COSTA RICAN SPANISH SPEAKERS'
PHONETIC DISCRIMINATION

LA DISCRIMINACIÓN FONÉTICA
DE LOS HABLANTES DEL ESPAÑOL COSTARRICENSE

WHITNEY CHAPPELL
The University of Texas at San Antonio
whitney.chappell@utsa.edu

ABSTRACT

Costa Rican Spanish listeners associate intervocalic [z] with specific social attributes in a matched-guise test (Chappell, 2016) but experience difficulty when explicitly asked to produce or even comment on the variant. Given this perception-production discrepancy, the present study seeks to determine how successfully listeners discriminate between allophonic differences like intervocalic [s] and [z] compared to other allophone pairs, phonemic contrasts, and identical stimuli. 106 Costa Rican listeners completed similarity rating and AX discrimination tasks in which they evaluated word pairs that were identical or differed only in one phoneme or allophone. Statistical analyses fitted to 2,862 tokens in the similarity rating task and 3,604 tokens from the AX discrimination task indicate that listeners perceive phonemic contrasts more successfully than allophonic differences, which, in turn, are perceived as more distinct than identity pairs. Interestingly, the [s] ~ [z] distinction is less successfully perceived than other allophone pairs including [n] ~ [ŋ] and [d] ~ [ð]. I contend that allophonic differences that encode linguistic information, e.g. the variable's position within the word, or are less expected given their low frequency are heard more successfully than [s] ~ [z]. However, even the least salient phonetic variants like [s] ~ [z] can encode local social meaning and contribute to listeners' evaluations of speakers' social qualities.

Keywords: *Costa Rican Spanish, voicing, intervocalic /s/, phonetic discrimination, allophony, perception.*

RESUMEN

Utilizando la técnica de par oculto, un trabajo reciente muestra que los costarricenses asocian la [z] intervocálica con ciertos atributos sociales (Chappell, 2016) pero tienen dificultad cuando se les pide que produzcan o comenten sobre la variante. Con el fin de resolver esta paradoja, el presente estudio indaga sobre este desajuste entre la percepción y la producción y busca determinar cuán exitosamente los oyentes costarricenses discriminan entre diferentes alófonos como la [s] y la [z] frente a otros alófonos, contrastes fonémicos y estímulos idénticos. 106 costarricenses completaron una tarea de evaluación de similitud y una tarea de discriminación AX en las que se escucharon pares de palabras idénticas o que difirieron en un sólo alófono o fonema. Los resultados de los análisis estadísticos de 2,862 evaluaciones de la tarea de similitud y 3,604 evaluaciones de la tarea de discriminación AX indican que los oyentes evalúan los

pares que contienen diferentes fonemas como significativamente más distintos que los pares que contienen diferentes alófonos y éstos, a su vez, se evalúan como significativamente más distintos que los estímulos idénticos. Dentro de la categoría alofónica, la distinción entre [s] y [z] se percibe con menos éxito que los otros pares alofónicos como [n] ~ [ŋ] y [d] ~ [ð]. Sostengo que las diferencias alofónicas que conllevan significado lingüístico sobre la posición de la palabra o que son menos probables dada su baja frecuencia se escuchan mejor que [s] ~ [z]. Sin embargo, incluso las variantes fonéticas menos sobresalientes pueden codificar significado social al nivel local e influir en las evaluaciones de los atributos sociales de los hablantes que se oyen.

Palabras clave: *Costa Rica, sonorización, /s/ intervocálica, discriminación fonética, alofonía, percepción*.

1. INTRODUCTION

Recent work on Costa Rican Spanish (Chappell, 2016) has established a connection between intervocalic /s/ voicing, e.g. *cosa* 'thing' as [koza], and indexical social meaning: a matched-guise test shows that Costa Rican listeners associate intervocalic [z] with less educated and lower-class speakers, and, when the speaker is male, intervocalic [z] also provokes higher evaluations of the speaker's niceness, confidence, Costa Rican-ness, and masculinity. That is, intervocalic [z] is interpreted by Costa Rican listeners as a meaningful marker of speakers' social attributes, though the gender of the speaker impacts the indexical social meanings evoked by nonstandard /s/ voicing.

