

# Perceptual Assimilation of Occluded Voiced Stops by Spanish Listeners

Marcos Rohena-Madrado  
Middlebury College

## 1. Introduction

Much research on cross-language speech perception has revealed that adults have difficulty distinguishing between phonetic pairs that are not contrastive in their native language (Goto 1971, Gillette 1980, Best & Strange 1992). However, many of these investigations have also found that not all non-native contrasts are perceived with equal difficulty (Best et al. 1988, Werker & Tees 1984b, Polka 1991). The performance of these cross-language listeners depends on several factors, among them the age of the listener (Werker 1995, Werker & Tees 1984a, Kuhl et al. 2006, Flege 1995), the type of testing conditions of the tasks (Werker & Tees 1984b, Werker & Logan 1985, Pisoni 1973, Davidson & Shaw 2012), the phonetic differences and similarities between the non-native phones and phonemes in the listener's native inventory (Best 1995, Flege 1995, Best et al. 1988, 2001), the phonotactic patterns in which the phones are presented (Dupoux et al. 1999, Davidson 2007, Hallé et al. 1998, 2003) as well as the phonological contexts in which the contrasts appear (Logan et al. 1991, Strange et al. 2001, Flege 1989). Several different models of cross-language speech perception have been developed in order to account for these differences in listener performance, for example, the Perceptual Assimilation Model (Best 1995, Best et al. 1988, 2001; Best & Tyler 2007) the Speech Learning Model (Flege 1989, 1995, 1999), and the Native Language Magnet Model (Kuhl 1993, Kuhl & Iverson 1995, Kuhl et al. 2008). These models all propose that the listeners do not perceive non-native contrasts in a native-like fashion because of the influence of their native sound systems. Analyzing the variable performance of listeners in cross-language perception tasks can shed light on the phonological organization of the listeners' native language. The listeners' native phonetic and phonological categories affect how they perceive novel sounds; however, the nature and extent of the effect of these categories is what differentiates the models of cross-language speech perception. For an overview of these and other models, see Strange (1995).

The present study will mainly be concerned with two of these models, Best's Perceptual Assimilation Model and Flege's Speech Learning Model. As will be discussed below, these two models make different predictions about the role that phonetic vs. phonological categories play in the perception and assimilation of non-native speech sounds. Using the perception of occluded voiced stops by Spanish listeners as a diagnostic case, this study will test each model's predictions with respect to the inter-relation between position-specificity and phonological categories in cross-language speech perception.

The remainder of this introductory section will describe and contrast the two models, present the variation in implementation of the Spanish stop voicing contrast, as well as discuss the results of previous perception studies. Section 2 will present the current experimental study consisting of a perceptual assimilation task and a goodness rating task of stop contrasts in native and non-native positions for Spanish listeners, which will serve to test the differing predictions of each model. Section 3 will present the results and discussion of these two experimental tasks. Finally, Section 4 will

---

\* I would like to thank the participants who volunteered their time and efforts. I am also grateful to the audience of the Hispanic Linguistic Symposium and to three anonymous reviewers for their comments, as well as to Lisa Davidson for her guidance and support during the various stages of this project. All errors remain my own.

provide some concluding remarks suggesting that when listeners assimilate native contrasts in non-native positions, the inter-relationship of phonetic to phonological categories, rather than to position-specific allophones, might be more influential in cross-language speech perception, as predicted by Best's Perceptual Assimilation Model.

## *1.1. Models of cross-language speech perception*

### *1.1.1. Perceptual Assimilation Model*

Best's Perceptual Assimilation Model (PAM) (Best 1995, Best et al. 1988, 2001, Best & Tyler 2007) posits that different types of assimilation patterns result from the articulatory similarities of the non-native phonemes to the native categories of the listeners. Best (1995) and Best et al. (1988, 2001) discuss several different types of assimilation that can occur depending on the comparison of the non-native phones and the listeners' native phonemic inventory. They propose various assimilation types, depending on the divergence from the native categories. The non-native sounds can be perceived as categorizable with respect to their native categories, in which case the accuracy in discrimination will vary depending on the goodness of fit to the native category that the non-native sounds are assimilated to. The assimilation types relevant for this study are mainly Two Category Assimilation (TC), Same Category Assimilation (SC) and Category Goodness Assimilation (CG).<sup>1</sup>

In the cases of TC assimilation, two non-native contrastive phonemes are mapped onto two natively contrastive phonemes. In this assimilation type PAM predicts high accuracy in discrimination, because, although the particular contrast does not exist in the native language, it is mapped onto a native pair, whose members are phonemically distinct, thus allowing an easier discrimination between the two. SC assimilation, on the other hand, occurs when the two non-native phonemes are mapped equally well onto the same single category in the native inventory. In these cases PAM predicts poor discrimination, since the contrasting non-native phones are perceived, essentially, as equally good exemplars of the same native category. In CG Assimilation the two non-native phonemes are mapped onto the same native category, but contrary to SC assimilation, one of the non-native phonemes is a better fit to the native category, i.e. it is a more prototypical exemplar of the native category. In these cases PAM predicts intermediate accuracy in discrimination rates, because, although the two non-native phones are mapped onto the same native category, the discrepancy in goodness of fit between the two phones is an aid in distinguishing them.

Given these assimilation types and the accuracy rates predicted for each, PAM posits the following accuracy scale: TC > CG > SC. According to this scale, listeners should be most accurate at discriminating Two Category Assimilation types, less accurate in Category Goodness Assimilations and least accurate in Same Category Assimilation types.

### *1.1.2. Speech Learning Model*

Another theory that attempts to explain and predict the patterns of cross-language speech perception is Flege's Speech Learning Model (SLM) (Flege 1989, 1995, 2003). Although discriminating between non-native categories becomes more difficult as the age of exposure increases, SLM proposes that the basic mechanisms for acquiring new categorical representations for speech sounds are essentially the same throughout the life span. According to SLM, perception of the non-native speech segments will depend on how they compare acoustically to the other native segments in the same position. The aspect of position is crucial, since SLM predicts that the listener should perceptually relate "positional allophones in the L2 to the closest positionally defined allophone (or "sound") in the L1" (Flege 1995:238-9). In other words, the non-native sounds are related to position sensitive allophones in the native language, rather than to a more abstract phonemic level.

Flege finds evidence for this sensitivity to position in several studies that have found discrepancies in the perception (and production) of non-native phonemes in initial vs. final position. For example,

---

<sup>1</sup> PAM also posits that non-native speech sounds can be perceived as uncategorizable and non-assimilable; however, these assimilation patterns will not be discussed in this paper. See Best et al. (1988, 2001) for predictions and results on these assimilation types.

several studies have found that native Japanese listeners have difficulty distinguishing between the American English phonemes /ɪ/ and /I/; however, the Japanese listeners can discriminate this non-native contrast better in word-final position than in word initial position (Logan et al. 1991, Strange 1992). These discrepancies in accuracy might be due to the different acoustic cues that /ɪ/ and /I/ present in the speech signal in final position that are comparatively more robust in final position than in initial position, particularly since there are acoustic cues present in the preceding vowel for final liquids but not for initial liquids (Sheldon & Strange 1982).

### 1.1.3. PAM vs. SLM

The basic hypothesis that non-native sounds will be assimilated by listeners into an already existing category in the native system is common to both SLM and PAM; however, these two models differ in several ways. One of these differences is that SLM was originally developed to predict and explain perception patterns of listeners involved in second language learning and acquisition, whereas PAM's main focus is on non-native speech perception by "naïve" listeners that have had no exposure to the language of the contrasts that they are trying to perceive. This might raise the question of whether these two models are entirely comparable, since they are not designed to explain the behavior of the same types of listeners. Although active language learners and completely "naïve" listeners are two different types of listeners, one could interpret "naïveness" as the beginning stage of the speech learning process, at least for the adult learners. If the "naïve" listeners were to begin to learn the stimulus language, they would then be solely within the scope of SLM and, presumably, they might begin to form new categories, depending on factors such as age of arrival, amount of L1/L2 input, etc. (Flege 1995). Furthermore, since SLM posits that the mechanisms of categorical representation remain the same throughout the lifespan, the underlying mechanisms of the L2 listener and the naïve listener should be comparable. Best & Tyler (2007) consider the similarities and differences between PAM and SLM and propose an extension of the tenets and predictions of PAM to the realm of second language learning (PAM-L2).

