An analysis of perceptual factors in the evolution of Spanish approximants

by

David Harper

B.A., Hiram College, 1994

M.A., University of Colorado at Boulder, 1997

M.A., University of Colorado at Boulder, 2007

A thesis submitted to the

Faculty of the Graduate School of the

University of Colorado in partial fulfillment

of the requirement for the degree of

Doctor of Philosophy

Department of Linguistics

2014

This thesis entitled:

An analysis of perceptual factors in the evolution of Spanish approximants

written by David Harper

has been approved for the Department of Linguistics

____

Dr. Rebecca Scarborough, Chair

____

Dr. Esther Brown

____

Dr. William Raymond

____

Dr. David Rood

____

Dr. Erin Shay

*Date*_______________

The final copy of this thesis has been examined by the signatories, and we find that both the content and the form meet acceptable presentation standards of scholarly work in the above mentioned discipline.

IRB protocol # 13-0083

Harper, David (Ph.D., Linguistics, Department of Linguistics)

An analysis of perceptual factors in the evolution of Spanish approximants

Thesis directed by Dr. Rebecca Scarborough

ABSTRACT

The role of perception in sound change is an open question. Perception may serve a functional purpose in sound change, for example, leading to enhanced distinctions, or it may affect sound change non-functionally as a result of misperception. This dissertation addresses this question by examining the stop/approximant alternation in Spanish. Traditional explanations of this alternation describe it as a lenition process motivated solely by articulatory factors. A widely accepted description of the core alternation is that the underlying stops /b,d,g/ are maintained after pauses and in post-nasal position, and through a lenition process these sounds lenite to [β̞,ð̞,ɣ̞] in intervocalic position and elsewhere. Detailed phonetic descriptions, however, have noted the appearance of approximants in non-leniting environments such as phrase-initial position, and cross-dialectal variation is common. This variation raises questions regarding how approximantization has spread to non-leniting environments, and suggests that a simple articulatory explanation is insufficient.

This dissertation explores how perceptual factors may be contributing to the spread of approximants across phonological environments and across dialects. To test whether enhancement or misperception has played a part in the spread of approximants across Spanish, four perception experiments and a production task were administered to speakers of two dialects of Spanish—one "conservative" (Colombian) and the other "innovating" (Mexican). The production data confirmed that variation is the norm within and across dialects: Approximants appear in non-leniting environments with some regularity, and speakers of Colombian Spanish showed lower rates of approximantization than Mexican speakers. Results from the perception experiments showed that confusability between the two segment types is high; however, there was no clear evidence of a correlation between confusability and likelihood of approximantization. In the experiments that investigated perceptual enhancement, there was no indication that approximants were perceptually advantageous with respect to place of articulation or voicing. As a result, perceptual enhancement was also ruled out as a motivation for this sound change. Overall, the role of perception as a mechanism of sound change was found to be negligible, but results of the experiments point to promising directions for future research concerning the perception/production loop and the nature of stored representations.

Acknowledgements

It has taken a village to complete this dissertation. There are many people from many walks of life that I would like to thank, and I will do my best to remember them all here.

My parents deserve the first mention. Aside from the obvious gift of life that they give me, they have also provided me with unwavering support as I found my way in the academic world. When I announced that I was going to be a Spanish major in college, I remember some gentle prodding about what profession that might lead to, but in the end they were fully supportive. This support was also consistent as I worked towards my dissertation, which must have seemed like a very, very long time. I will always appreciate that when you asked me about my dissertation, I felt like it was out of genuine interest.

Without encouragement from my academic mentors at CU I never would have made it this far. Esther Brown from the Spanish department deserves special mention for her great classes, and by involving me in her own research she made all this seem possible. The Linguistics department provided me not only with consistent funding but also many wonderful people and classes that kept me stimulated. Of course I'd like to especially thank my adviser, Rebecca Scarborough, whose willingness to meet with me weekly and her genuine faith in me were critical in keeping me on track. Also, her great feedback is strongly reflected in the document that follows.

Great departments usually have great students, and I'd like to also thank my fellow graduate students who provided encouragement in many forms. The academic variety of encouragement was provided by groups known only as SCULL and DAWGL. To those of you in these groups, I'd like to give my heartfelt thanks for creating an environment where nascent research could be explored without judgment. Strength was also drawn from the friendships I found in the Linguistics department and the activities that were pursued. In particular, the occasional softball game, ultimate match, and trips to the disc golf course kept me sane.

Lastly, I'd like to thank my lovely partner Maggie Barber, who without questioning watched me stare at my laptop too many evenings. I will always be thankful for your willing ear, your steady encouragement, and your warm embrace. I look forward to spending more time with you and our lovely daughter Carina in the near future.

Table of Contents

List of Tables

List of Figures

CHAPTER 1 — Introduction

1 Introduction

An interesting case of allophonic variation in Spanish is the stop/approximant alternation observable in words such as *dedo* [deð̞o], where in this example the dental /d/ is prescriptively realized as a stop phrase-initially and as an approximant intervocalically. Traditional descriptions of this alternation propose a unitary rule where stops occur in some environments and approximants in others, but detailed phonetic descriptions of this allophonic relation across varieties of Spanish suggest that the range of variation of this alternation is quite broad. Some dialects, such as Miami Spanish, permit approximants in virtually all phonological environments (Hammond 1976), while other dialects such as Colombian, (Eddington 2011, Amastae 1975) are much more restrictive. Traditional explanations of this alternation describe it as a lenition process motivated by articulatory factors; however, given the extent of variation across dialects and the appearance of approximants in non-leniting environments such as phrase-initial position, this explanation cannot be universal. For this reason, this dissertation explores a broader explanation that examines possible perceptual factors influencing the spread of approximants.

Dialectal variation suggests that the spread of approximants is a sound change in progress, and theories of sound change will be used to isolate the mechanisms that cause this spread. Specifically, given previous research's focus on articulatory explanations and—in this author's opinion—their inadequacy, this study will focus on perception-based theories of sound change that to date have been largely ignored for cases such as this. Questions concerning the role of perception in sound change are important for those of us interested in how phonological systems arise. In order to address the question of whether perception has played a role in this sound change, this dissertation was guided by the following questions:

1. **What is produced in the dialects examined?**

    The main focus of this dissertation is the effect of phonetic variation on perception and the role of such effects in sound change; however, production data was also elicited from this study's participants. The reasons for this are twofold:

- Data concerning phonetic variation in the Spanish stop series has been largely impressionistic; much of the data was collected and evaluated without the use of modern speech analysis tools such as Praat (Boersma & Weenick 2011). The collection of data through this project will contribute to the overall understanding of the dialectal variation of the stop series.
- Following Pierrehumbert, it is my belief that "…the classification of stimuli in perception provides data for the probability distributions controlling production" (2003:209). Consequently, it is hypothesized that speakers' production patterns will be correlated with their perceptual patterns. Pierrehumbert's statement may seem intuitively true, but additional research is necessary to confirm that it is true. This study should indicate whether the perception-production loop is as integrated as Pierrehumbert suggests.

2. **What is perceived by speakers?**

Most work on the approximant/stop alternation has focused on production, and little work has been done on allophonic perception. Several authors (e.g. Lavoie 2001**)** have claimed that native speakers do not perceive the difference between the stop and approximant alternants, but to my knowledge no researcher has provided empirical support for this claim. If it is true that speakers do not perceive a difference between these allophones, this would seem to suggest a disconnect between perception and production in the loop mentioned above. Delineating which factors influence perception will be critical in order to determine whether perceptual factors do indeed play a role in this sound change.

3. **How can evidence from the stop/approximant alternation in Spanish be used to test theories of sound change?**

Any sound change in progress can be examined to test theories of sound change. The range of dialectal variation across the Spanish speaking world—for this phenomenon and others—can be seen as a continuum, and at least one other author has proposed that the dialectal variation is indicative of a cline of sound change (Amastae 1995). Across the Linguistics literature, the primary explanations for sound change are articulatory, perceptual, or a combination of the two. Several competing explanations for sound change will be reviewed below, and in line with John Ohala's (1974, 1981, 1993) work the experiments employed here to test theories of sound change will be perceptual in nature.

The questions above have shaped the direction of this study. The remainder of this introduction is organized as follows. In section two I will review previous work on the description of this alternation. In section three I will outline the theoretical framework for this dissertation, focusing primarily on perception-based theories of sound change. Section four will outline the basic methodology for the study, and section five will briefly describe the specific experiments performed to answer the questions above.

2 Background

2.1 Descriptions of /b/, /d/, and /g/ variation

Early descriptive work on the Spanish voiced stop series described the non-stops as spirants, i.e. fricatives, and the phonological process was described as spirantization. Only in the past 20-30 years have authors begun to refer to this process as "approximantization." Before the more widespread use of modern measuring techniques, most authors believed that the continuant alternant was a spirant, but detailed phonetic studies have revealed that these sounds are more accurately described as approximants, and the phonological process is referred to as approximantization. A more detailed phonetic description of the approximants will be offered in the Production study in Chapter 2.

Much of the recent work (Harris 1969, Lozano 1979, Goldsmith 1981, Mascaró 1984, Barlow 2003, etc.) addressing the stop/approximant alternation in Spanish has centered on the question of whether the alternation is a fortition or lenition process. This question relates to the issue of which sound is the 'basic' sound, which in traditional approaches would be referred to as the "phonemic representation." Most traditional analyses of this phenomenon have described the voiced stop as the most basic sound and the approximant as an allophone resulting from a lenition process. A widely accepted description of the core alternation is:

Table 1-1: Traditional phonological description of Spanish spirantization

	Environments
[b, d, g]	After pauses; after nasals
[β, ð, ɣ]	Intervocalically and elsewhere[1]

The standard rule-based descriptions (Navarro Tomás 1967; Harris 1969; Lozano 1979; Mascaró 1984) have also claimed that /d/ is realized as a stop after laterals, but Eddington (2011) claims that /b/ and /g/ are no less stop-like than /d/ in this environment. Dialectal variation is significant and will be described below.

[1] Although the elsewhere condition is often used to determine which allophone is the basic sound, most Spanish philologists would say that the basic sound is the plosive variant.

Phonological processes cross word boundaries fairly easily in Spanish, and in the example below we can see how the same word can be pronounced two different ways. Approximants are most often realized in intervocalic environments, also occur across word boundaries. In the first example, we see a stop realization in phrase-initial position, but in the second example we see that the word-initial bilabial in the same word becomes approximantized when preceded by the clitic that ends in a low vowel:

1. *¿Bañas a la niña?* [**b**aɲas a la niɲa] 'Will you bathe the child?'
2. *¿La bañas?* [la **β**aɲas] 'Will you bathe her?'

While it is not the primary goal of this dissertation to address the question of whether this alternation is the result of a lenition or fortition process, it is relevant that linguists have come to different conclusions regarding this question. The fact that some linguists claim the alternation is the result of a fortition process, while other linguists claim just the opposite is an indication that the nature of the stored representation is controversial, and that more work needs to be done to address this question. Perception experiments may be able to shed light on the true nature of the representation of these sounds. If, for example, all approximantized alternants were perceived as stops, this would lend support to the idea that the underlying representation is a stop. Conversely, if all stops were perceived as approximantized alternants, then it could be argued that the underlying forms are approximants and the phonological process could indeed be described as a fortition process. It was not expected that the results of the experiments outlined below would yield such clear-cut answers, but a clearer picture of the nature of stored representations does emerge. It is my belief that one of the truly interesting questions is in fact what the specifics of the stored representations are, and not whether surface representations are the result of lenition or fortition processes, which is a secondary question.

The dialectal variation of this phenomenon is broad, and aside from a few cases of influence through language contact there are no obvious explanations as to why dialectal differences exist. It has not been explained, for example, why the Miami dialect of Spanish can have approximants in virtually every phonological environment while Honduran Spanish exhibits stops after semivowels and Salvadoran Spanish exhibits stops after all consonants. In Colombia, stops can be heard after all consonants and glides (Fernández 1982; Canfield 1981). Hualde has also said that utterance-initial stops are "sporadically…produced without full occlusion," i.e. as approximantized variants (2005: 141).

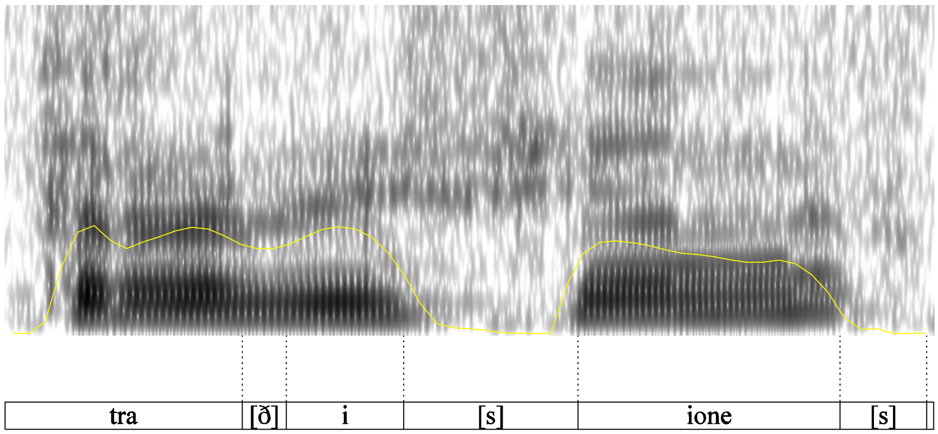
While phonemic splits and mergers across dialects have caught the eye of linguists, allophonic variation has been relatively understudied. Given that sound change most often starts insidiously and proceeds incrementally in the form of allophonic variation, studying the dialectal variation of this phenomenon in Spanish provides a rich testing ground for a number of interesting questions.

2.2 Physical parameters and articulatory motivations for change

In 1964, Peter Ladefoged was the first to coin the term 'approximant,' describing it as a "sound that belongs to the phonetic class vocoid or central resonant oral, and simultaneously to the phonological class consonant in that it occurs in the same phonotactic patterns as stops, fricatives and nasals" (1975:277). In the case of the Spanish approximants, evidence for the vocoid class can be seen in the realization of formant patterns characteristic of vowels, but phonotactically the sounds clearly follow consonantal patterns. In the latest edition of *A Course in Phonetics*, 'approximant' is defined as "an articulation in which one articulator is close to another but without the tract being narrowed to such an extent that a turbulent airstream is produced" (305). Martínez-Celdrán (2004) has suggested that the category of 'approximant' is really a series of at least three subcategories and not a unified category. The three subcategories are laterals, centrals and semi-vowels, but the subcategory that concerns us here is the non-rhotic central category, which in Spanish includes the bilabial [β̞], the dental [ð̞], and the velar [ɣ̞]. The tiny T diacritic under each symbol indicates openness, which places these sounds between the more open category of vowels, and the category of fricatives, which are closed enough to produce turbulence.

In Figure 1-1 below we can see a typical Spanish approximant in the word *tradiciones*. The approximant [ð̞] is marked by the slight dip in the yellow intensity line between the two vowels /a/ and /i/. Formant-like structure is visible, and there is little noise above 3000 Hz. Most studies up until 20-30 years ago described these sounds as fricatives, but if this were a fricative we would expect noise in the higher frequency regions and no formant structure. The first [s] in Figure 1-1 below is a typical fricative, evidenced by aperiodic energy in the higher frequency regions.

Figure 1-1: Spectrogram of *tradiciones*



Intensity ratios have been a favored means of measuring approximants, but for reasons that will be explained in Chapter 2, intensity will only be used as one of several diagnostic tools to determine whether sounds are approximants or stops. Intensity is a relative measure related to amplitude and is calculated in decibels. Compared to the surrounding vowels, the intensity of the approximant in the above spectrogram is slightly less, but will be much more than a true fricative; stops will have zero intensity before any release. The primary advantage of using an intensity measure is that it is a gradient measurement for a gradient variable: the greater the relative intensity level, the more approximant-like the sound. However, as will be described in Chapter 2, there are also drawbacks to relying solely on intensity.

In his article “Physical Parameters Behind the Stop-Spirant Alternation in Spanish,” (1997) Widdison proposes several articulatory and aerodynamic properties that have facilitated this alternation in Spanish. Given that the intervocalic environment is the environment most conducive to approximant realization, it is generally accepted that this is where the first approximants appeared. Widdison refers to the “tenuous relationship between voicing and oral constriction” (78) that makes it difficult to sustain voicing during a closure. One way to sustain voicing is through the venting of P_0 during the constriction phase, which thereby creates a situation of target undershoot—if the target is a stop—and results in approximantization.[2] Given that there is a contrast between voiced stops/approximants and voiceless stops in this position, e.g. /baka/~/baga/, maintaining the voicing contrast would be important.

Across dialects it has been observed that a preceding nasal reinforces a stop realization. Nasals are always homorganic with the subsequent stop, so a tight seal is already in place when the stop is to be realized. Representative examples are below:

Table 1-2: Post-nasal Stops

bilabial	*ambos* [ambos] ‘both’
dental	*ando* [ando] ‘I walk’
velar	*ángulo* [aŋgulo] ‘angle’

A similar argument has also been made for the case of the lateral, which has a dental articulation in Spanish, and a following /d/, which has the same dental place of articulation. Virtually every author (e.g. Amastae 1995, Face 2002) who has studied this phenomenon has claimed that the preceding lateral necessarily and without exception leads to the production of a stop /d/, but Eddington (2011) has claimed otherwise. In his study the lateral was not determined to be a factor in stop realization.

Given that the sounds in question are produced with different articulators, it would be expected that they behave differently. Historical records indicate that the elision of /d/ was frequent, while in the case of /b/ and /g/ elision was unusual (Lathrop 2003). More detailed

[2] As Widdison notes, “Potentiality in language does not always equal reality” (83). While approximantization is the norm in intervocalic position in Spanish, the closely related language of Portuguese maintains voiced obstruents in this position.

explanations of the research on articulatory and aerodynamic differences will follow in Chapter 2.

Other post-consonantal environments, i.e. after fricatives or [ɾ], show high variability, and glides also show significant variation. Amastae (1989) cites Highland Colombian Spanish as an example of a dialect where stops are common in these environments; however, across the Spanish speaking world he claims that spirants are more commonly found:

Table 1-3: Stops in Highland Colombian Spanish

	Highland Colombian Spanish	**More commonly**
post /s/	*desde* [desde] 'since'	*desde* [desðe] 'since'
post [ɾ]	*cargo* [kaɾgo] 'burden'	*cargo* [kaɾɣo] 'burden'
post-glide	*vaivén* [bajben] 'back-and-forth'	*vaivén* [bajβen] 'back-and-forth'

(Amastae 1989)

2.3 Dialectal variation

The breadth of dialectal variation in the Spanish speaking world is significant. Within the field of Spanish phonology, much research has been devoted to the variable realization of /s/, and for many Hispanists this is one of the defining characteristics of a dialect. The phenomenon in question here is much more subtle, and most laypersons would be hard-pressed to identify significant differences in the realization of /b,d,g/ across Spanish dialects. But descriptions from a variety of sources suggest the differences exist, and it is the insidious nature of the variation that makes it interesting. One of the aims of this dissertation is to capture a moment in time where a particular aspect of the language is in a state of flux—a snapshot of synchronic variation that subsequent research can use to chart this language's progress. While this snapshot will not provide a comprehensive picture of all dialects, it will serve as a complement to work that has already been done, while also aiming to improve the descriptive methods. Of the work that has been done, most descriptions appear to have relied mainly on the ear of the investigator, and it is not always clear how the investigators went about determining whether a sound was a stop, spirant, or approximant. This study outlines a comprehensive description of the methods followed to make such determinations. A chart summarizing the observations made in the Linguistics literature regarding the realization of /b,d,g/ in a wide variety of dialects is included in Appendix 1.

2.4 Sound change and dialectal variation

Many authors have commented on the parallels between diachronic change and synchronic variation (e.g. Ohala 1989, Penny 2000). In exploring dialectal variation of the voiced stop series in Spanish, it is the goal of this dissertation to explore links between perceptual motivations for diachronic change and synchronic variation, which may thereby contribute to our understanding of the manner in which dialects have evolved.

Amastae (1995), in his comparison of Colombian Spanish, Mexican Spanish, and Mexican-American Spanish, has proposed that the dialectal differences represent different stages of a sound change in progress. He describes the two Mexican dialects as "innovating" because the approximants are found in a wider variety of environments than in the Colombian dialect. In analyzing four environments that demonstrate wide variability across dialects, Amastae concludes the following:

Table 1- 4: Environments that favor or disfavor approximantization

Environment	Colombian (Bogotá)	Mexican	Mexican-American
Glide	Favors	Favors	Favors
/r/	Disfavors	Favors	Favors
/s/	Disfavors	Favors	Favors
/l/*	Disfavors	Favors	Favors

*Amastae found that the /ld/ sequence always results in a stop, so data for /l/ in the above chart only include /lb/ and /lg/ sequences.

Table 1-4 above serves to point out very general differences and similarities between three different dialects, but as Amastae points out, a closer look at individual segments within each dialect reveals other subtle differences. Amastae's study demonstrates how gradient and variable this phenomenon is, and also how difficult it is to arrive at an all-encompassing, unitary rule that neatly explains the historical trajectory of the phenomenon for all dialects. His study will be explored further in the next chapter.

2.5 The possible role of perception

The variability of the phenomenon and the encroachment of approximants into unexpected environments—specifically phrase-initial position where stops have a strong articulatory motivation—suggest that articulatory mechanisms may not be the only mechanisms at work. In Amastae's study (1995) of three dialects, he proposed that in the course of diachronic change, "…as rule application widens, the formal mechanism is first complication, then simplification" (265). Yet "complication" is not in itself a mechanism, but rather a descriptor of how messy the initial stages of diachronic change can be. What he discovered in comparing three dialects was that features and natural (articulatory) classes are unable to succinctly account for the observed variation, hence the "complication" problem. It is hypothesized here that perceptual factors can account for the "complications" observed at the nascent stages of this diachronic change, and that perception, or misperception, is in fact a mechanism of change.

3 Theoretical Framework

Early literature on sound change presumed a primarily articulatory-motivated approach, where the "lazy" speaker introduced subtle variants in casual speech that were eventually adopted by large groups of speakers. While "casual" speech is still considered by many the most likely source of sound change, efforts have been made in the last 30 years to move beyond the somewhat simplistic “lazy speaker” explanation. Current research aims to identify the mechanisms of sound change, but researchers are divided on a number of fronts. To expand the conversation on sound change, this dissertation will primarily test perception-based explanations for sound change.

Blevins and Garrett (2004) provide a succinct synopsis of the different approaches to sound change that make perception a focal point. The major divide among scholars concerns whether there is a functionalist mechanism to sound change. Blevins and Garrett label the functionalist mechanism “phonetic optimization,” i.e. sound change is driven by ease of articulation or facilitation of perception. Representatives of this camp include Flemming (1996, 2003), Boersma (1998), and Steriade (2001). Flemming, in discussing his Dispersion Theory model, claims that misperception has a very limited role in sound change, stating that it “can only hope to account for neutralization, not dispersion or enhancement… (where) speakers appear to take measures to increase the distinctiveness of contrasts” (2003). In her 2001 article, *Directional asymmetries in place assimilation: a perceptual account*, Steriade claims that assimilation can only take place if both “perceived similarity to the original form and improved functionality” result (242). While this statement is in reference to place assimilation patterns, I believe her views on the Spanish stop/approximant alternation would be similar. Applying her views to the phenomenon studied here, Spanish stops would have evolved into approximants because (a) they are similar enough perceptually, and (b) approximants have an “improved functionality,” either in terms of ease of articulation or perception.

Blevins and Garrett (1998) place themselves on the non-functionalist side of the debate, along with Ohala (1974, 1981, 1993). Blevins and Garrett believe that “reinterpretations of the ambiguities in real speech are the main force driving sound change" (2004:143). In their view, the dispersion effects observed by Flemming and others are a result of chance, as misperception often eventually leads to an "optimized" phonetics.

The most notable champion of the idea that perception—or more specifically, misperception—plays a role in sound change is John Ohala. In numerous publications, Ohala has claimed that the factor primarily responsible for sound change is the listener's mishearing or misparsing of their interlocutor's speech. High variability in the speech signal results in an ambiguous signal, which the listener may misperceive (Ohala 1981). Ohala has been particularly interested in changes involving alternants that are auditorily similar yet articulatorily distinct; in these cases he believes the alternation is motivated by perceptual factors. One such example is the dialectal alternation of English 'with' [wɪθ]~[wɪf]. Ohala would argue that speakers use both forms not due to any articulatory similarity, but rather to their acoustic similarity. In their discussion of metathesis, Blevins & Garrett (2004) discuss examples of metathesis that do not appear to facilitate articulation or perception and are driven instead by misperception. Such examples lead these researchers to conclude that sound change is perceptually driven and non-teleological, and both Ohala and Blevins & Garrett have been highly critical of the teleological assumptions of those from the functionalist camp.

Another crucial difference between the two approaches concerns the role of perception in the grammar itself. While Steriade (2001) has begun to develop a model of perception—the "P-map"—that she believes is part of the grammar, Ohala has not proposed that perception is a formal part of the grammar. For Ohala, innocent misperception contributes to the inception of the change, but no more (Mielke 2003).

The table below outlines in broad strokes the two main positions on the role of perception in sound change:

Table 1- 5: Perception-based Theories of Sound Change

Position	**Characteristics**	**References**
Innocent/Non-functionalist	non-teleological; no improvement in articulation or perception necessary; perception is not part of the grammar; listener-initiated sound change	Ohala (1981, 1990, 1993 *inter alia*); Blevins (2004)
Optimizing/Functionalist	teleological; improved functionality; perceptual knowledge part of a speaker's phonology; speaker-initiated sound change	Steriade (2001); Flemming (2003)

The main goal of this dissertation is to determine whether perceptual factors have led to the spread of approximants in Spanish, and an additional question that must be addressed is whether the perceptual mechanism is a functional one such as dispersion or enhancement, or a non-functional one such as innocent misperception. Articulatory factors have certainly enabled this alternation to arise, but what promotes or constrains the spread of approximants?

4 Methodology

Above I outlined two competing theories of perceptually-motivated sound change—innocent and optimizing. Finding conclusive evidence for either of the two positions has been difficult. In Steriade's 2001 article, she argues that misperception alone cannot be the sole root of the change because "the patterns of perceptual confusion observed in the laboratory do not exactly match attested sound changes." Certainly if there was a one to one match this would be strong evidence that misperception was the sole factor, but this does not imply that the reverse is true: the lack of an exact match does not necessarily entail that "optimization" is the sole motivating force. If we conclude that there are other factors involved, these could be non-optimizing factors such as language contact or sociolinguistic factors such as prestige. This study will attempt to first determine whether perception can be considered a factor in this sound change, and secondly determine whether the role of perception is innocent or optimizing. Some sociolinguistic or extralinguistic factors will be addressed, but because these factors are not the focus of this study they were minimized by employing carefully designed experiments.

Concerning the validity of using perception experiments to test theories of sound change, Ohala has claimed that "When listeners confuse these sounds in listening tests they are, in effect, duplicating sound change in the laboratory" (1989:184). Foulkes (1997) used perceptual experiments to show that misperception can be duplicated in the laboratory by examining the crosslinguistically common change of p>f>h. In particular, the results of his experiments suggest that the change from f>h was facilitated by the lack of acoustic cues on subsequent high back rounded vowels, which resulted in the acoustic similarity between [fu] and [hu]. For this reason he concluded that "the listener, and not the speaker, is primarily responsible" (271).[3] In this project, I will also use synchronic variation to try and duplicate sound change in the laboratory.

5 Subjects & Tasks

In Amastae's 1995 study of three dialects and their approximantization patterns, he concluded that the Colombian dialect was "conservative" while the two Mexican dialects were "innovating." I targeted these two groups for this study because of this contrast. In all, 31 Colombian speakers and 30 Mexican speakers were recruited for this study. A detailed

[3] Brown and Raymond (2012) found that extralexical phonetic context was a significant predictor of this lenition process. Contrary to Foulkes, they found no evidence that lexical phonology played a role.

description of these populations follows in Chapter 2, where results from the production task confirm that Amastae's assessment still holds.

In addition to the production task where subjects were asked to read words from a list, four perception experiments were used to test the hypotheses outlined above. Many perception experiments rely on there being a categorical difference between two items to achieve a clear result. This will generally not be the case when testing the perception of allophonic differences. In the case of allophonic alternants such as [d] vs. [ð̞], simple identification tasks are more complicated. Speakers will not be familiar with IPA, and Spanish orthography will have trained literate speakers to see these sounds as equivalent. The examples below illustrate how orthography can mask fine-grained phonetic distinctions:

- 'dedo' [deð̞o]
- 'bebe' [beβ̞e]
- 'gol' [gol], 'haga' [aɣ̞a]

In order to avoid biasing the subjects, I designed experiments that did not require instruction in IPA or explanations of the stop/approximant alternation. Below I will sketch the experiments that were administered; more detailed information is contained in each chapter.

5.1 Production Task

To enable a comparison between production and the results of the perception experiments, production data was collected from the participants by having them read words from a list; utterances were recorded and later analyzed. Given the results of previous research and descriptions of various dialects, it is clear that certain phonological environments promote and others inhibit approximantization. Not all possible environments could be elicited because this would have been too taxing for the participants, but the list of 84 words spanned the most frequently referenced environments in the literature and also included two variables less frequently included: stress and following vowel.

5.2 Perception Experiments

In this section I sketch out the theoretical motivations and basic experiment design of each experiment. Detailed descriptions of the methods and background are laid out in each individual chapter.

- ***Experiment 1:*** Perceptual discriminability in different phonological environments

One of the pre-conditions for a non-functionalist account of sound change is that sounds are confusable. In this experiment we examined whether certain phonological environments are more confusable than others through a same/different discrimination task. The effect of phonological environment on perception may be due to the presence or lack of cues, or masking effects. If phonological environment does affect perception, it would be expected that confusion patterns closely match production patterns. Such evidence would suggest that perceptual confusion is a contributing factor in the development of this sound change.

- ***Experiment 2:*** Improved functionality of approximants

In *The Phonetic Bases of Phonological Markedness* (2004), Hayes and Steriade discuss "scales of perceptibility." The fundamental idea is that internal or external cues cause feature distinctions to be better perceived in certain segments or phonological contexts than in others. For example, place of articulation is more easily identified in fricatives than stops in the same environment (Wright 2004). While neither Wright nor Hayes and Steriade address approximants directly, it was hypothesized that given approximants' similarity to fricatives and the additional information provided by approximants' formant structure, approximants would be more quickly and accurately repeated than stops in this repetition task. If approximants did elicit faster and more accurate responses, this would suggest that approximants have an “improved functionality” in the sense that Steriade has proposed, and such a result would support a functional explanation for approximantization.

- ***Experiment 3***: Place Identification

The purpose of this experiment was to determine whether place was more readily and more accurately identified in approximants as compared to voiced stops. While experiment 2 examined an overall effect of the functionality of stops vs. approximants, this experiment tested whether the identification of place of articulation, a critical contrastive distinction in Spanish, was improved. The task was a same-different discrimination task using nonce words, where the only difference in the pairs was place of articulation. The comparison of most interest is between stop pairs and approximant pairs. If approximant pairs were found to be more quickly and accurately identified than stop pairs, this would support a functional explanation for sound change.

