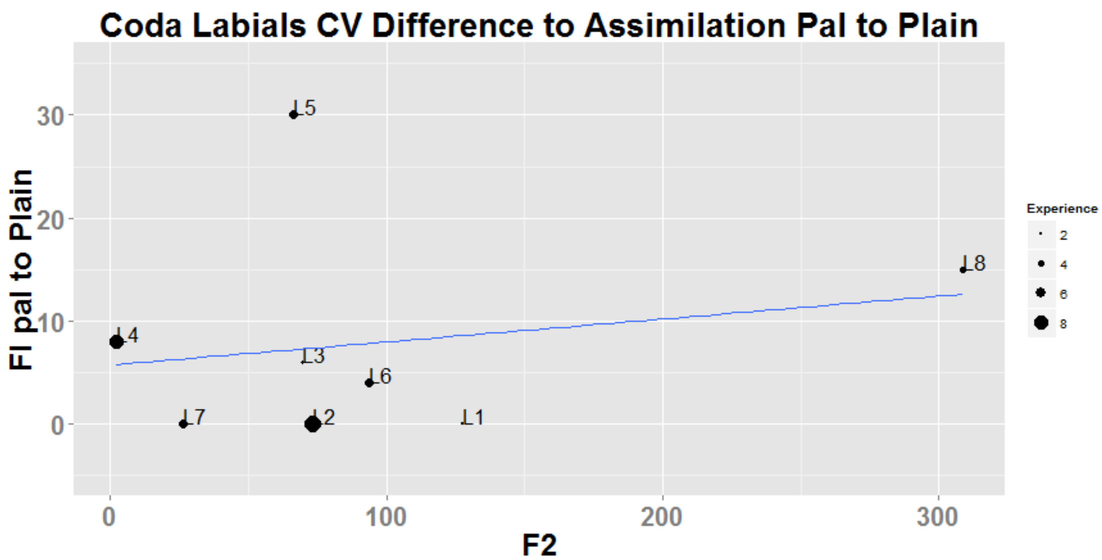

Figure 9.40: Coda labials comparison of the fit index assimilations of the palatalized consonant to the plain category and the mean difference in F2 between palatalized consonant and plain consonant at the CV transition for the learner subjects. The size of the dots represent the years of experience for each subject.

We move on to the analysis for coda labials. Figure 9.39 and Figure 9.40 above show a similar trend to that seen for the coda coronals. We see a negative relationship between the assimilation fit index to the plain category and difference in production at the VC transition, and a positive relationship for the CV transition. This means that as subjects increase the difference in production at the VC transition, so too do they decrease the tendency to assimilate the coda consonant to the plain category. This is an expected result, as exhibiting no difference in production of palatalized and plain consonants should correlate with also perceiving no difference between plain and palatalized consonants, which means assimilating the palatalized consonant to the plain category.

As for the coronals, we skip the analysis comparing perception to the consonant+glide category since no subjects assimilated the coda palatalized labial to this category. Recall that for the coronals, only one subject had any discernible assimilation pattern of the palatalized coronal coda to the glide+consonant category (L7). For the labials, two subjects exhibited some tendency to assimilate the coda labial to the glide+consonant category, (L1 and L8) both subjects with relatively low experience. Subject L1 at least exhibited a greater difference in production between the palatalized consonant and plain consonant at the VC transition. This means that for this subject, at least, our prediction holds, that is, the subject produced the palatalized consonant with a much higher mean F2 at the VC transition, thus producing this consonant in a manner close to a sequence of consonant+glide. However, this same subject also had a relatively high difference in production at the CV transition as well. This would, according to our predictions mean that this subject should have also had a relatively high perceptual assimilation rate of this consonant to the Russian palatalized consonant, this is not what we see, however. Subject L8 also had a high difference in production between palatalized and plain consonants at the CV transition. We now move on to examining the relation between perception to the palatalized category and production.