Another difference between these models is what each one posits as the unit of perception. While the perceptual objects in SLM are the acoustic cues in the non-native speech signal, PAM is based on "the ecological direct-realist premise that the focus of speech perception is on the information about the distal articulatory events that produced the speech signal" (Best & Tyler 2007:22). For this reason, PAM assumes the principles of Articulatory Phonology (Browman & Goldstein 1989, 1990, 1992), which posits that the timing and coordination of articulatory gestures necessary for speech function as the primitives for phonological organization.

The main difference between SLM and PAM, for the purposes of the present investigation, is each model's stance on the importance of the phonological position or context in which the non-native contrast is presented and its relationship to the native phonetic and/or phonological categories. In SLM the relationship of the non-native contrast to the native one is position-specific; whereas, in PAM the assimilation of the non-native phone takes into consideration phonetic-articulatory patterning "at both the abstract contrastive level and, importantly, at the level of non-contrastive gradient phonetic detail" in order to predict the assimilation patterns of the non-native sounds (Best & Tyler 2007:22). In SLM the acoustic similarity in the perceptual space will be based on a native phonetic category in the same position, while in PAM it is the similarities in the higher-order invariants, which should be relatively constant across positions, that are the targets for assimilation.<sup>2</sup> It is by no means the case that PAM disregards the importance of phonetic detail during assimilation, in fact, the phonetic variation between the gestures of the native and the non-native sound determine the goodness of fit of the non-native sound to the native category. Therefore, while SLM addresses equivalence only at the level of the phonetic category, PAM addresses equivalence at both the phonetic and phonological levels (Best & Tyler 2007:27).

Given the spirantization patterns of voiced stops found in Spanish, the perception of Portuguese stop voicing contrasts by Spanish listeners presents a useful test case to determine whether PAM or

<sup>2</sup> However, see Flege (1989), Hallé et al. (1998, 2003), and Hallé & Best (2007) for results that are problematic for both models' interpretation of the role of position.

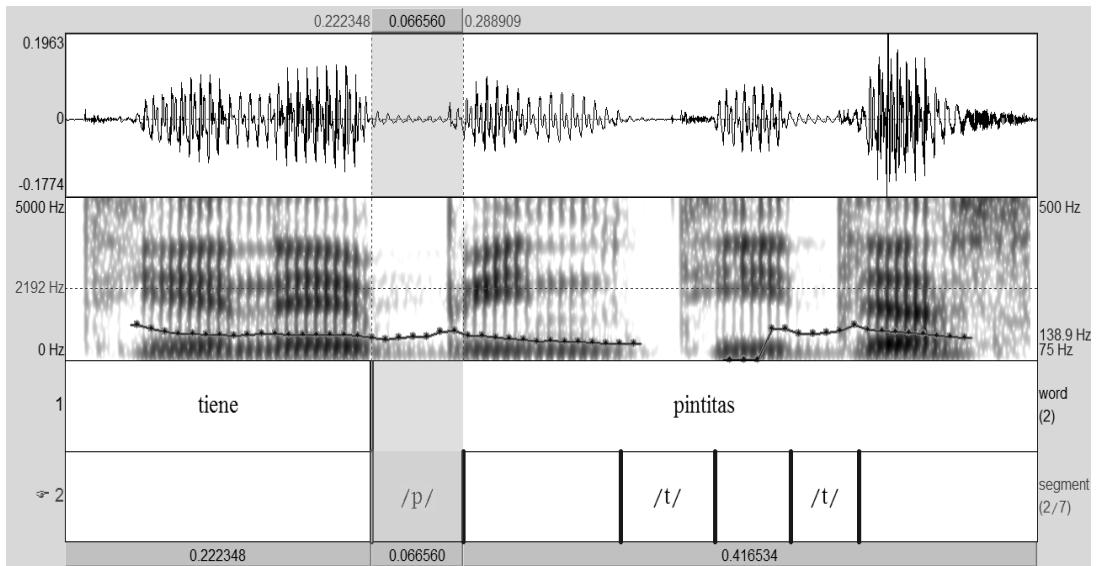
SLM make the right predictions for the role of position specificity vs. phonological categorization in the assimilation of non-native sounds.

### 1.2. *The implementation of the stop voicing contrast in Spanish*

Spanish presents a phonemic distinction between voiced and voiceless stops, where /p, t, k/ contrast with /b, d, g/. However, this voicing distinction is realized differently depending on the phonological context in which the stops occur. In phrase initial position (as well as after nasals and after /l/ for homorganic stops) the phonemic contrast is realized by a difference in voicing, where the phonologically voiceless stop is realized with short lag VOT and the phonologically voiced stop is realized with negative VOT, e.g. /tV/ vs. /dV/ is realized as [tV] vs. [dV] (Lisker & Abramson 1964, Martínez Celdrán & Fernández Planas 2007). On the other hand, the voicing contrast in intervocalic position is realized not only by a difference in voicing, but also by a difference in the manner of articulation, whereby the phonologically voiced stops spirantize and the phonologically voiceless stops retain their occlusion. For example, the contrast of /VtV/ vs. /VdV/ is realized as [VtV] vs. [VðV]. Traditionally, the allophones of the voiced stops in intervocalic position have been described as voiced fricatives [β, ð, γ] (Lapesa 1981, Quilis 1999); however, these phones are usually realized not as fricatives but as approximants [β̞, ð̞, γ̞] (Martínez-Celdrán et al. 2003, Santagada & Gurlekian 1989). Regardless of the constriction degree of the intervocalic segment, the realization of /VdV/ as \*[VðV] is not a licit realization for native Spanish speakers. For this reason, the canonical implementation of the voicing distinction for stops in Spanish is present only in phrase initial (or post-nasal) position, voiced stops will spirantize to approximants elsewhere: /b, d, g/ → [β̞, ð̞, γ̞].

The phenomenon of voiced stop spirantization has been described as a classic example of complementary distribution. However, several studies have shown that the spirantization of voiced stops in Spanish is in fact a gradient process, ranging from occlusion to deletion, and its realization is affected by factors such as speech rate (Soler & Romero 1999), lexical stress (Cole et al. 1999, Ortega-Llebaria 2004) duration of the stop gesture (Parrell 2010), place of articulation (Lavoie 2001, Colantoni & Marinescu 2010) nature of the flanking segments (Hualde et al. 2010, Cole et al. 1999, Kingston 2008), as well as dialect of Spanish (Amastae 1995, Lewis 2001). Similar spirantization patterns hold in cases where the voiced stop occurs across phrase-internal word boundaries (Cole et al. 1999).

There is also variation in the production of phonologically voiceless stops in Spanish as well. Across different varieties of Spanish, voiceless stops are often realized with partial or full voicing during closure (Machuca Ayuso 1997, Lewis 2001, Martínez Celdrán 2009, Parrell 2010, Torreira & Ernestus 2011). Machuca Ayuso (1997) found in her corpus of conversation speech more than 40% of phonologically voiceless stops /p, t, k/ were realized as [b, d, g]. Torreira & Ernestus (2011) found that in the Nijmegen Corpus of Spontaneous Spanish 32.7% of voiceless stops presented complete voicing during stop closure. This phenomenon of voicing of phonologically voiceless stops has been reported for many varieties of Spanish also; however, it also presents dialectal variation. Colantoni & Marinescu (2010) find relatively few cases of voicing of /p, t, k/ in their analyzed corpus of speech from the Argentinean provinces of San Juan and Corrientes. Speakers of Buenos Aires Spanish do exhibit cases of phonologically voiceless stops that are partially or fully voiced, as can be seen in Figure 1. Notice that the first /t/ in /pintitas/ is mostly voiceless, while the /p/ and the second /t/ are produced with complete voicing during closure, as evidenced by the presence of a voicing bar, periodicity in the waveform, and the continuous pitch tracker. The extent to which the frequency of this phenomenon in Buenos Aires Spanish differs to that found for the Spanish of Corrientes and San Juan by Colantoni & Marinescu (2010) is yet to be determined.