- ***Experiment 4:*** Confusability between voiced and voiceless pairs

One interesting result from Lewis’ (2001) dissertation on the weakening of voiceless stops is that Colombian speakers, who are known for their relatively conservative dialect in terms of approximant production, are much less likely than Peninsular Spanish speakers to exhibit voicing of intervocalic voiceless stops. This seems to suggest that approximantization and voiceless stop weakening may be linked. If the two processes are shown to be connected, one possible explanation could be that there is a sort of “consonantal chain shift,” where the intrusion of one consonant affects the production of another. Such a situation could lend support to advocates of Dispersion Theory if approximantization were seen as a means of creating perceptual distance from the encroaching voiceless stops that show increased degrees of voicing. By this logic, approximants would have an additional contrast from voiceless stops, and thereby have an improved functionality. To test this theory, this experiment turned back the clock to compare only stops in both phrase-initial and intervocalic positions. Subjects participated in a forced-choice identification task in which they heard a single word, e.g. *pago* [pago], and had to choose between the word they heard and its voiced or voiceless minimal pair, which in this case would be *Paco* [pako]. Stimuli were drawn from a position that favors stops, phrase-initial position, and a position that favors approximants, intervocalic position. Support for the chain shift hypothesis would come in the form of lower accuracy rates in approximant-friendly intervocalic position, indicating that in this position creating more distance between the two sounds would be advantageous.

6 Conclusion

This dissertation contributes insights on an intensely debated issue—the role of perception in sound change. Perceptual factors have often been ignored in research on sound change, but in the last 30 years a growing body of evidence suggests that perceptual factors play an important role in the initiation of sound change. Whether that role is “innocent,” “optimizing,” both or something else entirely remains to be conclusively shown, and the perception experiments outlined above will contribute to the ongoing debate. Related questions concerning the nature of mental representations, the extent to which the perception-production loop is integrated, and the phonetics/phonology interface will also be addressed.

CHAPTER 2 — Production

1 Introduction

While the main focus of this dissertation will concern the effect of phonetic variation on perception and its role in sound change, production data was also elicited from this study's participants. The reasons for this are threefold:

- Analysis concerning phonetic manner variation in the Spanish voiced stop series has been largely impressionistic; much of the data in past studies was collected and evaluated without the use of modern speech analysis tools such as Praat, which allows, for example, for the close inspection of intensity levels and formants. The collection of data through this project will contribute to the overall understanding of the phonetic realization and dialectal variation of the stop series.
- Following Pierrehumbert, it is my belief that "…the classification of stimuli in perception provides data for the probability distributions controlling production" (2003:209). Consequently, it is hypothesized that speakers' production patterns will be correlated with their perceptual patterns. While Pierrehumbert's statement may seem intuitively true, additional research is necessary to confirm that it is true. A first step in this direction is to collect production data.
- Many of the hypotheses for the perception experiments are contingent upon assumptions about production patterns—specifically which phonological characteristics affect approximantization. A detailed analysis of the speakers' production—both within and across dialects— is necessary to inform the experiments' hypotheses.

2 The Speakers

In the study 61 speakers participated—31 from Colombia and 30 from Mexico. One Mexican speaker's data was not included; it was determined that because of advanced age—she was 80— and limited literacy her responses were not reliable. Of the 31 Colombian speakers, 30 were raised in Bogota; the one that was not was born in Cali but had resided in Bogota for an extended period. The 29 Mexican speakers were mostly residents of the United States. The majority were speakers of central Mexican dialects, whose distribution is illustrated below:

Table 2-1: Geographical distribution of Mexican subjects

Zacatecas	Durango	Capital region	Other central cities	Other
15	4	3	3	4

Zacatecas is a municipality in north central Mexico. Durango borders the region and lies to the southwest. Speakers of both the Zacatecas and Durango regions affirmed that their dialects were highly similar, and speakers from near the capital and other central cities also claimed strong dialectal similarities with the most strongly represented regions in this study. The investigator believes that he has collected a dialectally homogenous group of speakers from central Mexico.

Although this study was not intended to be sociological in nature, some demographic data was collected as part of the study. This data is displayed in Table 2-2 below. The majority of the Colombian data was collected on a college campus, which explains why 90% of the Colombian speakers had educational experiences beyond high school. The vast majority of speakers ranged from 18-40 years of age, and for both dialects more women than men participated.

Table 2-2: Demographic data of speakers

	Age<30	Average Age	% of Females	Post-secondary education
Colombia (n=31)	23	25.8	58%	90%
Mexico (n=29)	16	32.5	69%	17%

3 Methods

3.1 The Data and Previous Work

In selecting words to include in the word list, the author focused on phonological environments most often mentioned in previous research as contributing to or inhibiting approximantization. There is an extensive body of work that describes individual dialects and their patternings; Lipski (1994) in particular provided descriptions of many dialects in the Americas. Comparisons between dialects based on previous research should be done judiciously; most older research does not include a detailed description of the methods used to determine how sounds should be classified, and descriptions are usually of the yes/no or favors/disfavors variety—i.e. the environment favors or disfavors approximantization—without qualification of what this means. More recent studies, e.g. Eddington (2011) and Carrasco et al. (2012) use modern measuring techniques and statistical methods to gain a more nuanced picture of differences among dialects and places of articulation. This study takes advantage of a modern speech analysis tool and state-of-the-art statistical modeling to gain a clear picture of how approximantization patterns are realized in both speaker groups.

Of the research that has been done on approximantization in various dialects, the one that most closely relates to this study is Amastae's (1995) comparison of Colombian Spanish, Mexican Spanish, and Mexican-American Spanish. In this study, Amastae proposed that the dialectal differences represent different stages of a sound change in progress, and he analyzed a wide variety of phonological environments to illustrate how each dialect might represent a different stage of the process. He describes the two Mexican dialects as “innovating” because the approximants are found in a wider variety of environments than in the Colombian dialect. Amastae found that four environments in particular showed a fairly dramatic split in whether they favored or disfavored approximantization:

Table 2-3: Environments that favor or disfavor approximantization

Environment	Colombian (Bogotá)	Mexican	Mexican-American
Glide	Favors	Favors	Favors
/r/	Disfavors	Favors	Favors
/s/	Disfavors	Favors	Favors
/l/*	Disfavors	Favors	Favors

*Amastae found that the /ld/ sequence always results in a stop, so data for /l/ in this chart only include /lb/ and /lg/ sequences.

Table 2-3 serves to point out very general differences between three different dialects, but as Amastae points out, a closer look at individual segments within each dialect reveals other interesting differences. For instance, while a voiced stop following a glide favors approximantization across dialects, this is not the case for orthographic ‘b’ in Colombian Spanish. Also, Amastae’s data suggest that there are ordering differences for the four environments in the different dialects. In other words, depending on the dialect and the segment, /r/ may be more likely than /l/ to induce approximantization, or vice versa.[4] Amastae’s study

[4] While Amastae’s study certainly highlights important differences among the three dialects, it does have a few shortcomings. It is never mentioned how the manner was determined, and it is not always clear whether putative

highlights how inadequate the Spanish spirantization 'rule' is, and was one of the motivations for investigating the mechanisms that control the variable realizations of this sound pattern.

In collecting the data, phonological environments were selected based primarily on the many previous descriptions of the allophonic variation of the Spanish stop series. It was not practical to elicit every possible phonological combination from the participants, so for ease of comparability with other studies the most commonly cited environments in other studies are included here. In Amastae's study, his analysis did not find significant differences for post-glide production in the three dialects, so this environment was not selected for analysis. In most places of articulation in Amastae's study, the lateral and the alveolar flap pattern together; because the lateral is often described in other studies it was chosen over the alveolar. Subsequent vowel—a factor uncommonly included in other studies—was included here because of the potential aerodynamic effects caused by differences between high and non-high vowels. Sandhi pairs were also included to test whether word boundaries affect approximantization rates. The preceding environments that were ultimately examined are:

- phrase-initial position, e.g. ***v**oté*
- intervocalic position (consonant always preceded by /a/ or /e/), e.g. *na**d**ar*
- post-lateral position, e.g. *col**g**ar*
- post-nasal position (i.e., [mb], [nd], [ŋg]), e.g. *tu**m**ba*
- post-fricative position, i.e. post /s/, e.g. *des**d**én*

Of the 84 words used as elicitation stimuli, 73 were disyllabic, 9 had three syllables, and 2 had four syllables. All of the phones considered in this study appear in syllable-initial position, which is always the syllable position for bilabials and velars in Spanish. Dentals will occasionally be

categorizations for the dialects are significantly different. In some cases, the factor weights in the rankings are extremely close.

found in word-final, i.e. syllable-final, position, but not in this study. As mentioned above, vowels following the target consonant were also included as a variable in this study. Spanish has five vowels, and all were included in post-consonantal position. The five vowels are /i/, /e/, /a/, /o/, and /u/.

Sandhi pairs, e.g. *el **d**año/la **d**ama*, were created by including the male and female definite articles, i.e. *el* and *la*, before word-initial consonants. The motivation for including sandhi contexts was twofold:

- The majority of words in a language are not phrase-initial in natural speech, and as a result most word-initial consonants are affected by the word-final environment of the preceding word. Spanish is known for allowing phonological processes to cross word boundaries. By including sandhi pairs, I hope to make this small study somewhat more generalizable to natural language.
- Studies have shown mixed results as to whether sandhi contexts differ at all from environments found word-internally. Carrasco (2008) found no statistically significant differences when comparing word-internal and word-initial environments in Madrid Spanish, but did find differences in Costa Rican Spanish. This study hopes to contribute to that discussion.

Stress has also been shown to influence approximantization. Carrasco et al. (2012) found that all three phonemes showed increased approximantization rates in unstressed positions. The authors in Carrasco's study do, however, caution that if approximantization is determined by using intensity ratios the higher intensity values of stressed vowels may distort ratios somewhat. This argument against using intensity ratios will be considered later in the Annotation Guidelines section.

Two variables that were included in analyses but were not found to be significant are gender (of speaker) and word frequency. Gender has been shown to affect production patterns (Labov 1990), but for these two dialects there were not significant differences. Word frequency or the frequency in which sounds occur in leniting environments has been shown to affect approximantization rates (Brown 2013). Frequent affixes, e.g. the past participle suffix *-ado*, have also been shown to affect approximantization rates (Bybee 2002); for this reason none of the phones appear in affixes. In this study word frequency was addressed a priori by selecting words that did not differ greatly in frequency, but finding words of equal frequency across all phonological environments is difficult. To determine whether the word frequency of words selected for this task was a factor in approximantization, this variable was included in the initial regression analysis that will be discussed below. A complete word list and corresponding frequencies taken from the Corpus del español are listed in Appendix 2.

3.2 Data Collection

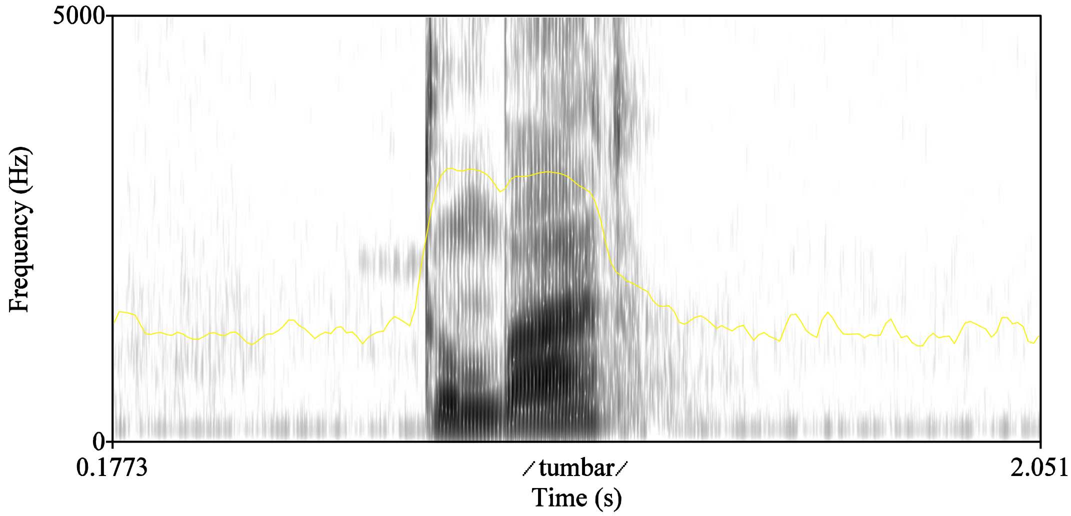
Data were elicited using PsychoPy (Peirce 2007), an open source application that was used both for data collection and administration of the four experiments. Participants were shown a stimulus on the screen and had 3.5 seconds to say the word aloud. Collection was via a head-mounted microphone and their responses were recorded with the PsychoPy software. Incorrect pronunciations were not included in the analysis. The majority of the incorrect pronunciations involved issues of stress. Spanish has many pairs of words where stress determines meaning, e.g. *beso* ‘I kiss’ vs. *besó* ‘he/she kissed.’ The lack of context in the word list presented more problems than anticipated, and speakers occasionally misread these words.

3.3 Annotation Guidelines

In order to determine the production patterns of approximantization, the investigator analyzed each individual token in Praat and determined whether the token was either a stop or an approximant. Many studies (e.g. Carrasco et al. 2012, and Hualde et al. 2011) have made use of the measure of relative intensity to determine degree of approximantization. This measure is calculated by measuring the difference between the intensity at its lowest point in the consonant and at its highest point in the subsequent vowel. Approximants are generally characterized by much smaller intensity differences than stops. While this measure works well for intervocalic position, it works less well for consonants that follow nasals and laterals, whose own intensity levels seemed to strongly affect the intensity level of following consonants. In post-fricative environments it was also difficult to determine the manner when only considering relative intensity. Also, as mentioned above, whether the subsequent vowel is stressed or unstressed can affect relative intensity given that stressed vowels have a higher intensity value than unstressed vowels. It was ultimately decided that relative intensity could be used as supporting evidence for determination of manner, but that other factors should also be considered and in fact prioritized. Also, exact measurements of intensity were not deemed necessary for this portion of the study, as it was determined that patterns of approximantization and differences between dialects could be established without employing relative intensity measurements. The series of criteria that were established are outlined below, but first I will provide two images—the first an example of a post-nasal stop and the second an example of a phrase-initial approximant.

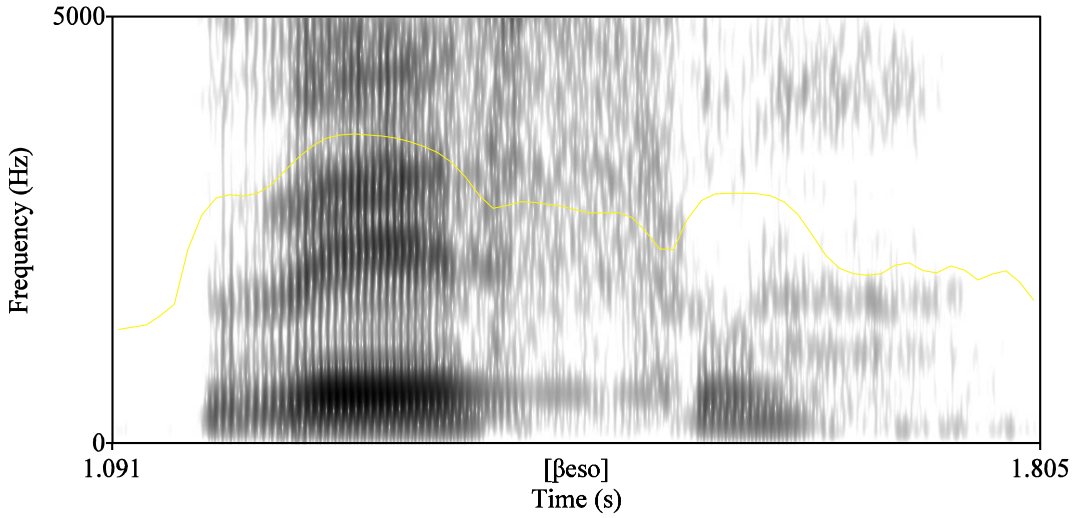
The first image in Figure 2-1 is a spectrogram of a Mexican speaker's pronunciation of *tumbar*. There is clear evidence of a stop bar for the bilabial, but the intensity contour shown by the yellow line hardly dips in the transition from the nasal to the stop. Using a relative intensity measure for such an example would suggest that this is an approximant, but both audibly and visibly it is unmistakably a stop.

Figure 2-1: Mexican speaker's production of *tumbar* [tumbar]



In the second image in Figure 2-2 we see an example of a phrase-initial approximant in the word *beso* produced by a Mexican speaker. The formants are very salient in the approximant, and the intensity contour is at 2900 Hz. Notice that this intensity level is almost exactly the same as the unstressed /o/ at the end of the word. In this instance, relative intensity can be used as supporting evidence that the phrase-initial phone is in fact an approximant.

Figure 2-2: Mexican speaker's production of *beso* [β̞eso]



For the reasons outlined above, relative intensity was only used as supporting evidence in this study, and a series of criteria were established to categorize the production data:

1. Look for stop release. If stop release is present→ stop.
2. If stop bar is present but it sounds like a fricative→ stop. It was determined that some stops were made with the tongue in forward position. This resulted in a laminal stop that had some frication.
3. If in intervocalic position, look for closure with no energy. If there is closure with no energy→ stop.
4. Look for energy or formants. Formants are a clear indication of approximantization. If formants are present in the consonant→ Approximant.
5. If there is no stop release and little energy during phonation, this is most likely indicative of fricativization. In this case, the phone will be classified as "approximant."

Regarding point five, there were clearly some tokens that were neither stops nor approximants, but more resembled fricatives. This is to be expected of a phenomenon that has a wide range of variation, and was most commonly observed in words that contained orthographic ‘v’. Even though “continuant” might be the more appropriate all-encompassing term, I will use “approximant” because this is the standard term in the field.

4 Results

In the following sections I will first present descriptive statistics of the data before exploring through statistical analysis whether the perceived differences are significant, insignificant, or are modulated by other variables. Linguists have long been aware of the fact that in linguistic data there are multiple factors affecting a single process, and the field has seen continuous advancement in its methods to try and capture the factors that shape the inherent variability of language. Descriptive statistics will give us some important first impressions, but ultimately the only way to analyze data that has multiple factors is to use advanced statistics.

Analysis in this and other chapters will be carried out with regression models that include mixed effects. Regression models can be used to analyze both discrete and continuous data, and in this chapter I will predict a binary dependent variable with values of *approximant* or *stop*. Binary variables can be modeled using logistic regression with mixed effects. In this case, the manner—approximant or stop—is the dependent variable. The fixed effects are predictors that may or may not affect approximantization. Mixed effects models have become increasingly popular because they address a common problem in Linguistics and Psycholinguistics: Not all observations are independent. Truly independent subjects and items, i.e. words or stimuli, would not be repeated in a perfect study, but the nature of linguistic experimentation makes this necessary. For both my production data and my experiments the assumption of independence is violated for subjects and items: subjects repeated tasks, and items were repeated across subjects. For this reason, subjects and items will be included as random effects in all statistical models. By introducing random effects, fewer Type 1 errors, i.e. false positives, will result, and the model will in effect be more robust (Hay 2011). By capturing both by-subject and by-item variation the model actually has a higher threshold for significance, and we can be more confident in the results.

Below, I will first show that each of the three phonemes /b,d,g/ have different patterns of approximantization and are best analyzed separately. I will also discuss whether the sandhi pairs behave differently from word-internal environments. The majority of the exposition will be dedicated to describing the factors that contribute to approximantization for each place of articulation.

4.1 Place of Articulation

In Table 2-4 below, I present the overall approximantization rates for each place of articulation and dialect. Overall, bilabials have the highest rate of approximantization and velars have the lowest rate. Dentals fall squarely in between. The Colombian speakers approximantize much less frequently than the Mexican speakers when producing bilabials and velars, but in the case of dentals the difference is smaller. This result suggests that the effect of dialect varies according to place of articulation.

Table 2-4: Overall approximantization rates by place of articulation and dialect

Phone	Dialect	Approximant	Stop	# of Tokens
b	Colombian	40.29%	59.71%	839
	Mexican	51.39%	48.61%	755
b Total		**45.55%**	**54.45%**	**1594**
d	Colombian	36.91%	63.09%	829
	Mexican	38.58%	61.42%	744
d Total		**37.70%**	**62.30%**	**1573**
g	Colombian	25.97%	74.03%	828
	Mexican	37.45%	62.55%	721
g Total		**31.31%**	**68.69%**	**1549**
Grand Total		**38.25%**	**61.75%**	**4716**

Given that the three phonemes in question employ different articulators, it is not surprising that they behave differently. Historical records indicate that the elision of /d/ was frequent, while in the case of /b/ and /g/ elision was unusual (Lathrop 2003). Widdison attributes this to the nature of the articulators: while the lips and the back of the tongue are slow-moving, the tongue tip is quick and mobile. The slower articulators maintain a longer constriction than the tongue tip, which means that the dental's gesture may be "superficial at best and the auditory impression quite difficult to resolve" (80). In an oft-cited study, Ohala and Riordan (1979) point out that the size of the oral cavity behind the constriction also affects the aerodynamics of voicing. The bilabial plosive /b/, having the largest cavity behind the constriction, is voiced the longest. The velar plosive /g/ is voiced the shortest length of time, and the dental /d/ falls in between. Given the difference in articulators, it would be surprising if the frequency of approximants were the same among the three phone pairs. This appears to be the case in this study, as will be shown below.

4.2 Phonological Environment

Table 2-5 below focuses on approximantization rates in five different phonological environments. Intervocalic position and post-nasal position are at the two ends of the approximantization spectrum, as intervocalic position has by far the highest rates of approximantization while post-nasal position has the lowest. Note, however, that variability is the norm and that only one cell—velars in post-nasal position for Colombian speakers—has a value of 0%.

Table 2-5: Effect of Phonological Environment on Approximantization Rates

Place	Dialect	Overall	Phrase-initial	Intervocalic	Post-lateral	Post-nasal	Post-fricative
/b/	Colombian	40.29%	14.69%	85.47%	16.48%	5.56%	33.87%
	Mexican	51.39%	26.02%	84.23%	66.67%	3.33%	75.86%
/d/	Colombian	36.91%	8.72%	89.96%	6.67%	1.67%	27.42%
	Mexican	38.58%	13.95%	85.36%	10.34%	8.93%	44.44%
/g/	Colombian	25.97%	9.42%	58.66%	12.09%	0.00%	9.68%
	Mexican	37.45%	18.37%	69.41%	38.37%	3.45%	26.79%

Phrase-initial position shows a very low rate of approximantization, but like all categories there is variation. For all places of articulation the Mexican dialect shows a higher rate of approximantization than the Colombian dialect. Bilabials show the highest rate of approximantization, although this is complicated by the fact that orthographic ‘v’ was pronounced as a fricative or approximant much more frequently than orthographic ‘b’. Orthographic ‘v’ was found in intervocalic and phrase-initial positions, but not in post-lateral, post-nasal, or post-fricative positions. The role of orthography in the production of bilabials will be described below.

Intervocalic position shows the highest rate of approximantization, but again we see variation by place of articulation. Bilabials and dentals both show about the same rate, and in this environment Colombian speakers approximantized at a slightly higher rate than the Mexican speakers. Velars show a much lower rate of approximantization, and in this case Mexican speakers show a higher rate of approximantization.

Post-lateral position shows very different patterns for the two dialects. Colombian speakers are unlikely to approximantize across places of articulation, while Mexican speakers show significant variability. Dentals show relatively similar low levels of approximantization, while the rate of approximantization for bilabials and velars is quite different. Mexican speakers approximantize bilabials more often than not, and velars are approximantized much more frequently in Mexican Spanish than in Colombian.

Post-nasal position shows the most consistency across places of articulation and dialects. Nasals are the strongest inhibitors of approximantization for all places of articulation and all dialects.

Post-fricative position shows significant differences between the three places of articulation and the dialects. The Mexican dialect strongly favors approximants for bilabials, and the Colombian dialect also shows the highest rate of approximantization for bilabials. Dentals also show a fairly high rate of approximantization compared to other non-intervocalic positions, and velars show a moderate rate of approximantization.

In table 6 below, stress does not appear to affect dentals or bilabials, while velars show higher approximantization rates for unstressed syllables. This table can actually serve as a cautionary tale for why raw percentages are not a good way of doing analysis. In this table the conflation of variables and the unbalanced nature of the data set lead to a misleading representation of whether stress affects approximantization. The logistic regression analyses that follow will show that for all places of articulation, stress is a significant predictor in the model. This will be discussed further below.

Table 2-6: Effect of Stress on Approximantization rates

Place	Dialect	Overall	Stressed	Unstressed
/b/	Colombian	40.29%	40.91%	39.52%
	Mexican	51.39%	50.97%	51.91%
/d/	Colombian	36.91%	36.32%	37.60%
	Mexican	38.58%	38.46%	38.72%
/g/	Colombian	25.97%	24.16%	28.08%
	Mexican	37.45%	35.52%	40.00%

Following vowel is rarely mentioned in the literature as a contributing factor to approximantization. Given that vowel height is a distinguishing factor in the 5-vowel Spanish system, it seems plausible that raising and lowering of the tongue in anticipation of the vowel might affect approximantization. Anticipatory lowering for /a/, for example, could result in more airflow and consequently increase the likelihood of approximantization. For ease of comparison I include below all five vowels with the two environments that appear most frequently and where all vowels are represented, intervocalic position and phrase-initial position. Eliciting words for each environment configuration would have greatly elevated the number of words, and some sequences, for example /ldu/, are very infrequent in Spanish. Dialect is not included in the chart below for ease of presentation and because the rates in both dialects did not differ from those described above.

Table 2-7: Approximantization and following vowel

Following Vowel	Position	/d/	/b/	/g/	Totals
/i/	Intervocalic	83.47%	77.05%	46.79%	69.89%
	Phrase-initial	3.36%	20.34%	1.72%	8.50%
/u/	Intervocalic	87.67%	72.03%	60.68%	71.43%
	Phrase-initial	1.69%	26.67%	5.66%	15.14%
/e/	Intervocalic	90.00%	93.62%	59.14%	81.18%
	Phrase-initial	7.50%	5.08%	2.06%	5.07%
/o/	Intervocalic	84.62%	88.54%	68.32%	80.40%
	Phrase-initial	2.65%	23.81%	8.49%	11.42%
/a/	Intervocalic	93.33%	95.80%	82.20%	90.48%
	Phrase-initial	26.07%	18.69%	30.00%	26.27%

Table 2-7 suggests that vowel height may be a predictor of approximantization. The low vowel /a/ has the highest average approximantization rate for both intervocalic and phrase-initial position, and /i/ and /u/—the two high vowels—have the lowest approximantization rates. In the logistic regression models below, I will examine whether this difference is significant.

4.3 Summary

The data presented above reveal differences along several parameters: Place of articulation, phonological environment, and stress all appear to affect approximantization for these two dialects. In order to determine exactly which variables play significant roles, more sophisticated analysis is required. I will now turn to the logistic regression analysis to further explore how all of these variables interact with one another.

5 Statistical Analysis of Results

It is the nature of linguistic data that multiple factors affect a single process. The data described above suggest that there are multiple factors affecting approximantization, but the interaction of factors cannot be determined by looking at mere percentages. It should also be mentioned that the data itself is unbalanced, i.e. not all possible combinations of factors are represented equally. Post-nasal, post-lateral, and post-fricative environments, for example, are limited to a single following vowel, while intervocalic position has all five. Also, there are slight variations in the number of accepted tokens for each word. While these imbalances can distort results like those above, modern statistical methods like logistic or linear regression deal with this problem without difficulty. Descriptive statistics do, however, serve to gain initial impressions of the validity of hypotheses, provide a basis for building regression models, and are generally more digestible than advanced statistical analyses.

Binary dependent variables can be modeled using logistic regression with mixed effects. In this study on production, the manner—approximant or stop—is the dependent variable. The explanatory variables are all the predictors that may or may not affect approximantization. Mixed effects models have become increasingly popular because they address a common problem in Linguistics and Psycholinguistics: Not all observations are independent.

In the following sections, I will present the results from mixed effects logistic regression models. Overall, many different theoretically- and linguistically-motivated models were run, but ultimately only the models that best fit the data are presented here.

For those not familiar with regression models, the output from a logistic regression model is not readily interpretable. In order to facilitate the basic understanding of the models below, I provide here an example with some guidelines for interpretation. The sample output from below is the first few lines of the model that is primarily concerned with whether the three places of articulation are different from one another. Essentially, regression models are comparisons between categories of a variable; they tell us whether the categories are significantly different from one another while holding all other variables constant. Here are the first few lines from the output of a logistic regression model run with the lme4 package (Bates et al. 2011) in R (R Core Team 2013). There are other variables in the model and this output is only for illustrative purposes:

Table 2-8: Sample Logistic Regression Table

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	4.4105	0.4688	9.409	<2e-16 ***
d	-1.0506	0.3361	-3.1250	.00178 **
g	-1.7119	0.3492	-4.9029	.48e-07 ***

The intercept is always provided and is not generally interpreted; it is the value of the dependent variable when all independent variables are zero. The two categories listed here, “d” and “g”, are two categories from the single variable place of articulation, which also includes “b”. The bilabial is the default category in this case; all defaults are dummy-coded with a value of zero and are never listed in the model. The dental and velar categories are both interpreted in relation to the default using the coefficients, which are listed under the column “Estimate.” The coefficient indicates the increase or decrease in log odds of approximantization. For “d,” approximantization decreases by a log odds of 1.0506 compared to the default “b,” and in the rightmost column we can see that this difference is significant because the p-value is less than .05 (p=.00178). In the case of the velar, its coefficient of -1.7119 indicates an even greater difference from the bilabial.

One reason why logistic regression models are somewhat opaque is because the coefficients are expressed in log odds. An easier way to view the results of a logistic regression model is described by Hay (2011), where any configuration of independent variables can be calculated by summing the estimated log odds and then using the inverse logit function (exp(x)/(1+exp(x)) to derive probabilities. Relevant probabilities will be presented in charts to better visualize differences between different configurations.

In the following two sections I will assess whether two variables should be included or not in subsequent models. Many studies separate place of articulation because the patterns of each are somewhat different, and I will show that this is the case in this study as well. I will also look at whether including sandhi pairs as a separate variable, or whether joining these pairs with post-lateral or intervocalic position, makes the most statistical sense. Finally I will analyze each place of articulation separately.

5.1 Place of articulation

The descriptive data above suggest that each place of articulation has somewhat different approximantization patterns. This is not surprising given the very different nature of the articulators: lips for the bilabials, tip of the tongue for the dentals, and back of the tongue for the velars. The results of the model presented in Table 2-9 below address the question of whether there are significant differences between the three different places of articulation in this study. By setting /b/ as the default value for place of articulation we can see that it differs significantly from both /d/ ($p=.00178$) and /g/ ($p <.001$) while holding all other variables constant.