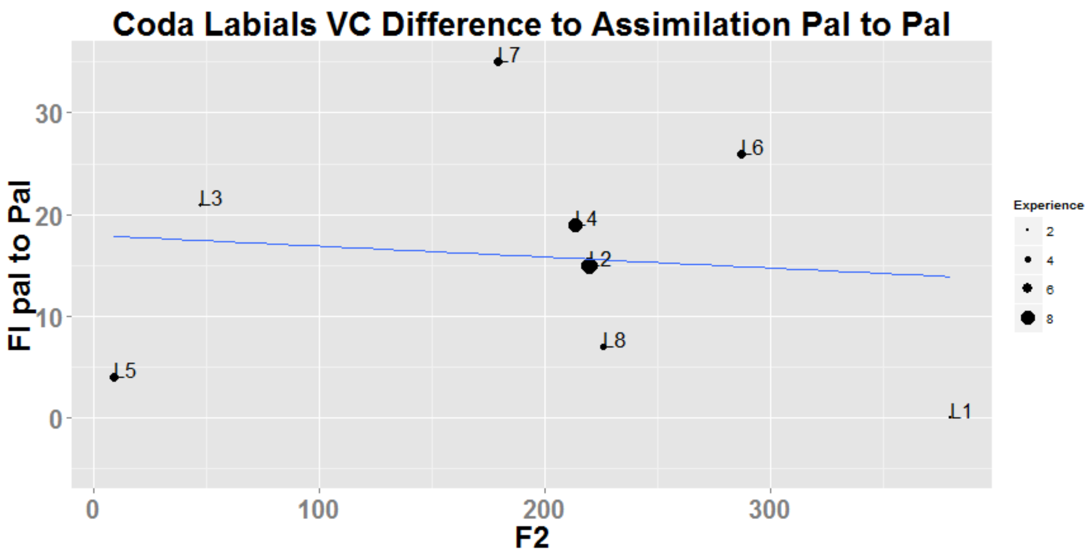

Figure 9.41: Coda labials comparison of the fit index assimilations of the palatalized consonant to the Russian palatalization category and the mean difference in F2 between palatalized consonant and plain consonant at the VC transition for the learner subjects. The size of the dots represent the years of experience for each subject.

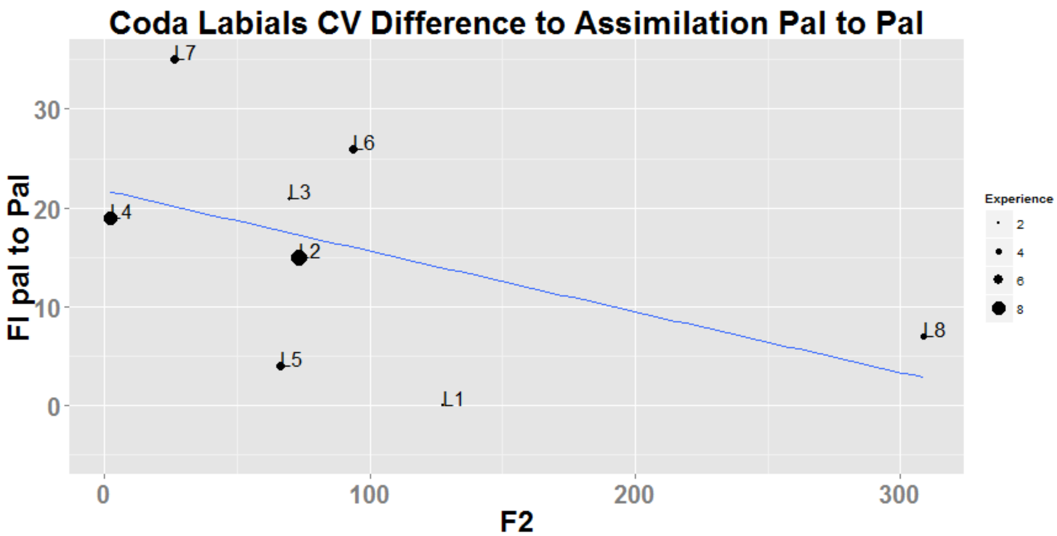

Figure 9.42: Coda labials comparison of the fit index assimilations of the palatalized consonant to the Russian palatalization category and the mean difference in F2 between palatalized consonant and plain consonant at the CV transition for the learner subjects. The size of the dots represent the years of experience for each subject.

The pattern of the relationship between perception and production for the coda labials to the palatalized category is not the same as that for the coda coronals. We see a slightly negative relationship at the VC transition where we saw a slightly positive relationship for the coronals. We see a negative relationship at the CV transition where we also saw a negative relationship at the CV transition for the coronals. As for the coda coronals, we do not see a clear pattern between experience and performance in production. The native Russian speakers exhibited greater differences in production between plain and palatalized consonants at the VC transition than did the learner subjects (see Figure 9.21-Figure 9.26 earlier in the chapter), so we can see that the learner group did not produce the coda labial consonants accurately, and subjects with greater experience such as L2 and L4 are no exception.