**Figure 1** – The phrase *tiene pintitas*, “has spots” spoken by a 27-year-old male from Buenos Aires during a sentence reading task.

In addition to voicing, many studies have also found cases of spirantization of voiceless stops as well, where they are realized as voiceless or voiced approximants (Machuca Ayuso 1997, Lewis 2001, Parrell 2010, Torreira & Ernestus 2011). This is potentially problematic for the description of the voicing distinction of stops in Spanish, since there is an overlap in the realization of stops such that both voiced and voiceless ones can be realized as [β, ð, γ] or as [b, d, g].<sup>3</sup>

Considering the evidence above, it cannot be said that voiced stop spirantization is categorical when not in phrase initial or postnasal context, or that phonologically voiceless stops in Spanish are always phonetically voiceless. However, what can be gleaned from most of the studies is that phrase initial position and intervocalic positions seem to be the two poles in the continuum. What is common to all varieties of Spanish is that voiced stops in phrase initial position (#CV) favor occlusion and voiced stops in intervocalic (VCV) position favor spirantization. Similarly, for voiceless stops, #CV position should favor voiceless realizations, whereas in VCV position is where lenition to a voiced stop (or even approximant) would be more likely to occur.

### 1.3. Testing predictions of PAM vs. SLM: previous studies

Given the wide range of variation in voiced and voiceless stop production in Spanish, the question arises of how Spanish listeners will perceive intervocalic occluded voiced stops. The voicing discrepancy in the obstruent inventory in Spanish, as well as the positionally determined process of spirantization of voiced stops, presents a test case for the predictions made by SLM and PAM in terms of the position specificity of non-native speech perception. The Perceptual Assimilation Model and the Speech Learning Model each make different predictions for the accuracy rates of the Spanish listeners making these discriminations of canonical stop voicing contrasts in intervocalic position, such as the ones produced in Brazilian Portuguese (BP) as well as many other languages (Barbosa & Albano 2004).

For the stop voicing contrast in initial position SLM would predict high accuracy in discrimination, since the acoustic cues of the voicing contrast in BP are similar to the ones in Spanish for that position. For the stop voicing contrasts in intervocalic position, the voiceless stops in BP will be related to the voiceless stops in intervocalic position in Spanish; however, the voiced stops in intervocalic position in Brazilian Portuguese [VdV] are incongruous with the spirantized positional

<sup>3</sup> See Martínez-Celdrán (2008) and Parrell (2010) for alternative analyses of the stop voicing contrast in Spanish.

allophone of /VdV/ that is most common in Spanish: [VðV].<sup>4</sup> The BP [VdV], although voiced, because of its occlusion and burst, would be acoustically more similar to the positional allophone for Spanish /VtV/. Furthermore, as discussed above, phonologically voiceless stops in intervocalic position often present voicing during closure, which presents another reason why BP [VdV] would be associated with the positional allophone [VdV] of the Spanish sequence /VtV/. This would increase the perceived acoustic similarity of BP [VtV]-[VdV] for Spanish listeners, thereby making the discrimination harder. For these reasons, the Speech Learning Model would predict that Spanish listeners would have lower accuracy rates for the BP stop voicing contrast in intervocalic position.

The predictions that Best's Perceptual Assimilation Model would make for the phrase initial stops are similar to the ones made by the Speech Learning Model: initial stops in BP, [tV]-[dV], should assimilate to the analogous categories in Spanish as a Two Category Assimilation, /t/-/d/. For this reason, listeners should exhibit high levels of accuracy. For the stop voicing contrast in intervocalic position, however, PAM's predictions differ from those of SLM. PAM proposes an interaction of the phonetic and phonological levels, such that, while listeners might perceive phonetic discrepancies between the phone and the position in which it is presented, they are also attending to higher-order invariants that distinguish the articulation of the voiced and voiceless stops, which would be constant across positions. Therefore, the BP [VdV] should be assimilated to the voiced stop Spanish category /d/, regardless of the fact that occluded allophones [VdV] for native /VdV/ do not occur in that position in Spanish. In these cases, PAM would predict levels of accuracy in discrimination similar to those for the stop voicing contrast in initial position, which is analogous in Portuguese and Spanish. This type of assimilation would also be a Two Category Assimilation, where BP [VtV] maps onto Spanish /t/ and BP [VdV] maps onto Spanish /d/; therefore, listeners should exhibit high accuracy rates. The obstruent in [VdV] perhaps might be a worse fit for the Spanish /d/ category in that position; however, since the stimulus that the listeners are comparing it to belongs to another category, this should not affect their accuracy in discrimination between /t/ and /d/.

Rohena-Madrado (2011) tested these predictions of SLM and PAM with respect to the effect of position specificity on the accuracy of discrimination of non-native obstruent voicing contrasts by Buenos Aires Spanish listeners in cross-language speech perception experiments. He used both AX and ABX discrimination tasks, with BP as the non-native stimulus language. In AX tasks participants listen to two consecutive stimuli (A and X) and have to determine whether the second stimulus is the same as the first or whether it is different. In the ABX tasks, the participants listen to three consecutive stimuli (A, B, and X), two of which are explicitly from different categories (A vs. B). Then the participants need to determine the category to which the X stimulus belongs, either to A or to B.

Rohena-Madrado (2011) found that the listeners' discrimination of stop voicing contrasts was not affected by whether it occurred in a native or a non-native context. For both AX and ABX discrimination tasks, the listeners performed at ceiling levels both in the native initial position (e.g. [tV]-[dV]) as well as in the non-native intervocalic position (e.g. [VtV]-[VdV]). These results seem to support the Perceptual Assimilation Model, which predicts that the higher-order invariants, constant across positions, would influence the assimilation patterns of the stops in the non-native sequences to the native voiced stop categories.

One potential issue with Rohena-Madrado's (2011) AX and ABX discrimination task results is that, although they prove that Spanish listeners can very accurately discriminate native voicing contrasts in non-native positions, they do not provide direct evidence on the type of assimilation of the non-native stimuli to the native categories. It is unlikely that the [VtV]-[VdV] stimuli were associated equally well to the same category (SC assimilation) given their ceiling accuracy rates. It is not clear, however, whether they were assimilated to the two corresponding native categories (/t/-/d/: TC assimilation) or whether it was a case of CG assimilation where the considerably poor fit of the [VdV] stimuli, either to /VtV/ (as SLM would predict) or to /VdV/ (as PAM would predict), aided them in with discrimination, resulting in the very high discriminability. Furthermore, in the AX and ABX discrimination tasks the listeners always had at least one other stimulus to compare non-canonical

---

<sup>4</sup> Since SLM and PAM would not make differing predictions for the stop pairs in different places of articulation, in the discussion of position, "t", "d", and "ð" will be used to symbolize voiceless occlusion, voiced occlusion, and spirantization, respectively, for all stop voicing pairs.

tokens to; therefore, there was always a basis for comparison which might have also aided in the discrimination. The question remains whether the non-native [VdV] stimuli would be associated with /VtV/ or /VdV/ when presented in isolation.

## 2. The current study

### 2.1. Research questions

The main goals of this study are to answer the following questions:

1. Do Spanish listeners assimilate non-natively occluded intervocalic voiced stops, V[d]V, to their native voiced category, /d/, or to their native voiceless category, /t/?
2. How good of an exemplar of the chosen category do they consider this occluded intervocalic voiced stop?

The answers to these questions will help determine which model, SLM or PAM, makes the correct predictions in terms of whether non-native sequences are associated with position-specific allophones or whether the higher order invariants in the non-native sequences permit an assimilation of the stop to a more abstract phonological category. Furthermore, examining the goodness of fit of the non-native tokens to the native category will help determine whether these are cases of TC assimilation or CG assimilation. In order to answer these research questions, a perceptual assimilation task and a goodness rating task were conducted and are detailed below.