Table 2-9: Place of articulation and differences in production

```
Number of obs: 4716, groups: Word, 84; participant, 59

Fixed effects:         Estimate   Std. Error z value  Pr(>|z|)
(Intercept)              4.4105     0.4688   9.409    < 2e-16 ***
targetd                 -1.0506     0.3361  -3.125    0.00178 **
targetg                 -1.7119     0.3492  -4.902    9.48e-07 ***
Colombian               -0.9853     0.2528  -3.898    9.69e-05 ***
Stress                  -0.4280     0.2471  -1.732    0.08331 .
Vowele                  -0.5558     0.4139  -1.343    0.17933
Voweli                  -1.1571     0.4350  -2.660    0.00781 **
Vowelo                  -0.2558     0.4357  -0.587    0.55712
Vowelu                  -1.3735     0.4191  -3.277    0.00105 **
Context1Lateral         -4.0243     0.4771  -8.435    < 2e-16 ***
Context1Nasal           -7.1838     0.6261 -11.474    < 2e-16 ***
Context1PostFricative   -3.2014     0.5039  -6.353    2.11e-10 ***
Context1PhraseInitial   -6.3089     0.3944 -15.997    < 2e-16 ***
```

The fact that both /d/ and /g/ are significantly different from the default value /b/ shows that both the dental and velar have different patterns of approximantization from the bilabial. That the coefficients for /d/ and /g/ are fairly far apart— -1.0506 and -1.7119 respectively—suggests that /d/ and /g/ are also different from one another. Also, as will be discussed below, bilabials were affected by orthography, and this does not pertain to the dentals or velars. For these reasons, I will analyze each place of articulation separately and later evaluate their similarities and differences.

5.2 Sandhi

In the production data, each place of articulation included two word pairs that addressed sandhi contexts. It was thought that rates of approximantization in sandhi environments might differ from word-internal rates of approximantization in the same environment. As mentioned above, previous studies have indicated that this effect may be dependent on dialect. Pairs were created that highlighted an extremely common sequence in Spanish—the feminine and masculine definite articles followed by a noun. These pairs are listed in Table 2-10:

Table 2-10: Sandhi Words

	Intervocalic Sandhi	Post-lateral Sandhi
/b/	*la burra*	*el burro*
/d/	*la dama*	*el daño*
/g/	*la gata*	*el gato*

Table 2-11 below shows the rates of approximantization for intervocalic and post-lateral positions without including the sandhi word pairs, and the approximantization rates for the above sandhi word pairs.

Table 2-11: Sandhi approximantization rates

	Intervocalic	Intervocalic Sandhi	Post-Lateral	Post-Lateral Sandhi
/b/ Colombian	87.69%	65.52%	22.58%	3.45%
/b/ Mexican	86.21%	67.86%	68.97%	62.07%
/d/ Colombian	89.60%	93.10%	8.20%	3.45%
/d/ Mexican	83.81%	96.55%	12.07%	6.90%
/g/ Colombian	56.69%	75.86%	14.52%	6.90%
/g/ Mexican	68.14%	79.31%	37.93%	39.29%

Table 2-11 suggests there may be some differences among preceding environments involving sandhi environments and word-internal environments—in particular for bilabials. The difference between intervocalic word-internal environments and intervocalic sandhi pairs is ~20% for each dialect. To test whether including an additional variable improves the model enough to justify its inclusion, two separate models differing by a single factor can be run and then they can be compared using an ANOVA (Raudenbusch & Bryk, 2002). For each place of articulation two models were run—one with Sandhi as an explanatory variable and one without. The two models were then compared by running an ANOVA to see if the more complex model that included the Sandhi variable resulted in an improvement in log likelihood, which would indicate a better fit. Despite the apparent differences in the table above, the resulting ANOVAs were all insignificant, which suggests that Sandhi is not a useful predictor. For this reason, Sandhi pairs have been incorporated into all subsequent models as either post-lateral or intervocalic tokens.

The effect of following vowel is one possible explanation for why the sandhi pairs did not merit separate inclusion. In intervocalic position, all possible following vowels were included, yet in sandhi pairs bilabials were followed by /u/ and dentals and velars by /a/. In Table 2-12 below we compare intervocalic rates of approximantization from Table 2-7 to intervocalic sandhi approximantization rates, and we can see that these rates are all quite similar. This suggests that the lower rates for bilabial sandhi pairs may be attributable to the use of /u/ as a following vowel, while for intervocalic /d/ and /g/ the higher rates for the sandhi pairs may be due to the following low vowel /a/. As will be shown below, low vowels were found to be a significant promoter of approximantization.

Table 2-12: Sandhi approximantization rates and following vowel

Following Vowel	Position	/d/	/b/	/g/	Totals
/i/	**Intervocalic**	83.47%	77.05%	46.79%	69.89%
	Intervocalic Sandhi	n/a	n/a	n/a	n/a
/u/	**Intervocalic**	87.67%	**72.03%**	60.68%	71.43%
	Intervocalic Sandhi	n/a	**66.7%**	n/a	n/a
/e/	**Intervocalic**	90.00%	93.62%	59.14%	81.18%
	Intervocalic Sandhi	n/a	n/a	n/a	n/a
/o/	**Intervocalic**	84.62%	88.54%	68.32%	80.40%
	Intervocalic Sandhi	n/a	n/a	n/a	n/a
/a/	**Intervocalic**	**93.33%**	95.80%	**82.20%**	90.48%
	Intervocalic Sandhi	**94.8%**	n/a	**77.6%**	n/a

5.3 /d/

The factors shown in Table 2-13 below include four independent variables. In creating a mixed model, random intercepts for subject and word, and by-word random slopes for dialect and by-subject random slopes for Stress, Vowel and Context were also included. We followed the recommendation by Barr et al. (2013) to "keep it maximal" when including random effects. By this they mean that all relevant random effects that are associated with the fixed effects should be included, and this was the course followed here. In this experiment, random intercepts for item and subject were included, as well as random slopes for dialect grouped by item, and stress, vowel and context grouped by subject. All of the variables are dummy-coded, which means that the default is coded as 0 and other categories within the variable are coded 1. Table 2-13 below outlines the four variables and states what their default values are:

Table 2-13: /d/ variables

Variable Name	**Default Category**	**Other Categories**
Dialect	Mexican Dialect	Colombian Dialect
Stress	Unstressed	Stressed
Vowel	/a/	/e/, /i/, /o/, /u/
Context	Intervocalic	Post-Lateral, Post-Nasal, Post-Fricative, Phrase-Initial

Both Dialect and Stress are two-way variables, while the other variables, Vowel and Context, are each 5-way variables. All of the variables are dummy-coded, which means that the default is coded 0 and other categories within the variable are coded 1. In the case of Vowel and Context, the significance of each of the four categories listed as fixed effects can only be interpreted as different from the default values, /a/ and Intervocalic respectively.

Using Raudenbusch & Bryk's (2002) technique of comparing models with an ANOVA, it was ultimately determined that the best-fitted model included all of the variables listed above as well as an interaction between Dialect and Context. The model displayed and analyzed below was compared against several models, e.g. the simple model with no interactions, a model that included interactions between dialect and all other variables, a model that included an interaction between Stress and Context, and several others. One of the outcomes of these tests is that the variables Stress and Vowel are not involved in any significant interactions. Below I will describe the results of the model and provide interpretations of each variable. For variables involved in interactions I will provide visual representations that show exactly where the interactions take place.

Table 2-14: Logistic Regression Results for Dental Production

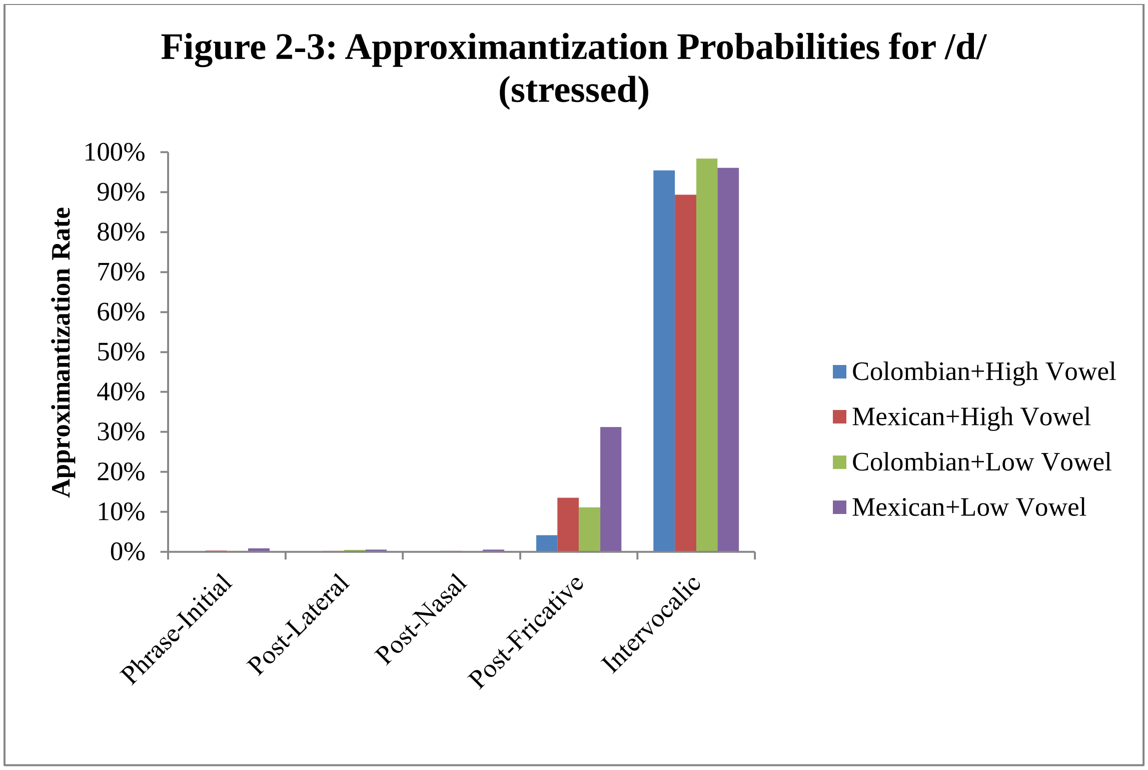
```
Number of obs: 1573, groups: participant, 59; Word, 31

                         Estimate   Std. Error  z value    Pr(>|z|)
(Intercept)               3.8355     0.5477      7.003     2.50e-12 ***
Stress                   -0.6356     0.3284     -1.936     0.05293
Vowele                    0.1424     0.5526      0.258     0.79673
Voweli                   -1.0678     0.5391     -1.981     0.04763 *
Vowelo                   -0.9805     0.5458     -1.797     0.07240
Vowelu                   -1.0383     0.5504     -1.886     0.05923
Dialect(Colombian)        0.9231     0.6033      1.530     0.12602
ContextLateral           -8.4413     1.0995     -7.677     1.63e-14 ***
ContextNasal             -8.4061     1.0997     -7.644     2.11e-14 ***
ContextPostFricative     -3.9898     0.7179     -5.558     2.73e-08 ***
ContextPhraseInitial     -8.0001     0.7304    -10.953     < 2e-16 ***
Dialect:Lateral          -1.1253     1.5095     -0.746     0.45596
Dialect:Nasal            -2.8365     1.8837     -1.506     0.13212
Dialect:Fricative         -2.2154     0.8447     -2.623     0.00872 **
Dialect:PhraseInitial    -2.1547     1.2597     -1.710     0.08718
```

I present above the full model of the fixed effects as it is presented in R. The two columns of primary interest are "Estimate," which indicates the value of the coefficient of the indicated variable, and Pr(>|z|), which is the p-value. P-values of less than .05 indicate that the given category is significantly different from the default category. Because the dependent variable, approximantization, was dummy coded (0=stop, 1=approximant), negative coefficients should be interpreted as having an inhibitory effect on approximantization.

Both Dialect and Stress are two-way variables, while the other variables, Vowel and Context, are each 5-way variables. In the case of Vowel and Context, the significance of each of the four categories listed as fixed effects can only be interpreted as different from the default values, /a/ and Intervocalic respectively. I will first discuss the two variables that are not involved in interactions before turning to Context and Dialect, which must be discussed in tandem with the interaction effects.

- **Stress**: Stress is a marginally significant predictor in this model (p=.05293). When /d/ is found in a stressed syllable, the log odds of approximantization decreases .6356.
- **Vowel**: The high front vowel, /i/ (p=.04763), is the only vowel that is below the significant threshold of .05, and can be considered significantly different from /a/. However, both of the back vowels /o/ (p=.07240) and /u/ (p=.05923), are marginally significantly different from the default vowel /a/. The vowel /e/ (p=.79673) is not significantly different from /a/, and given that its coefficient is very near zero, can be considered very similar to /a/. The high vowel and the two back vowels have very similar negative coefficients, which indicates that they all have a similar inhibitory effect on approximantization. To take one example, the vowel /u/ decreases the log odds of approximantization by 1.0383. Given that /a/ and /e/ are not significantly different from one another and the remaining three vowels all have negative coefficients, it would seem that the high and back vowels are less likely to approximantize than the other vowels.
- **Dialect & Context**: For dental /d/, all of the Context categories are significantly different from the default category, intervocalic position. The high negative coefficients for post-lateral, post-nasal and phrase-initial positions indicate that they are all very different from intervocalic position, and as will be shown below they exhibit a very low likelihood of approximantization. Post-fricative position shows more variability, and the fact that it has a significant interaction with Dialect will be discussed below.

In the simple model that contained only main effects and no interactions, there was not a significant effect of Dialect, and in the model displayed here there is not a significant simple effect of Dialect. However, this model was selected because it outperformed the simple model, and as we can see above there is a significant interaction between Dialect and Post-fricative Context. This implies that the difference in Post-fricative position depends on dialect, while for the other contexts there is not an interaction with Dialect. This can be most easily visualized by taking the above coefficients and calculating probabilities for certain configurations. Because all of these variables act simultaneously, values for each variable must be chosen. In Figure 2-3 below I include configurations for high and low vowels in stressed position. We can also see the significant interaction between post-fricative position and dialect. While all other contexts are very close to zero probability for both dialects, the difference between dialects in post-fricative position is significant. For this configuration, Mexican speakers approximantize at a rate of 31% if the vowel is /a/, and Colombians approximantize at a rate of 11%.



In intervocalic position Colombian speakers had slightly higher rates of approximantization, but this effect was not in fact significant.

5.4 /g/

The logistic regression model for /g/ includes the same four independent variables as for /d/. Random intercepts for subject and word, and by-word random slopes for dialect and by-subject random slopes for Stress, Vowel and Context were also included. The chart below outlines the four variables and states what their default values are:

Table 2-15: /g/ variables

Variable Name	Default Category	Other Categories
Dialect	Mexican Dialect	Colombian Dialect
Stress	Unstressed	Stressed
Vowel	/a/	/e/, /i/, /o/, /u/
Context	Intervocalic	Post-Lateral, Post-Nasal, Post-Fricative, Phrase-Initial

In the discussion on dentals above, I briefly described the process of choosing a statistical model by performing a likelihood ratio test with ANOVA. The same procedure was followed with velars, and the result was that the simple model, i.e. the model without interactions, performed better than any model with interactions. For this reason no interactions are included in this model. Below I will evaluate each of the variables in turn before presenting a chart to illustrate the probabilities for different configurations.

Table 2-16: Logistic Regression Results for Velar Production

Number of obs: 1549, groups: participant, 59; Word, 28

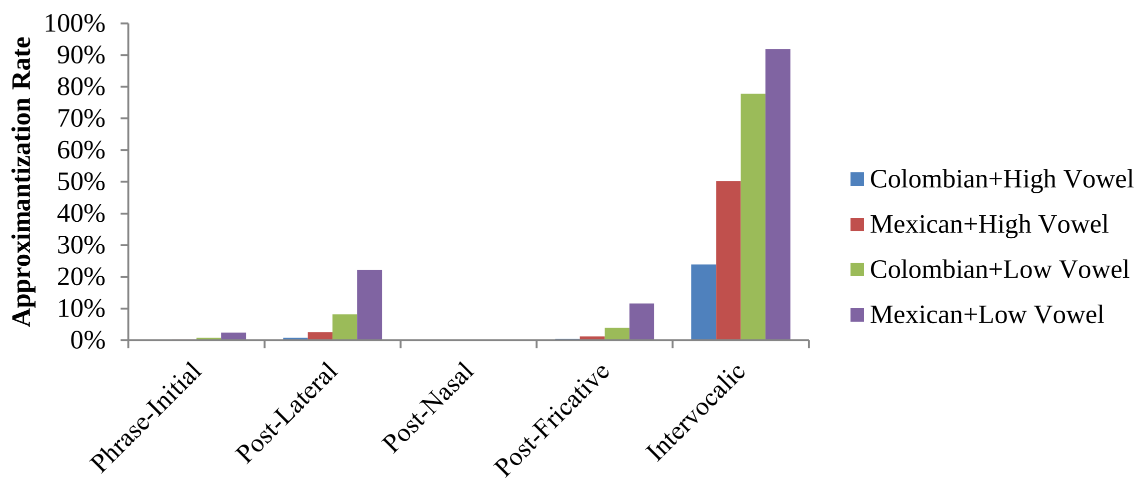
Fixed effects:

	Estimate	Std. Error	z value	Pr(>|z|)	
(Intercept)	3.1451	0.4368	7.200	6.00e-13	***
Dialect	-1.1704	0.3224	-3.630	0.000283	***
Stress	-0.7183	0.2688	-2.672	0.007545	**
Vowele	-1.5038	0.4951	-3.037	0.002388	**
Voweli	-2.4161	0.4866	-4.965	6.85e-07	***
Vowelo	-0.5852	0.4343	-1.348	0.177809	
Vowelu	-1.3546	0.4991	-2.714	0.006651	**
ContextLateral	-3.6810	0.4764	-7.727	1.10e-14	***
ContextNasal	-9.1288	1.4868	-6.140	8.26e-10	***
ContextPostFricative	-4.4546	0.5333	-8.353	< 2e-16	***
ContextPhraseInitial	-6.1454	0.5251	-11.704	< 2e-16	***

- **Dialect**: For velar /g/, the two dialects are significantly different from one another (p=.00283). In this respect, velars are different from the dentals described above. The log odds of approximantization decreases by 1.1704 when the speaker is Colombian. In other words, Colombian speakers approximantize /g/ significantly less than Mexican speakers.
- **Stress**: Stress is a significant predictor in this model (p=.007545). When /g/ is found in a stressed syllable, the log odds of approximantization decreases .7183. Dentals showed a very similar pattern.

- **Vowel**: The only vowel that is not significantly different from /a/ is /o/ (p=.177809). The remaining three vowels are all significantly different from /a/ and all have negative coefficients. This indicates that when followed by /e/, /i/ and /u/, velars are less likely to approximantize than when followed by /a/.
- **Context**: All of the Context categories are significantly different from the default category, Intervocalic. This is not surprising given that intervocalic position is the only phonological environment in which approximantization occurs at a high rate. The large coefficients of all the other contexts indicate that they are strong inhibitors of approximantization, i.e. they have the opposite effect of intervocalic position.

Figure 2-2: Approximantization Probabilities for /g/ (stressed)



The chart shows that the probability of approximantization for post-nasal environments is well below 1% for both high and low vowels in both dialects. Phrase-initial position shows some variation, but the highest probability is only 2.4% for the Mexican low vowel. For the post-fricative and post-lateral environments there is a marked increase in the probability of approximantization, and in the chart it is clear that the probability is conditioned by both dialect and vowel height. In post-lateral position the rate of approximantization goes from .8% and 2.5% for the high vowel to 8.1% and 22.2% for the low vowel for Colombian and Mexican speakers respectively. In post-fricative position the rate of approximantization goes from .4% and 1.2% for the high vowel to 3.9% and 11.6% for the low vowel for Colombian and Mexican speakers respectively. Given that there are no interactions, intervocalic position shows the same basic pattern, albeit with much higher rates of approximantization. The probability of approximantization for high vowels is 23.9% and 50.3% for Colombian and Mexican respectively. The probability of approximantization for low vowels is 77.8% and 91.9% for Colombian and Mexican respectively.

5.5 /b/

In the bilabial model of production I have included a predictor that was not in the dental or velar models. The predictor "Orthography" is coded for whether the word is written with a "v" or a "b". While these letters represented contrastive sounds in earlier stages of Spanish, the researcher believed that in modern Spanish dialects—including Colombian and Mexican—these letters represented equivalent sounds. During data collection it was easily observable to the researcher that the varying orthography for /b/ was affecting speakers' pronunciation. After recordings I asked several speakers whether they were aware of this, and they confirmed that they had learned in school that these letters were pronounced differently. Given the task—reading a word list—in retrospect it is not surprising that the subjects adopted a hypercorrect speaking style and were affected by the orthography. However, this result was somewhat unexpected and required the addition of this variable for bilabials. Table 2-17 below outlines the five variables and states what their default values are:

Table 2-17: /b/ variables

Variable Name	Default Category	Other Categories
Dialect	Mexican Dialect	Colombian Dialect
Stress	Unstressed	Stressed
Vowel	/a/	/e/, /i/, /o/, /u/
Context	Intervocalic	Post-Lateral, Post-Nasal, Post-Fricative, Phrase-Initial
Orthography	‘v’	‘b’

Dialect, Stress, and Orthography are two-way variables. The other variables, Vowel and Context, are each 5-way variables. In the case of Vowel and Context, the significance of each of the four categories listed as fixed effects can only be interpreted as different from the default values, /a/ and Intervocalic respectively. All of the variables are dummy-coded, which means that the default is coded 0 and other categories within the variable are coded 1.

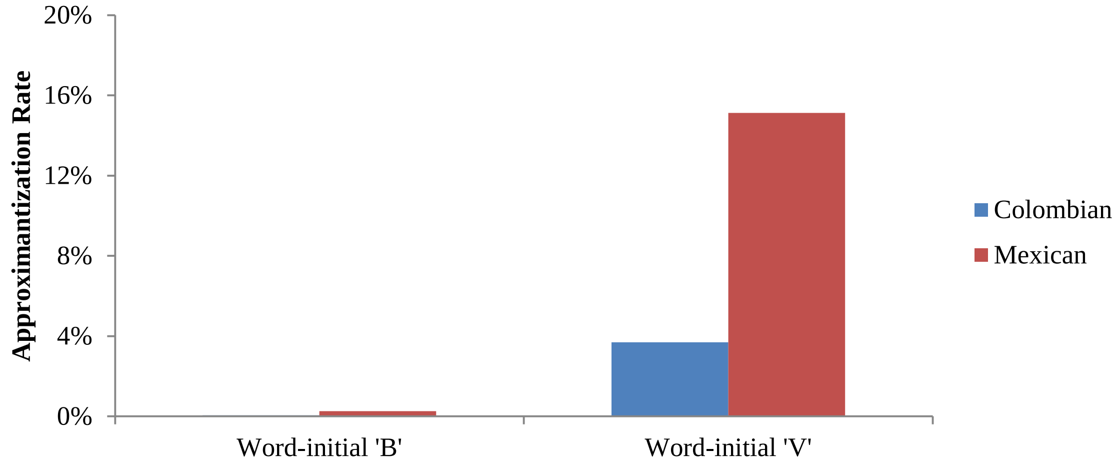
The results of the likelihood ratio test with ANOVA showed that including an interaction between Dialect and Context resulted in the best-fitted model. I will first discuss the three variables that are not involved in interactions before turning to Context and Dialect, which must be discussed alongside the interaction effects.

Table 2-18: Logistic Regression Results for Bilabial Production

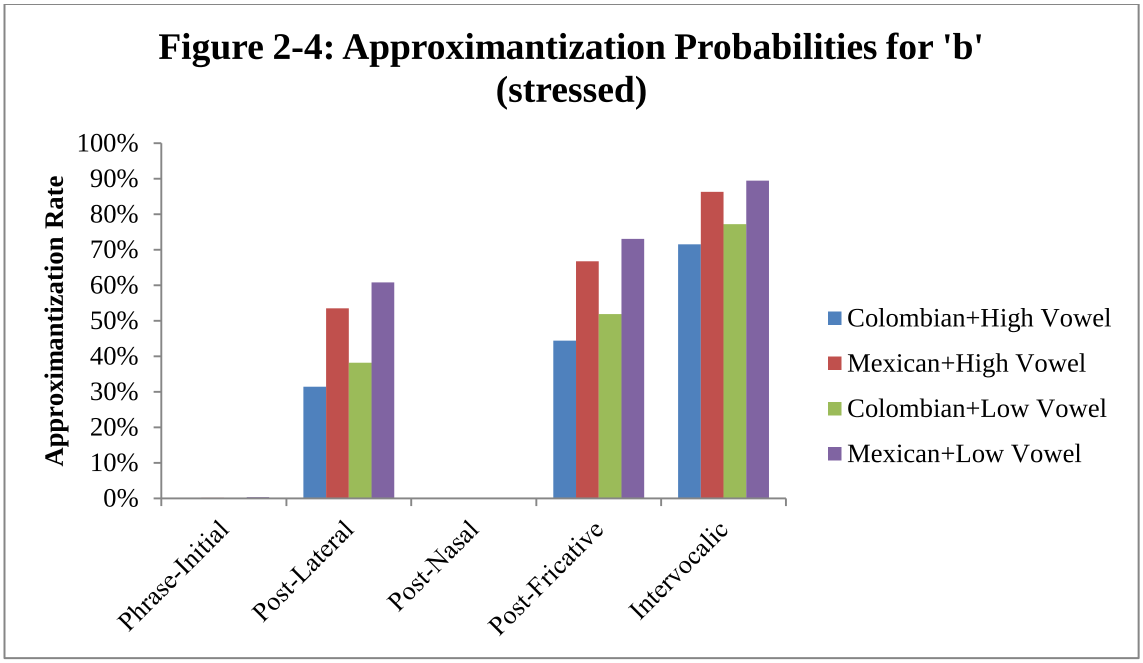
```
Number of obs: 1594, groups: participant, 59; Word, 31

Fixed effects            Estimate Std. Error z value   Pr(>|z|)
(Intercept)               7.13424    0.65706  10.858    < 2e-16 ***
Vowele                   -1.70892    0.72989  -2.341   0.019215 *
Voweli                   -0.54260    0.35384  -1.533   0.125163
Vowelo                   -2.41017    0.57735  -4.175   2.99e-05 ***
Vowelu                   -1.06086    0.35724  -2.970   0.002982 **
Orthography              -4.07742    0.50147  -8.131   4.26e-16 ***
Dialect                   0.14821    0.54933   0.270   0.787316
ContextLateral           -1.06029    0.45477  -2.331   0.019729 *
ContextNasal             -7.80457    1.16920  -6.675   2.47e-11 ***
ContextPostFricative     -0.09506    0.78652  -0.121   0.903805
ContextPhraseInitial     -7.40951    0.69093 -10.724    < 2e-16 ***
Stress                   -0.90267    0.24857  -3.631   0.000282 ***
Dialect:Lateral          -3.61126    0.63210  -5.713   1.11e-08 ***
Dialect:Nasal             0.18841    1.42672   0.132   0.894937
Dialect:PostFricative    -4.15847    0.98929  -4.203   2.63e-05 ***
Dialect:PhraseInitial    -0.61177    0.87309  -0.701   0.483493
```

- **Stress**: Stress is a significant predictor in this model (p=.000805). When /b/ is found in a stressed syllable, the log odds of approximantization decreases .9057. Both dentals and velars demonstrated essentially the same pattern.
- **Vowel**: The only vowel that is not clearly significantly different from /a/ is /i/ (p=.125163). The remaining three vowels are all significantly different from /a/, and all have negative coefficients. This indicates that when followed by these four vowels, bilabials are less likely to approximantize than when followed by /a/. This follows the general pattern seen in dentals and velars where there is a significant difference between the low vowel and other vowels, although in this case the high front vowel does not achieve significance.
- **Orthography**: This new variable is also significant, which indicates that words containing orthographic 'b' and 'v' are significantly different from one another. The log odds of approximantization for words that contain 'b' is reduced by 4.1987. This was true for both Mexican and Colombian subjects. The chart below illustrates differences found in phrase-initial position before stressed high vowels. When the letter is 'v', Mexican speakers approximantize at a rate of 15.1%, while Colombian speakers approximantize at 3.7%. Both of these rates are significantly greater than when the letter is 'b'.

Figure 2-3: Approximantization of Phrase-initial ‘b’ and ‘v’



- **Dialect & Context**: In the model above there is not a significant simple effect for Dialect, but there are several significant interactions that involve Dialect and Context. This implies that the differences we see in rates of approximantization depend on Dialect, and that there is not a uniform effect of dialect across all contexts. Notice also that several of the simple effects for Context are still significant. This implies that there is a difference between Intervocalic and post-lateral, post-nasal, and phrase-initial when the interacting term is at zero, i.e. when Dialect is Mexican. We can see how the interaction is significant by comparing the differences between post-fricative position and intervocalic position below. The difference between the two positions is actually much greater for Colombian subjects than for Mexican subjects, and the same is true for post-lateral position. Phrase-initial position and post-nasal position have very similar patterns in both dialects—in particular when the word begins with ‘b’. Both of these environments strongly disallow approximantization for bilabials.



5.6 Production Summary

In this summary I will first present some general descriptions of the analysis from above. I will then highlight differences and overlaps between each of the three places of articulation.

5.6.1 General summary of contributing factors to approximantization

One of the goals of looking closely at the production data of my subjects was to identify which environments "promote" approximantization. For binary variables the question is whether the difference in approximantization is significantly different for the two categories. We can generalize the results of the binary variables as follows in Table 2-19:

Table 2-19: Summary of Effects for Binary Variables

	Stress	**Dialect**	**Orthography**
/d/	phones in stressed syllables approximantize less often	significant interaction in post-fricative position where Colombians approximantize less; otherwise no significant differences across variables	n/a
/g/	phones in stressed syllables approximantize less often	Colombian speakers approximantize significantly less often	n/a
/b/	phones in stressed syllables approximantize less often	Colombian speakers approximantize significantly less often in post-lateral and post-fricative positions	'b' decreases approximantization as compared to 'v'

The comparisons for the variables with multiple categories are more complex, as the comparison is always with whatever the default value is. A summary of the results is below in Table 2-20.

Table 2-20: Summary of Effects for Variables with Multiple Categories

	Vowel	Context
/d/	Only the high vowel /i/ is significantly different from /a/, although the two back vowels are nearly significant. The negative coefficients associated with /i/, /o/, and /u/ indicate that dentals followed by these vowels are less likely to approximantize than those followed by /a/. The mid-front vowel /e/ was not different from /a/.	Phrase-initial, post-nasal, and post-lateral environments almost categorically impede approximantization. Intervocalic position strongly promotes approximantization. Post-fricative position is significantly different from all other positions: It approximantizes more frequently than the "inhibitors" but not nearly as much as in intervocalic position.
/g/	All vowels except for /o/ are significantly different from /a/. The negative coefficients associated with /i/, /e/, and /u/ indicate that phones followed by these vowels are less likely to approximantize than those followed by /a/.	Phrase-initial and post-nasal environments almost categorically impede approximantization. Intervocalic position strongly promotes approximantization. Post-fricative and post-lateral positions are significantly different from all other positions: They approximantize more frequently than the "inhibitors" but not nearly as much as in intervocalic position.
/b/	The only vowel that is not clearly significantly different from /a/ is /i/ (p=.125163). The remaining three vowels are all significantly different from /a/, and all have negative coefficients. This indicates that when followed by these four vowels, bilabials are less likely to approximantize than when followed by /a/.	Phrase-initial and post-nasal environments almost categorically impede approximantization. Intervocalic position strongly promotes approximantization. Post-lateral and post-fricative positions are significantly different from all other positions; depending on the configuration approximantization rates can be greater than 50% for both post-lateral and post-fricative positions (See Figure 2-4: Approximantization Probabilities for 'b').