9.3.6. Labials Standard Deviation

LEARNERS ONSET LABIALS							
Subject	SD diff pal and plain (VC)	SD diff pal and plain (CV)	Perception plain-pal (axb)	Perception pal-glide (axb)	Perception FI Pal to Plain	Perception FI Pal to CJ	Perc FI Pal to Pal
L1	144	21	0.96	1.00	0	3	5
L2	43	81	1.00	1.00	0	7	3
L3	34	114	0.96	1.00	0	17	11
L4	62	68	0.96	1.00	0	4	27
L5	2	20	0.96	0.96	0	30	4
L6	12	114	1.00	0.88	0	13	18
L7	33	203	1.00	0.96	0	0	33
L8	-16	16	1.00	1.00	0	1	30

Table 9.9: The table above presents for each subject in the learner group for onset labials the mean difference in standard deviation between palatalized and plain consonants at the two transitions, and for perception the percent correct discrimination from the AXB tests for both plain-palatalized and palatalized-glide contrasts, as well as the fit index assimilation of the palatalized consonant to the plain category, the consonant+glide category and the palatalized category.

LEARNERS CODA LABIALS							
Subject	SD diff pal and plain (VC)	SD diff pal and plain (CV)	Perception plain-pal (axb)	Perception pal-glide (axb)	Perception FI Pal to Plain	Perception FI Pal to JC	Perc FI Pal to Pal
L1	244	0	1.00	1.00	0	27	0
L2	-10	-19	0.96	1.00	0	0	15
L3	59	-15	0.71	1.00	6	0	21
L4	-107	6	0.88	1.00	8	0	19
L5	9	6	0.79	1.00	30	0	4
L6	18	-11	0.83	1.00	4	0	26
L7	3	9	0.83	0.92	0	0	35
L8	202	416	1.00	1.00	15	5	7

Table 9.10: The table above presents for each subject in the learner group for coda labials the mean difference in standard deviation between palatalized and plain consonants at the two transitions, and for perception the percent correct discrimination from the AXB tests for both plain-palatalized and palatalized-glide contrasts, as well as the fit index assimilation of the palatalized consonant to the plain category, the glide+consonant category and the palatalized category.

For onset labials, as for other comparisons we have seen there is again variation in the relation between standard deviation differences and perception. We see a few subjects, namely L2, L3, L6 and L7 exhibiting high standard deviation differences for the onset codas at the CV transition, but we see only two of these subjects, L2 and L3 also exhibiting a low fit index of the palatalized consonant to the palatalized category. As seen above, there is no clear pattern as to level of experience with these findings. As for the naïve subjects, for coda labials we also see in general larger differences in standard deviation than for the coda coronals for the VC transition.

Subjects such as L1 and L8 exhibited high average difference in standard deviation. Interestingly, both of these subjects had perfect discrimination of the plain-palatalized contrast.

In fact, they were the only two subjects in the learner group to exhibit this result. However, neither subject exhibited a high fit index value of the palatalized consonant assimilating to the palatalized category, instead they tended to prefer assimilating this consonant to the plain or the JC category. The following chapter will conclude with closing remarks and further discussion about the relationships between perception and production.

Chapter 10 General Discussion Perception and Production:

To recap, in 0 we introduced several goals for this dissertation and in this section we summarize how this dissertation has attempted to address these goals. The first goal was to investigate perception and production of an example of non-native gestural timing, namely palatalization. We have examined the properties of palatalization in general and in Russian, and provided three separate perception studies as well as a production study to examine both the differences in production of palatalization by English and Russian speakers. It was argued that one of the main differences between English and Russian with respect to palatalization was the relative coordination of the gestures involved in Russian palatalization. Through two of the perception studies presented above we endeavored to tackle the second major goal, which was to test the Perceptual Assimilation Model (PAM) of non-native speech by extending it to new data: non-native perception of gestural timing (palatalization). We performed both a discrimination study and a categorization study and compared the results both at the group level but also at the subject level. These data demonstrated that the PAM was generally supported but not for all subjects.

We also introduced interesting data to the PAM in the sense that we made predictions about possible multi-category assimilations. That is, some of the predictions (and the data subsequently confirmed this) posited that speakers of English would assimilate the gestural constellations of Russian palatalization (constituting one category in Russian) into multiple categories (e.g. CJ or JC) in English. The importance of Articulatory Phonology came into play here, in that if we assume articulatory gestures and their coordination as primitives in phonology, as opposed to segments, it makes it easy to posit that speakers of one language which coordinates gestures a certain way into a single segment could assimilate the same gestures, but coordinated in a more nativelike manner, which happens to result in multiple native segments.