### 2.2. Stimuli

A subset of the Brazilian Portuguese stimuli used by Rohena-Madrado (2011) in the ABX experiment was used for the present transcription and goodness rating tasks. The stimuli contained one obstruent of the following: 6 stops /p, b, t, d, k, g/ and 6 fricatives /f, v, s, z, ʃ, ʒ/ (in this paper only the stop data will be discussed). The obstruents were presented in the following two templates: /Cala/ and /Cili/ for initial position and /laCa/ or /liCi/ for intervocalic position. These two types of templates were chosen so that the intervocalic and initial tokens would both appear in similar contexts in terms of length and syllabic structure. BP was used as the stimulus language because it is relatively similar to Spanish, i.e. most of the phonemes of Spanish are contained in the phonemic inventory of Brazilian Portuguese (Mattoso Câmara 1972). Most important for this study, BP voiceless stops are described as having short-lag VOT and voiced stops are described as prevoiced, i.e. with negative VOT, in both initial and intervocalic position (Barbosa & Albano 2004). In other words, phonologically voiced stops in intervocalic position are always realized as occlusives, they do not spirantize as in Spanish. See the Appendix for the acoustic data on the BP stop stimuli.

The stimuli were recorded by a female native speaker of Brazilian Portuguese, 27 years old, from São Paulo, Brazil. She had been living and working in Buenos Aires for several years and was bilingual in Spanish.<sup>5</sup> The stimuli were recorded in a quiet room using a Sony MZ-RH910 Hi-MD recorder and Shure WH20XLR head-mounted microphone. The stimuli were then checked by 3 other native BP speakers in order to make sure that each was an appropriate exemplar of the corresponding BP category.<sup>6</sup> There was a total of 48 stimuli (24 stops and 24 fricatives) and were presented in two different blocks, one intervocalic and one initial position. The stimuli in the blocks were pseudo-randomized and the blocks were counterbalanced. The results from the fricative stimuli will not be discussed in this paper.

<sup>5</sup> The only stimuli which merited special attention were those containing /ti/ and /di/ sequences. In most varieties of Brazilian Portuguese dental plosives are realized as affricates before high front vowels, e.g. /ti/ > [tʃi], /di/ > [dʒi] (Barbosa & Albano 2004). For the stimuli containing /ti/ and /di/ sequences, the speaker was instructed to produce the consonants as plosives ([ti] and [di]), both in initial and intervocalic position. Although this is not common in Spanish or in BP, the speaker was able to produce [ti] and [di] stimuli that were comparable to the other stops (see Appendix).

<sup>6</sup> The stimuli containing /ti/ and /di/ sequences were checked by the investigator to ensure that the obstruents were produced as [ti] and [di] and not affricated, since nonaffricated [ti] and [di] are not common in most varieties of Brazilian Portuguese.

### 2.3. Participants

38 native speakers of Spanish participated in this study. They were all born and raised in the Buenos Aires metropolitan area of Argentina. 2 participants were excluded on account of having previous knowledge of Portuguese and 2 were excluded for having near-native knowledge of another language (in these cases, English). Of the remaining 34 subjects, none had previous exposure to Portuguese; therefore, they were all “naïve” listeners. There were 18 males and 16 females; their ages ranged from 20-67, with a mean age of 43.

### 2.4. Procedure

The instructions stated that the participants would be listening to some words in a foreign language and that they should transcribe each word as they would write them in Spanish. All of the letters corresponding to the vowels and the /l/ in the stimulus templates were already provided, so that they would only have to focus on the transcriptions of the obstruent in question.

After every word there was a space for comments where the participants could write whatever information they wanted about the particular sounds. After transcribing the obstruent and writing any comments, they were asked to look back at the letter (or letters) that they transcribed and think about how they are most commonly pronounced in Spanish and then determine whether the sound that they had heard was a good example of that most usual pronunciation. They were provided with a 5-point Likert scale and were asked to give a rating depending on how good of an example of that most usual pronunciation it was (1 = very bad example, 5 = very good example).

Since rating non-native exemplars with respect to a native phoneme could prove confusing to people without linguistics background, two example stimuli were provided at the beginning of each block. The example phoneme was /r/ in either initial or intervocalic position depending on the block. The example nonce words were either <rele> /rele/ or <lerru> /leru/ for each position. The first example was produced as [rele] or [leru], with an alveolar trill, and it was aimed that the listener would rate this as a 5. The second example was produced as [ɾele] or [leru] with a uvular trill, to which they would give a much lower rating. These examples were both produced and recorded by the investigator, a native speaker of Puerto Rican Spanish. This dorsal pronunciation, although not a phoneme of Spanish, is a variant of /r/ that speakers of Spanish are aware of, since it is considered to be an idiolectal variant of /r/ produced by people who, for physiological reasons, cannot produce an alveolar trill. It is also the variant associated French-accented speakers of Spanish. Although this sound is not “native” to the Spanish phonemic inventory,<sup>7</sup> speakers are aware that it is not an uncommon pronunciation in idiosyncratic or nonnative speech (and several participants noted this in their comments). Furthermore, <r> or <rr> would be the only ways of representing [ɾ] orthographically in Spanish as a category of /r/, thereby familiarizing them with the concept of better and worse examples of the same category. These example stimuli with canonical [r] and non-canonical [ɾ] exemplars of /r/ served as practice before transcribing and rating the experimental stimuli. After hearing, transcribing, and rating the example stimuli, the participants were allowed to ask any questions and then they continued with the rest of the 24 experimental stimuli for each block.

The stimuli were played as an experiment file in Praat (Boersma & Weenik 2010) on a Dell XPS M1330 laptop computer over Sennheiser HD202 circumaural headphones in a quiet room. Once the stimulus was played, the participants had to transcribe the obstruent, give a rating, and write comments if they so chose. The interstimulus interval was not fixed in order to allow the participants time for comments. Once the participant signaled to the investigator that they were finished with that trial, the investigator played the subsequent one. Participants could take a break between the blocks if they wanted.

---

<sup>7</sup> Although a dorsal rhotic is common in many Western European languages, e.g. French, German, Southern Swedish, Lisbon Portuguese, etc., in the Spanish-speaking world this is an established dialectal variant only in Puerto Rico, where [ɾ] is a native sociolinguistic variant of /r/ for some speakers (Hualde 2005:187, Lipski 1994:333-4). Given the relatively low familiarity of Argentineans with Puerto Rican Spanish, however, it is highly unlikely that the Buenos Aires participants in this study ever encountered [ɾ] as a sociolinguistic variant.

## 2.5. Predictions

In initial position, the BP stops /p, t, k, b, d, g/ should be identified as their corresponding stops in Spanish, since they are realized as occlusives; however, in intervocalic position there could potentially be more chance for confusion, since phonologically voiced stops are not occluded natively, but rather spirantized. If we assume that listeners assimilate the occluded intervocalic stops to the corresponding native stop category, regardless of the lack of spirantization, then they should identify them as such. However, as was discussed in section 1.2, in connected speech it is not uncommon for phonologically voiceless stops to be realized as voiced occluded stops. Given this information, it could be the case that subjects listen to cases of occluded voiced stops in BP and associate them to phonologically voiceless stops, since they share many of the acoustic details. This confusion might not have affected the accuracy in Rohena-Madrazo's (2011) discrimination experiments, because listeners in the AX and the ABX tasks always had at least one other stimulus to compare non-canonical tokens to; however, in the transcription task subjects are presented with isolated stimuli and this might impair their identification of potentially non-canonical tokens. If the listeners assimilate BP [VdV] to Spanish /VtV/ it would confirm SLM's predictions that assimilation of non-native phones depends on the association of the acoustic cues to the closest positional allophone. However, an assimilation of BP [VdV] to Spanish /VdV/ would confirm PAM's predictions that the presence of the higher-order invariants in the phonetically deviant non-native phone will influence its assimilation to a more abstract phonological category.