5.6.2 Differences and Overlaps

In summarizing the effects of each variable, I will highlight differences and overlaps that are expected to bear on results from the production experiments.

- **Dialect & Context**: Because in two of the models above Dialect and Context were involved in an interaction, I will discuss these two variables jointly. Each place of articulation showed some effect of dialect, but all in different ways:
  - Dentals only exhibited an effect of dialect for the interaction involving post-fricative environment, and there was no simple effect of dialect. This result was foreshadowed in the descriptive statistics, where there were few differences between approximantization rates for Mexican and Colombian subjects. Specifically, approximantization rates were negligible for both dialects in phrase-initial, post-lateral, and post-nasal position, while in post-fricative position Mexican speakers approximantized significantly more often than Colombians. For intervocalic position, both dialects showed high rates of approximantization.
  - The model for velars did not include the interaction between Dialect and Context because the addition of the interaction term did not significantly improve the model fit. Velars showed a clear significant main effect for dialect, meaning that the effect of dialect was similar across all variables. In this case, the effect was that Colombians approximantize significantly less than Mexican speakers. Post-lateral and post-fricative positions showed some approximantization, while phrase-initial and post-nasal positions were near zero. Intervocalic position showed the highest rates of approximantization, but this was also moderated by other variables.
  - In the bilabial model, there is no simple effect for dialect, but several Dialect by Context interactions were significant. The negative coefficients for the post-lateral and post-fricative interaction terms indicate that the rate of approximantization for

Colombian speakers is significantly less in these positions; intervocalic position showed a similar effect. Phrase-initial and post-nasal positions showed negligible approximantization rates for both dialects. For Context, several simple effects were significantly different from the default, intervocalic. Post-nasal and Phrase-initial positions showed significant differences from intervocalic position; their large negative coefficients indicate that they both strongly inhibit approximantization. Post-lateral position is also different from intervocalic position, but its coefficient of -1.06029 indicates that the difference is not nearly as great. Post-fricative position does not have a simple effect that is different from intervocalic, but it is involved in a significant interaction.

Direct comparisons between the three places of articulation should be made with caution. We established at the beginning that the three places of articulation are different from one another, and this resulted in separating them into separate analyses. The consequence of this decision is that none of the models discussed above actually involve a direct comparison between places of articulation. Also, in the case of bilabials there is even an additional variable that affects approximantization. Nonetheless, there are at least two differences that stand out and should be noted:

- The virtual absence of approximantization in post-lateral position for dentals differs from the other two places of articulation. This difference has long been attributed to the fact that the lateral and the dental are homorganic. This explanation gains traction when considering that nasals—which are also homorganic with the following voiced stops—also strongly inhibit approximantization.

    - The effect of dialect for dentals was quite limited, whereas for velars and bilabials it was more widespread.

- **Stress** is a significant predictor for all three places of articulation. When phones are found in a stressed syllable, the likelihood of approximantization decreases. These results are in line with Lavoie's conclusions about stress and lenition (2001). No interactions with dialect were significant, which implies that the effect of stress was not significantly different across these two dialects.

- **Vowels** generally showed the same patterns across dialects, although not all differences were significant. The low vowel /a/ was used as the default value in the models above, and with one exception—/e/ for dentals— all other vowel coefficients had a negative value. This implies that the low vowel is approximantized at a higher rate than the other vowels. No interactions between vowels and dialect were found to be significant.

5.7 Implications of Production Results for Experiments

The analysis above lays the foundation for this study in several important ways. First, it has established that the approximantization patterns are different in each dialect. In the configurations above, approximantization rates were always higher for the Mexican speakers. Interactions involving Dialect and Context were found for dentals and bilabials, indicating that the effect of context differs from dialect to dialect. This information serves to direct the hypotheses detailed in each of the experiments in coming chapters.

CHAPTER 3 — Experiment One: *Perceptual discriminability in different phonological environments*

1 Introduction

In the production study (Chapter 2) it was shown that rates of approximantization are affected both by phonological environment and place of articulation. This confirms what many other studies have found, but previous research has assumed that approximantization can be explained solely to articulatory factors. The experiment described in this chapter aims to establish whether the ability of native Spanish speakers to discriminate between stops and approximants is affected by phonological environment in a discrimination task. It is anticipated that some phonological environments are better perceptual environments than others for discriminating between approximants and stops, and that these differences will be reflected in subjects' ability to determine whether pairs of words are identical or different. The effect of phonological environment on perception may be due to the presence or lack of cues, or masking effects. If it is found that perception is affected by phonological environment, it would be expected that confusion patterns closely match both production patterns and the hypothesized direction of sound change. Specifically, it is predicted that environments that promote approximantization will be correlated with higher rates of confusability. Such evidence would suggest that perceptual confusion is driving the development of this sound change. From a theoretical standpoint, this study specifically addresses non-functionalist explanations for sound change like Ohala's claim that innocent misperception is at the root of the change.

2 Methods

2.1 The Experiment

This experiment is a discrimination task where participants decided whether pairs of stimuli are “same” or “different.” The two auditory stimuli were either identical, e.g., [todo]–[todo], or allophonically related, e.g., [todo]–[toð̞o]. The task was performed in a variety of phonological environments. It was expected that phonological environment would affect listeners’ ability to accurately determine whether stimuli were identical or not.

Subjects were presented with two variants of a single, disyllabic word. Over the course of the experiment they heard all four possible combinations in a randomized order. An example is below:

Table 3-1: Paired stimuli combinations

Combination	**Correct Response**
stop~stop, e.g. [todo]~[todo]	identical
stop~approximant, e.g. [todo]~[toð̞o]	not identical
approximant ~approximant, e.g. [toð̞o]~[toð̞o]	identical
approximant ~stop, e.g. [toð̞o]~[todo]	not identical

The phonological environments that were tested are listed below:

- phrase-initial position
- post-lateral
- post-nasal
- post-fricative, i.e. after /s/
- intervocalic
- following vowel: /a/, /e/, /i/

Subjects were instructed at the beginning of the task that each pair of words is in effect "the same word" and that we are interested in fine-grained distinctions. In all, there were 264 pairs of stimuli that were created for analysis. Because of the repetitive nature of this task, the experiment was administered in three different segments. Between each segment, participants completed one of the other tasks.

Both accuracy and response times were measured. If misperception or confusability were indeed a contributing factor in the progression of this sound pattern, we would expect listeners to be more likely to incorrectly identify non-identical pairs as identical in environments where the sound change is most progressed, i.e. in environments where approximants are significantly more likely to be found in production. The environment that most clearly favored approximantization is intervocalic position, but there were other variables that also proved significant. To recap from the production chapter, the following environments showed significant differences in terms of approximantization rates:

- Stress: Phones in stressed syllables approximantize less often.
- Dialect: Overall, Colombian speakers approximantize bilabials and velars less often than Mexican speakers. Dentals did not exhibit a significant main effect of dialect.
- Vowel Height: For dentals and velars, the low vowel /a/ approximantized significantly more often than the high vowel /i/. In the case of the bilabials, there was not a significant difference between the high and low vowel.
- Environments for each PoA:
    - Dentals: Phrase-initial, post-nasal, and post-lateral environments almost categorically impede approximantization. Intervocalic position strongly promotes approximantization. Post-fricative position is significantly different from all other

positions: It approximantizes more frequently than the “inhibitors” but not nearly as much as in intervocalic position. Also, in post-fricative position Mexican speakers approximantized significantly more often than Colombians.

- Velars: Phrase-initial and post-nasal environments almost categorically impede approximantization. Intervocalic position strongly promotes approximantization. Post-fricative and post-lateral positions are significantly different from all other positions: They approximantize more frequently than the “inhibitors” but not nearly as much as in intervocalic position.
- Bilabials: Phrase-initial and post-nasal environments almost categorically impede approximantization. Intervocalic position strongly promotes approximantization. Post-lateral and post-fricative positions are significantly different from all other positions; depending on the configuration, approximantization rates can be greater than 50% (See Figure 2-4: Approximantization Probabilities for 'b'). In intervocalic, post-lateral, and post-fricative positions, the rate of approximantization for Colombian speakers is significantly less.

Previous research has shown that in discrimination tasks where the task is difficult, there is a bias towards “identical” (McGuire 2010). This is the result expected here, but reaction times were also recorded to see where subjects experience uncertainty. Longer reaction times reflect the difficulty of the task, and it was expected that phonological environments that cause difficulties in perception would have longer reaction times.

Before administering this experiment, several predictions were made:

- For phonological environments that promote the realization of approximants, e.g. intervocalic position, higher rates of confusability will also be evident. This will be referred to as the **confusability hypothesis.**
- For phonological environments that promote the realization of approximants, longer reaction times in the task will be displayed in non-identical pairs.
- Subjects' ability to perceive differences will depend on place of articulation.
- Subjects' ability to correctly distinguish between a stop and an approximant realization will depend on the variety of Spanish they speak. Specifically, the dialect with the highest approximantization rate—the Mexican dialect—will show lower discrimination accuracy.

2.2 The Stimuli

In creating the stimuli for this experiment, precautions were taken to ensure that stimuli sounded as natural as possible. A male native Spanish speaker from Spain was asked to produce words that would later be used to create the stimuli. The Castilian dialect was chosen because it was believed that Colombian and Mexican Spanish speakers would be equally familiar or unfamiliar with this dialect. Stimuli were then created from two sets of words—words that were the basis for the stimuli and words that were not. The words that were not used as stimuli contained the phone of interest that would later be spliced into the stimuli. For each phonological environment, one stimulus containing an approximant and one containing a stop were created. In selecting the approximant, multiple words were elicited to capture a token with typical approximant properties—notably a slight dip in intensity level and visible formants in the spectrogram. Stops were relatively easier to obtain from phrase-initial position. The stress pattern and the following vowel of both the base and the non-base words always matched. For example, in order to create the two stimuli for *nadar,* the Castilian speaker produced the words *nadar,* *dato,* and *enfadar*. Using Praat, the /d/ was extracted and either a stop or an approximant was inserted in its place:

Stop stimulus:	[nað̞ár] →	[na_ar]	→ [na_ar] + [d] = [nadár]
	[dáto] →	[d]	
Approximant:	[nað̞ár] →	[na_ar]	→ [na_ar] + [ð̞] = [nað̞ár]
	[enfað̞ár] →	[ð̞]	

To make the sounds as natural as possible, all splices were concatenated at the waveforms' zero-crossings. In pairs that were different, the amplitude was normalized by scaling the intensity in Praat to 70 dB so that amplitude could not be a discriminating factor. For each approximant/stop pair, the stimulus was created by first extracting the stop or approximant from the original and then inserting the stop or approximant that was extracted from a different word. In the example above, the stop and approximant were inserted into the same [na_ar] sequence. The length of each of the inserted phones was made equivalent so that length could not be a discriminating factor. In post-nasal position, it was found that the length of closure ranged between 10 and 30 milliseconds, so closures were kept to within this range.

The same phones that were spliced into words like *nadar* above were also spliced into other words that matched in stress and following vowel. For example, the same /d/s that were spliced into *nadar* were also spliced into the stimuli *andar* and *saldar*. This method was employed to ensure that any differences in perception across phonological environments can thus be attributed to the environment and not to some internal quality of the phone.

Ideally, all words would have been disyllabic and of equal word frequency. However, the nature of the Spanish lexicon does not allow for such convenience; two of the words (out of 65) were trisyllabic, and words had differing frequencies. Word frequency was included as a predictor in the models and will be discussed below. The complete list of words and their oral frequencies are listed in Appendix 3. Because it has been shown that frequent affixes like –*ado* (Bybee 2002) affect approximantization rates, words were chosen that did not have the phone of interest in an affixed position.

3 Speakers and data

The same group of speakers that participated in the production data collection also participated in this experiment. Most participants seemed fully engaged in the experiment, and only 3.1% of responses were eliminated due to responding too quickly (< 300 milliseconds) or not quickly enough (>7 seconds). In all, 14,356 responses were accepted for analysis.

3.1 Variables

For the most part, the same variables that appeared in the production chapter also appear here. Variables in the vowel space do not include back vowels, but the important distinction of vowel height is maintained with the three vowels /a/, /e/, and /i/. While orthography was found to be a significant contributor to approximantization for bilabials, it was not found to affect perception in this experiment. For this reason orthography will not be included in any of the models described in this chapter.

Table 3-2: Experiment 1 Variables

Variable Name	**Default Category**	**Other Categories**
Dialect	Colombian Dialect	Mexican Dialect
Unstressed	Stressed	Unstressed
Vowel	/a/	/e/, /i/
Context	Intervocalic	Post-Lateral, Post-Nasal, Post-Fricative, Phrase-Initial

4 Analysis & Results

In order to analyze results from the experiments I have chosen two approaches. The first involves the same method used to analyze the production data, namely logistic regression. This approach will be used to evaluate what factors influence "correct" responses. As mentioned above, in discrimination tasks such as this where the task is difficult, there is a bias towards "identical." Of the data accepted for analysis, 50.15% of the pairs were classified as same, yet 80.5% of the responses were "same." This result is indicative of a task that is difficult.

The second approach takes advantage of the reaction time data that was collected in this experiment; to analyze these data, linear regression was employed. Reaction times can reveal the difficulty involved in processing, and longer reaction times may also be symptomatic of perceptual challenges—even if perception is successful. If it were true that approximantization and perceptual difficulty are linked, it would be expected that phonological environments that have higher rates of approximantization would have longer reaction times.

4.1 Accuracy Results

This section will analyze what factors contribute to the incorrect or correct identification of the different pairs, as these pairs task the subject with perceiving the subtle alternations that may occur or may have occurred in natural speech. In presenting the results, I will relate them to predictions made for this experiment and the production data.

4.1.1 Predictions Involving Phonological Environment and Place of Articulation for the Overall Model

Before administering the experiment, I predicted that phonological environment would affect confusability rates. I also predicted that confusability patterns would mirror the production data: For phonological environments that promote the realization of approximants, higher rates of confusability should also be evident. In order to test these hypotheses, several mixed effects logistic regression models were run.

Recall that there were four different pairs that were presented to the subjects:

- approximant/approximant
- stop/stop
- approximant/stop
- stop/approximant

In a first pass of the mixed effects logistic regression model the coefficients for the two orderings of "different" pairs—i.e. approximant/stop or stop/approximant—were nearly identical. This indicates that the ordering of the pairs is not relevant, so these categories were grouped together. In order to examine confusability between stops and approximants, I first ran a model on just the different pairs with all places of articulation included in the model, using the following variables:

Table 3-3: Complete Experiment 1 Variables

Variable Name	**Default Category**	**Other Categories**
Dialect	Colombian Dialect	Mexican Dialect
Unstressed	Stressed	Unstressed
Following Vowel	/a/	/e/, /i/
Context	Intervocalic	Post-Lateral, Post-Nasal, Post-Fricative, Phrase-Initial
Target PoA	/b/	/d/, /g/

I also examined whether word frequency should be included as a variable in the model by using the method established by Raudenbusch & Bryk (2002) to determine which predictors contribute significantly to the model. Using this method, the procedure for choosing the best-fitted model was to run models with linguistically-motivated variables and then compare one model with x variables to another with x-1 variables in an ANOVA. If the difference is significant, i.e. $p<.05$, the model with more variables is favored. In the case of Word Frequency, adding this variable to the model ultimately analyzed below showed no significant improvement in the model: ($\chi2(2)=3.7288$, $p=0.155$).

As in the analysis of the production data, all random effects were included in the model. In this experiment, random intercepts for item and subject were included, as well as random slopes for dialect grouped by item, and stress, vowel and context grouped by subject. In a model that analyzed the different pairs and included all places of articulation, there were 7156 observations, 125 different items, and 60 subjects (31 Colombian, 29 Mexican). Table 3-4 below provides the output for this model.

Table 3-4: Accuracy regression model of different pairs

```
Number of obs: 7156, groups: File, 125; Participant, 60

Fixed effects:
                     Estimate Std. Error z value Pr(>|z|)
(Intercept)          -1.14018    0.22092   5.161 2.46e-07 ***
Unstressed           -0.27856    0.11309   2.463 0.013771 *
Vowele                0.05517    0.12815  -0.431 0.666824
Voweli               -0.28178    0.17912   1.573 0.115698
Dialectmexican        0.07789    0.22398  -0.348 0.728024
ContextNasal         -0.30828    0.20493   1.504 0.132499
ContextPostFricative -0.99886    0.26434   3.779 0.000158 ***
ContextPostLateral   -0.78083    0.21626   3.611 0.000305 ***
ContextPhraseInitial -0.11047    0.17475   0.632 0.527305
Targetd               0.34898    0.15197  -2.296 0.021655 *
Targetg               0.00812    0.14624  -0.056 0.955719
```

The coding for the dependent variable was 0=Incorrect and 1=Correct. Positive coefficients, therefore, indicate a tendency to answer correctly. The first significant variable in the list is “Unstressed” (p=.013771). If the phone is found in an unstressed syllable, the log odds of a correct answer decreases by .27856. In other words, the participants were more likely to correctly perceive that the stimuli in a pair were different if the phone was in a stressed syllable.

In the case of the context variable, we see that two categories are significantly different from the default category, intervocalic position. Both the post-fricative and post-lateral environments significantly increase the log odds of incorrect answers—by .99886 and .78083 respectively. Post-nasal position shows the same tendency, but it is not significantly different from intervocalic position. Unexpectedly, the contexts that inhibit approximantization seem to show a strong confusability tendency.

Finally, the target variable tells us whether there are any differences between the three places of articulation. With /b/ as the default place of articulation, it is apparent that the bilabial differs significantly from /d/ but not from /g/. The coefficient of 0.34898 (p=.021655) indicates that subjects are more likely to correctly determine that a pair of dental stimuli are different from one another than when presented with bilabial stimuli. Given that the velar's coefficient is very near zero and it is not significantly different from the default /b/, it can be said that the velar behaves very similarly to the bilabial and is likewise significantly different from dentals. Because the three places of articulation appear to have different perceptual patterns, below I will isolate each one of the places of articulation to determine exactly how their perceptual patterns are affected by the other variables.

The result of this model affirms that phonological environment affects confusability, but the more refined hypothesis does not appear to hold. The phonological environment that most promotes approximantization—namely intervocalic position—actually shows a significantly lower rate of confusability than both post-fricative and post-lateral environments.

Several important hypotheses are addressed in this overall model that looks at accuracy in the different pairs. Before looking at each place of articulation individually, I sum up the findings below:

- Subjects' ability to perceive differences depends on place of articulation. Dentals are correctly identified as "different" at a significantly higher rate than bilabials or velars. This suggests that confusability may be higher for bilabials and velars than for dentals.
- Before administering the experiments, it was predicted that for phonological environments that promote the realization of approximants, higher rates of confusability would also be evident. This holds true for stress, but it is not the case for any of the other variables. Neither dialect or vowel height play a significant role in perception in this overall model, and in the case of the context variables the two categories significantly different from intervocalic position—post-fricative and post-lateral positions—the effect is actually in the opposite of the predicted direction. Compared to intervocalic position, these two positions induce significantly higher rates of incorrect answers. While it appears here that post-nasal position is not different from intervocalic position, a more detailed look at each place of articulation will reveal a more complicated story.

I will now look at each individual of place of articulation separately to determine whether the overall model reflects the perceptual patterns of velars, bilabials, and dentals.

4.1.2 Velars

Table 3-5 below shows accuracy rates for the velar different pairs. The low overall accuracy rates suggest that confusability is quite high, regardless of environment. The regression analysis below will determine whether there are significant differences between these environments.

Table 3-5: Velar Accuracy Rates for Different Pairs

Context	Correct	Incorrect	N
Intervocalic	27.07%	72.93%	676
Nasal	21.58%	78.42%	431
Post-Fricative	18.63%	81.37%	451
Post-Lateral	12.42%	87.58%	451
Phrase-Initial	26.98%	73.02%	708
Grand Total	**22.34%**	**77.66%**	**2717**

The overall model (Table 3-4) was used to motivate separate analyses of each place of articulation. In the sections that follow, I will use the method established by Raudenbusch & Bryk (2002) to determine which predictors should be included in the final model. In the case of Word Frequency, adding this variable to the velar model ultimately analyzed below showed no significant improvement in the model: ($\chi 2(2)=1.1756$, $p=0.5555$). This result, however, should not be interpreted as a reflection on frequency effects on approximantization, but is rather a reflection on the homogeneity of the stimuli's word frequency.

We also examined whether including an interaction between Context and Dialect improved the model. The result for this ANOVA was: ($\chi 2(4)=8.834$, $p=0.06408$). This result is nearly significant, but upon examining the coefficients for both models it was determined that little benefit would be derived from interpreting the model with interactions; in essence the two models had the same result. For this reason, the model without interactions will be analyzed below.

Velars follow the general pattern of the overall model. All of the inhibitors of production—post-nasal, post-fricative and post-lateral positions—motivate a significantly higher error rate, i.e. in these positions, subjects are more likely to incorrectly choose that a stimulus pair is the same when they are in fact different. The effect of phrase-initial position does not differ from intervocalic position. The high vowel environment also results in significantly more incorrect responses than the low vowel environment, while /e/ is not significantly different from /a/.

Table 3-6: Accuracy: Regression Results for Velars

```
Number of obs: 2717, groups: Participant, 60; File, 48

Fixed effects:
                      Estimate Std. Error z value Pr(>|z|)
(Intercept)           -0.84074    0.21803   3.856 0.000115 ***
Unstressed            -0.10146    0.12074   0.840 0.400725
Vowele                 0.17619    0.15071  -1.169 0.242355
Voweli                -1.03952    0.24237   4.289 1.79e-05 ***
Dialectmexican        -0.16924    0.23600   0.717 0.473293
ContextNasal          -0.81140    0.25795   3.146 0.001658 **
ContextPostFricative  -1.11124    0.27089   4.102 4.09e-05 ***
ContextPostLateral    -2.07759    0.31736   6.546 5.89e-11 ***
ContextPhraseInitial  -0.02894    0.18617   0.155 0.876450
```

In the previous chapter on production I described the variables that influence production in some detail. It is possible to group these variables into two general categories: inhibitors and facilitators of approximantization. These general categories can then be matched up with two general patterns from the perceptual experiments: inhibitors and facilitators of perception. The comparisons relate to categories within each group, e.g. speakers of the Mexican dialect approximantize significantly more than speakers of the Colombian dialect. In this case the Colombian dialect will be classified as an “inhibitor” of approximantization while the Mexican dialect will be classified as a “facilitator.” To be classified as inhibitor or facilitator, the variables must be significantly different from one another; it does not imply that rate of approximantization exceeds an arbitrary amount. The prediction made at the beginning of this study claims that facilitators of approximantization should be inhibitors of perception. Velars break down as follows:

Table 3-7: Velars: Summary of Production and Confusability Effects

	Approximantization	Perception
Stressed syllables	Inhibitor	no significant effect
Colombian dialect	Inhibitor	no significant effect
Vowels /e/ and /i/	Inhibitor	/i/ inhibits and /e/ has no effect (compared to /a/)
post-nasal; post-lateral; post-fricative	Inhibitor	Inhibitor
phrase-initial	Inhibitor	no significant effect (compared to intervocalic)
Unstressed syllables	Facilitator	no significant effect
Mexican dialect	Facilitator	no significant effect
/a/	Facilitator	Facilitator
intervocalic	Facilitator	Facilitator

4.1.2.1 Link between confusability and production

For velars, the above chart does not show a clear link between production patterns and results from this perception experiment. The central hypothesis—that environments that facilitate approximantization would show more confusability—does not hold for any of the variables above. In two cases the opposite is true: The low vowel and intervocalic position, both facilitators of approximantization, are also perceptual facilitators.

4.1.3 Dentals

Table 3-8 below shows accuracy rates for the dental different pairs. Dentals have the highest overall rate of correct responses at 28.07%. Similar to velars, intervocalic position shows a relatively high rate of accuracy. Unlike velars, nasal position shows a very high accuracy rate. The low overall accuracy rates again suggest that confusability is quite high, although here the difference between post-fricative and post-nasal position is striking. The regression analysis below will determine whether there are significant differences between these environments.

Table 3-8: Accuracy Rates for Dental Different Pairs

Context	Correct	Incorrect	N
Intervocalic	30.06%	69.94%	346
Nasal	39.40%	60.60%	401
Post-Fricative	16.67%	83.33%	222
Post-Lateral	26.28%	73.72%	392
Phrase-Initial	25.21%	74.79%	702
Grand Total	**28.07%**	**71.93%**	**2063**

The model shown in Table 3-9 below includes the variables that are of greatest theoretical interest and excludes both Word Frequency and the interaction between Dialect and Context because likelihood ratio tests using the ANOVA technique showed that the inclusion of either of these variables did not result in a better fit. In the case of Word Frequency, the ANOVA that compared a model including Word Frequency to the model ultimately described below clearly shows that this variable had no influence on the outcome of the model: ($\chi 2(2)=1.0734$, $p=0.5847$. As said before, this result has more to do with the selection of a homogenous set of stimuli than it does with actual effects of word frequency on approximantization. The inclusion of the interaction between Dialect and Context also showed no improvement over the model described below: ($\chi 2(4)=6.962$, $p=0.1379$.

Recall that in the overall model dentals were deemed significantly different from bilabials. One salient difference between the two lies in post-nasal position. The fact that post-nasal position is not significantly different from intervocalic position in the overall model is attributable to the opposite tendencies in velars and dentals. While velars—and as we will see below, bilabials to a somewhat lesser degree—induce fewer correct responses in post-nasal position, post-nasal dentals are correctly identified as different at a very high rate. In the regression model output below, dentals in post-nasal position increase the likelihood of a correct answer by a log odds of .99419 as compared to the default, intervocalic position. Post-fricative position is the only other context variable that is significantly different from intervocalic position; when found in post-fricative position, accuracy decreases by a log odds of 1.20845. Given that post-fricative and post-nasal positions have opposite effects we can conclude that they are also significantly different from one another. The effect of post-lateral and phrase-initial positions do not differ from intervocalic position. The mid vowel environment also results in significantly more incorrect responses than the low vowel environment, while /i/ is not significantly different from /a/.

Table 3-9: Accuracy: Regression Results for Dentals

```
Number of obs: 2063, groups: Participant, 60; File, 36

Fixed effects:
                     Estimate Std. Error z value Pr(>|z|)
(Intercept)           -0.87446    0.31369   2.788  0.00531 **
Unstressed            -0.43447    0.18755   2.317  0.02053 *
Vowele                -0.73731    0.25056   2.943  0.00325 **
Voweli                -0.03400    0.29969   0.113  0.90968
Dialectmexican        -0.20625    0.26438   0.780  0.43532
ContextNasal           0.99419    0.37575  -2.646  0.00815 **
ContextPostFricative  -1.20845    0.53362   2.265  0.02354 *
ContextPostLateral     0.30471    0.37283  -0.817  0.41376
ContextPhraseInitial   0.07562    0.29684  -0.255  0.79893
```

I summarize below in Table 3-10 the effect of the different variables under consideration in terms of their effect on approximantization and perception of dentals in different pairs.

Table 3-10: Dentals: Summary of Production and Confusability Effects

	Approximantization	**Perception**
Stressed syllables	Inhibitor	Facilitator
Colombian dialect	no significant effect	no significant effect
Vowels /e/ and /i/	/i/ inhibits and /e/ has no effect (compared to /a/)	/e/ inhibits and /i/ has no effect (compared to /a/)
post-nasal; post-lateral; post-fricative	Inhibitors	post-nasal facilitates perception; post-fricative is an inhibitor; post-lateral shows no significant effect (all compared to intervocalic)
phrase-initial	inhibitor	no significant effect (compared to intervocalic)
Unstressed syllables	Facilitator	Inhibitor
Mexican dialect	no significant effect	no significant effect
/a/	/a/ facilitates (compared to /i/) and has no effect (compared to /e/)	/a/ facilitates (compared to /e/) and has no effect (compared to /i/)
intervocalic	Facilitator	facilitator (when compared to post-fricative); inhibitor when compared to post-nasal; no significant effect when compared to other environments

4.1.3.1 Link between confusability and production

For dentals, the link between factors that facilitate approximantization and perceptual confusability is different from velars but no clearer. There are three areas where the prediction matches the results:

- Dentals were the only place of articulation where dialect did not play a role in production, and likewise dialect does not play a role in perception.

- Post-nasal position, an almost categorical inhibitor of approximantization, results in perceptual accuracy significantly greater than intervocalic position.
- Stressed syllables inhibit approximantization but facilitate perception.

Two categories show the opposite of the predicted perceptual effect:

- Post-fricative position is more confusable than intervocalic.
- The low vowel /a/ is a perceptual facilitator compared to /e/.

All other categories showed no significant perceptual patterns.

4.1.4 Bilabials

Table 3-11 below shows accuracy rates for the bilabial different pairs. Like the dentals, intervocalic and phrase-initial environments show the highest accuracy rates. This place of articulation shows the least variation amongst the five context variables in comparison to the other places of articulation, and again the low overall accuracy rates suggest that confusability is quite high, regardless of environment. The regression analysis below will determine whether there are significant differences between these environments.

Table 3-11: Accuracy Rates for Bilabial Different Pairs

Row Labels	Correct	Incorrect	N
Intervocalic	25.65%	74.35%	581
Nasal	17.11%	82.89%	456
Post-Fricative	20.69%	79.31%	232
Post-Lateral	21.16%	78.84%	397
Phrase-Initial	26.20%	73.80%	710
Total	**22.94%**	**77.06%**	**2376**

Like the other two places of articulation, bilabials show no significant effect of word frequency. The ANOVA results from the likelihood ratio test show that this variable had no significant influence on the model: ($\chi 2(2)=3.0796$, $p=0.2144$). The nearly significant result

($\chi2(4)=8.9093$, $p=0.06341$) from the ANOVA comparing the model containing the Dialect by Context interaction and the simpler model suggests that the model with the interaction should be considered, but like the velars, the coefficients for the main effects were very similar to one another. For this reason only the simple model will be discussed.

In the chapter on Production we saw that bilabial orthography, i.e. ‘b’ or ‘v’, influenced production patterns. This experiment was not designed taking orthography into account, and while both orthographic variants of /b/ are represented here, they are not well-distributed. For example, all post-fricative bilabials are written with ‘v’, while all post-lateral bilabials are written with ‘b’. Also, given that subjects do not actually see the words, it is believed that orthography plays a lesser role in this task. For these reasons, Orthography was not included as a variable in this experiment.

Like velars, bilabials follow the general pattern of the overall model, but only one of the context variables—post-nasal position—shows error rates that are significantly different from intervocalic position. In post-nasal position, the log odds of a correct answer decrease by .8603. The effect of post-fricative, post-lateral, and phrase-initial positions do not significantly differ from intervocalic position. Bilabials followed by /e/ or /i/ result in significantly fewer incorrect responses than the default /a/.

Table 3-12: Accuracy: Regression results for Bilabial Different Pairs

```
Number of obs: 2376, groups: Participant, 60; File, 41

Fixed effects:
                     Estimate Std. Error z value Pr(>|z|)
(Intercept)           -1.4634     0.2793   5.240 1.61e-07 ***
Unstressed            -0.5999     0.1881   3.188 0.001431 **
Vowele                 0.6928     0.1857  -3.730 0.000191 ***
Voweli                 0.6331     0.2640  -2.398 0.016473 *
Dialectmexican        -0.3339     0.2791   1.196 0.231577
ContextNasal          -0.8603     0.2879   2.989 0.002803 **
ContextPostFricative  -0.4281     0.3276   1.307 0.191252
ContextPostLateral    -0.0163     0.2813   0.058 0.953799
ContextPhraseInitial   0.1479     0.2367  -0.625 0.532116
```

As with the other two places of articulation, I summarize below in Table 3-13 the effect of the different variables under consideration in terms of their effect on approximantization and perception.