The third major goal was to investigate the extent of a link or correlation between perception and production of non-native timing. The predictions for the relationship between perception and production for the naïve English group differed for each syllable position and palatalization type. We examined the perception of the palatalized consonants in both onset position and coda position and compared subjects' perception of these palatalized consonants to the subsequent production of the same. This was done by examining the assimilation fit index of each palatalized consonant to each of the perceptual categories and comparing this fit index to the difference in F2 at each consonantal transition between the palatalized consonant and the plain consonant. The general prediction was that a low value for the difference in production would indicate that the subject produced the palatalized and plain consonants similarly and as such should correspond to an assimilation fit index to the plain category. Another prediction was that a high difference in F2 between the palatalized and plain consonants at the VC transition should correspond to a high fit index of assimilation for the palatalized consonant to the glide+consonant category. A third prediction was that higher variability in production (as measured by standard deviation) for the palatalized consonants would correspond to less accurate perception. Conversely, a high difference in F2 between the palatalized and plain consonants at the CV transition was predicted to correspond to a high fit index to the consonant+glide category. These predictions assume a close relationship between perception and production, in that a subject who perceives a palatalized consonant as closest to a certain category (plain, consonant+glide, or glide+consonant) is predicted to produce the consonant in a similar way. As seen in the exposition above, this prediction did not hold for all subjects.

The predictions did not hold specifically for the relation between production and the assimilation tendency to the consonant+glide category, specifically for onset consonants. A positive relationship was predicted for difference in F2 at the CV transition and assimilation fit index to the consonant+glide category for these cases. This means that we expected for subjects who exhibited a greater difference in F2 at the CV transition between plain and palatalized consonants a higher tendency to assimilate these consonants to the consonant+glide category. This was not observed for all subjects for either coronals or labials.

One prediction which seemed to hold better was the relation between the assimilation of palatalized codas to the glide+consonant category and a higher F2 difference between plain and palatalized coda consonants at the VC transition. The relationship between assimilation of the palatalized consonant to the glide+consonant category and a higher F2 value at the VC transition for coda palatalized consonants was the strongest amongst all the comparisons. That is, for both labials and coronals, subjects who produced palatalized codas with a higher F2 at the VC transition tended to perceive these consonants as sequences of glide+consonant. We also saw further evidence of this by further examining a smaller subset of subjects who behaved differently in systematic ways, some of which perceive the palatalization timed one way (glide+consonant) and some of which perceive it a different way (consonant+glide). This finding is interesting because it seems to validate the idea posited in this dissertation that speakers of a language (English) which lacks the gestural timing configuration present in a different language (Russian) when confronted with speech sounds with such a timing configuration (synchronous coordination of tongue body and consonantal gesture of Russian palatalization) would use strategies present in their own language. One of the posited strategies speakers could use was re-interpreting the palatalized consonants as sequences of glide and consonant. This seemed to have been observed for a subset of naive English speakers, and the relation between perception and production seemed also to hold for a subset of speakers.

One of the main findings in previous chapters is the inter-subject variability in both perception and production. Subjects, especially in both the English groups, differed in their tendencies to assimilate palatalized consonants to different categories. We saw that there were tendencies for subjects to follow in the predictions outlined.

One final remark to be made concerns the predictions outlined in Chapter 3, specifically concerning the differences predicted between coronals and labials. In the early chapters of this dissertation we discussed possible different strategies that English speakers could take in producing or perceiving palatalized coronals and labials. One of the possibilities in perception was that English speakers could more readily assimilate onset palatalized labials to the consonant+glide sequence perceptual category than onset coronals. The reason for this is due to the existence of such sequences in American English (e.g. ‘beauty’, ‘puny’) alongside a lack of such sequences for coronals. We did see a tendency for naive English speakers to assimilate onset palatalized labials with a higher fit index to the consonant+glide sequences than onset coronals. In terms of production, we also predicted the possibility of English speakers producing palatalized onset as sequences of consonant+glide due to the existence of these sequences in their dialect of English. Indeed we did see this, as English speakers had much a higher difference in F2 between palatalized labials and plain labials than between palatalized coronals and plain coronals. These results taken together seem to suggest that naive English speakers treated the onset palatalized labials consonants more readily as sequences of consonant+glide than they did the onset palatalized coronals.

Finally, the fourth and fifth major goals were to test the effect of learning or L2 experience in perception and production of palatalization and to examine individual differences in the link between perception and production in non-native speech. In Section 10.1 we discuss in more detail the fourth goal, and throughout this discussion in general we have seen how inter-subject variability is a clear result that is seen in this data.

One last remark has to be made about the nature of the speech materials examined here. Namely, we have considered the Russian consonants interesting because they are palatalized. We have not said much about the opposing members of the contrasts (the non-palatalized variants) and in many cases use the term "plain" to describe them. In reality, much more work needs to be done to paint a more complete picture on the nature of this contrast. It isn't simply a plain-palatalized contrast, but probably actually a velarized-palatalized contrast. However, the data presented here is still valuable because we are still testing a non-native contrast and because especially in the context of back vowels (which is the case for the data here), velarization seems to be less salient than in other contexts (e.g. Bennett et al. to appear).