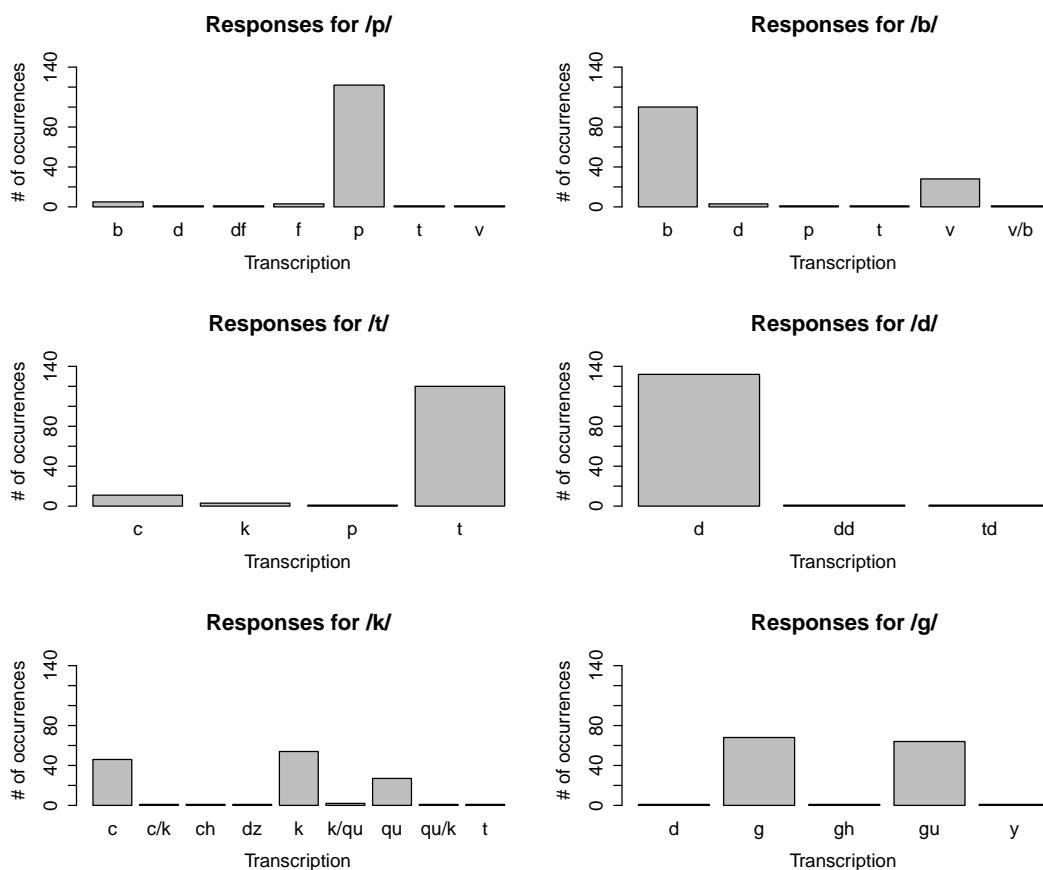
In terms of the goodness ratings, the voiceless stop stimuli should have high ratings in initial as well as in intervocalic position, but for the voiced stop stimuli there should be some discrepancy in the ratings for initial position vs. intervocalic position. The non-spirantized voiced stops of BP should be given lower ratings than their corresponding voiceless stop, regardless of whether they are identified as an example of a BAS voiced stop or a voiceless stop, since they would be non-canonical in either case.

## 2.6. Statistical Procedures

Since it is not expected that there be much variation in the transcription of the voiceless stop stimuli, they will be pooled together and analyzed in a two-way ANOVA, in order to confirm that neither type of stimulus stops nor position have an effect on their goodness ratings. However, the statistical tests of the voiced obstruents will depend on how they are assimilated and will likely require separate tests of independence (i.e. chi-square tests) between frequent responses and position before performing individual t-tests to determine whether there is an effect of position in the goodness ratings for the voiced stop stimuli.

## 3. Results and Discussion

The transcription responses and ratings for each participant were pooled for the analysis. The overall responses for the orthographic transcription of the stop stimuli can be seen in Figure 2 below.



**Figure 2** – Distribution of the transcription responses for each of the stop stimuli collapsed across position and vowel contexts.

From the inspection of the graphs in Figure 2, it seems that in most cases there are one or more clear options that were clearly favored by most listeners as the response for each stop stimulus, and the rest of the responses were very infrequent.

The responses for BP /p/ were overwhelmingly <p>, so the other responses were eliminated from the analysis. Similarly, the responses for BP /t/ were overwhelmingly <t>, so the other responses were also eliminated from the analysis. However, for the BP /k/ stimuli, there is more than one predominant response, <k>, <c>, and <qu> being the most frequent. This diversity in responses is likely due to the orthographic conventions of Spanish. Recall that the /k/ in the stimuli was always followed either by an <a> or an <i>. In Spanish, the orthographic sequence <ca> can be used to represent /ka/; however, <ci> does not represent /ki/ but rather /si/. The Spanish spelling convention to represent /ki/ is <qui>. The grapheme <k> is relatively infrequent in Spanish, occurring most commonly in foreign proper names or borrowings, also in the metric compounds with *kilo-*; however, <k> always represents /k/, regardless of the following vowel, which might account for its high use here. For these reasons, all responses <c>, <k>, <qu>, or any combination thereof, were combined and the remaining responses were eliminated from the analysis.

The three subsets for each voiceless stop stimulus were combined and a two-way ANOVA was run with goodness of fit ratings as the dependent variable and stimulus obstruent (/p, t, k/) and position (initial vs. intervocalic) as independent factors, with subject as a random factor. The results yielded no significant main effect of obstruent,  $F(2,28)=0.222$ ,  $p=.802$ , no significant main effect of position,  $F(1,28)=2.326$ ,  $p=.139$ , and no significant interaction,  $F(2,28)=1.043$ ,  $p=0.366$ . This result was expected, since the BP stimuli correspond with the canonical voiceless stops in Spanish in both initial and intervocalic positions.

Among the responses for BP /b/, <b> was the most frequent identification; however, <v> was also a considerably frequent response as well. In Spanish orthography both <b> and <v> are representations of the phoneme /b/, so either would be a valid transcription. Responses other than <b> or <v> were eliminated from further analyses. Interestingly, the distribution of <b> and <v> responses is not uniform across position (see Table 1 below).

	<b>	<v>
Initial	57	7
Intervocalic	43	21

**Table 1** – Distribution by position of <b> and <v> responses for BP /b/.

A chi-square test of independence shows that the relation between the listeners' responses (<b> vs. <v>) and the position of the stimuli is significant,  $\chi^2(1, N=128)=7.726, p=0.005$ . Initial /b/ stimuli were less likely to be identified as <v> than as <b>, whereas intervocalic stimuli of /b/ were more likely to be identified as <v> than as <b>. Although, there seems to be a bias towards identifying intervocalic [b] as <v>, this has no effect on the ratings of the stimuli. To investigate the effect of this bias on the subjects' ratings of BP /b/, a linear mixed effects regression model<sup>8</sup> was conducted with *response* (<b> vs. <v>) and *position* (initial vs. intervocalic) as fixed effects, and *subject* as a random factor, (shown in Table 2 below).

Response	Predictor	$\beta$	S.E.	t	p <sub>MCMC</sub>
rating	(Intercept)	4.367			
	response	-.410	.344	-1.194	.235
	position	-.844	.148	-5.685	<.001
	response:position	.337	.382	.881	.380

**Table 2** – Results of the linear mixed effect regression model adjusted to the ratings of BP /b/, with *response* and *position* as fixed factors and *subject* as a random factor. (Intercept = <v>, intervocalic).

The results yielded a significant main effect for position, such that the average rating for the stimuli with BP [b] was significantly higher for initial position (M=4.30, SD=0.81) than for intervocalic position (M=3.52, SD=1.10). The main effect of response was not significant and the interaction was not significant either, indicating that the effect of position holds, regardless of the subject's response.