Table 3-13: Bilabials: Summary of Production and Confusability Effects

	Approximantization	**Perception**
Stressed syllables	Inhibitor	Facilitator
Colombian dialect	inhibitor (marginally significant, p=.073526)	no significant effect
Vowels /e/ and /i/	no significant effect	/e/ and /i/ facilitate (compared to /a/)
post-nasal	inhibitor (compared to intervocalic)	post-nasal inhibits perception
post-lateral; post-fricative	post lat and post-fric are inhibitors, but to what degree depends on the totality of factors; both are significantly different from the default, intervocalic, so they don’t approximantize as much as intervocalic	no significant effect (compared to intervocalic)
phrase-initial	inhibitor	no significant effect (compared to intervocalic)
Unstressed syllables	Facilitator	Inhibitor
Mexican dialect	facilitator (marginally significant, p=.073526)	no significant effect
/a/	no significant effect	/a/ inhibits (compared to /e/ and /i/)
intervocalic	Facilitator	facilitator (when compared to post-nasal); no significant effect when compared to other environments

4.1.4.1 Link between confusability and production

Like the other two places of articulation, bilabials do not show an obvious link between factors that facilitate approximantization and factors that impede perception. Only one of the variables that facilitates approximantization—unstressed syllables—shows heightened perceptual confusability. There are also two areas where the opposite is true, i.e. factors that inhibit approximantization also inhibit perception:

- Bilabials found in stressed syllables approximantize less and also inhibit perception.
- Like velars, bilabials in post-nasal position approximantize very infrequently and are significantly more confusable than the default category, intervocalic position.

All other categories showed no significant perceptual patterns.

4.1.5 Perceptual Discriminability Accuracy Summary

The overall model from above showed that perceptual patterns in dentals differ from the other two places of articulation, and the detailed analyses of each individual place of articulation show that no place of articulation is exactly alike. For this reason, the analysis of each individual place of articulation is favored over just presenting the overall model. When looking at the models for the individual places of articulation it can be seen that both significant and insignificant variables were driven by one or two places of articulation. For the context variable, post-fricative and post-lateral positions acted as perceptual inhibitors and were both significantly different from intervocalic position in the overall model. However, this pattern only held for velars. In the case of dentals, only post-lateral position was significantly different from intervocalic, and in the case of bilabials neither was significantly different. Sounds found in post-nasal position showed no overall effect, but in each of the three places of articulation post-nasal position was found to be significantly different from intervocalic position. This apparent oddity can be attributed to the opposing directional effects of dentals compared to velars and bilabials. Dentals in post-nasal position showed the highest accuracy rate of any context variable, while in velars and bilabials post-nasal position proved much more detrimental to perception (compared to intervocalic position).

Stress was one of the more consistent predictors in the models above, and it follows the confusability hypothesis: Unstressed syllables, which facilitate approximantization, inhibit perception. In the overall model, stress was the only variable that followed this prediction, and in all three places of articulation the same tendency held, although this tendency was only significant for bilabials and dentals.

Although dialect showed clear differences in production, it was not deemed a significant variable in the overall model and was not significant in any of the individual models. Vowels had no overall effect and showed no consistent patterns when analyzed in each individual place of articulation.

The accuracy results show no clear link between perceptual confusability and approximantization patterns. The prediction that environments favoring approximantization—specifically intervocalic position—would show higher rates of confusability is not borne out in the analysis to this point. One result that is common across environments and places of articulation is the low accuracy rate: Out of the 15 environments examined, only two dental environments—intervocalic and post-nasal—even exceeded a 30% accuracy rate. This suggests that across the board it is difficult for Spanish speakers to perceive differences between approximants and stops in synthesized speech.

4.2 Reaction Times: Linear Regression Results

Reaction time data is used to measure processing load. In this experiment, it was hypothesized that when the different pairs have their difference in a more confusable environment, longer reaction times will result because it will take subjects longer to process that a difference was in fact perceived.

4.2.1 Methods

The analysis was limited to correct different pairs for two reasons. First, the low accuracy rate for different pairs allows us to be fairly confident that subjects actually perceived a difference and were not just mechanically responding “same” to stimuli that sound similar. Secondly, reaction times can be measured from the point at which the difference was presented in the second word of the pair.

One issue with reducing the data to correct different pairs is that it greatly reduces the amount of data. The low overall accuracy rate leaves only 1731 total tokens. Eight participants had less than 10 correct answers, which made it difficult to “keep it maximal” when running models. Attempts to run models with random slopes resulted in errors because of overfitting; to create slopes more than one data point would be necessary, and this was not always the case. In the end only models with random intercepts were run.

Reaction time was calculated from the offset of the changed phone in the second word of the pair, as in the example below. The pipe symbol indicates the time from which the reaction time was measured:

(1) [nadár nað̞|ár]…(button press)

The time between the pipe symbol and the button press is the recorded reaction time. Reaction times were log-transformed to address skewness, and all coefficients below reflect this.

4.2.2 Results

The cross-tab below shows the average reaction times in milliseconds for each phonological environment and each dialect. Intervocalic position has noticeably longer reaction times than the other word-medial environments, and it would appear that the Mexican subjects took notably longer to respond in each context.

Table 3-14: Average Reaction Times (ms)

	Dialect		
Context	**Colombian**	**Mexican**	**Overall**
Intervocalic	1147	1296	1222
Nasal	1116	1155	1135
Post-Fricative	1064	1148	1103
Post-Lateral	1055	1198	1111
Phrase-Initial	1159	1265	1205
Grand Total	1147	1296	1222
N	**935**	**796**	**1731**

The best-fitted model for the reaction time analysis proved to be much simpler than the logistic regression analyses above. The ANOVA results from the likelihood ratio test showed that Place of Articulation, Stress, and Vowel Height were not useful predictors. Also, the interaction between Dialect and Context was not useful. In the end, the model below including only Dialect and Context was selected as the best fit.

Table 3-15: Reaction Times Regression Results

```
Number of obs: 1731, groups: File, 125; Participant, 60

Fixed effects:
                      Estimate Std. Error        df t value Pr(>|t|)
(Intercept)            3.02337    0.02344  70.77000 129.008  < 2e-16 ***
Dialectmexican         0.04262    0.03172  55.92000   1.344  0.18445
ContextNasal          -0.03651    0.01384 106.64000  -2.637  0.00961 **
ContextPostFricative  -0.04035    0.01671 160.72000  -2.415  0.01687 *
ContextPostLateral    -0.02804    0.01485 137.89000  -1.889  0.06105 .
ContextPhraseInitial   0.01597    0.01215 103.77000   1.315  0.19148
```

In the model, the positive coefficient for Dialect shows that Mexican speakers take slightly longer, but the difference between Mexican and Colombian speakers was not significant (p=.18445). Two of the three word-medial environments are significantly different from intervocalic position, and the third—post-lateral position—is marginally significant (p=.06105).

All three show reaction times faster than intervocalic position. This result suggests that the perceptual processing load is greater in positions where approximantization is facilitated, and is in line with the confusability hypothesis: Longer reaction times were expected in environments that are more confusable, and it was predicted that the more confusable environments would show higher approximantization rates. The one caveat that should be mentioned here is that phrase-initial position does not pattern like the approximantization inhibitors, but rather has average reaction times much like intervocalic position. This result, however, may be in part due to measuring techniques. While word-medial positions were all measured from approximately the same point in the word, the measurement for phrase-initial position started a fraction sooner. This small difference may account for the discrepancy here.

5 Summary

We have now taken two important steps in evaluating whether perception plays a role in sound change. In the previous chapter on production patterns, we confirmed that several phonological factors play a role in approximantization and that not all places of articulation behave the same. In the study in this chapter, we examined whether perceptual confusability could be linked to approximantization patterns outlined in the first chapter. The confusability hypothesis stated that for phonological environments that promote the realization of approximants, higher rates of confusability would also be evident, measured either categorically or in terms of reaction times. The low accuracy rate across places of articulation and across environments indicates that confusability is frequent, and that native speakers of both Mexican and Colombian Spanish are likely to perceive stops and approximants as “same.” While the analysis of accuracy in specific phonological environments did not yield conclusive results, the reaction time analysis fell in line with the hypothesis that longer reaction times, which can indicate increased difficulty of the task, are more common where approximantization is likely.

From a theoretical standpoint, this experiment lends some support to non-functionalist approaches to sound change:

- Confusability between stops and approximants is high. Equating two phones is a necessary component of Ohala's idea of innocent misperception.
- Unstressed syllables facilitate approximantization and inhibit perception. This is the lone phonological predictor from the accuracy models that supports the confusability hypothesis.
- Reactions times for intervocalic position—where approximantization is frequent—are significantly higher than other word-medial positions that inhibit approximantization. This result suggests that the perceptual processing load in intervocalic position is higher and may result in more confusability, although it should be mentioned that the accuracy analysis did not find that more errors occurred in intervocalic position.

CHAPTER 4 — Experiment 2: *Scales of Perceptibility*

1 Introduction

This experiment addresses whether improved functionality could be a causal mechanism for the spread of approximantization. The theoretical foundation for this experiment comes from Hayes and Steriade's notion of "scales of perceptibility." In *The Phonetic Bases of Phonological Markedness* (2004), Hayes and Steriade describe how internal or external cues can cause feature distinctions to be better perceived in certain segments or phonological contexts than in others. For example, place of articulation is more easily identified in fricatives than stops in the same environment. While Hayes and Steriade do not address approximants directly, given approximants' similarity to fricatives and the additional information provided by formant structure, approximants should be more quickly identified than stops. If it is shown that approximants are more readily and more accurately identifiable than stops, this would suggest that approximants have an "improved functionality" in the sense that Steriade has proposed, and it would provide another reason for the progression of approximantization to new phonological environments in Spanish.

In order to test whether improved functionality may be a causal mechanism for this sound change, I tested whether stops or approximants are more readily and accurately perceived in a position where both sounds are occasionally found in some dialects—phrase-initial position. Subjects were asked to repeat individual words as quickly as possible that began with one or the other variant—approximants or stops. Below are examples of words with both variant pronunciations that were used in the experiment:

- *vete* [bete]~[β̞ete]
- *dato* [dato]~[ð̞ato]
- *guiso* [giso]~[ɣ̞iso]

Each word was given twice as a stimulus—once in its stop form and a second time in its approximant form. Subjects then had approximately 3 seconds to repeat the word.

1.1 Research Hypotheses

Hayes and Steriade have asserted that because of internal cues, place of articulation in fricatives is more easily identified than in stops. In this chapter we first examine whether approximants have an improved functionality in a general sense before evaluating individual features, and in experiment three in Chapter 5 we examine specifically whether place of articulation is more easily identified in approximants. The following hypotheses for this experiment assume that approximants have an improved functionality over stops:

- Correct repetition rates will be higher for approximant-initial variants than for stop-initial variants.
- Response times for approximant-initial variants will be faster than for stop-initial variants.

2 Methods

The same group of 60 speakers—31 Colombian and 29 Mexican— that participated in the production data collection also participated in this experiment. Of the 1,737 responses collected, all but six responses were accepted for analysis. These six responses were not accepted because an error in the data collection software rendered them uninterpretable. Both accuracy and reaction time data were collected.

In order to facilitate comparisons across places of articulation, an ideal study design would have had three-way minimal pairs with equal word frequencies for each. The structure of the Spanish lexicon did not allow this, so three-way near minimal pairs were used. To ensure that the results are not affected by differences in familiarity between words, word frequency was included in the models below. Word frequencies were taken from the online Corpus del Español (Davies), using the 5,113,249 word corpus of 20th century oral Spanish. The word frequencies of each word in this corpus are listed in Appendix 4. Word frequencies were incorporated into the model by taking the log of the raw count + 1 of this corpus. The addition of 1 allows logs to be calculated if the frequency is zero (New et al. 2007), and the log itself eliminates the skewness commonly found in word frequency distributions (Baayen et al., 2006). All words have the same CVCV pattern with penultimate stress, and all five vowels were used as following vowels. The complete list is given in Table 4-1 below:

Table 4-1: Stimuli for Repetition Task

Following Vowel	**/d/**	**/b/**	**/g/**
i	*dicho*	*bicho*	*guiso*
e	*dedo*	*vete*	*gueto*
a	*dato*	*bata*	*gato*
o	*dote*	*vote*	*gota*
u	*ducha*	*bula*	*gula*

For each word, two stimuli were created—one with a phrase-initial approximant and one with a phrase-initial stop. The Spanish speaker who produced the stimuli for this and other experiments produced phrase-initial stops in all of these words, and these recordings were used for the stop stimuli. Approximants were created by removing the stop and splicing in the approximant at a zero crossing. Small adjustments were then made to ensure that the stop and approximant had the same length, with length measurements being made from the first indication of voicing for both.

3 Results

3.1 Accuracy

Overall accuracy rates were different for each place of articulation, and very different for the given manner. Given manner refers to the manner of the phrase-initial phone of the stimulus, which was either an approximant or stop. To be “accurate” in this experiment implies that the subject repeated the word correctly with either a phrase-initial approximant or stop, independent of what the given manner was. For example, if the stimulus was [dato] *dato*, either [dato] or [ð̞ato] would be considered “accurate.” Overall, subjects accurately repeated words 1310/1731 times, i.e. 75.68%. Dentals and bilabials had virtually identical overall accuracy rates, while velars were a bit lower. Approximant stimuli had higher error rates in all cases, and dentals showed the highest error rate at 45.52% for approximant stimuli. Conversely, dentals had the highest accuracy rates for stop stimuli at 93.2%. Bilabials and velars followed fairly similar patterns; each have approximately 30% inaccuracy rates for approximants, and bilabials have somewhat higher inaccuracy rates for stops at 19.45%, while the inaccuracy rate for velars is lower at 11.86%.

Table 4-2: Overall Accuracy

PoA	Given Manner	Incorrect	Correct
d	approximant	45.52%	54.48%
	stop	6.80%	93.20%
d Total		**26.03%**	**73.97%**
b	approximant	32.76%	67.24%
	stop	19.45%	80.55%
b Total		**26.11%**	**73.89%**
g	approximant	30.45%	69.55%
	stop	11.86%	88.14%
g Total		**20.68%**	**79.32%**
Grand Total		**24.32%**	**75.68%**
N		**421**	**1310**

The effects of place of articulation do not appear to be consistent across dialects, as shown in Table 4-3. While Colombian subjects repeated approximant and stop dental stimuli more accurately, Mexican speakers repeated approximant and stop bilabial stimuli slightly more accurately. By far the largest difference was for velar approximant stimuli; Colombian subjects accurately repeated these stimuli at a rate of 81.97%, while Mexican subjects were much lower at 59.03%. Across places of articulation, Colombian subjects recorded a slightly higher rate for both given manners.

Table 4-3: Dialect and Accuracy

PoA	Given Manner	Colombian Correct	Mexican Correct
d	approximant	57.93%	51.03%
	stop	94.67%	91.67%
b	approximant	66.22%	68.28%
	stop	79.05%	82.07%
g	approximant	81.97%	59.03%
	stop	90.00%	86.21%
Overall	**Approximant**	**67.95%**	**59.45%**
Overall	**Stop**	**87.95%**	**86.64%**
Grand Total		**78.33%**	**73.04%**

There were 421 incorrect responses in this experiment. Table 4-4 lists the different errors produced when the stimulus was repeated incorrectly. There were essentially three different types of errors: incorrect place of articulation, incorrect manner, and incorrect voicing. Also, items coded as “null” were instances where the subject failed to respond, while “elided” indicates that only the phone of interest was deleted. Elided tokens were more common in velars, most of which were the approximantized form of *guiso* [ɣ̞iso] being reproduced as the more frequent *hizo* [iso]. Of the total number of errors, approximant stimuli comprise 73.24% (309/421) of the errors. The majority of the incorrect responses were stops with a different place of articulation, e.g. subjects that heard the stimulus [ð̞ato] may have said [gato]. Laterals comprise the second highest number of incorrect responses, and the fact that many dental stimuli resulted in a lateral production is most likely because the Spanish dental and lateral are homorganic—and therefore quite similar. Voicing errors were also a common error. When the stimulus was a velar approximant, subjects responded with /k/ in 12 instances. In the case of stop stimuli, 26 velar stimuli were produced as /k/, and 46 bilabial stimuli were produced as /p/. Voicing errors were the most common error when the stimulus was a stop, while for approximants the most common error involved place of articulation. Given that the approximants are voiced throughout the phone, it would be surprising to see voicing errors involving approximants. It is, however, somewhat surprising to see so many place of articulation errors. It was expected that approximants’ internal cues—specifically the formants—would provide even more place cues than stops and show higher rates of accuracy, but this is not the case. In the case of the stop stimuli, voicing errors—actually limited to velars and bilabials—was most likely a result of VOT lags that were short. Had they been a fraction longer, the error rate for stops would have most likely been even lower.

Table 4-4: Incorrect Responses

		Given PoA			
Given Manner	**Produced Phone**	**/d/**	**/b/**	**/g/**	**Grand Total**
approximant	g	49	60		109
	l	49	8	11	68
	elided	4	6	25	35
	b	23		7	30
	d		6	11	17
	k			12	12
	r	3	1	8	12
	w	1	5		6
	h		4	1	5
	p		3	1	4
	m	1	2		3
	j			2	2
	x		1	1	2
	bl	1			1
	pɾ	1			1
	t			1	1
	null			1	1
approximant Total		**132**	**96**	**81**	**309**
stop	p		46		46
	k			26	26
	g	6	5		11
	b	9		2	11
	d		2	5	7
	l	4			4
	w		2		2
	m		2		2
	elided			1	1
	n	1			1
	null			1	1
stop Total		**20**	**57**	**35**	**112**
Grand Total		**152**	**153**	**116**	**421**

Accuracy Rates for individual words are included in Appendix 5.

For the statistical analysis all of the variables below were included as predictors of repetition accuracy. Word frequency was also included as a continuous variable. Following vowel was not included in the final accuracy model because it did not contribute significantly according to the results from the likelihood test ($\chi 2(19)=14.798$, $p=0.7353$). However, it will be included below in a different model that uses the manner produced by the subject as the dependent variable.

Table 4-5: Experiment 2 Categorical Variables

Variable Name	Default Category	Other Categories
Dialect	Colombian	Mexican
Given Manner	Approximant	Stop
Vowel	/a/	/e/, /i/, /o/, /u/
Place of Articulation	/b/	/d/, /g/

Table 4-6 below reflects the best fit for these data. While simple effects for Place of Articulation and Dialect are not significant, there is a significant interaction between the two that will be discussed below. Given Manner, the variable of most interest in this study, is significant ($p=0.000157$). This variable indicates that if the given manner is a stop, the likelihood of a correct response increases by a log odds of 2.0950. This result indicates that subjects were not as accurate if the stimulus began with an approximant. Word frequency also contributed significantly to this model; the positive coefficient indicates that higher frequency equates to a significantly higher accuracy rate in this task.

Table 4-6: Accuracy rates For Repetition Task (with interaction)

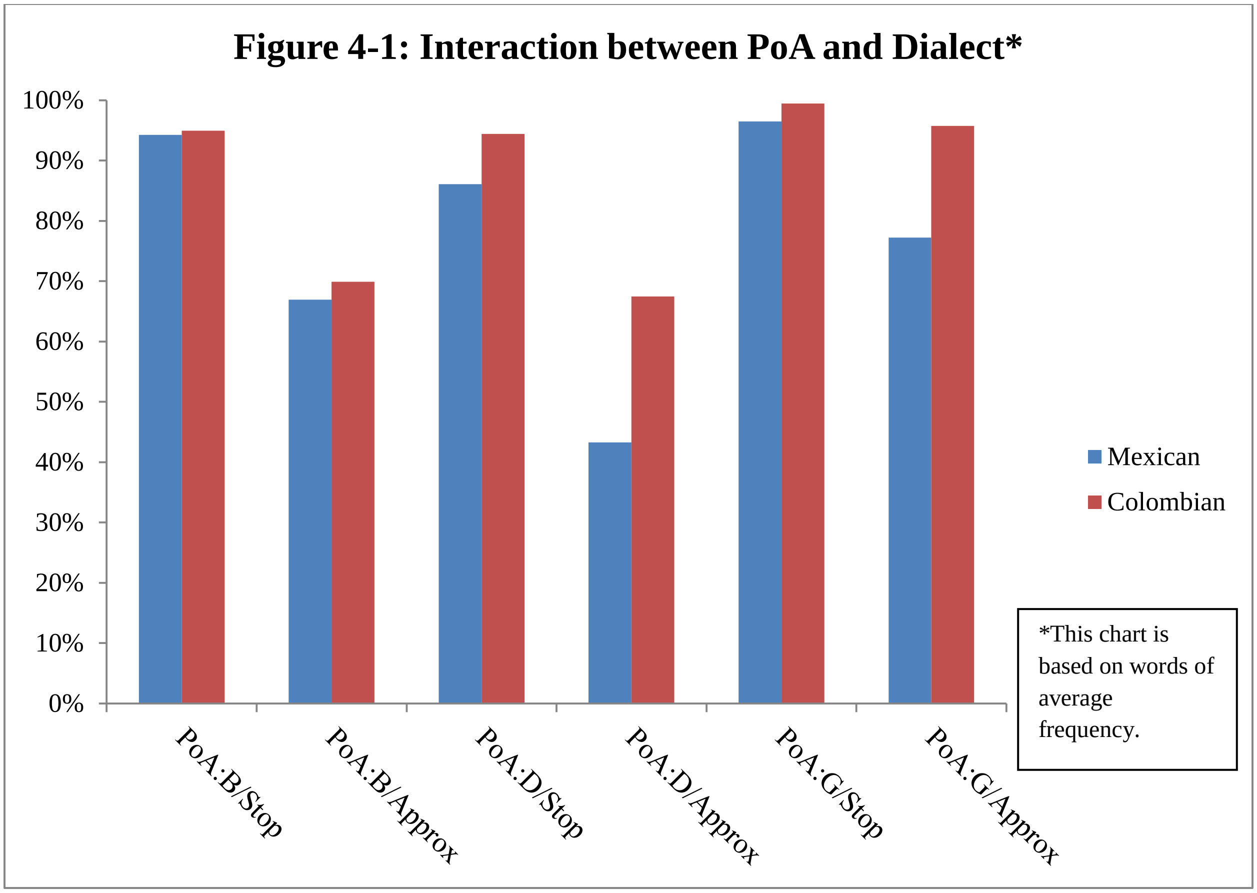
Number of obs: 1731, groups: participant, 60; File, 30

Fixed effects:

	Estimate	Std. Error	z value	Pr(>|z|)	
(Intercept)	-0.6032	0.9341	-0.646	0.518401	
PoAd	-0.1146	1.0773	-0.106	0.915259	
PoAg	2.2684	1.0988	2.064	0.038977	*
DialectMexican	-0.1375	0.4352	-0.316	0.752129	
Given Mannerstop	2.0950	0.5542	3.780	0.000157	***
Word Frequency	1.1321	0.4154	2.725	0.006428	**
PoAd:DialectMexican	-0.8620	0.6624	-1.301	0.193135	
PoAg:DialectMexican	-1.7528	0.7353	-2.384	0.017136	*

Table 4-3 showed that the overall difference in accuracy rates between the two dialects was not great—78.33% and 73.04% for Colombian and Mexican respectively. There were, however, several differences across places of articulation. This difference is reflected in the significant interaction between Place of Articulation and Dialect. In the chart below we see that for dentals and velars, there is a difference between the two dialects when the stimulus' given manner is approximant. In both cases the Mexican speakers' accuracy was lower than Colombian speakers'. For dentals, the probability of Mexican subjects responding correctly to the phrase-initial approximant stimuli was 43.3%, while for Colombians it was 67.5%; this difference was nearly significant (p=0.06085)5. For velars, the probability of Mexican subjects responding correctly to the phrase-initial approximant stimuli was 77.2%, while for Colombians it was 95.7%; this difference was significant (p=0.00121)[6].

[5] This p-value was calculated from another model where dentals were set as the default.
[6] This p-value was calculated from another model where velars were set as the default.



3.2 Reaction Time Results

Reaction times can be used to indicate whether certain stimuli make greater cognitive processing demands. In this experiment, faster reaction times could show that a particular stimulus is perceptually advantageous, i.e. easier to hear (and reproduce). In Table 4-7 below, I show the average RT of correct responses in seconds, measured from the end of the stimulus to the beginning of the subject's response. If the subject responded with a voiced obstruent, reaction time measurement stopped at the first sign of voicing. If the obstruent was voiceless, the stop release was used as the end point of the measurement.

Table 4-7: Average RTs by PoA and Given Manner

PoA	Given Manner	Average of RT (seconds)
d	approximant	0.770
	stop	0.686
d Total		**0.717**
b	approximant	0.861
	stop	0.799
b Total		**0.827**
g	approximant	0.831
	stop	0.666
g Total		**0.734**
Grand Total		**0.759**

Table 4-7 suggests that reaction times are slower for approximants than stops, and it also looks like bilabial responses are slower overall when compared to dentals and velars.

In the reaction time model shown below in Table 4-8, all of the variables discussed in the accuracy model above were considered. Only correct responses were included in the analysis. In analyzing reaction times, several variables were eliminated by using the likelihood ratio test with the ANOVA function. Vowel and all interactions did not contribute significantly to explaining reaction time. Word Frequency was found to improve model accuracy ($\chi2(2)=14.305$, $p=0.0007828$) although the predictor itself does not achieve statistical significance (*t* value<2). In the case of Place of Articulation, the ANOVA produced a result that was not significant ($\chi2(8)=11.408$, $p=0.1796$), which suggests that the simpler model without Place of Articulation should be favored. In the model below, only Given Manner is significant (*t* value=2.415). When the given manner is stop, reaction times are significantly faster. This is indicated by the negative coefficient. The two dialects do not differ from one another.

Table 4-8: RT without PoA

```
Number of obs: 1303, groups: participant, 60; File, 30

Fixed effects:
                    Estimate Std. Error t value
(Intercept)        -0.078996   0.032212  -2.452
DialectMexican      0.001611   0.022313   0.072
Given Mannerstop   -0.061986   0.025663  -2.415
Word Frequency_log -0.017452   0.016902  -1.032
```

3.3 Manner Produced

Of the total correct responses, subjects produced approximants 7.75% of the time, but as Table 4-9 below suggests, approximantization depends on the manner of the stimulus. Overall, the rate of approximantization for approximant stimuli is 10.65%, while for stops it is 5.73%. For each place of articulation the given manner of approximant elicits a higher rate of produced approximants. In the case of the bilabials, however, it should be noted that of the 57 total approximants produced, 31 were from a single word, *vete*. Of those 31 tokens, 22 came when the approximant was the given manner, and 9 came when the stop was the given manner. As was mentioned in the production chapter, orthography most likely plays a role here, despite the fact that the subjects did not see the word during this experiment.

Table 4-9: Manner Given/Manner Produced: Correct results

		Manner Produced	
PoA	**Given Manner**	**Approximant**	**Stop**
d	approximant	8.23%	91.77%
	stop	5.47%	94.53%
d Total		**6.48%**	**93.52%**
b	approximant	17.01%	82.99%
	stop	7.66%	92.34%
b Total		**11.89%**	**88.11%**
g	approximant	6.01%	93.99%
	stop	4.25%	95.75%
g Total		**4.98%**	**95.02%**
Overall	**Approximant**	**10.65% (N=57)**	**89.35% (N=535)**
Overall	**Stop**	**5.73% (N=44)**	**94.27% (N=768)**
Grand Total		**7.75%**	**92.25%**

In the following model I investigate whether Given Manner affects Manner Produced. While this question is not directly related to the research hypotheses laid out in the beginning of this chapter, it does relate to the more general question of whether native Spanish speakers are able to perceive the difference between stops and approximants—in this case in phrase-initial position. If the given manner significantly affects production, we can reasonably say that native speakers have a certain level of awareness of this variation, and that their production may be affected as a result. I have changed the dependent variable to Manner Produced and included other variables mentioned in the accuracy analysis. Only correct responses were included in this model, which means that the dependent variable is dichotomous, coded approximant (=0) or stop (=1). Word frequency was not found to be a significant predictor in this model and was dropped from the final analysis.

Table 4-10: Manner Produced Results for Repetition task

```
Number of obs: 1303, groups: participant, 60; File, 30

Fixed effects:
                 Estimate Std. Error z value Pr(>|z|)
(Intercept)        3.1071     0.9405   3.304 0.000954 ***
PoAd               3.9271     1.1825   3.321 0.000897 ***
PoAg               3.5757     1.1722   3.050 0.002285 **
Given Mannerstop   2.2723     0.8053   2.822 0.004778 **
DialectMexican    -1.2005     0.7473  -1.607 0.108154
Vowele            -0.1134     0.7944  -0.143 0.886458
Voweli             3.4223     1.0430   3.281 0.001033 **
Vowelo             2.0897     0.8218   2.543 0.010994 *
Vowelu             1.9253     1.2966   1.485 0.137585
```

Place of Articulation is a significant predictor in this model, as both dentals (p=0.000897) and velars (p=0.002285) are significantly different from bilabials. As mentioned above, however, this difference may in fact be attributable to the ‘v’ found in words like *vete*. Dentals and velars are not significantly different from one another. The variable Given Manner is a significant predictor of Manner Produced. This means that if the given manner is stop, the log odds of a stop being produced is increased by 2.2723. The Vowel variable, which was not included in the accuracy model, is included here because of the significant difference between the low vowel, /a/, and the high vowel and back vowels. The near-zero coefficient of the mid-front vowel /e/indicates that it patterns much like /a/. The general pattern of the vowels is much like what was seen in production, where the low vowel promoted approximantization across all places of articulation. The vowel /e/ also followed this pattern in the case of dentals in the production analysis.

While the effect of Dialect in this model is not significant (p=0.108154), it is worth noting the trend indicated by the coefficient. The coefficient indicates that when the speaker is Mexican, the log odds of that speaker producing an approximant increases by 1.2005. Evidence of this can be seen in Table 4-11 below, where Mexican speakers approximantize at a higher rate regardless of the manner given.

Table 4-11: Manner production rates

Given Manner	**Manner Produced**	**Colombian**	**Mexican**	**Grand Total**
Approximant	approximant	3.87%	4.91%	4.37%
	stop	37.80%	35.50%	36.68%
Stop	approximant	2.23%	4.60%	3.38%
	stop	56.10%	54.99%	55.56%
Grand Total		**100.00%**	**100.00%**	**100.00%**

4 Summary of Accuracy Results and Manner Produced Results

The results from above clearly show that approximants are not perceptually advantageous compared to stops in Spanish. While higher word frequency led to greater accuracy, this variable was not involved in any significant interactions, thereby allowing us to be confident in the results of other variables. Dialect differences in the production task do not manifest themselves in this experiment, as in general Dialect played little role apart from the interaction between Place of articulation and Dialect described in section 3.1. In the interactions outlined in Table 4-1 above, Mexican speakers—despite producing approximants at a higher rate—did not reproduce approximants as accurately as Colombians in the case of velars and dentals. In the model that addressed what factors affected the manner produced by subjects, Given Manner had a significant effect on Manner Produced. This suggests that speakers are indeed sensitive to these subtle allophonic differences, and that input does affect output. The effect of Dialect was not significant, but there was some indication that Mexican speakers approximantize more often regardless of the given manner. Vowel was found to be a significant predictor in this model, as the phones followed by /a/ and /e/ approximantized significantly more often than the others.