10.1. Discussion Learning:

In this section we summarize the varied findings with respect to learning and relate these observations to the predictions outlined earlier in this chapter and in this dissertation. In terms of production, one major difference between the learner subjects and the naive subjects is a smaller range in the difference between F2 between palatalized and plain consonants. This is true for both onsets and coronals. For example, at the VC transition for coda coronals, the naive group exhibited a greater range of mean F2 difference, with some subjects exhibiting values as high as 200Hz or 500Hz, while for the learner group at this transition for coda coronals the highest value was near 100Hz. This demonstrates that learners behaved more uniformly in terms of production than naive speakers did. This finding was also corroborated by specifically measuring average standard deviation for each place of articulation and consonant in production (however, for coda coronals the reverse is true, naïve subjects in general exhibited on average lower variation at the VC transition than did learners).

In production, even though the learners behaved more similar to each other than the naive speakers did, they did not behave more similarly to the native Russian speakers. This is evident in the plots presented in this chapter, especially at the VC transition. The native Russian speakers exhibited higher values for the difference in F2 between palatalized and plain consonants at the VC transition. This brings us now to the relation between production and perception and the effect of learning.

Unlike the naive group subjects in the learner group did not assimilate the coda palatalized consonants to the glide+consonant category to the same degree, one subject (L1) exhibited relatively high fit index values to the glide+consonant category for the codas, compared to several naive subjects. This subject had 2 years of experience. Compared to the naive group the learner group also did not assimilate the palatalized consonant to the plain category to the same degree. One subject (L5) exhibited a high assimilation fit index for both coda coronals and labials to the plain category. This subject had 5 years of experience. This brings us to the concept of learning and experience and its effect on the relationship between production and perception. For the onset coronals we saw no real association with experience and production and perception for the assimilation to the consonant+glide category. For onset labials in general perception and production were not linked, that is, independent of experience we saw no real association between perception and production. When we consider the one perceptual category the learner group could assimilate the palatalized consonant to which the naive group lacked -- the Russian palatalization category, we also found that experience is not a predictor of assimilating to the category. We saw subjects with differing degrees of experience assimilate the palatalized consonants to the palatalized categories to different extents.

In terms of the difference between place of articulation, learner subjects differed to the degree to which they performed more like native Russians depending on perception or production. For example, specifically comparing coda consonants, learner subjects managed to produce the palatalized labials more distinctly than the plain labials to a greater degree than the coronals. That is, subjects with higher experience produced mean F2 differences of between 200Hz and 300Hz, while the differences in the coronals at the VC transition were much lower. However, the same subjects who produced the palatalized labials in coda position with a raised F2 did not also consistency assimilate the palatalized labials to the palatalized category as well as they did the palatalized coronals to the palatalized category. This means that learner subjects tended to be better at producing a difference between palatalized and plain coda labial consonants than the coronal counterparts, but where at the same time worse at perceiving palatalized labial consonants as clear exemplars of palatalized consonants than coronal consonants. There are a few possible explanations for this discrepancy between coda labials and coronals. First, for production, recall from Chapter 3 that Kochetov (2002, 2004) found differences in the phasing of the consonantal (LA, TT) and palatal (TB) gestures between Russian palatalized labials and coronals in different syllable positions. It is also the case that Russian palatalized coronals are more frequent than palatalized labials in the language. Kochetov's (2002) study discovered that even native Russian speakers were better at identifying coda palatalized coronals as the proper category than coda palatalized labials.

Taken together it is possible that learners are better perceiving the coda coronals as palatalized coronals either due to their increased frequency in the language, and thus greater familiarity, or perhaps due to some other increases perceptual saliency attributable to palatalized coronals as opposed to palatalized labials (Kochetov 2002). It is possible that learners are producing palatalized coda labials as more distinct from plain consonants than coronals because they could simply be easier to produce given the fact that the palatalization gesture and consonantal gesture occur on completely separate organs (lips and tongue). This is just speculation, but it is interesting that the relation between perception and production differs depending on the place of articulation and syllable position.

10.2. Conclusion

To conclude these studies show that we do see evidence that generally supports the PAM model (Best e al 1988) in that when the group on average perceived the individual consonants in a contrast by assimilating them to two separate (or at least not the same) categories, in general discrimination of that contrast was better than when there was a greater tendency to assimilate both consonants of a contrast to the same category. This finding held across groups and at the group level. In terms of relationship between perception and production, we did not find strong patters at the group level but we did see some patterns when investigating at the subject level. Some groups of subjects behaved as expected in our predictions, others did not. Perhaps the most important finding in this dissertation is that humans vary from individual to individual in their performance when perceiving and producing a second language. Another important finding is the result observed for the learner group.