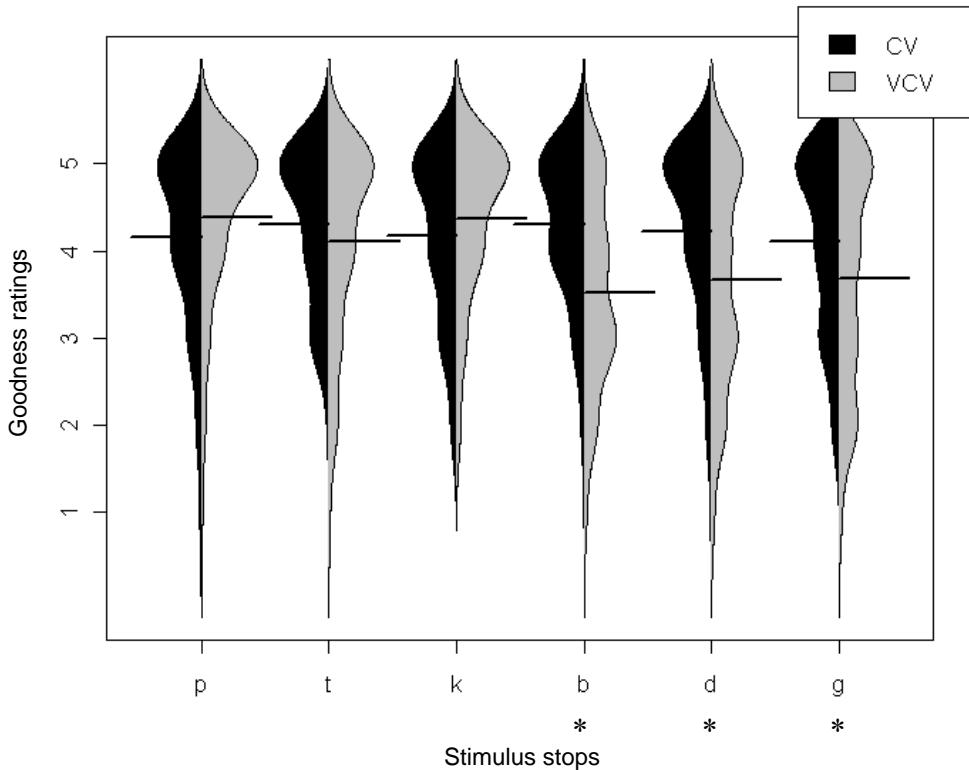
It might be the case that the listeners notice a discrepancy in the non-native occluded voiced stops and they identify it with <v>, the non-prototypical spelling of BAS /b/. However, the occluded voiced stops are not a better exemplar of <v> than of <b>. Given these results, it is warranted to collapse the responses of <b> and <v>; for both responses, [VbV] is rated lower than [bV].

The responses for BP /d/ were overwhelmingly <d>, so the other responses were eliminated from the analysis. The results of a t-test show that the goodness ratings for /d/ as <d> in intervocalic position (M=3.662, SD=1.189) are significantly lower than in initial position (M=4.224, SD=0.997),  $t(124)=2.939, p=0.004$ .

Most of the responses for the BP stimuli /g/ are distributed between <g> and <gu>. Similar to the case of /k/, this is likely due to the same spelling rule that affects the orthographic representation of /g/ depending on the following vowel. In Spanish, the orthographic sequence <ga> can be used to represent /ga/; however, <gi> does not represent /gi/ but /xi/. The Spanish spelling convention to represent /gi/ is <gui>. For these reasons, the responses <g> and <gu> were collapsed and the remaining responses were eliminated from the analysis. The results of a t-test show that ratings for intervocalic /g/ stimuli are significantly lower (M=3.682, SD=0.994) than the ratings for the initial /g/ stimuli (M=4.106, SD=1.242),  $t(124)=2.166, p=0.032$ .

<sup>8</sup> A mixed-effect regression model was conducted rather than a two-way ANOVA, because of the unbalanced distribution of the responses by subject, i.e. not every subject gave both orthographic variants in their responses. For this reason subject could not be included as a random factor in the ANOVA; however, a mixed-effect model can take subject as a random factor under these conditions.

The beanplots (Kampstra 2008) in Figure 3 below show the rating distribution for each of the stimulus stops, pooled across all subjects for initial and intervocalic positions (after the infrequent transcriptions have been removed).



**Figure 3** - Beanplots showing the rating distribution for each stop stimulus for initial and intervocalic position. Horizontal lines indicate the mean rating. Asterisks indicate a significant difference between CV and VCV position.

One very salient finding of this analysis of the stop stimuli is the effect of position for each of the voiced stops. There was no significant difference in the goodness ratings between the voiceless stops in initial vs. intervocalic position. In contrast, all the BP intervocalic occluded voiced stops were rated significantly lower than the voiced stops in initial position. Recall that Rohena-Madrado (2011) did not find any effect of position in the AX and ABX discrimination experiments. However, in this goodness rating task, participants gave significantly worse ratings for the intervocalic voiced stops than to initial voiced stops.

Judging from the orthographic responses, it seems that the listeners do assimilate the occluded voiced stops to their native voiced stop category, regardless of the lack of spirantization. Furthermore, listeners do not seem to associate occluded intervocalic voiced stops with the relatively common positional variant of voiceless stops in intervocalic position. This is precisely the point where the Perceptual Assimilation Model's predictions would differ from the Speech Learning Model's predictions. Considering the realization patterns of stops in Spanish discussed in section 1.2, [ada] as a realization of Spanish /ata/ would be much more frequent in casual speech than [ada] as a realization of Spanish /ada/. Under SLM, if the obstruent in BP [ada] were to relate to a positional variant, it would be much more plausible for it to relate to /ata/, since the acoustic characteristics of it are similar to how /ata/ is usually produced<sup>9</sup>. However, this does not seem to be the case. The data support PAM's

<sup>9</sup> However, it should be noted that the stimuli were all isolated words and thus evidently not casual, connected speech, which might have inhibited the association to the casual speech variant. Neither PAM nor SLM address how stylistic or environmental effects might affect the patterns of association of the non-native sounds; however, it has been shown that this does affect listeners' perceptual boundaries (see Hay et al. 2006).

predictions in that the non-native [ada] seems to assimilate to the higher order invariants associated with Spanish /ada/, but also taking into consideration the phonetic positional information and qualifying it is a worse fit to that category, at least as compared to a phrase initial tokens.

#### 4. Conclusions

The results from these transcription and goodness rating tasks provide the necessary information in order to understand and fully interpret the results from the AX and ABX discrimination experiments. All the BP stops were assimilated to their corresponding stop categories in BAS, even in intervocalic position, in accordance with Best's PAM. Despite the occlusion in intervocalic position, the listeners were able to perceive the articulatory invariants in the production of the voiced stop category rather than rely solely on the acoustic cues, which were more similar to those of phonologically voiceless (but phonetically voiced) occluded stops in Spanish.

To answer the research questions that were posited earlier, Spanish listeners do seem to assimilate occluded intervocalic voiced stops to their native voiced category, despite the fact that occlusion in this position is non-native in Spanish (or at least extremely infrequent). The findings from this investigation shed light on the AX and ABX results from Rohena-Madrado (2011), confirming that, despite the native variation in Spanish stop production, the contrast implemented in a non-native position, [VtV]-[VdV], is perceived as a Two Category assimilation type. Although the intervocalic occluded voiced stop is perceived as a less prototypical exemplar of the native voiced stop category, this does not seem to interfere with the discrimination of the contrast, which is also in line with PAM's predictions.

It would be warranted to corroborate these findings on assimilation of stop voicing contrasts in Spanish with future discrimination, perceptual assimilation, and rating tasks by manipulating other cues to the stop voicing contrast, e.g. duration, VOT, or onset of F1. Furthermore, the frequency and distribution of voiceless stop lenition in Buenos Aires Spanish and its pervasiveness across speech styles needs to be investigated further, particularly considering the fact that the current stimuli were produced in isolation and in a careful style. Further studies would benefit from the use of stimuli taken from relatively less careful speech, such as the examples presented in Figure 1.

Taking into consideration the results of this study with those of Rohena-Madrado (2011), Flege's Speech Learning Model, with its assimilation depending solely on positionally defined phonetic categories, does not make the right predictions; rather, the results of this study support Best's Perceptual Assimilation Model, which takes into account the interrelationship of phonetic categories with phonological categories in the identification of the higher order invariants as the targets of assimilation. These findings suggest that the mapping of non-native contrasts to native phonological categories, not just to position-specific allophones, plays a determining role in cross-language speech perception.

## 5. Appendix

This appendix contains a more detailed acoustic description of the BP stop stimuli used in the perceptual assimilation task described in Section 2.3.