5 Discussion

This experiment has addressed the question of whether improved functionality—specifically perceptual functionality— could be a causal mechanism for the spread of approximantization in phrase-initial position. Hayes and Steriade (2004) found that place of articulation in fricatives is more easily perceived than in stops on a scale of perceptibility, and it was predicted that in the case of Spanish, approximants would be more easily perceived than stops. Such a result would have been evidence of the improved functionality of approximants.

The analyses above showed no evidence that approximants have an improved functionality over stops. To the contrary, stops are more accurately and more rapidly repeated than approximants, and the majority of the errors involving approximant stimuli were place of articulation errors. Given the rates of approximantization we have seen in both the production task and this repetition task, this result may not seem so surprising. Table 4-12 shows the rates of approximantization in both tasks:

Table 4-12: Approximantization Rates in Two Tasks

Place	Dialect	Phrase-initial (Production)	Experiment 2
/b/	Colombian	14.69%	10.85%
	Mexican	26.02%	12.90%
/d/	Colombian	8.72%	5.75%
	Mexican	13.95%	7.28%
/g/	Colombian	9.42%	2.14%
	Mexican	18.37%	8.17%

The lower approximantization rates in Experiment 2 may seem somewhat surprising given that subjects were primed with phrase-initial approximants half of the time, but a fair number of the responses to phrase-initial approximant stimuli—36.4%—were eliminated because they were incorrect. As a result, 58.78% of all correct tokens were from phrase-initial stops, and of these a very high percentage were produced as stops. In the case of bilabials, the elevated approximantization rates are due to orthographic ‘v’, which accounts for 40% of the bilabial stimuli in this experiment. Results from both tasks indicate that approximantization is still relatively infrequent in both dialects in phrase-initial position. This is important when considering Hayes and Steriade’s claim that fricatives—and in my analysis, approximants—are ranked higher on the scale of perceptibility than approximants. While this may be valid in a universal sense where all other factors are equal, in the case of a single L1, linguistic factors are always unequal. In the chart above, there is a clear preference for producing stops in phrase-initial position. In this case, subjects would also be much more familiar with perceiving stops in this position, and any marginal advantage gained in perceptual salience is thereby trumped by familiarity. This is the reason it is believed that Dialect was not a significant predictor in the models above: Although in the production task Mexican speakers produced significantly more approximants than Colombian speakers, the relatively infrequent appearance of approximants in phrase-initial position is still the determining factor. We can also add that phrase-initial position is probably the most perceptually salient position for a voiced stop. In this position, the stop is not subject to any environmental masking as described by Wright (2001), so our choice of word position may also have been a mitigating factor.

There are two other interesting results from this experiment that merit revisiting: the significance of the low vowel and the importance of manner in the stimulus. That the low vowel was a significant predictor of manner in this experiment is in line with results from the production task, and also concurs with the results from Brown & Brown (2012) concerning /s/ reduction in Cali Spanish. In both cases, the low vowel in post-consonantal position was a significant predictor of approximantization. Given that this is true in both the production task and the perceptual repetition task, this points to an articulatory explanation. Stimuli in this experiment were not more or less approximant-like because of the following vowel, so this suggests the explanation is not driven by perception. In anticipation of the low vowel, the tongue is lowered and more airflow passes through the oral cavity, resulting in more approximants than when other vowels are in post-consonantal position.

The manner of the stimulus was also a significant predictor of the manner produced by the speaker. This means that if subjects heard a stop they were more likely to produce a stop, and if they heard an approximant they were more likely to produce an approximant. While it is true that approximant-initial stimuli had significantly higher error rates, it is also true that speakers responded correctly 63.6% of the time if the stimulus began with an approximant. Table 8 shows that the stimulus' manner is influential, but it also shows an extremely strong preference for obstruents in phrase-initial position. Even when the stimulus was approximant-initial, subjects responded with a stop in 89.35% of the correct responses. This suggests that there is a strong link in the mental representations of these two alternants, but that the obstruent is the dominant representation in phrase-initial position.

Chapter 5 — Experiment 3**:** ***Place Identification and Improved Functionality***

1 Introduction

In the previous experiment, we tested the hypothesis that approximants have increased in frequency over time due to their improved functionality. Results strongly suggest that approximants did not have an improved functionality over stops. In fact, stops were more accurately and more rapidly perceived than approximants. The experiment described in this chapter continues the investigation of whether approximants could have an improved functionality by directly examining whether place of articulation identification is affected by manner. This question was addressed to some degree in the previous chapter, but it is the sole focus in this chapter. The purpose of the experiment described in this chapter is to determine whether place is more readily and more accurately identified in approximants as compared to voiced stops.

In this experiment we have limited the context to phrase-initial position. This environment was chosen because articulatory explanations for approximantization in this position are more difficult to sustain, and subjects did indeed produce approximants in this position in the production task. Previous articulatory explanations (cf. Ohala 1983; Widdison 1997) have claimed the approximants result from natural phonetic processes—specifically target undershoot where the stop is not fully realized as a stop because of the articulatory openness of surrounding vowels. This explanation works fine for intervocalic position, but not for phrase-initial approximants that are also phrase-initial—or read in a word list like in this study. For this reason other explanations should be considered, and this experiment will examine whether one of the perceptually-motivated explanations—namely perceptual optimization of place—is viable.

Given that perceptual optimization has been proposed (cf. Flemming 1996, 2003, Boersma 1998, and Steriade 2001) as a possible explanation for the initiation of certain sound change patterns, determining whether approximants or stops have an improved functionality is an important question when considering perception's role in this sound change. The task is a same-different discrimination task, where the only potential difference in the pairs is place of articulation. Both accuracy and reaction times were recorded. In addition to identical pairs for each place of articulation, each possible mismatched pair was included. In this experiment, manner was held constant to isolate place as the variable of interest. An example matrix of the stimulus pairs is below. The same matrix was used for both approximants and stops, thereby allowing us to determine whether differences in place of articulation are more easily perceived when the pair is an approximant (pair) or stop (pair). One-third of the pairs were same pairs while all others were different:

Table 5-1: Place of articulation experiment pairs

	Dental	Bilabial	Velar
Dental	dental~dental	dental~bilabial	dental~velar
Bilabial	bilabial~dental	bilabial~bilabial	bilabial~velar
Velar	velar~dental	velar~bilabial	velar~velar

1.1 Predictions

The driving question behind this experiment is whether approximantization serves a functionalist agenda—namely that this sound change results in phonetic optimization. For this experiment we have isolated place of articulation to see if an approximantized form of the voiced stop series proves more perceptually optimal or salient than the stop form. Predictions based on a functionalist approach to sound change would include the following:

- Approximants are more often correctly identified than stops.
- Reaction times for correct identification of approximants are shorter than for stops.

2 Methods

2.1 Experimental Design

This experiment is a discrimination task where participants must decide whether pairs of stimuli are "same" or "different." Different pairs differ solely by place of articulation; all other variables are held constant. While in experiment one perceptual discriminability in many phonological environments were evaluated, this experiment only isolates phrase-initial position, and nonce words were chosen instead of real words. The focus on phrase-initial position in this experiment is justified because articulatory explanations for approximantization in this position are weak—in particular in the word-list style used for elicitation in the production chapter.

2.2 The Stimuli

The choice of nonce words was necessary in order to keep all other elements of the words constant. While minimal pairs are relatively easy to find between any two phones in word-initial position in the stop series, three-way minimal pairs for all place of articulation and vowel combinations are not available in Spanish. With nonce words, however, it was possible to find a Spanish-sounding sequence that could be used across all five vowels. For each word listed in the table below, both a stop-initial and approximant-initial stimulus were created. All words have the same CVfo pattern with penultimate stress.

Table 5-2: Place of articulation experiment stimuli

Post-consonantal vowel	/b/	/d/	/g/
/a/	/bafo/	/dafo/	/gafo/
/e/	/befo/	/defo/	/gefo/
/i/	/bifo/	/difo/	/gifo/
/o/	/bofo/	/dofo/	/gofo/
/u/	/bufo/	/dufo/	/gufo/

Using nonce words has an additional advantage in that it militates against the possible role of differing word frequencies or familiarity effects. Over the course of the experiment, all possible pairs were presented to the subject. For each iteration of the experiment a different randomized order was used to ensure that any significant effects were not due to sequencing.

Despite being nonce words, every effort was made to make these words sound as natural as possible. For every word, the initial phone was extracted and then either a stop or approximant was spliced onto the word. The same Castilian speaker who provided the samples in previous experiments also provided words for this experiment. All target phones came from real words produced by this speaker, and the nonce bases were also elicited from the speaker. Each consonant was added on to the same [Vfo] sequence, so any potential difference is attributable to the initial consonant. For example, for all the low vowel words, all six variants—3 stops and 3 approximants— were appended to the same [afo]. Thus, when comparing a different sequence such as [ɣ̞afo ð̞̌afo], the only difference lay in the initial consonant. To ensure that transitions from consonant to vowel were smooth, when all of the initial phones were recorded they had the same following vowel as the eventual stimulus. For example, the initial [g] for [gafo] was taken from [gato] while all of the other [gV] sequences were taken from different words. All splices were made at the waveforms' zero-crossings, and the spliced consonant included the first vowel peak from the original. In pairs that were different, so that amplitude could not be a discriminating factor, the amplitude was normalized by scaling the average intensity to 70 dB in Praat. For each of the 6 phones associated with the same subsequent vowel, the length of each of the inserted phones was made equivalent so that length could not be a discriminating factor. Length was measured from the first indication of voicing.

3 Speakers and Data

The same group of speakers that participated in the production data collection also participated in this experiment. Only 3.4% of responses were eliminated due to responding too quickly (< 300 milliseconds) or not quickly enough (>7 seconds). Of the 5,400 responses collected, 5,216 responses were accepted for analysis. Both accuracy and reaction time data were collected.

In this experiment, the 60 speakers (31 Colombian; 29 Mexican) had a much higher overall accuracy rate than in the first discrimination task that measured confusability in different phonological environments; however, there was still a high rate of inaccuracy. As Table 5-3 shows, when the pairs were approximants the overall accuracy rate was only 55.75%, while for stop pairs it was 75.83%. Overall, it was much more difficult for subjects to correctly identify different pairs as different, which indicates a fairly strong response bias towards “same.”

Table 5-3: Accuracy rates for Approximants and Stops

Manner	Correct Button	Correct	Incorrect	N
Approximant	different	44.55%	55.45%	1780
	same	78.11%	21.89%	891
Total		**55.75%**	**44.25%**	**2671**
Stop	different	71.16%	29.63%	1713
	same	85.46%	14.54%	832
Total		**75.83%**	**24.81%**	**2545**
Grand Total		**65.55%**	**34.65%**	**5216**

When comparing accuracy rates across dialects, approximant pairs show very little difference. In the case of stop pairs, Colombian subjects appear to be more successful at identifying stop pairs as same or different. The regression analyses below will show whether these differences are significant.

Table 5-4: Accuracy rates by dialect

Dialect	Manner	Correct Button	Correct	Incorrect	N
Colombia	**Approximant**	different	44.31%	55.69%	923
		same	78.19%	21.81%	463
	Total		**55.63%**	**44.37%**	**1386**
	Stop	different	76.32%	23.68%	891
		same	87.91%	12.09%	430
	Total		**80.09%**	**19.91%**	**1321**
Colombia Total			**67.57%**	**32.43%**	**2707**
Mexico	**Approximant**	different	44.81%	55.19%	857
		same	78.04%	21.96%	428
	Total		**55.88%**	**44.12%**	**1285**
	Stop	different	65.57%	34.43%	822
		same	82.84%	17.16%	402
	Total		**71.24%**	**28.76%**	**1224**
Mexico Total			**63.37%**	**36.63%**	**2509**
Grand Total			**65.55%**	**34.45%**	**5216**

4 Data Analysis

The answer to the most basic question—whether approximants are more often correctly identified than stops—seems fairly clear from the descriptive statistics above, but we will look at a logistic regression model to confirm that stops are indeed more perceptually optimal. We will also look at the interaction between manner and dialect that looks to be significant above. For this experiment—both accuracy and reaction time models— the explanatory variables under consideration are:

Table 5-5: Experiment 3 explanatory variables

Variable	Default	Other Categories
Manner	Approximant	Stop
Dialect	Colombian	Mexican
Vowel	/a/	/e/,/i/,/o/,/u/

Two variables that have appeared in all of the other tasks are word frequency and place of articulation, but because of the structure of this experiment neither will be used as a variable here. Word frequency has no bearing on this experiment given that all words are nonce words, and place of articulation could only be analyzed in terms of the combinations of the stimuli, e.g. bilabial/dental, dental/velar, etc. It is not clear how this analysis could be interpreted.

For the accuracy analysis I will be reporting on mixed effects logistic regression models. For the reaction time data I will be using mixed effects linear regression models.

4.1 Accuracy analysis

As in other chapters in this study, the set of predictors used in the model was determined by using Raudenbusch & Bryk's (2002) technique of comparing models with an ANOVA. The descriptive statistics above suggested that Dialect may be involved in an interaction with Manner, and in an ANOVA test this model was in fact significantly better than the simple model ($\chi2(1)=9.4659$, $p=0.002093$). It was determined that this model was the best fit for the data, and was the model selected for analysis. In the mixed model shown in Table 5-6, we follow Barr et al. (2013) and include random intercepts for subject and word, and by-word random slopes for dialect and by-subject random slopes for Vowel and Manner. The dependent variable in Table 5-6 below is Accuracy, coded 0=Incorrect, 1=Correct.

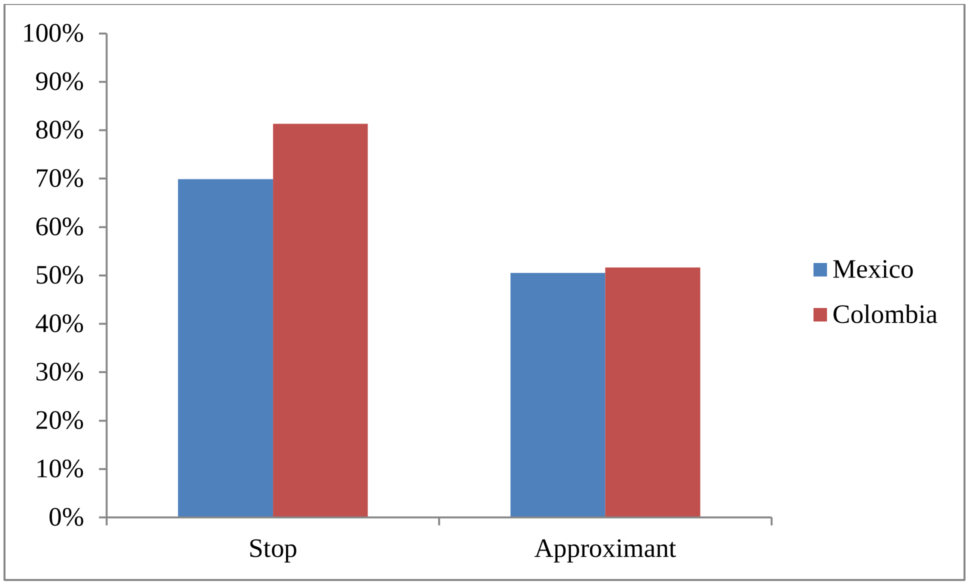
Table 5-6: Accuracy results

```
Number of obs: 5216, groups: File, 88; Participant, 60

Fixed effects:
                      Estimate Std. Error z value Pr(>|z|)
(Intercept)            0.06490    0.23821  -0.272  0.78527
MannerStop             1.40541    0.22436  -6.264  3.75e-10 ***
Dialectmexico         -0.04371    0.12229   0.357  0.72077
Vowele                 0.03672    0.28783  -0.128  0.89850
Voweli                 0.01971    0.29268  -0.067  0.94631
Vowelo                 0.63433    0.28465  -2.228  0.02585 *
Vowelu                 0.49999    0.28280  -1.768  0.07706 .
Manners:Dialectmex    -0.58564    0.17994   3.255  0.00113 **
```

The descriptive statistics above showed that accuracy rates are higher for stop pairs than approximant pairs. The logistic regression analysis shows that this difference is statistically significant for Manner. The Manner variable is dummy coded, with approximants as the default (=0) and stops as the other category (=1). The coefficient for Manner indicates that correct answers increase by log odds of 1.40541 when the pair's manner is stop. The difference between accuracy rates of stop and approximant pairs is significant ($p<.001$). There is also a significant difference between back vowels and non-back vowels. The default for Vowel is /a/, meaning that all other vowels are compared to the low vowel /a/. The vowel /o/ is significantly different from /a/ ($p=.0268$) while /u/ is nearly significant ($p=.0818$). The two front vowels have coefficients near zero, which indicate that they are very similar to /a/, and therefore also dissimilar to the two back vowels. The positive coefficients for the two back vowels indicate that accuracy rates are significantly higher when the phones precede these vowels.

As you can see in the model above, the interaction between manner and dialect is statistically significant ($p=.00113$). By exponentiating the coefficients and then calculating probabilities for a specific configuration, the differences found between the two dialects can be seen graphically. The chart below shows the probability of getting a response correct when the phone is followed by the low vowel, which in this case happens to be the vowel that results in the most incorrect answers. The other two front vowels are very similar, while the two back vowels would show slightly higher accuracy rates than those in the chart below.

Figure 5-1: Accuracy differences for Dialect and Manner



While the accuracy rate for approximants is virtually identical for the two dialects, there is a significant difference for stops. For this configuration, Mexican speakers correctly identified stops as same or different 70% of the time, while Colombian speakers answered correctly 81% of the time. Importantly, accuracy is higher when subjects are presented stop stimuli for both dialects.

4.1.1 Accuracy Analysis Summary

In this section we have shown that approximantized phones are not more easily perceived than stop phones, thereby calling into question the original hypothesis. In fact, stops demonstrated a much higher accuracy rate. Accuracy rates were also shown to be higher when stops or approximants were followed by back vowels. The interaction shows that Colombian speakers perceive stops in phrase-initial position significantly more accurately than Mexican speakers. Approximants, however, were perceived accurately at virtually identical rates.

In the section that follows I will analyze reaction times to see if processing load was affected by manner. After concluding that section I will discuss the implications of these results on the predictions related to perception and sound change made in the introduction.

5 Reaction Time Analysis

While the accuracy data from above have already painted a fairly clear picture concerning which manner of articulation is perceptually optimal, reaction time can contribute to this discussion by indicating which manner incurs the greatest processing load. Longer reaction times would correlate with greater processing load, and can be loosely associated with the amount of "effort" that a speaker must make in order to determine whether a pair of words is identical or different.

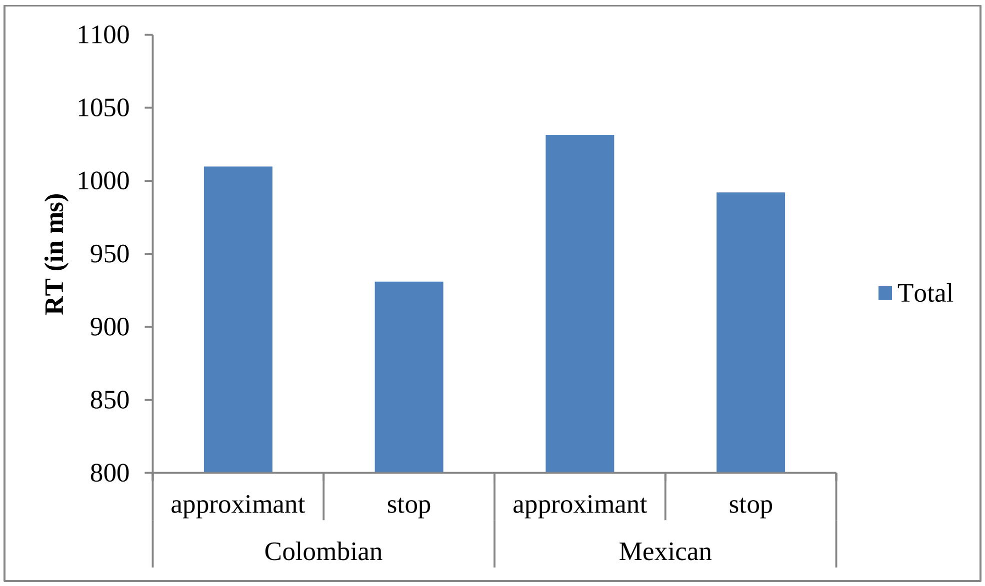
5.1 Methods

To assess reaction time, I employ a mixed effects linear regression model using the same explanatory variables as in the logistic regression above. This model includes random intercepts for subject and item, by-subject random slopes for Vowel and Manner, and by-item random slopes for Dialect. Measurement of reaction times for these stimuli began at the offset of the second phrase-initial stop or approximant in the pair and ended when the subject pressed a button on the button box. Only correct answers whose pairs were different were admitted to this portion of the study. Like in experiment 1, the reason only different pairs are used is because the point of measurement is clear: Subjects who correctly respond to different pairs are presumably responding to the difference they have perceived in the second pair, whereas with same pairs it is unclear from which point they are responding. Reaction times were log-transformed to address skewness, and all coefficients below reflect this.

5.2 Results

The analysis of reaction time corroborates what was seen above in the accuracy analysis. Figure 5-2 below displays average reaction times in milliseconds for approximants and stops and is also separated out by Dialect. In both dialects, reaction times are longer for approximant pairs. While it appears there may be a significant interaction between Dialect and Manner in the chart below, this is not borne out by the regression analysis.

Figure 5-2: Average Reaction Times in Experiment 3



In the mixed effects linear regression model presented in Table 5-7 below, when the manner variable is stop, log reaction time significantly decreases, as indicated by the negative coefficient and the *t* value > 2. In contrast to the accuracy model above, Vowel shows no significant differences between the five categories, and the ANOVA results from the likelihood ratio test showed that this category should not be included in the model ($\chi2(19)=19.935$, $p=0.3985$). This leaves only Manner and Dialect as predictors, and the log likelihood ratio test revealed that the possible interaction between the two was not a significant contributor to the model ($\chi 2(1)=0.0712$, $p=0.7895$). Even Dialect can be eliminated in this fashion as well ($\chi 2(4)=2.0662$, $p=0.7236$), leaving us with Manner as the only significant predictor.

Table 5-7: Reaction Time Results

```
Number of obs: 2012, groups: Participant, 60; File, 59

Fixed effects:
             Estimate Std. Error t value
(Intercept)  2.983204   0.015602  191.20
MannerStop  -0.024257   0.006297   -3.85
```

The results of this reaction time analysis show that the stimuli's manner is solely responsible for changes in reaction time, and that approximant stimuli take longer to process than stop stimuli. While Chart 2 suggests that there may be an interaction between Dialect and Manner, the actual difference between the two dialects for stops in the above configuration is only about ~60ms, while for approximants it is ~21ms.

6 Discussion

Of the several perception-motivated explanations for sound change, this experiment addresses the most specific optimization question: Could approximantization in phrase-initial position be motivated by an improvement in place of articulation cues? The answer to this question appears to be no. Accuracy rates were significantly better for stops, and reaction times were significantly longer for approximants, so it seems clear that approximants are more difficult to perceive in phrase-initial position. This result in and of itself does not mean that perception has no role in this sound change, but we can conclude that a functionalist, optimization-motivated explanation is not supported by the data in this experiment.

If Pierrehumbert's assertion that "the classification of stimuli in perception provides data for the probability distributions controlling production" (2003: 209) is true, we would expect to see correlations between the productive and perceptual patterns of subjects. For this study, the most important takeaway from Pierrehumbert's claim is that perception shapes production, thereby opening the door for perception-motivated sound change. In the production chapter we learned that approximantization in phrase-initial position is rare, but even in reading word lists—where we can expect hypercorrect speech—approximants appeared with some frequency. In fact, this position was not even the most conservative. As shown in Table 5-8, approximantization was even more restricted in post-nasal position. And across all environments, Colombians approximantized significantly less than Mexicans.

Table 5-8: Approximantization rates in Production

Place	Dialect	Phrase-initial	Post-nasal
/b/	Colombian	14.69%	5.56%
	Mexican	26.02%	3.33%
/d/	Colombian	8.72%	1.67%
	Mexican	13.95%	8.93%
/g/	Colombian	9.42%	0.00%
	Mexican	18.37%	3.45%

If Pierrehumbert is correct about perception influencing the probability distributions in production, we should see some parallels between production and perception in this experiment. Approximantization rates were quite low in phrase-initial position, and the ability of subjects to perceive approximantized pairs as same or different was also low as compared to stop pairs: 76% of stop pairs were judged same or different correctly, while only 56% of approximant pairs were judged correctly. In this case, then, the variant that is produced most frequently correlates with the most perceptually optimal variant.

The results of this experiment showed that the only significant dialectal difference was that Colombian subjects correctly identified stop pairs as same or different more frequently than

Mexican subjects. Accuracy rates for approximantized pairs were virtually identical for the two dialects. While the difference in stop accuracy was not great—81% vs. 70% in the configuration above—it is worth considering whether the higher rate of stops in the Colombian dialect could contribute to better judgments regarding place for stops in this perception experiment. Here we again see parallels with production, as the Colombian speakers produced significantly more stops in phrase-initial position. The inverse pattern, however, is not true: There is no correlation between a higher rate of approximantization for Mexican subjects and a higher rate of accuracy for these subjects when judging approximant pairs.

It was also shown above that accuracy rates improve when the consonants are followed by back vowels; no such effect was found for reaction times. In the analysis on production, each place of articulation had slightly different effects for different following vowels, but the one characteristic that was true for all was that the high vowels were less likely to approximantize than the low vowel /a/. There were not consistent effects for front or back vowels, and it is unclear why having a lower F2 value—the main distinguisher between front and back vowels—would be a more conducive perceptual environment.

This experiment aimed to evaluate whether place of articulation was more easily identifiable in stops or approximants, and from a theoretical perspective on sound change sought to assess whether perceptual optimization could be a motivation for phrase-initial stops ceding ground to approximants. The results of the experiment allow us to reject the strong hypothesis that approximantization is motivated by perceptual optimization of place of articulation in phrase-initial position for the Spanish stop series. While a perception-based, functional explanation for the appearance of approximants in phrase-initial position is unsatisfactory, this opens the door for other explanations. This experiment was limited to evaluating the phonetic optimization approach, and we will reserve final judgment on other explanations, e.g. Ohala's non-functionalist approach.

CHAPTER 6 — Experiment 4: ***Enhancement as a motivator for approximantization in intervocalic position***

1 Introduction

This experiment is motivated both by diachronic developments from Latin to Spanish and also by current research that has documented the voicing of voiceless stops in current varieties of Spanish. Given that we have been tracking how and why approximants may have developed in different environments and dialects of Spanish, it is important to consider how approximants came to be in the first place. The a priori explanation has always been articulatory, but this experiment will entertain a perceptual explanation by exploring a pre-approximant phase of development where the only contrast was between voiced and voiceless obstruents. The experiment design will allow us to compare perceptual confusability of voiced and voiceless minimal pairs, e.g. ***p****eso/****b****eso* and *va**c**a/va**g**a*, in phrase-initial position and intervocalic position. These two environments are at opposite ends of the approximantization scale, thereby facilitating a comparison of distinct environments. If perception plays a part in the spread of approximants, one possible explanation for the appearance of approximants in intervocalic position could be that the voiced obstruents were not sufficiently distinct from their voiceless counterparts in the same environment. Such an explanation would be in line with Flemming's Dispersion Theory of Contrast, which has three functional goals:

1. Maximize the distinctiveness of contrasts
2. Minimize articulatory effort
3. Maximize the number of contrasts

(2004)

For example, Flemming has claimed that the sound change from voiced stops to prenasalized voiced stops was motivated by a “strategy” to enhance voicing contrasts, i.e. maximize their distinctiveness. Citing Herbert (1986), Flemming argues that prenasalization of voiced stops only occurs when there is an existing contrast with voiceless stops. Absent this contrast, prenasalization does not occur. For Flemming, this pattern suggests that the motivation is not articulatory, but rather perceptual—specifically to enhance the contrast between voiced and voiceless minimal stop pairs (2005). Following this logic, approximantization in Spanish would have served to increase the distinctiveness of the contrast by adding a contrastive manner element to the minimal pairs. Whereas in Latin the contrast only involves voicing, in modern Spanish there is also a contrast in manner between the stop and approximant, e.g.:

Latin	**Old Spanish**
[vaca] →	[vaca]
[vaga] →	[vaɣ̞a]

Below, I will first give an overview of where today’s approximants came from before moving on to the details of the experiment.

2 Historical Background: From Latin to Spanish

Thanks to the many written records from Latin to modern Spanish, the evolution of the voiced obstruents is fairly clear. In intervocalic environments, Latin had a three-way contrast between voiceless geminates and singleton voiced and voiceless pairs:

Table 6-1: Latin Stops

	Bilabials	**Dentals**	**Velars**
voiceless geminate	-pp-	-tt-	-kk-
voiceless obstruent	-p-	-t-	-k-
voiced obstruent	-b-	-d-	-g-

The evolution of each of these sounds follows a similar path into old Spanish: The voiceless geminates become voiceless singletons; the voiceless obstruents become voiced intervocalically; and the voiced obstruents are either lost or spirantized:

Table 6-2: From Latin to Spanish

Latin phoneme	Old Spanish phoneme	Example
-pp-	/p/	CUPPA > *copa* 'wine glass'
-p-	/b/	CŪPA > *cuba* 'wine vat'
-b-	/β/	CIBU > *cevo* 'food'
-tt-	/t/	GUTTA > *gota* 'drop'
-t-	/d/ (= [ð])	CATĒNA > *cadena* 'chain'
-d-	Ø	SEDĒRE > *ser* 'to sit, be'
-kk-	/k/	SICCU > *seco* 'dry'
-k-	/g/ (= [ɣ])	SECURU > *seguro* 'sure'
-g-	Ø	LĒGĀLE > *leal* 'loyal'

(Penny, 72-81)

The contrast between voiced and voiceless singletons was maintained from Latin to Spanish in word-initial position. In the case of bilabials, Old Spanish maintained a distinction between /b/ and /v/ in word-initial position, but this contrast was already neutralized in Latin in intervocalic position. Penny (1991) claims that the full merger of these two phones was not complete until the late medieval and early modern periods, but it should be noted that in the production task speakers in both dialects were affected by orthographic 'b' and 'v'.

Modern Spanish exhibits voiced and voiceless stop pairs in several phonological environments. Examples are below in Table 6-3:

Table 6-3: Voiced and voiceless minimal pairs

	Bilabials		**Dentals**		**Velars**	
	voiceless	voiced	voiceless	voiced	voiceless	voiced
Phrase-initially	*pan* [pan] 'bread'	*van* [ban] 'they go'	*tos* [tos] 'cough'	*dos* [dos] 'two'	*col* [kol] 'cabbage'	*gol* [gol] 'goal'
Intervocalically	*lapa* [lapa] 'limpet'	*lava* [laβ̞a] 'wash'	*nata* [nata] 'cream'	*nada* [nað̞a] 'nothing'	*Paco* [pako] 'Paco'	*pago* [paɣ̞o] 'I pay'

If phones appear word-initially and are not involved in sandhi phonological processes, the contrast is generally between two obstruents, e.g. [p] vs. [b], although it was demonstrated in the production chapter that even when reading a word list, subjects occasionally approximantize phrase-initial voiced consonants. Intervocalically, the contrast also includes a contrast in manner, e.g. [p] vs. [β̞], as voiced stops are almost universally pronounced as approximants across dialects and voiceless stops are most often realized as voiceless stops.