Taking into account experience and learning, we do not see strong evidence that a differing degree of experience beyond two years is strongly associated with improvement in abilities of perception and/or production of Russian palatalized consonants in learners. Some possible explanations for this could include the age of learning, or the fact that speakers of higher experience lose the ability or motivation to accurately perceive/produce such contrasts. It is possible that such speakers are no longer motivated to accurately produce or perceive such consonants in an experimental setting as exposed to them here. In conclusion, it seems that having any degree of experience in the target language causes speakers to improve their production and perception in general, but increasing experience does not seem to affect perception and production of Russian palatalization.

References

Beddor, P. S., & Gottfried, T. L. 1995. Methodological issues in cross-language speech perception research with adults. In *Speech Perception and Linguistic Experience: Theoretical and Methodological Issues in Cross-Language Speech Research* (W. Strange, editor), pp. 207-232. Timonium, MD: York Press.

Beddor, P.S. 2015. The relation between language users' perception and production repertoires. *Proceedings of the 18th International Congress of Phonetic Sciences*, Glasgow, UK, ed. The Scottish Consortium for ICPhS 2015, University of Glasgow.

Bennett, Ryan, Máire Ní Chiosáin, Jaye Padgett & Grant McGuire. to appear. An Ultrasound Study of Connemara Irish Palatalization and Velarization. *Journal of the International Phonetic Association*.

Best, Catherine T. 1995. A Direct Realist View of Cross-Language Speech Perception. In *Speech perception and linguistic experience: Issues in cross-language research,* ed. Strange, W., 171-204. Timonium, MD:York Press.

Best, Catherine T., and Winifred Strange. 1992. Effects of phonological and phonetic factors of cross-language perception of approximants. *Journal of Phonetics* 20:305-330.

Best, Catherine T., and Michael D. Tyler. 2007. Nonnative and second-language speech perception: Commonalities and complementarities. In *Second language speech learning: The role of language experience in speech perception and production.* eds. Munro, M. J., and O. S. Bohn, 13-34. Amsterdam: John Benjamins.

Best, Catherine T., Pierre A. Hallé, Bohn Ocke-Schwen, & Alice Faber. 2003. Cross-language perception of nonnative vowels: Phonological and phonetic effects of listeners' native languages. In *Proceedings of Proceedings of the 15th International Congress of Phonetic Sciences. Barcelona, Spain.* Anonymous 2889-2892. Barcelona, Universitat Autonoma de Barcelona.

Best, Catherine T., Anthony Traill, Allyson Carter, K. David Harrison, & Alice Faber. 2003. !Xoo click perception by English, Isizulu, and Sesotho listeners. In *Proceedings of the 15th International Congress of Phonetic Sciences. Barcelona, Spain.* Anonymous 853-856. Barcelona, Universitat Autonoma de Barcelona.

Best, Catherine T., G. W., McRoberts, & E. Goodell. 2001. Discrimination of non-native consonant contrasts varying in perceptual assimilation to the listener's native phonological system. *Journal of the Acoustical Society of America*, 109, 775–794.

Best, Catherine T., Pierre A. Hallé, & Jennifer S. Pardo. 2007. English and French Speakers' Perception of Voicing Distinctions in Non-Native Lateral Consonant Syllable Onsets. In *Proceedings of Interspeech 2007: International Speech Communication Association,* Anonymous 27-31. Antwerp.

Best, Catherine T., & Pierre A. Hallé. 2010. Perception of initial obstruent voicing is influenced by gestural organization. *Journal of Phonetics*, 38: 109-126.

Bhat, D. N. S. 1978. A General Study of Palatalization. In J. H. Greenberg (ed.), *Universals of Language*, Vol. 2: Phonology. Stanford, California: Stanford UP. 47-91.

Boersma, Paul & David Weenink. 2007. Praat: Doing phonetics by computer. http://www.praat.org/.

Bolla, Kálmán. 1981. A conspectus of Russian speech sounds, vol. Bd. 32. Köln: Böhlau.

Browman, C., & L. Goldstein. 1986. Towards an articulatory phonology. *Phonology yearbook*, 3, 219-252.

Browman, C., & L. Goldstein. 1988. Some notes on syllable structure in articulatory phonology. *Phonetica,* 45, 140-155.

Browman, C. P., & Goldstein, 1. 1991. Gestural structures: Distinctiveness, phonological processes, and historical change. In 1. G. Mattingly & M. Studdert-Kennedy (Eds.), Modularity and the Motor Theory ofSpeech Perception (pp. 313-338). Hillsdale NJ: Lawrence Erlbaum Associates.