<i>configuration</i>	<i>stimulus</i>	<i>stop</i>	<i>closure</i>	<i>burst</i>	<i>aspiration</i>	<i>VOT</i>
CVLC	bala	b	-	0.003	0	-0.041
CVLC	bili	b	-	0.005	0	-0.052
CVLC	dala	d	-	0.006	0	-0.076
CVLC	dili	d	-	0.006	0	-0.055
CVLC	gala	g	-	0.015	0	-0.077
CVLC	gili	g	-	0.005	0	-0.084
CVLC	kala	k	-	0.014	0.036	0.045
CVLV	kili	k	-	0.005	0.065	0.071
CVLC	pala	p	-	0.006	0.008	0.014
CVLC	pili	p	-	0.008	0.004	0.010
CVLC	tala	t	-	0.012	0.008	0.017
CVLC	tili	t	-	0.010	0.019	0.023
LVCV	laba	b	0.091	0.009	0	fully voiced
LVCV	libi	b	0.099	0.009	0	fully voiced
LVCV	lada	d	0.076	0.008	0	fully voiced
LVCV	lidi	d	0.070	0.010	0	fully voiced
LVCV	laga	g	0.070	0.008	0.014	fully voiced
LVCV	ligi	g	0.077	0.006	0.006	fully voiced
LVCV	laka	k	0.103	0.016	0.010	0.025
LVCV	liki	k	0.084	0.02	0.047	0.058
LVCV	lapa	p	0.128	0.009	0	0.009
LVCV	lipi	p	0.137	0.010	0	0.010
LVCV	lata	t	0.114	0.009	0.009	0.018
LVCV	liti	t	0.117	0.009	0.008	0.017

**Table 3** – Acoustic detail of the perceptual assimilation stimuli. All duration measurements are in seconds. Closure durations are not provided for the phrase initial stimuli, because there is no way of identifying the exact onset of the stop closure from the acoustic signal, since they are always preceded by silence. It can be assumed that the duration of the initial voiced stop closures are at least as long as their negative VOT.

## References

- Amastae, Jon. 1995. Variable spirantization: constraint weighting in three dialects. *Hispanic Linguistics*, 6(7): 267-285.
- Barbosa, Plinio, and Eleonora Albano. 2004. Brazilian Portuguese. *Journal of the International Phonetic Association*, 34(2): 227-232.
- Best, Catherine T. 1995. A Direct Realist View of Cross-Language Speech Perception. In *Speech Perception and Linguistic Experience: Issues in Cross-language Speech Research*, ed. Winifred Strange, 171-203. Timonium, MD: York.
- Best, Catherine T., Gerald W. McRoberts, and Nomathemba M. Sithole. 1988. Examination of Perceptual Reorganization for Nonnative Speech Contrasts: Zulu Click Discrimination by English-Speaking Adults and Infants. *Journal of Experimental Psychology: Human Perception and Performance*, 14(3): 345-360.
- Best, Catherine T., Gerald W. McRoberts, Elizabeth Goodell. 2001. Discrimination of non-native consonant contrasts varying in perceptual assimilation to the listener's native phonological system. *Journal of the Acoustical Society of America*, 109(2): 775-794.
- Best, Catherine T., and Winifred Strange. 1992. Effects of phonological and phonetics factors on cross-language speech perception on approximants, *Journal of Phonetics*, 20: 305-330.
- Best, Catherine T., and Michael D. Tyler. 2007. Nonnative and second-language speech perception: Commonalities and complementarities. In *Language experience in second language speech learning. In honor of James Emil Flege*, ed. Ocke-Schwen Bohn and Murray J. Munro, 13-34. Amsterdam: John Benjamins.
- Boersma, Paul, and David Weenink. 2010. Praat: doing phonetics by computer. Version 5.1.26. <http://www.praat.org>.

- Browman, Catherine, and Louis Goldstein. 1989. Articulatory Gestures as Phonological Units. *Phonology*, 6: 201-252.
- Browman, Catherine, and Louis Goldstein. 1990. Tiers in articulatory phonology, with some implications for casual speech. In *Papers in Laboratory phonology I: Between Grammar and the Physics of Speech*, ed. John Kingston and Mary E. Beckman, 341-76. Cambridge: Cambridge University Press.
- Browman, Catherine, and Louis Goldstein. 1992. Articulatory phonology: an overview. *Phonetica*, 49: 155-180.
- Colantoni, Laura, and Irina Marinescu. 2010. The Scope of Stop Weakening in Argentine Spanish. In *Selected Proceedings of the 4th Conference on Laboratory Approaches to Spanish Phonology*, ed. Marta Ortega-Llebaria, 100-114. Somerville, MA: Cascadilla Proceedings Project.
- Cole, Jennifer, José I. Hualde, and Khalil Iskarous. 1999. Effects of Prosodic and Segmental Context on /g/-Lenition in Spanish. In *Proceedings of the Fourth International Linguistics and Phonetics Conference*, ed. Osamu Fujimura, Brian D. Joseph and Bohumil Palek, 575-589. Prague: The Karolinum Press.
- Davidson, Lisa. 2007. The relationship between the perception of non-native phonotactics and loanword adaptation. *Phonology*, 24(2): 261-286.
- Davidson, Lisa, and Jason Shaw. 2012. Sources of illusion in consonant cluster perception. *Journal of Phonetics*, 40: 234-248.
- Dupoux, Emmanuel, Kazohiko Kakehi, Yuki Hirose, Christophe Pallier, and Jacques Mehler. 1999. Epenthetic vowels in Japanese: A perceptual illusion? *Journal of Experimental Psychology: Human Perception and Performance*, 25: 1568-1578.
- Flege, James E. 1989. Chinese subjects' perception of the word-final English /t/-/d/ contrast: Performance before and after training. *Journal of the Acoustical Society of America*, 86: 1684-1697.
- Flege, James E. 1995. Second language speech learning: Theory, findings and problems. In *Speech Perception and Linguistic Experience: Issues in Cross-language Speech Research*, ed. Winifred Strange, 233-272. Timonium, MD: York.
- Flege, James E. 1999. Age of learning and second-language speech. In *New Perspectives on the Critical Period Hypothesis for Second Language Acquisition*, ed. David P. Birdsong, 101-132. Hillsdale, NJ: Erlbaum.
- Flege, James E. 2003. Assessing constraints on second-language segmental production and perception. In *Phonetics and Phonology in Language Comprehension and Production, Differences and Similarities*, ed. Niels Olaf Schiller and Antje Meyer, 319-355. Berlin: Mouton de Gruyter.
- Gillette, Susan. 1980. Contextual variation in the perception of L and R by Japanese and Korean Speakers. *Minnesota Papers in Linguistics and Philosophy of Language*, 6: 59-72.
- Goto, Hiromu. 1971. Auditory perception by normal Japanese adults of sounds "L" and "R". *Neuropsychologia*, 9(3): 317-323.
- Hallé, Pierre A, and Catherine T. Best. 2007. Dental-to-velar perceptual assimilation: A cross-linguistic study of the perception of dental stop+/l/ clusters. *Journal of the Acoustical Society of America*, 5(1): 2899-2914.
- Hallé, Pierre, Catherine T. Best, and Asaf Bacharach. 2003. Perception of /dl/ and /tl/ clusters: A cross-language perceptual study with French and Israeli listeners. *Proceedings of the 15<sup>th</sup> International Congress of Phonetic Sciences*. Barcelona, Spain, 2893-2896.
- Hallé, Pierre A., Juan Seguí, Uli Frauenfelder, and Christine Meunier. 1998. The processing of illegal consonant clusters: A case of perceptual assimilation? *Journal of Experimental Psychology: Human Perception and Performance*, 24: 592-602.
- Hay, Jennifer, Paul Warren, and Katie Drager. 2006. Factors influencing speech perception in the context of a merger-in-progress. *Journal of Phonetics*, 34(4): 458-484.
- Hualde, José I. 2005. *The sounds of Spanish*. Cambridge: Cambridge University Press.
- Hualde, José I., Miquel Simonet, Ryan Shosted, and Marianna Nadeu. 2010. Quantifying Iberian Spirantization: Acoustics and articulation. Paper presented at LSRL 40, Seattle, March 26-28, 2010.
- Kampstra, Peter. 2008. Beanplot: A Boxplot Alternative for Visual Comparison of Distributions. *Journal of Statistical Software*, 28(1): 1-9.
- Kingston, John. 2008. Lenition. In *Selected Proceedings of the 3rd Conference on Laboratory Approaches to Spanish Phonology*, ed. Laura Colantoni and Jeffrey Steele, 1-31. Somerville, MA: Cascadilla Proceedings Project.
- Kuhl, Patricia K. 1993 Early linguistic experience and phonetic perception: implications for theories of developmental speech perception. *Journal of Phonetics*, 21: 125-139.
- Kuhl, Patricia K., Barbara T. Conboy, Sharon Coffey-Corina, Denise Padden, Maritza Rivera-Gaxiola, Tobey Nelson 2008. Phonetic learning as a pathway to language: new data and native language magnet theory expanded (NLM-e). *Philosophical Transactions of the Royal Society B*, 363: 979-1000.