In other environments, minimal pairs are less common, but examples of both voiced and voiceless sounds occurring in similar environments are not hard to find:

Table 6-4: Voiced and voiceless pairs in other environments

	Bilabials		Dentals		Velars	
	voiceless	voiced	voiceless	voiced	voiceless	voiced
post-nasal	*lámpara* [lampaɾa] 'lamp'	*ambas* [ambas] 'both'	*manta* [manta] 'blanket'	*manda* [manda] 'command'	*manca* [manka] 'one-armed'	*manga* [manga] 'sleeve'
post-lateral	*milpa* [milpa] 'corn field'	*alba* [alba] 'dawn'	*alto* [alto] 'tall'	*caldo* [kaldo] 'stew'	*calco* [kalko] 'calque'	*algo* [algo] 'something'

The contrast in voicing is constant across environments, but as the chapter on production showed, an additional contrast in manner can be found across virtually all environments to varying degrees. This is especially true in casual, rapid speech, where approximantization is extremely likely in intervocalic position. In both the Mexican and Colombian dialects, approximants can be found even in phrase-initial position, so in effect there can be a two-way voiced alternation that contrasts with the voiceless obstruent: *ves* [bes] or [β̞es] 'you see' vs. *pez* [pes] 'fish'.

In this experiment we will limit the study to phrase-initial and intervocalic position because minimal pairs are relatively easy to find in these positions and the two positions have opposite approximantization patterns, thereby facilitating a comparison between the two environments.

2.1 Voicing of /p, t, k/

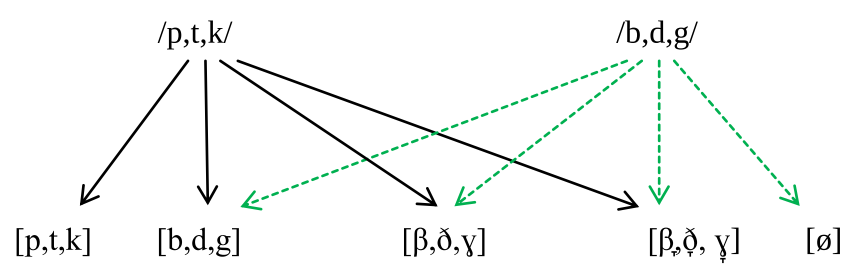
Several authors have noted that voiceless plosives have a tendency to become voiced across Spanish dialects, although like their voiced counterparts there is a great deal of variation. This tendency follows the historical tendency, where voiceless singletons became voiced intervocalically. Lewis (2001), in an unpublished doctoral dissertation, analyzed the speech of four speakers from Bilbao in Northern Spain and four speakers from Colombia—two from Bogotá and two from Medellín. His results showed that in an informal conversational style, Bilbaoans voiced /p, t, k/ 55% of the time. On the other hand, Colombians only exhibited voicing in approximately 10% of cases. Torreira and Ernustus (2011) analyzed the conversational Spanish of 52 speakers from Madrid and found that in intervocalic position voicing of voiceless stops was common: 33% of the stops were completely voiced and 62% of the stops were voiced for at least half of the duration. 25% of the stops were realized as approximants, and (it is assumed) the remaining 75% were realized as actual stops. Hualde (2011) cites several studies of individual dialects that show voicing of voiceless obstruents to varying degrees.

The fact that voiceless obstruents can be realized as voiced approximants in intervocalic position—and possibly elsewhere—raises the question of whether phonological neutralization happens in Spanish. What are the potential consequences for minimal pairs such as *la vaca* 'the cow' ~ *la vaga* 'the lazy person' losing the contrast and converging on [la β̞aɣ̞a]? It is certainly possible that—aside from sentential context—other cues, e.g. degree of constriction, length of adjacent vowels, or length of the phone itself, may aid in identifying the sound as a member of the voiced or voiceless set of phonemes, but there is unquestionably the potential for neutralization.

2.2 Potential for Confusability

The design of this experiment aims to address specifically whether the threat of neutralization in intervocalic position motivated an "enhancement" of the voiced alternations. If voiceless obstruents are once again encroaching on the domain of voiced alternants, it is reasonable to ask whether this may result in perceptual difficulties for the listener, and whether this could eventually have some effect on the phonological system of Spanish. The chart below illustrates how the varied production of the stop series could give rise to overlap—and potential confusion—in the intervocalic position:

Figure 6-1: Potential for confusability in intervocalic position



(Lewis:11)

To date, studies have produced varied results concerning the frequency of substitution of phonetic voiced approximants for voiceless stops. Most research seems to suggest that the stop closure is likely to be maintained even if voicing occurs, but it is difficult to make conclusive statements that apply to all dialects of Spanish. One interesting result from Lewis' (2001) dissertation is that Colombian speakers, who are known for their relatively conservative dialect in terms of approximant production, are much less likely than Peninsular Spanish speakers to exhibit voicing of intervocalic voiceless stops. This seems to suggest that the two processes may be linked in a sort of consonantal chain shift, where the intrusion of one consonant affects the production of another. Such a situation could lend support to advocates of Dispersion Theory if approximantization were seen as a means of creating perceptual distance from the encroaching voiceless stops that show increased degrees of voicing.

3 Experiment 4: Enhancement as a motivator for approximantization in intervocalic position

Examining the perceptual consequences of the possible domino effect described above—where the voicing of intervocalic voiceless obstruents leads to approximantization of voiced obstruents in the same position—can be approached in a variety of ways. The approach selected here is to turn back the clock to a time before approximants were found intervocalically. This study has been designed to have minimal pairs that differ by voicing in phrase-initial and intervocalic position—two positions on the opposite end of the approximantization continuum. It is hypothesized that confusability between voiced and voiceless stop pairs will be greatest in intervocalic position—the position where approximantization is most common. To simulate the pre-approximantization state of the language, only stops will be compared in this experiment. For a complete list of the words used in this experiment, see Appendix 5.

This experiment is a timed, forced-choice identification task where the subjects were asked to listen to a word and select the word they hear from a choice of two words on the screen. The two words differed by a single sound. For example, the subject heard *denso* and chose either ‘denso’ or ‘tenso’ by pressing the correct button on the button box. To ensure that results were not unduly influenced by ordering of the words on the screen, each word was used twice as a stimulus, with the ordering reversed a second time. For example, in the case of *denso*, this stimulus was offered twice—in one instance with the ordering of selections as “A. denso B. tenso” and in the other instance it was reversed to “A. tenso B. denso”. Selection of stimuli was randomized automatically by PsychoPy. Reaction times were also recorded.

The following predictions are guided by the idea that approximantization is an example of chain shift, motivated by a need to enhance the distinctions between voiced and voiceless stops in intervocalic position:

- Confusability between voiced and voiceless stop pairs will be greater in intervocalic position than in phrase-initial position. Given that approximantization is relatively rare in phrase-initial position, the need for enhancement would seem to be minimal; in effect, stops in phrase-initial position are already distinct enough, while in intervocalic position they are more confusable and in need of enhancement.
- Place of articulation will affect confusability rates. The production chapter showed, for example, that velars had the lowest rate of approximantization, and consequently we would expect velars to show less confusability than the other places of articulation in intervocalic position.
- Reaction Times will be longer in intervocalic position due to the higher confusability rates.

3.1 Subjects

The same group of speakers—31 Colombian and 29 Mexican—that participated in the other tasks also participated in this experiment.

3.2 Stimuli and Administration of Experiment

In terms of the comparisons that we want to make in this experiment, stimuli can be thought of in sets of 4 stops—two sets of minimal pairs that differ in voicing and word position. The example matrix below illustrates how the data for bilabials preceded by a low vowel was organized. Minimal pairs are in each column:

Table 6-5: Example matrix of data in Voicing Experiment

	Word-initial	**Intervocalic**
Voiced	*vago* [bago]	*cava* [kaba]
Voiceless	*pago* [pago]	*capa* [kapa]

All of the words except *cava* [kaba] in the above matrix occur naturally as a stop. In the production chapter we learned that voiced stops rarely appear in intervocalic position, so these stimuli had to be spliced together. For phrase-initial environments, no splicing was necessary, but I did ensure that the phrase-initial pairs were of a naturally comparable length. Previous studies, e.g. Colantoni (2010), have shown that voiced stops are shorter than voiceless stops, and this was also evident when comparing the phrase-initial stops in this study. On average the voiced stops were 20-30% shorter than their voiceless counterparts, and this was taken into consideration when creating the intervocalic voiced stops. All of the voiced stops used in intervocalic position were taken from stops occurring in natural phrase-initial positions; they were later spliced into intervocalic position by inserting a stop closure and including the first vowel peak and matching it at a zero crossing. For phrase-initial stops, essentially the same procedure was used, as voiced stops were spliced onto the voiceless stops. The diagram below in Figure 6-2 shows how the voiced minimal pairs were created.

Figure 6-2: Stimulus Creation for Voicing Experiment

Word-Initial stimulus:	[pago] →	[_ago]	→ [_ago] + [b] = [bágo]
	[bágo] →	[b]	
Intervocalic Stimulus:	[kapa] →	[ka_a]	→ [ka_a] + [b] = [kába]
	[bagó] →	[b]	

The advantage of inserting the voiced stops into the voiceless minimal pair is that it ensures the only difference is between the voiced and voiceless stops. Stress of the spliced voiced stops was also matched to the resulting stimulus to ensure maximum naturalness.

Following vowel height was also included as a factor in this experiment. For both dentals and bilabials three-way voiced and voiceless pairs were followed by /i/, /e/, and /a/. For velars there are no di-syllabic minimal pairs for the mid-vowel, so only the high and low vowels were included as following environments. The complete list of 34 words can be found in Appendix 6.

Speakers of Spanish will note that the words in the appendix have stress patterns associated with word position. With the exception of the words *repela/revela*, all intervocalic stimuli are onsets of unstressed syllables and all phrase-initial stimuli are onsets of stressed syllables. Results from experiment one indicated that accuracy rates improved for stimuli in stressed environments, and this will have to be taken into consideration when analyzing the results below.

In all, 3,999 responses were submitted for analysis. As in previous experiments, reaction time was used as a means of culling invalid responses. In all, 6 responses were eliminated because of response times that were too fast (<300ms), and 7 responses were removed because of unreasonably lengthy response times (>7,000ms).

3.3 Accuracy Results

The subjects correctly distinguished between voiced and voiceless minimal pairs at a high rate. Overall, 87% of the responses were correct. The regression model (reported below) found no significant differences between place of articulation, i.e. between bilabials, dentals, and velars, so the summary below does not separate these. The highest accuracy rate for both dialects was achieved in phrase-initial position when the phone in question was voiced: 98% for Colombian speakers and 91% for Mexican speakers. The lowest accuracy rate for both dialects was phrase-initial position with voiceless phones—79% and 81% for Colombian and Mexican speakers respectively.

Table 6-6: Accuracy

Dialect	Environment	Voiced	Correct	Incorrect	N
Colombia	**Intervocalic**	Voiceless	94.06%	5.94%	539
		Voiced	86.30%	13.70%	540
	Intervocalic Total		**90.18%**	**9.82%**	**1079**
	Word-initial	Voiceless	78.96%	21.04%	480
		Voiced	98.13%	1.88%	480
	Word-initial Total		**88.54%**	**11.46%**	**960**
Colombia Total			**89.41%**	**10.59%**	**2039**
Mexico	**Intervocalic**	Voiceless	85.49%	14.51%	517
		Voiced	83.98%	16.02%	518
	Intervocalic Total		**84.73%**	**15.27%**	**1035**
	Word-initial	Voiceless	81.17%	18.83%	462
		Voiced	90.50%	9.50%	463
	Word-initial Total		**85.84%**	**14.16%**	**925**
Mexico Total			**85.26%**	**14.74%**	**1960**
Grand Total			**87.37%**	**12.63%**	**3999**

Of the 34 words that were included in this experiment, only five were responded to incorrectly greater than 20% of the time. These words include three voiceless phrase-initial words, one voiced intervocalic word, and one voiceless intervocalic word. The list of these words is below in Table 6-7, and the complete list of words and their error rates can be found in Appendix 6.

Table 6-7: Five Most Often Incorrectly Identified Words

Word	Incorrect
quiso [**k**iso]	50.8%
pibe [pi**b**e]	51.3%
peso [**p**eso]	33.1%
peque [pe**k**e]	23.3%
tique [**t**ike]	22.2%

To see what factors significantly affected accuracy results, a mixed effects logistic regression was run on the 3,999 responses. The variables that were included are listed below:

Table 6-8: Experiment 4 Variables

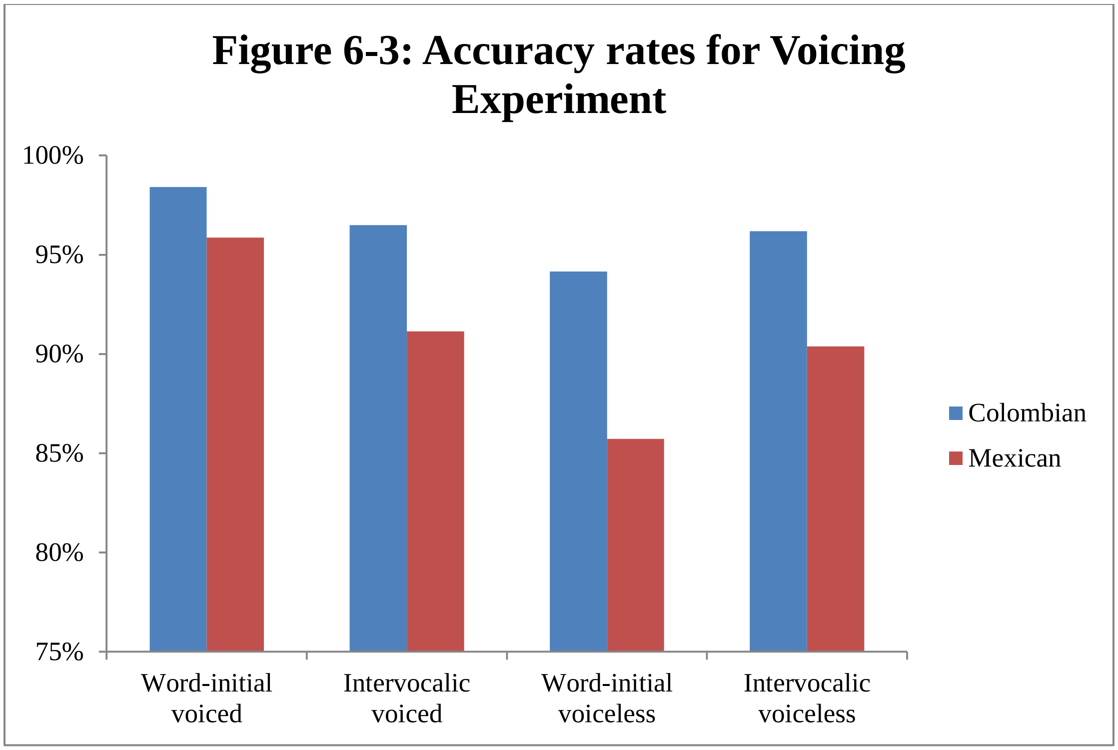
Variable Name	Default Category	Other Categories
Dialect	Colombian	Mexican
Voicing	Voiceless	Voiced
Vowel	/a/	/e/, /i/
Position	Intervocalic	Word-Initial
Place of Articulation	/b/	/d/, /g/

It was predicted above that place of articulation would be a significant predictor in this model, but this was not found to be true. Using the technique of comparing two models by using an ANOVA, it was shown that the addition of place of articulation as an explanatory variable did not result in an improved model over the simpler model without place of articulation ($\chi2(2)=1.7019$, $p=0.427$). This suggests that the effect of voicing on word recognition is the same across all three places of articulation, and for this reason place of articulation was ultimately left out of the models analyzed below. The inclusion of vowel height also did not improve the model ($\chi2(2)=1.4574$, $p=0.4825$), and this was also left out of the final model that will be analyzed below. Random intercepts for word and subject were also included in the final model, as well as by-word random slopes for dialect and by-subject random slopes for Voicing and Context were also included.

Table 6-9: Enhancement Experiment Regression Results

```
Number of obs: 3999, groups: Participant, 59; File, 34

Fixed effects:
                      Estimate Std. Error z value Pr(>|z|)
(Intercept)            3.22625    0.44993   7.171 7.47e-13 ***
DialectMexico         -0.98564    0.39661  -2.485   0.0129 *
Voiced                 0.08991    0.43771   0.205   0.8373
PhraseInitial         -0.44701    0.40882  -1.093   0.2742
Voiced:PhraseInitial   1.25795    0.58539   2.149   0.0316 *
```

The dependent variable is accuracy (Incorrect=0, Correct=1), so coefficients with a positive sign indicate that they increase accuracy over the default category, while negative signs indicate that accuracy decreases as compared to the default category. In the model there is a significant interaction between word position and voicing. To more easily see this interaction, I include a chart below that illustrates the accuracy rates for the four combinations of word position and voicing for the two dialects. There was not a significant interaction with Dialect, but this variable is included below to demonstrate the significant main effect it has on the model.



We can see the significant main effect of Dialect in this model reflected in the overall lower accuracy rate for Mexican speakers: The log odds of selecting a correct answer was decreased by .986 if the speaker was Mexican (p=.0129). In the scenarios above in Figure 6-3, Mexican speakers' accuracy rates were between 2.5-8.5% lower than Colombian speakers. The other significant variable in the model was the interaction between Position and Voicing, which has a coefficient of 1.258. That this interaction is significant indicates that the effect of Position depends on Voicing (p=.0316). There are essentially four different comparisons that can be made based on this interaction—two each for voicing and place of articulation. For the voiced category, the log odds of a correct response is higher for phrase-initial than for intervocalic, although this difference is only marginally significant (p=.0828)[7]. This difference is reflected in the first two pairs of columns above. For the voiceless category, the log odds of a correct

[7] This p-value was calculated by setting the defaults to different values from those in the model above.

response is lower for phrase-initial than for intervocalic, although this is not considered a significant difference (p=.2742). For the phrase-initial category, the log odds of a correct response is significantly higher for voiced stimuli than for voiceless stimuli (p=.00428)[8], and this difference is reflected by the first and third pairs of columns above. In intervocalic position, whether the stimulus was voiced or voiceless had no bearing on accuracy rates (p=.8373).

3.4 Reaction Time Results

The measuring of reaction time can reveal the difficulty involved in processing, and it would be expected that more perceptually challenging tasks would have longer reaction times. In this experiment it was expected that longer reaction times could be expected in the environment where perception was most difficult. Based on the chain shift hypothesis, this environment was expected to be intervocalic position.

Reaction times were recorded by measuring from the moment of the offset of the phone of interest to the moment when the subject responded. All reaction times less than 300ms and greater than 7 seconds were deemed outliers and excluded from analysis. Prior to analysis all RTs were log transformed to normalize the natural skewing of reaction time data. Only the correct responses—3,494 tokens— were admitted for this analysis.

One potential source of imbalance in the way reaction times are measured is their position in the word. When a stimulus contains a phrase-initial stop, e.g. *peso* [**p**eso], the subject hears the critical contrast sooner than when the stimulus is an intervocalic stop, e.g. *peque* [pe**k**e]. It was ultimately determined that this was not an issue for calculating reaction times because the subject could view the two choices for .5 second on the screen before responding. It

[8] This p-value was calculated by setting the defaults to different values from those in the model above.

is therefore fair to assume the subject knew at which point in the word to look for the contrast. This point will be discussed below when looking at the model results.

As in the case of the logistic regression analysis above, the variables Place of Articulation and Vowels were not included in this model because likelihood ratio tests using the ANOVA technique revealed that the inclusion of either of these variables did not result in a better fit. The model below also does not include the interaction term between Voicing and Context for the same reason, as its inclusion also did not result in a better fitted model.

Table 6-10: Reaction time results

```
Number of obs: 3494, groups: Participant, 59; File, 34

Fixed effects:
                  Estimate Std. Error t value
(Intercept)      2.9428742  0.0274686  107.14
DialectMexico    0.1517160  0.0281330    5.39
Voiced          -0.0007523  0.0222070   -0.03
PhraseInitial   -0.0076482  0.0219198   -0.35
```

The dependent variable for the above model is log (RT/ms). Positive coefficients indicate that reaction time of the listed category takes longer than the default category. Neither Voicing nor Position were significant predictors of reaction time. The coefficient for Position, whose listed category is phrase-initial position, is -.008 and is not significant, indicating there is virtually no difference between the two contexts. Overall, the average reaction time for phrase-initial position was 1.103 seconds, while for intervocalic it was 1.154 seconds. The original hypothesis predicted longer reaction times for intervocalic position because it was believed processing would be more taxing in intervocalic position. Reaction times were measure from the offset of the distinguishing stop in phrase-initial or intervocalic position. To use the example words from above, having equivalent reaction times implies that the reaction times measured from the two offsets are essentially the same:

Figure 6-4: Calculating RTs for Word-initial and Intervocalic Positions

p | e s o ……………….[button press]

p e **k** | e ………………….[button press]

The two double arrows represent the same reaction time. In these two examples, speakers would have seen two options for each word:

A. peso B. beso
A. pegue B. peque.

The options appear .5 second before the stimulus is played, and I believe it is reasonable to assume that subjects would have noted the orthographic difference before hearing the stimulus. For this reason I believe the points of measurement are equivalent, and I can therefore conclude that there is no significant difference in reaction times between phrase-initial and intervocalic positions.

4 Discussion

It was established in the production chapter that there are significant differences between approximantization rates in intervocalic and phrase-initial positions. While the three voiced stops in the stop series approximantize at a very high rate in intervocalic position, the opposite is true for phrase-initial position. In the chain shift hypothesis described above, it was hypothesized that a possible motivation for approximantization in intervocalic position—beyond the traditional articulatory explanation—was to increase the contrast between voiced and voiceless stops. Following this hypothesis, the addition of the manner contrast could be motivated by a need to maintain contrasts in intervocalic position—a position that has been encroached upon historically by partial voicing of voiceless stops. The same motivation would be absent in phrase-initial position, where voiceless stops have not shown a tendency to become voiced. If the chain hypothesis were true, we would expect a correlation between perceptual difficulty, word position, and rate of approximantization, i.e. we would find lower rates of accuracy and slower reaction times in intervocalic position than in phrase-initial position.

The predictions made at the beginning of this chapter—in line with the chain shift analysis described above—do not hold when looking at the results. The first prediction—that place of articulation would be a significant predictor in this experiment—was not upheld by initial analysis and was not included in the models that were ultimately analyzed. It was also initially predicted that confusability between voiced and voiceless stop pairs would be greatest in intervocalic position, but the results from the model analysis above do not support this prediction. There were no significant differences between phrase-initial and intervocalic for either voiced or voiceless, although for voiced stimuli the difference in accuracy rates was almost significant. In this case, phrase-initial position had a higher rate of accuracy, i.e. the predicted effect of the prediction was nearly significant. While there was a significant difference involving voicing in phrase-initial position—voiced phones were more accurately perceived—this does not reflect on the more important claim regarding word position. It should also be noted that the stress confound mentioned earlier does not affect this result. Even though all but one intervocalic pair of stimuli was unstressed and all phrase-initial pairs were stressed, word position still did not affect accuracy rates in the expected direction. Recall that stress had a positive effect on perception in experiment 1, so the fact that phrase-initial position—always stressed—does not generate significantly more correct answers than intervocalic position—almost always unstressed—is actually further confirmation that word position does not factor into confusability for voiced and voiceless minimal stop pairs. The initial prediction that intervocalic position is a more confusable position must therefore be rejected.

Reaction times were not affected by any of the word-specific properties of Voicing or Position, thereby indicating that intervocalic position was not more demanding in terms of processing, and that voiced and voiceless items were equally demanding. The Dialect variable was found to be the only significant predictor in the reaction time model, where Mexican subjects took significantly longer than Colombian subjects to respond to the stimuli. This was true regardless of word position or voicing properties. The effect of Dialect was not a significant predictor of reaction time in any of the other experiments, so in order to explain the significance of this variable we should start by looking at how this experiment differs from the others. Two of the other experiments were forced choice same/different tasks, where subjects listened to pairs of words and selected "same" or "different." The other experiment was a repetition task; subjects simply repeated the word that they heard. The experiment discussed in this chapter was different because subjects had to listen to a word and choose from two words on the screen. For all but two of the word pairs there was only a single letter that differed in orthography, but this detail may have been enough to slow down subjects with lower literacy levels. The two populations under investigation here had dramatically different attained levels of education: The majority of the Colombian speakers were either in the process of completing a college degree or had already, while many of the Mexican speakers had not advanced to secondary school. As a result, it is believed that the significant effect of Dialect in this experiment is attributable to the divergent literacy rates in the two populations.

5 Conclusion

This chapter has addressed one of the perceptually-motivated explanations for sound change—namely Flemming's notion of enhancement, where speakers "appear to take measures to increase the distinctiveness of contrasts" (2003:173-4). If speakers were motivated by enhancement, we would have expected to see significant differences in perceptual difficulty between phrase-initial position—where approximantization rarely takes place—and intervocalic position—where approximantization is very frequent. The results of this experiment suggest that enhancement does not motivate approximantization in intervocalic position, but other perceptually-motivated explanations for approximantization remain in play.

CHAPTER 7 — Conclusions

This study has sought to investigate the possible role of perception in the production of a phonological alternation in two Spanish dialects. Past studies described the alternation between stops and approximants as a purely articulatory-driven process, but these accounts cannot explain why approximants appear, for example, in phrase-initial position. When approximants are recorded in phrase-initial position when speakers read from a word list—a task that should encourage hypercorrect speech—this suggests that more than just articulatory processes are at work. For this reason, I have explored alternative explanations for approximantization throughout this study. In this final chapter, I will revisit the major findings concerning production of voiced stops and approximants and any links with perception. I will also relate these findings to the perception-based theories of sound change that motivated the design of the experiments, and, given the inadequacy of the perception-based theories, I will offer an alternative hypothesis. Finally, I will highlight the major accomplishments of this dissertation and propose future avenues of research.

1 Production Results

In Amastae's (1995) study on three Spanish dialects—Colombian, Mexican, and Mexican-American—he found variability in approximantization rates both within and across dialects. Despite collecting 42,000 tokens collected from a word list, the effects of different phonological environments within the same dialect differed for each place of articulation, and the ordering of effects across dialects was not consistent. For example, the effect of preceding /r/, /s/, or /l/ was dependent on both place of articulation and dialect; in some cases the preceding trill caused more approximantization than the other two segments, whereas in other instances the trill caused the least among the three. Despite the significant variation within and across dialects in his study, two conclusions reached in Amastae's study were also supported here:

1. Colombian speakers approximantize less than Mexican speakers.
2. The three places of articulation are significantly different from one another.[9]

In terms of the production patterns described in this study, findings were largely in line with other descriptions. The points below highlight the most significant findings that relate to the perception experiments and the following discussion:

- Overall, Colombian speakers approximantized less often than Mexican speakers. However, in the case of dentals, lower approximantization by Colombians was only recorded for post-fricative position.
- Stress inhibits approximantization, i.e. phones in stressed syllables approximantize less often.
- The following low vowel /a/ promotes approximantization as compared to the other vowels—most specifically the high vowel /i/.
- Post-nasal position is the strongest inhibitor of approximantization.
- Word-initial position strongly impedes approximantization, but it should be noted that about 15% of all phrase-initial tokens were approximantized.
- Post-lateral position strongly impedes approximantization of dentals. Only 8.5% of all post-lateral dental tokens were approximantized. The rates for bilabials and velars were 41% and 24.8% respectively.

[9] Amastae actually had 4 "segments" because he treated orthographic '*v*' as a distinct variable.

- Of all post-consonantal environments, post-fricative position was the most permissive of approximants. Dentals approximantized at a rate of 35.3%, bilabials at 54.2%, and velars at 17.8%.
- A small sample of words paired with the definite articles *el* and *la* were elicited to investigate sandhi contexts. It was determined that sandhi contexts behaved no differently from post-lateral or intervocalic contexts.

2 Links between Production and Perception

In each of the four experiment chapters, I discussed whether there was a link between production and perception. In this section I will consolidate the findings from the separate discussions.

The first experiment tested the confusability hypothesis, which predicted that phonological environments that promote the realization of approximants, e.g. intervocalic position, would exhibit higher rates of confusability. It was believed that if confusability was a mechanism causing approximantization, higher rates of approximantization should be found in the more confusable environments. High rates of confusability were evident in the experiment, as both Colombian and Mexican subjects incorrectly identified non-identical pairs as identical at a very high rate. However, of all the variables that promoted approximantization, only unstressed position aligned with the confusability hypothesis, i.e. it was the only condition that showed both increased approximantization and increased confusability. That intervocalic position—by far the most likely environment to approximantize—did not show this tendency certainly casts doubt on the hypothesis. Dialect also did not show the expected tendencies to support the hypothesis. For these reasons, it was decided that the confusability hypothesis should be rejected.

In experiment two, subjects repeated words that began either with a stop or an approximant. This study was limited to phrase-initial position, which has a low approximantization rate. It was shown that subjects were sensitive to the manner of the stimulus, as the given manner significantly affected the manner produced in correct responses. If the stimulus was an approximant, for example, subjects were more likely to produce an approximant. However, subjects from both dialects overwhelmingly preferred stop pronunciations. In terms of response time and accuracy, the most commonly produced variant in the production task was also the quickest and most accurate: Stops were correctly repeated more often and more rapidly than approximants. The error analysis showed that the most common error involving approximant stimuli was place of articulation, while the most common error for stop stimuli involved voicing.

The third experiment looked directly at place identification in phrase-initial position using nonce stimuli, and the results largely confirmed what the error analysis in the previous experiment suggested. It was easier for subjects to discriminate the place of articulation differences of stop stimuli than approximant stimuli in this position. And like the repetition task, subject accuracy was significantly better when the stimulus was the more frequently produced variant, i.e. when the stimulus was a stop.

In experiment 4, we compared whether voiced or voiceless stops were more easily perceived in phrase-initial and intervocalic positions. The comparison in this experiment was focused on the difference in accuracy between intervocalic and phrase-initial positions, and the expectation was that the very different approximantization rates in these two positions would be reflected in the perception experiment. In particular, it was expected that intervocalic position would show lower accuracy rates. Results showed, however, that there were no significant differences between phrase-initial and intervocalic for either voiced or voiceless variants.

In the following sections, I will relate the above results to the theoretical motivations that inspired the experiments.