Browman, C., & L. Goldstein. 1992. Articulatory phonology: An overview. *Phonetica*, 49, 155-180.

Browman, C. & L. Goldstein. 1995. Dynamics and Articulatory Phonology. In T. vans Gelder and R. F. Port (Eds.), *Mind As Motion.* Cambridge, MA: MIT Press.

Burnham, D. K. 1986. Developmental loss of speech perception: Exposure to and experience with a first language. *Applied Psycholinguistics* 7: 207-40.

Davidson, Lisa. 2006. Phonology, phonetics, or frequency: Influences on the production of nonnative sequences. *Journal of Phonetics* 34:104-137.

Davis, S. and Hammond, M., 1995. On the status of onglides in American English. *Phonology*, *12*(02), pp.159-182.

Diehm, Erin. 1998. Gestures and linguistic function in learning Russian: Production and perception studies of Russian palatalized consonants. Ph. D. Dissertation, The Ohio State University.

Flege, James Emil. 1987a. The production of "new" and "similar" phones in a foreign language: evidence for the effect of equivalence classification. *Journal of Phonetics* 15:47-65.

Flege, James Emil. 1987b. Cross-Language switching in stop consonant perception and production by Dutch speakers of English. *Speech Communication* 6:185.

Flege, James Emil. 1990. Age of learning effects and the authenticity of voice-onset time (VOT) in stop consonants produced in a second language. *Journal of the Acoustical Society of America* 89:395-411.

Flege, James Emil., Murray J. Munro, & Laurie Skelton. 1992. Production of the word-final English /t/-/d/ contrast by native speakers of English, Mandarin, and Spanish. *Journal of the Acoustical Society of America* 92:128-143.

Flege, James Emil, Murray J. Munro, and Robert Allen Fox. 1994. Auditory and categorical effects on cross-language vowel perception. *Journal of the Acoustical Society of America* 95:3623-3641.

Flege, James Emil. 1995. Second language speech learning: theory, findings, and problems. In *Speech Perception and Linguistic Experience: Theoretical and Methodological Issues in Cross-Language Speech Research* (W. Strange, editor), pp. 229-273. Timonium, MD: York Press.

Flege, James Emil. 1999. The relation between L2 production and perception. In *Proceedings of the XIVth International Congress of Phonetic Sciences,* ed. by Ohala, John J., et al, 1273. San Francisco,

Flege, James Emil. 2003. Assessing constraints on second-language segmental production and perception. In *Phonetics and Phonology in Language Comprehension and Production,* eds. Schiller, Niels O., and Antje S. Meyer, 319-355. Berlin, New York:Mouton de Gruyter.

Fowler, Carol A. 2003. Speech Production and Perception. In *Handbook of Psychology, Vol. 4, Experimental Psychology,* 238-261. New Jersey:John Wiley & Sons, Inc.

Fowler, Carol A., & Galantucci, B. 2005. The Relation of Speech Perception and Speech Production. *The Handbook of Speech Perception*, Blackwell Publisher, pp. 633-652.

Fowler, Carol A., Valery Sramko, David J. Ostry, Sarah A. Rowland & Pierre Hallé. 2008. Cross language phonetic influences on the speech of French-English bilinguals. *Journal of Phonetics* 36:649-663.

Fox, Robert Allen, James E. Flege, & Murray J. Munro. 1995. The perception of English and Spanish vowels by native English and Spanish listeners: A multidimensional scaling analysis. *Journal of the Acoustical Society of America* 97:2540.

Gafos, A. 2002, A grammar of gestural coordination. *Natural Language and Linguistic Theory* 20:2, pp. 269-337.

Gafos, A. & S. Benus. 2006. The dynamics of phonological cognition. *Cognitive Science* 30, 5, 905-943.

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

Goldstein, Louis, & Carol A. Fowler. 2003. Articulatory phonology: A phonology for public language use. In N. Schiller, & A. Meyer (Eds.), *Phonetics and phonology in language comprehension* (pp. 159-208). Berlin: Mouton-de Gruyter.

Guion, Susan G., James Emil Flege, Reiko Akahane-Yamada, and Jessica C. Pruitt. 2000. An investigation of current models of second language perception: The case of Japanese adults' perception of English consonants. *Journal of the Acoustical Society of America* 107:2711-2724.

Halle, Morris. 1959. The sound pattern of Russian. The Hague: Mouton.

Hamilton, William S. 1980. *Introduction to Russian Phonology and Word Structure.* Bloomington: Slavica Publishers.