- Kuhl, Patricia K., and Paul Iverson. 1995. Linguistic experience and the perceptual magnet effect. In *Speech Perception and Linguistic Experience: Issues in Cross-language Research*, ed. Winifred Strange, 121-154. Baltimore, MD: York.
- Kuhl, Patricia K., Erica Stevens, Akiko Hayashi, Toshisada Deguchi, Shigeru Kiritani, and Paul Iverson. 2006. Infants show a facilitation effect for native language phonetic perception between 6 and 12 months. *Developmental Science*, 9: F13-F21.
- Lapesa, Rafael. 1981. *Historia de la Lengua Española*. Madrid: Gredos.
- Lavoie, Lisa. 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. Doctoral Dissertation, University of Illinois at Urbana-Champaign.
- Lipski, John. 1994. *Latin American Spanish*. New York: Longman.
- Lisker, Leigh, and Arthur S. Abramson. 1964. A cross-language study of voicing in initial stops: Acoustical measurements. *Word*, 20: 384-422.
- Logan, John S., Scott E. Lively, and David B. Pisoni. 1991. Training Japanese listeners to identify English /r/ and /l/: A first report. *Journal of the Acoustical Society of America*, 89: 874-886.
- Machuca Ayuso, María Jesús. 1997. Las obstruyentes no continuas del español: relación entre las categorías fonéticas y fonológicas en el habla espontánea. Doctoral Dissertation, Universitat Autònoma de Barcelona.
- Martínez-Celdrán, Eugenio. 2008. Some Chimeras of Traditional Spanish Phonetics. In *Selected Proceedings of the 3rd Conference on Laboratory Approaches to Spanish Phonology*, ed. Laura Colantoni and Jeffrey Steele, 32-46. Somerville, MA: Cascadilla Proceedings Project.
- Martínez Celdrán, Eugenio. 2009. Sonorización de las oclusivas sordas en una hablante Murciana: problemas que presenta. *Estudios de Fonética Experimental*, XVIII: 253-271.
- Martínez Celdrán, Eugenio, and Ana Fernández Planas. 2007. *Manual de fonética española: Articulaciones y sonidos del español*. Barcelona: Ariel.
- Martínez-Celdrán, Eugenio, Ana María Fernández-Planas, and Josefina Carrera-Sabaté. 2003. Castilian Spanish. *Journal of the International Phonetic Association*, 33(2): 254-259.
- Mattoso Câmara, Joaquim. 1972. *The Portuguese Language*. Chicago: University of Chicago Press.
- Ortega-Llebaria, Marta. 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 *Laboratory approaches to Spanish phonology*, ed. Timothy Face, 237-253. Berlin: Mouton de Gruyter.
- Parrell, Benjamin. 2010. How /b d g/ differ from /p t k/ in Spanish: A dynamic account. Paper presented at LabPhon 12, Albuquerque, New Mexico, July 2010.
- Pisoni, David B. 1973. Auditory and phonetic codes in the discrimination of consonants and vowels. *Perception & Psychophysics*, 13: 253-260.
- Polka, Linda. 1991. Cross-language speech perception in adults: Phonemic, phonetic and acoustic contributions. *Journal of the Acoustical Society of America*, 89(6): 2961-2977.
- Quilis, Antonio. 1999. *Tratado de fonología y fonética españolas* (2nd edition). Madrid: Gredos.
- Rohena-Madrado, Marcos. 2011. Sociophonetic Variation in the Production and Perception of Obstruent Voicing in Buenos Aires Spanish. Doctoral Dissertation, New York University.
- Santagada, Miguel, and Jorge Gurlekian. 1989. Spanish voiced stops in VCV contexts: Are they fricative variants or approximants? *Revue de phonétique appliquée*, 91-93: 363-75.
- Sheldon, Amy, and Winifred Strange. 1982. The acquisition of /r/ and /l/ by Japanese learners of English: Evidence that speech production can precede speech perception. *Applied Psycholinguistics*, 3: 243-261.
- Soler, Antonia, and Joaquín Romero. 1999. The role of duration in stop lenition in Spanish. *Proceedings of the 14<sup>th</sup> International Congress of Phonetic Sciences*, San Francisco, 483-486.
- Strange, Winifred. 1992. Learning non-native phoneme contrasts: Interactions among subject, stimulus, and task variables. In *Speech Perception, Production, and Linguistic Structure*, ed. Yoh'ichi Tohkura, Eric Vatikiotis-Bateson and Yoshinori Sagisaka, 197-219, Tokyo: Ohmsha.
- Strange, Winifred. 1995. *Speech Perception and Linguistic Experience: Issues in Cross-language Speech Research*. Timonium, MD: York.
- Strange, Winifred, Reiko Akahane-Yamada, Rieko Kubo, Sonja A. Trent, and Kanae Nishi. 2001. Effects of consonantal context on perceptual assimilation of American English vowels by Japanese listeners. *Journal of the Acoustical Society of America*, 109(4): 1691-1704.
- Torreira, Francisco, and Mirjam Ernestus. 2011. Realization of voiceless stops and vowels in conversational French and Spanish. *Laboratory Phonology*, 2(2): 331-353.

- Werker, Janet. 1995. Age-related Changes in Cross-Language Speech Perception: Standing at the Crossroads. In *Speech Perception and Linguistic Experience: Issues in Cross-language research*, ed. Winifred Strange, 155-167. Baltimore, MD: York.
- Werker, Janet, and John S. Logan. 1985. Cross-language evidence for three factors in speech perception. *Perception & Psychophysics*, 37(1): 35-44.
- Werker, Janet & Richard Tees. 1984a. Cross-Language Speech Perception: Evidence for Perceptual Reorganization During the First Year of Life. *Infant Behavior and Development*, 7: 49-63.
- Werker, Janet & Richard Tees. 1984b. Phonemic and phonetic factors in adult cross-language speech perception. *Journal of the Acoustical Society of America*, 75(6): 1866-1878.

# Selected Proceedings of the 15th Hispanic Linguistics Symposium

edited by Chad Howe, Sarah E. Blackwell,  
and Margaret Lubbers Quesada

Cascadilla Proceedings Project Somerville, MA 2013

## Copyright information

Selected Proceedings of the 15th Hispanic Linguistics Symposium  
© 2013 Cascadilla Proceedings Project, Somerville, MA. All rights reserved

ISBN 978-1-57473-457-7 library binding

A copyright notice for each paper is located at the bottom of the first page of the paper.  
Reprints for course packs can be authorized by Cascadilla Proceedings Project.

## Ordering information

Orders for the library binding edition are handled by Cascadilla Press.  
To place an order, go to [www.lingref.com](http://www.lingref.com) or contact:

Cascadilla Press, P.O. Box 440355, Somerville, MA 02144, USA  
phone: 1-617-776-2370, fax: 1-617-776-2271, [sales@cascadilla.com](mailto:sales@cascadilla.com)

## Web access and citation information

This entire proceedings can also be viewed on the web at [www.lingref.com](http://www.lingref.com). Each paper has a unique document # which can be added to citations to facilitate access. The document # should not replace the full citation.

This paper can be cited as:

Rohena-Madrazo, Marcos. 2013. Perceptual Assimilation of Occluded Voiced Stops by Spanish Listeners.  
In *Selected Proceedings of the 15th Hispanic Linguistics Symposium*, ed. Chad Howe et al., 140-156. Somerville, MA: Cascadilla Proceedings Project. [www.lingref.com](http://www.lingref.com), document #2881.