3 Perception-based Theoretical Explanations

Each of the experiments was motivated by one or more aspects of the perceptually-based theories of sound change outlined in the introduction. Following Blevins & Garrett (2004), I assigned each theory either a functionalist or non-functionalist designation. As a reminder, the main characteristics of each approach are:

Table 7-1: Summary of perception-based theories of sound change

Theoretical Position	**Characteristics**	**References**
Innocent/Non-functionalist	non-teleological; no improvement in articulation or perception necessary; perception is not part of the grammar; listener-initiated sound change	Ohala (1981, 1990, 1993 *inter alia*); Blevins (2004)
Optimizing/Functionalist	teleological; improved functionality; perceptual knowledge part of a speaker's phonology; speaker-initiated sound change	Steriade (2001); Flemming (2003)

Three of the four experiments specifically addressed functionalist motivations for sound change. In experiment two, we examined whether perceptual cues could be a factor. Wright (2004) claims that place of articulation is easier to recover in fricatives than in stops, and I hypothesized that this notion should be extendable to approximants. In both fricatives and approximants, listeners can rely on the more lengthy internal cues to make judgments about voicing and place, and it was believed that formants in the approximant would provide valuable place cues. For stops, subjects must rely on the relatively short release burst and formant transitions. Results from the experiment showed that voicing was very rarely misperceived when the stimulus was an approximant, while for stops—although errors were less common overall—voicing errors were the most common mistake. Conversely, mistaken place identification was much more common for approximant stimuli than for stop stimuli. That place identification in stops is easier for Spanish speakers was confirmed in experiment three. In this experiment, speakers of both dialects were able to discriminate place of articulation differences more easily when the stimuli were obstruents.

Given that the results of experiments 2 and 3 show that stops are perceived more quickly and accurately, this suggests that phonetic optimization cannot be a motivation for approximantization. It also calls into question the original assumption that the internal cues in approximants would be better place cues than stops. At least for the Spanish speakers in these two dialects, place cues were better in stops. It is suspected that the transitional cues from stop to vowel are in some way better than the transitional cues from approximant to vowel. Further research would be necessary to determine whether the original assumption that approximants pattern like fricatives in terms of scales of perceptibility was justified.

Experiment 4 sought to evaluate whether perceptual enhancement was a motivation for approximantization, but there was no evidence that this was the case. The hypothesis was that intervocalic position would be a more confusable position than phrase-initial position for the voicing contrast, thereby explaining why the addition of the manner contrast occurs preferentially in intervocalic position. However, accuracy rates were not significantly different in the two positions, and as a result the chain shift hypothesis was rejected.

Experiments 2, 3 and 4 show conclusively that perceptual optimization does not motivate approximantization. In looking at non-functionalist explanations, one of the pre-conditions of sound change for Ohala is having a situation of confusability between two sounds. Previous descriptions (e.g. Lavoie 2001) of Spanish have stated that the stop and approximant alternants were indistinguishable for native speakers, but these claims lacked empirical support. Experiment one supported this assertion to some degree. Speakers of both dialects had significant difficulty detecting when the stimuli pairs were different. However, for approximantization to be associated with confusability, higher confusability rates needed to be associated with environments that promote approximantization. There were two results that demonstrate this association:

- Unstressed syllables facilitate approximantization and inhibit perception. This is the lone phonological predictor from the accuracy models that supports the confusability hypothesis, which predicted that for phonological environments that promote the realization of approximants, higher rates of confusability would also be evident.
- In experiment 1, reactions times for intervocalic position—where approximantization is frequent—are significantly higher than other word-medial positions that inhibit approximantization. This result suggests that the perceptual processing load in intervocalic position is higher and might lead to more accommodations.

The evidence in support of the confusability hypothesis is not overwhelming. In the accuracy model, variables such as Dialect, Vowel, and Context did not show significant differences in the predicted direction, and for this reason the confusability hypothesis cannot be accepted.

On the whole, the experiments did not demonstrate that either misperception or optimized perception is a mechanism of change in Spanish approximantization. It is likely that the high rate of perceptual equivalence shown in experiment 1 may be a facilitating factor to the spread of approximants, but this can only be hypothesized at this point.

4 An alternate explanation

While neither the functionalist nor non-functionalist perception-based explanations for sound change seem to apply to the case of Spanish approximantization, a purely articulatory explanation cannot account for all of the variation in Spanish either. Even in the most prohibitive environment for approximantization, post-nasal position, approximantization occurs in 4% of the words produced in the production task, and for phrase-initial position the rate is almost 15%. In casual speech, these percentages are likely to be higher. In both of these positions, traditional descriptions of the “rule” expect stops. In the case of post-nasal position, this is the expectation because the seal is already in place. For phrase-initial position, without a preceding vowel speakers should have no trouble making a closure and hitting the “target.” Absent a satisfactory articulatory explanation, how do we explain the appearance of approximants in phrase-initial position? Returning to the production task, recall that the sandhi pairs, *el* or *la* + NOUN, had the same patterns as word-internal post-lateral and intervocalic positions. In other words, as long as an individual word is not phrase-initial, phrase-initial /b/, /d/, and /g/ will most likely be subject to the same phonological processes as when they appear word-medially. And if we consider that relatively few words actually occur phrase-initially in natural speech, native Spanish speakers will have plenty of experience hearing and producing phrase-initial approximants. This suggests that future research should incorporate corpus-based studies or production that more closely resembles actual usage.

Two of the experiments focused on phrase-initial position because it was suspected that this was an environment resistant to change but with some variation. I predicted that the perceptual “advantages” of approximants could be what is swaying approximantization to increase in this position. This hypothesis was disproven by results from the experiments. The results, however, follow a common thread: Subjects perceived and reproduced most accurately those variants that they have the most experience with in phrase-initial position. In the production task, subjects produced stops more than 85% of the time, and in the perception tasks where phrase-initial position was the focus, stops were clearly better. The repetition task showed that the given manner was not lost on subjects, but their own production patterns were far more influential.

The idea that “experience” is an influential aspect of production and perception is not new. Joan Bybee has made experience, or language use, the foundation of her theories on phonology and language change (e.g. Bybee 2001, 2002). And as mentioned in the introduction, Pierrehumbert believes that “the classification of stimuli in perception provides data for the probability distributions controlling production” (2003:209). Both Bybee and Pierrehumbert would agree that input shapes output, and that the perception/production loop is continuously shaping language. In this study, in the two experiments that pitted approximants and stops against one another—the repetition task and the nonce word task—subjects were quicker and more accurate with the stimuli they were most familiar with, i.e. stops in phrase-initial position. In addressing similar issues on lenition in Spanish, Brown & Raymond (2012) have shown that experience with particular discourse contexts—including extralexical contexts like the sandhi word pairs in this study—is a strong predictor of reduction rates. In short, this study on Spanish approximantization has suggested that the usage patterns of a particular language may even override putative universal factors such as scales of perceptibility. And it is believed that these usage patterns continue to shape the dialects that were the object of this study.

5 Final Points & Future Directions

The role of perception in this sound change in Spanish was found to be negligible, but nonetheless this dissertation still makes significant contributions to the field and paves the way for future work. In the following, I will highlight the most significant insights and most promising directions for future research.

5.1 Insights on Production

Production patterns of the two dialects for the most part followed the patterns as described in other studies. Colombian speakers are still "conservative" in their use of approximants, while Mexican speakers approximantize at a higher rate. Dental patterns, however, showed very little difference between the two dialects. Why bilabials and velars would show the predicted dialect differences but dentals would not was not clear from this study. To further explore this discrepancy, a corpus-based analysis of conversation data would be best. Also, comparisons with other dialects would be fruitful to see if this pattern is seen across the many Spanish dialects.

Two variables that are not always considered were shown to affect approximantization in this study. The low vowel /a/ induced significantly more approximantization than other vowels in both the word-list task and the repetition task. Stress was also a significant predictor, as phones found in stressed syllables were significantly less likely to approximantize than those found in unstressed syllables. These two factors should be considered in future studies on approximantization.

In addition to validating the previous descriptions of these two dialects and their general trends, the chapter on production highlighted some complications with relying solely on intensity ratios to measure approximantization. While intensity ratios work fairly well for intervocalic position, it works less well for measuring post-nasal consonants and does not account for the influence of stress. Future work on approximantization will need to take these factors into account.

5.2 Insights on Stored Representations

At the beginning of this study, I mentioned that much of the debate surrounding the approximant/stop alternation centers on whether this is a process of lenition or fortition. This debate naturally entails a position on what the “base” form is. I stated that the truly interesting question should center on the nature of the stored representations. Some insight can be gained on this question as a byproduct of experiment two, where subjects repeated the word that they heard. The stimuli varied by manner in phrase-initial position, and there were equal numbers of approximant-initial and stop-initial stimuli. Subjects overwhelmingly repeated the words using stops, suggesting that in phrase-initial position the strongest representation is a stop. This result motivates an expansion of this type of experiment in other phonological environments that showed different approximantization patterns. For example, given the same study design but focusing on intervocalic or post-fricative position, what would subjects produce?

5.3 Insights on the Perception/Production Loop

Pierrehumbert has proposed that perceptual classifications control production patterns (2003). In the repetition task, subjects showed they were sensitive to sub-phonemic detail by producing more approximants when the given manner of the stimulus was an approximant. In effect, this would indicate that they “classified” the stimuli as approximants and reproduced them as such. However, as mentioned above, they still strongly favored stops in phrase-initial position. The fact that subjects were affected by the given manner in this experimental setting implies a certain degree of integration in the perception/production loop, but more research is necessary to determine whether the ability to perceive sub-phonemic details plays a role in sound change.

5.4 Insights on Confusability and the Role of Perception in Sound Change

The confusability experiment confirmed what had previously only been expressed anecdotally: Native speakers have difficulty distinguishing stops from approximants. Although they are not completely indistinguishable, there is a high rate of perceptual equivalence. Establishing perceptual equivalence, or at least a high rate of confusability, was a critical pre-condition to explore non-functionalist, perceptually-motivated theories of sound change in the spirit of Ohala. There is strong evidence that this pre-condition has been met, but given that confusability did not show the expected correlation with approximantization, we cannot say that perceptual equivalence is the mechanism of this sound change. I do not wish to suggest, however, that perceptual equivalence plays no role, and it may in fact be a pre-condition or even a facilitator of sound change. Further research will be necessary to determine the exact role that perceptual equivalence plays.

6 Concluding Remarks

Spanish stops have a long history of change, and the synchronic variation observed in these two dialects suggests that change is ongoing. A full account of the mechanisms that propel this change onward remains to be written, but this study has served to narrow the possibilities and point to new avenues of investigation.

References

Amastae, Jon. (1989). The intersection of s aspiration/deletion and spirantization in Honduran Spanish. *Language variation and change* 1. 169–183.

Amastae, Jon. (1995). Variable Spirantization: Constraint Weighting in Three Dialects. *Hispanic Linguistics* 6.1-2. 1-20.

Barlow, Jessica A. (2003). The stop-spirant alternation in Spanish: Converging evidence for a fortition account. *Southwest Journal of Linguistics* 22, 51-86.

Barr, D. J., Levy, R., Scheepers, C., & Tilly, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, 68, 255–278.

Bates, D., Maechler, M., & Bolker, B. (2011). *lme4: Linear mixed-effects models using S4 classes*. R package version 0.999375-41. http://CRAN.R-Project.org/package=lme4.

Beddor, P. S. (2009). A Coarticulatory Path to Sound Change. *Language* 85 (4), 785–821.

Blevins, J. (2004). *Evolutionary Phonology: The Emergence of Sound Patterns*. Cambridge, Cambridge University Press.

Blevins, Juliette. (2007). Interpreting misperception: beauty is in the ear of the beholder. In Maria-Josep Solé, Patrice Streeter Beddor & Manjari Ohala (eds.) *Experimental approaches to phonology*, 144-154. Oxford: Oxford University Press.

Blevins, Juliette, and Andrew Garrett. (2004). The evolution of metathesis. In Bruce Hayes, Robert Kirchner, and Donca Steriade, eds., *Phonetically Based Phonology*. Cambridge: Cambridge Univ. Press, 117-156.

Blevins, Juliette and Andrew Garrett. (1998). The origins of consonant-vowel metathesis. *Language* 74:3. 508-556.

Boersma, Paul & Weenink, David (2011). Praat: doing phonetics by computer [Computer program]. Version 5.2.28, retrieved 12 September 2011 from http://www.praat.org/

Boersma, Paul. (1998). *Functional phonology: Formalizing the interactions between articulatory and perceptual drives*. Doctoral dissertation, University of Amsterdam.

Brown, Esther and Earl Brown. (2012). Syllable-final and Syllable-initial /s/ Reduction in Cali, Colombia: One Variable or Two? In Richard File-Muriel & Rafael Orozco, eds. *Linguistic Studies in Colombian Varieties of Spanish*. Madrid: Iberoamericana, 89–106.

Brown, Esther & William D. Raymond (2012). How Discourse Context Shapes the Lexicon: Explaining the Distribution of Spanish f- / h- words. *Diachronica*, 29.2, 139–161.

Brown, Esther. (2013).Word Classes in Studies of Phonological Variation: Conditioning Factors or Epiphenomena? *Selected Proceedings of the 15th Hispanic Linguistics Symposium*. Ed. by Chad Howe, Sarah Blackwell & Margaret Lubbers Quesada. Somerville, MA: Cascadilla Proceedings Project. 179-186.

Bybee, Joan. (2001). *Phonology and language use*. Cambridge: Cambridge University Press.

Bybee, Joan. (2002). "Word Frequency and Context of Use in the Lexical Diffusion of Phonetically Conditioned Sound Change". *Language Variation and Change* 14. 261–290.

Canfield, D. Lincoln. (1981). *Spanish pronunciation in the Americas*. Chicago, IL: University of Chicago Press.

Carrasco, P. (2008). *An acoustic study of voiced stop allophony in Costa Rican Spanish*; doct. diss. University of Illinois at Urbana-Champaign.

Carrasco, Patricio & José Ignacio Hualde. (2009). *Spanish voiced obstruent allophony reconsidered*. Manuscript, Roanoke College.

Carrasco, P., J.I. Hualde & M. Simonet. (2012). Dialectal differences in Spanish voiced obstruent allophony: Costa Rican vs. Iberian Spanish. *Phonetica* 69: 149-179.

Colantoni, Laura & Irina Marinescu. (2008). The scope of stop weakening in Argentine Spanish. Paper presented at *Fourth Conference on Laboratory Phonology.* University of Texas, Austin.

Colantoni, L. (2011). Laboratory approaches to sound variation and change. In Diaz-Campos, M (ed.), *The Handbook of Hispanic Sociolinguistics*. UK: Wiley-Blackwell. 9-35.

Cole, J., Hualde, J. I., & Iskarous, K. (1999). Effects of prosodic and segmental context on /g/ deletion in Spanish. In O.Fujimura, B. D. Joseph, & B. Palek (Eds.), *Proceedings of the fourth linguistics and phonetics conference.* 575-589.

Davies, Mark. (2002-) *Corpus del Español: 100 million words, 1200s-1900s*. Available online at http://www.corpusdelespanol.org.

Eddington, David. (2011). What are the contextual phonetic variants of /β,ð,γ/ in colloquial Spanish? *Probus* 23:1-19.

Face, Timothy L. (2002). Disentangling the necessarily entangled: The phonology and phonetics of Spanish sprirantization. *Southwest journal of linguistics* 21. 55–71.

Fernandez, Joseph. (1982). The allophones of /b, d, g/ in Costa Rican Spanish. *Orbis* 31. 121–146.

Flemming, Edward. (1995). *Auditory Representations in Phonology.* Ph.D. Dissertation, UCLA.

Flemming, Edward. (2004). Contrast and perceptual distinctiveness. In Hayes et al. *Phonetically Based Phonology*. 232–276.

Flemming, Edward. (2005). Speech perception in phonology. In David B. Pisoni & Robert E. Remez (eds.) The handbook of speech perception, 156–182. Malden, MA & Oxford: Blackwell.

Foulkes, P. (1997). Historical laboratory phonology – Investigating /p/ > /f/ > /h/ changes. *Language and Speech*, 40 (3). 249-276.

Goldsmith, John. (1981). Subsegmentals in Spanish phonology. In William Cressey & Donna Jo Napoli (eds.), *Linguistic symposium on Romance languages 9*, 1–16: Washington, D.C.: Georgetown University Press.

González, Carolina. (2002). Phonetic variation in voiced obstruents in North-Central Peninsular Spanish. *Journal of the international phonetic association* 31. 17–31.

Hammond, Robert M. (1976) Phonemic restructuring of voiced obstruents in Miami-Cuban Spanish. In *1975 colloquium on Hispanic linguistics*, ed. by Frances M. Aid, Melvyn C. Resnick & Bohdan Saciuk, 42-51. Washington D.C.: Georgetown University Press.

Harris, James W. (1969). *Spanish Phonology*. Cambridge, MA: MIT Press.

Hayes, Bruce & Donca Steriade (2004). The phonetic basis of phonological markedness. Hayes, Bruce, Robert Kirchner & Donca Steriade (eds.), *Phonetically based phonology*, Cambridge University Press, Cambridge, pp. 1–33.

Hualde, José I. (2011). Sound change. In Marc van Oostendorp, Colin J. Ewen, Elizabeth Hume & Keren Rice (eds.), *The Blackwell Companion to Phonology*. London and New York: Blackwell Publishing.

Hualde, José I. (2011). Lenición de obstruyentes intervocálicas en español. V Congreso de Fonética Experimental, Cáceres.

Hualde, José I. (2005). *The sounds of Spanish*. Cambridge: Cambridge Univ. Press.

Jun, Jongho (2004). Place assimilation. In B. Hayes, R. Kirchner and D. Steriade (Eds.) *Phonetically Based Phonology,* 58-86.Cambridge University Press.

Kirchner, R. (2004). Consonant lenition. In B.Hayes, R. Kirchner, & D. Steriade (Eds.), *Phonetically based phonology.* Oxford: Oxford University Press. pp. 313-345.

Ladefoged, Peter. (1975). *A Course in Phonetics.* Orlando: Harcourt Brace.

Lathrop, Thomas A. (2003)*, The Evolution of Spanish*, Newark, Delaware: Juan de la Cuesta

Lavoie, L. (2001). *Consonant strength: phonological patterns and phonetic manifestations*. New York: Routledge.

Lewis, Anthony. (2001). *Weakening of Intervocalic /p, t, k/ in Two Spanish Dialects: Toward the Quantification of Lenition Processes*. Unpublished doctoral dissertation. University of Illinois at Urbana-Champaign.

Lipski, John M. (1994). Spanish stops, spirants, and glides: From consonantal to [vocalic]. In Michael Mazzola (ed.), *Issues and theory in Romance languages,* 67–86. Washington D.C: Georgetown University Press.

Lipski, John M. (1999). The many faces of Spanish /s/-weakening: (re)alignment and ambisyllabicity. Advances in Hispanic linguistics, ed Javier Gutierrez-Rexach and Fernando Martinez-Gil, 198-213. Somerville, MA: Cascadilla Press.

Lozano, María del Carmen. (1979). *Stop and spirant alternations: Fortition and spirantization processes in Spanish phonology*. Bloomington, IN: Indiana University Linguistics Club.

Martinez- Celdrán, E. (1991). Sobre la naturaleza fonética de los alófonos de /b, d, g/ en el español y sus distintas denominaciones. *Verba* 18, 235-253.

Martínez-Celdrán, E. (2004). Problems in the classification of approximants. *Journal of the International Phonetic Association,* 34. pp. 201-210.

Mascaró, Joan. (1984). Continuant spreading in Basque, Catalan, and Spanish. In Mark Aronoff & Richard T. Oehrle (eds.), *Language sound structure*, 287–298. Cambridge, MA: Cambridge University Press.

Mazzaro, Natalia. (2011*). Experimental Approaches to Sound Variation: a Sociophonetic Study of Labial and Velar Fricatives and Approximants in Argentine Spanish.* Unpublished doctoral dissertation. University of Toronto.

Mielke, Jeff. (2003). The Diachronic Influence of Perception: Experimental Evidence from Turkish. *Proceedings of the 29th annual meeting of the Berkeley Linguistics Society*.

Montes Giraldo, J. J. (1975). Breves notas de fonética española. *BICC*, XXX, 1975, 340-1.

Navarro Tomás, Tomás. (1967). *Manual de pronunciación española*, 6th ed. New York: Hafner.

New, B., Brysbaert, M., Veronis, J., & Pallier, C. (2007). The use of film subtitles to estimate word frequencies. *Applied Psycholinguistics*, 28, 661-677.

Ohala, John. (1974). Experimental historical phonology. In John M. Anderson & Charles Jones (eds.), *Historical linguistics*, vol. 2, 353–389. Amsterdam: North Holland.

Ohala, John. (1981). Articulatory constraints on the cognitive representation of speech. In: T. Myers, J. Laver, & J. Anderson (eds.), *The cognitive representation of speech*. Amsterdam: North Holland. 111 - 122.

Ohala, John. (1983). The origin of sound patterns in vocal tract constraints. In Peter F. MacNeilage (ed.), *The Production of Speech*, 189–216. New York: Springer.

Ohala, John. (1989). Sound change is drawn from a pool of synchronic variation. In L. E. Breivik and E. H. Jahr (eds.), *Language Change: Contributions to the Study of its Causes.* New York: Mouton de Gruyter.

Ohala, John. (1992). What's cognitive, what's not, in sound change. In Kellermann, G. and M.D. Morrissey (eds.), *Diachrony within Synchrony: Language History and Cognition.* Frankfurt am Main. Peter Lan Verlag, pp. 309-355.

Ohala, John. (1993). The phonetics of sound change. In Charles Jones (ed.), *Historical Linguistics:Problems and Perspectives.* London: Longman. 237-278.

Ohala, John & Carol J. Riordan. (1979). Passive vocal tract enlargement during voiced stops. In Jared L. Wolf & Dennis H. Klatt, (eds.), *Speech communication papers*, 89–92. New York: Acoustical Society sof America.

Ortega-Llebaria, M. (2004). Interplay between phonetic and inventory constraints in the degree of spirantization of voiced stops: Comparing intervocalic /b/ and intervocalic /g/ in Spanish and English. In T.Face (Ed.), *Laboratory Approaches to Spanish phonetics and phonology*. The Hague: Mouton de Gruyter.

Peirce, JW (2007). PsychoPy - Psychophysics software in Python. Journal of Neuroscience Methods, 162(1-2):8-13.

Penny, Ralph. (1991). *The History of the Spanish Language*, 2nd ed. Cambridge: Cambridge University Press.

Penny, Ralph. (2000) *Variation and Change in Spanish*. Cambridge: Cambridge University Press.

Pierrehumbert, J. (2003) Probabilistic Phonology: Discrimination and Robustness. In R. Bod, J. Hay and S. Jannedy (eds.) *Probability Theory in Linguistics*. The MIT Press, Cambridge MA, 177-228.

Pierrehumbert, J. (2002) Word-specific phonetics . *Laboratory Phonology VII*, Mouton de Gruyter, Berlin, 101-139.

R Core Team (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

Raudenbush, S. W., & Bryk, A. S. (2002). Hierarchical Linear Models: Applications and Data Analysis Methods, Second Edition. Newbury Park, CA: Sage.

Steriade, Donca (2001). Directional asymmetries in place assimilation: a perceptual account. Hume, Elizabeth & Keith Johnson (eds.), Perception in Phonology, Academic Press, pp. 219–250.

Torreira, F., & Ernestus, M. (2011). Realization of voiceless stops and vowels in conversational French and Spanish. Laboratory Phonology, 2(2), 331-353.

Trask. L., (2007). Trask's Historical Linguistics. (Revised by McColl Millar, R.) , Hodder Arnold Press.

Widdison, Kirk A. (1997). "Physical parameters behind the stop-spirant alternation in Spanish". *Southwest Journal of Linguistics* 16.73-84.

Wright, Richard. (2001). Perceptual Cues in Contrast Maintenance. In *The Role of Speech Perception in Phonology*, ed. by Elizabeth Hume and Keith Johnson, 251-277. San Diego: Academic Press.

Wright, R. (2004) A review of perceptual cues and cue robustness. Hayes, Bruce, Robert Kirchner & Donca Steriade (eds.), *Phonetically based phonology*, Cambridge University Press, Cambridge, pp. 34-57.

Appendix 1: Stop/Approximant Alternation Geographic Distribution

Region	Location	Defining Characteristics	Classification, i.e. conservative (few approximants) vs. innovating	Method of Measurement	Reference
North America and Mexico	Yucatan, Mexico	Stops between vowels	Very conservative	not provided	Face (from Lipski 1994)
	Mexico (Chihuahua)	Spirantize in post-consonantal environments	innovating	not provided	Amastae (1995)
	USA (Mexican-American, El Paso)	Spirantize in post-consonantal environments	innovating	not provided	Amastae (1995)
Central America	El Salvador	Stops after consonants	conservative	not provided	Face (from Lipski 1994)
	Honduras	Stops after semivowels	conservative	not provided	Face (from Lipski 1994);Amastae (1986)
	Panama	Elision of /d/ phrase-finally	innovating	not provided	Face (from Lipski 1994)
South America	Bolivia	Elision between vowels	innovating	not provided	Face (from Lipski 1994)
	Colombia	Stops after glides	conservative	not provided	Montes Giraldo (1975)
	Colombia (Bogotá)	Conserves stops in post-consonantal environments; Favors stops after /r/, /l/, and/s/ but only favors stops after glides in the case of /b/	conservative	not provided	Amastae (1975)
Peninsular	Castillian	Approximants after nasals	innovating	not provided	Face (from Aguilar 1993)
	North-central peninsular	-voiced stops in initial position or after nasals	innovating		González (2002)

		-voiced approx after vowels, approximants, fricatives and liquids -/ld/ pronounced as stop; -become voiceless fricatives in coda position			
Non-specific	Most conservative: 1. Nicaragua 2. Spain 3. Colombia Most innovating 1. Venezuela			Intensity differences; only one speaker per location	Eddington (2011)

Appendix 2: Lemma frequency for Production Task

The following frequencies were taken from the Corpus del español (Davies). All frequencies are raw frequencies based on 20th century counts. Overall frequency refers to the frequency of each word in all contexts—written and oral—while oral frequency refers only to oral contexts. The number of words for each are 20,540,030 and 5,113,249, for overall and oral, respectively.

Word	Overall Frequency	Oral Frequency
disco	722	130
dato	354	155
dedo	576	65
doce	1494	632
duque	413	33
dictó	97	14
danés	150	3
decir	12599	7435
doné	0	0
duchar	3	3
adicto	18	6
nadar	136	39
nadé	0	0
nadó	7	0
caduco	19	6
adictivo	4	1
cada	14956	2940
nade	1	0
nado	24	6
maduró	10	0
visa	118	13
vaso	535	36
beso	407	35
vote	30	22
busque	70	24
billar	108	11
basé	2	1
besé	44	0
voté	10	4
bufón	24	0

bebí	32	2
cabal	125	52
lave	21	2
lavó	37	4
tabú	42	14
débil	529	65
lava	174	13
lave	9	3
lavo	14	7
abusar	42	8
guiso	29	1
gato	702	66
gueto	24	7
gota	196	13
gusta	2971	1999
guiar	70	22
gastó	47	12
Guejar	0	0
gozó	53	2
gustó	514	339
aguije	0	0
pagar	1207	355
pagué	29	5
pagó	161	23
aguda	197	19
águila	127	5
paga	406	144
pague	80	21
pago	766	96
agujero	319	16
andar	763	152
tumbar	18	6
vengar	23	2
saldar	26	2
silbar	22	3
colgar	79	10
desdén	63	2
resbala	42	0

rasgar	8	0
banda	769	76
tumba	408	23
manga	143	14
falda	265	32
silba	21	0
cuelga	87	4
desde	25486	4803
resbalar	52	2
rasga	12	0
la burra	34	8
el burro	152	44
la dama	506	60
el daño	601	183
la gata	67	13
el gato	702	66

Appendix 3: Word Frequency for Experiment One

The following frequencies were taken from the Corpus del español (Davies). All frequencies are raw frequencies based on 20th century counts. Overall frequency refers to the frequency of each word in oral contexts. The total number of words is 5,113,249.

Word	**Oral Frequency**
aguije	0
águila	5
alba	5
anda	290
andar	152
ande	14
andé	0
base	716
basé	1
bebé	25
bebí	2
besé	0
beso	35
billar	11
cabal	52
cada	2940
colgar	10
colgué	3
cuelga	4
cuelgue	8
daba	441
danés	3
dato	155
debe	2265
deber	94
débil	65
dedo	65
desde	4803
desdén	2
desván	3
dictó	14
disco	130

falda	32
gastó	12
gato	66
Guejar	0
gueto	7
guía	90
guiar	22
manga	14
mangar	2
mangue	0
mangué	1
moldé	0
nadar	39
paga	144
pagar	355
pague	21
pagué	5
pedí	74
rasga	0
rasgar	0
rasgue	2
rasgué	0
saldar	2
silbar	3
silbe	0
silbé	0
tilde	0
trasver	0
tumba	23
tumbar	6
tumbe	0
tumbé	0
visa	13

Appendix 4: Word frequency for Repetition Task

The following frequencies were taken from the Corpus del español (Davies). All frequencies are raw frequencies based on 20th century counts. Overall frequency refers to the frequency of each word in all contexts—written and oral—while oral frequency refers only to oral contexts. The number of words for each are 20,540,030 and 5,113,249, for overall and oral, respectively

	/d/	all/oral	/b/	all/oral	/g/	all/oral
i	*dicho*	4159/1860	*bicho*	114/23	*guiso*	29/1
e	*dedo*	576/65	*vete*	149/45	*gueto*	24/7
a	*dato*	354/155	*bata*	147/13	*gato*	702/66
o	*dote*	25/6	*vote*	30/22	*gota*	196/13
u	*ducha*	116/25	*bula*	43/3	*gula*	16/0

Appendix 5: Individual Word Accuracy Rates

Word	Given Manner	Correct	Incorrect
bata	approximant	57.89%	42.11%
	stop	41.38%	58.62%
bicho	approximant	86.44%	13.56%
	stop	96.55%	3.45%
bula	approximant	27.12%	72.88%
	stop	88.33%	11.67%
dato	approximant	18.64%	81.36%
	stop	85.96%	14.04%
dedo	approximant	96.49%	3.51%
	stop	98.28%	1.72%
dicho	approximant	68.42%	31.58%
	stop	93.33%	6.67%
dote	approximant	13.33%	86.67%
	stop	93.33%	6.67%
ducha	approximant	78.95%	21.05%
	stop	94.92%	5.08%
gato	approximant	75.00%	25.00%
	stop	95.00%	5.00%
gota	approximant	93.22%	6.78%
	stop	100.00%	0.00%
gueto	approximant	73.33%	26.67%
	stop	93.22%	6.78%
guiso	approximant	48.33%	51.67%
	stop	100.00%	0.00%
gula	approximant	61.02%	38.98%
	stop	51.72%	48.28%
vete	approximant	100.00%	0.00%
	stop	96.61%	3.39%
vote	approximant	65.00%	35.00%
	stop	79.31%	20.69%

Appendix 6: Experiment 4 accuracy rates

Word	Correct %
pibe	48.7%
quiso	49.2%
peso	66.9%
peque	76.7%
tique	77.8%
vaga	81.2%
taca	83.1%
psique	83.1%
pide	83.9%
revela	84.7%
tenso	87.2%
atas	87.2%
vaca	88.1%
cava	88.1%
cana	89.0%
rete	88.9%
pegue	91.5%
guiso	91.5%
pino	92.4%
vago	92.4%
daca	92.4%
pipe	93.2%
denso	93.2%
repela	94.9%
pago	94.9%
sigue	94.9%
beso	95.8%
vino	95.8%
dique	96.6%
hadas	96.6%
rede	96.6%
gana	97.5%
pite	98.3%
capa	98.3%
Grand Total	**87.4%**