Jones, Daniel, & Dennis Ward. 1969. *The Phonetics of Russian.* Cambridge University Press.

Kedrova, G., Anisimov, N., Zaharov, L., & Pirogov, Y. 2008. Magnetic resonance investigation of palatalized stop consonants and spirants in Russian. *Journal of the Acoustical Society of America* 123: 3325.

Kochetov, Alexei. 1998. Labial palatalization in Polish dialects: A gestural account of phonetic implementation*. Proceedings of the Annual meeting of Canadian Linguistic Association*, May 1998. 38-50.

Kochetov, Alexei. 2002. Production, perception, and emergent phonotactic patterns: A case of contrastive palatalization. ed. Glowacka, Dorota. London: Routledge.

Kochetov, Alexei. 2006a. Syllable position effects and gestural organization: Articulatory evidence from Russian. In *Papers in Laboratory Phonology VIII*, Louis Goldstein, Douglas Whalen, and Catherine Best (eds.), 565-588. Berlin/New York: Mouton de Gruyter.

Kochetov, Alexei. 2009. Phonetic variation and gestural specification: Production of Russian consonants. In *Variation in phonetics and phonology,* C. Féry, R. van de Vijver, and F. Kügler (eds.).

Krakow, R. A. 1999. Physiological organization of syllables: A review. *Journal of Phonetics* 10, 160-167.

Labov, W., Ash, S., & Boberg, C. 2006. The atlas of North American English: Phonetics, phonology, and sound change

Ladefoged, Peter & Ian Maddieson. 1996. The sounds of the world's languages. Oxford; Cambridge, MA: Blackwell.

Larsen-Freeman, D. 1997. Chaos/complexity science and second language acquisition. *Applied Linguistics* 18, 2, 141–65.

MacKain, K.S., Best, C.T. and Strange, W., 1981. Categorical perception of English/r/and/l/by Japanese bilinguals. *Applied Psycholinguistics*, *2*(04), pp.369-390.

Marin, S. & Pouplier, M. 2010. Temporal organization of complex onsets and codas in American English: Testing the predictions of a gestural coupling model. *Motor Control* 14(3), 380-407

Nam, Hosung, Louis Goldstein, Elliot Saltzman & Dani Byrd. 2004. TADA: An enhanced, portable Task Dynamics model in MATLAB. *Journal of the Acoustical Society of America* 115(2). 2430.

Nava, Emily. 2010. Connecting phrasal and rhythmic events: evidence from second language speech. Ph.D. Dissertation, University of Southern California.

Ní Chiosáin, Máire & Jaye Padgett. 2012. An acoustic and perceptual study of Connemara Irish palatalization. *Journal of the International Phonetic Association* 42.2, 171-191.

Piske, Thorsten, Ian R. A. MacKay, & James E. Flege. 2001. Factors affecting degree of foreign accent in an L2: a review. *Journal of Phonetics* 29:191-215.

Proctor, Michael. 2009. Gestural Characterization of a Phonological Class: the Liquids. PhD, Yale.

Sancier, M., & Carol A. Fowler,. 1997. Gestural drift in a bilingual speaker of Brazilian Portuguese and English. *Journal of Phonetics*, 25, 421–436.

Saltzman, E. L., & Munhall, K. G. 1989. A dynamical approach to gestural patterning in speech production. *Ecological Psychology,* 1, 333-382.

Schertz, J., T. Cho, A. Lotto, and N. Warner. 2015. Individual differences in phonetic cue use in production and perception of a non-native sound contrast. *Journal of Phonetics* 52.183-204.

Sheldon, A. and Strange, W., 1982. The acquisition of/r/and /l/ by Japanese learners of English: Evidence that speech production can precede speech perception. *Applied Psycholinguistics*, *3*(03), pp.243-261.

Strange, W. 1995. Phonetics of second-language acquisition: Past, present, future. In K. Elenius and P. Branderud (eds) *Proceedings of the ICPhs 95*. Arne Stombergs: Stockholm

Timberlake, Alan. 2004. A reference grammar of Russian. Cambridge; New York: Cambridge University Press.

Zsiga, Elizabeth C. 2000. Phonetic Alignment constraints: consonant overlap and palatalization in English and Russian. *Journal of Phonetics* 28:69-102.

Zsiga, Elizabeth C. 2003. Articulatory timing in a second language. Studies in Second Language Acquisition 25, 399–432.

Zsiga, Elizabeth C. 2011. External sandhi in a second language: The phonetics and phonology of obstruent nasalization in Korean and Korean-accented English. Language 87: 289 – 345

Whalen, D. H. 1999. Three lines of evidence for direct links between production and perception in speech. In Anonymous 1257-1260. San Francisco.