In spite of the robust tendency for listeners to evaluate speakers who produce intervocalic [z] differently than the same speakers producing intervocalic [s], the Costa Rican participants in Chappell (2016) experienced difficulty when asked to produce or explicitly comment on intervocalic [z]. These participants were instructed to listen to the researcher's production of [s] and [z], e.g. [keso] and [kezo] for *queso* 'cheese', and imitate these pronunciations as closely as possible. Even speakers who consistently voiced intervocalic /s/ in spontaneous speech struggled with the task and generally relied on one of two initial strategies: they either produced voiceless [s] two times, e.g. [keso] and [keso] or produced [s] and

[θ], e.g. [keso] and [keθo],[1] assuming that the researcher had instructed them to produce the phonemic contrast (*distinción*) between /s/ and /θ/ found in central and northern Spain.[2] This phonemic contrast does not exist in Latin America; /s/ and /θ/ were historically merged and are now realized as a single alveolar sibilant (Morgan, 2010:283-286), making the speakers' production of [θ] when asked to produce [z] highly unexpected.

This perception-production discrepancy presents an interesting paradox. On the one hand, Costa Rican listeners hear intervocalic [z] as a meaningful marker of a speaker's social attributes and provide consistent evaluations of the qualities indexed by intervocalic [s] and [z]. On the other hand, Costa Rican speakers are not able to produce and comment on intervocalic [z] without a great deal of coaching, suggesting that the variation is below the level of consciousness. There is some support in the literature that different types of phonemic contrasts are perceived differently (Hume and Johnson, 2003), but little is known about the ways in which non-contrastive, allophonic variants are perceived. The present study seeks to investigate Costa Rican listeners' allophonic discrimination using three pairs of allophones, [s] ~ [z], [n] ~ [ŋ] and [d] ~ [ð], with the goal of comparing how successfully each pair is perceived as compared to phonemic contrast pairs, e.g. [ð] ~ [ɾ], and identical stimuli, e.g. [d] ~ [d]. A detailed exploration of these intervocalic phones' perceptibility will help determine how Costa Ricans discriminate between variants and if all allophonic differences are perceived in the same way.

2. LITERATURE REVIEW

Speech perception has fascinated linguists for decades, with early research focusing on the perception of phonemic contrasts. For instance, in *Grundzüge der Phonologie* (1939) Trubetzkoy makes three principle predictions about phonemic discrimination. First, Trubetzkoy posits that an individual's experience with his/her native language will influence his/her success in discriminating between sounds.

---

[1] It should be noted that all speakers recorded for Chappell (2016) were ultimately able to produce the intervocalic [z], but the successful production of the phones did require a great deal of repetition and correction.

[2] Peninsular *distinción* involves the phonemic contrast between /s/ and /θ/, and the contrast is reflected orthographically in the distinction in pronunciation between <s>, on the one hand, and <z>, <ce>, and <ci>, on the other. For example, *casa* 'house' would be pronounced as [kasa], but *caza* 's/he hunts' is realized as [kaθa].

Next, the phonological contrast between two sounds in the listener's native language will influence his/her ability to perceive the sounds. Finally, the type of phonological contrast will affect speech perception as well.

Experience with one's native language and the phonological status of two sounds in the native language have been robustly shown to influence perception in the literature, supporting Trubetzkoy's first two hypotheses. Listeners are more adept at identifying contrasts in their native language than a second language (Best *et al*., 1998; Dupoux *et al*., 1997; Francis and Nusbaum, 2002; Polka and Werker, 1994; Strange, 1995). For example, L1-Japanese L2-English speakers are less successful at identifying the /r/ and /l/ contrast in English, as the sounds are not contrastive in Japanese (Goto, 1971; MacKain *et al*., 1981). Similarly, L1-Spanish L2-English speakers often experience difficulty identifying contrastive vowels in English that are not contrastive in Spanish (Fox *et al*., 1995), and American English speakers experience difficulty when discriminating between Hindi dental-retroflex stop contrasts that are not contrastive in American English (Pruitt, 1995).

Trubetzkoy's first two hypotheses are further supported by Boomershine *et al*. (2008), who explore the perception of [d], [ð], and [ɾ] for English and Spanish speakers. While these three phones exist in both English and Spanish, these languages place the sounds "…in very different positions in the linguistic system of contrasts" (Boomershine *et al*., 2008:3). That is, [d] and [ð] are contrastive in English but not in Spanish, e.g. *dare* and *there* form a minimal pair in English while [ada] and [aða], meaning 'fairy' in Spanish, do not. On the other hand, [d] and [ɾ] are contrastive in Spanish but not in English, e.g. *oda* 'ode' and *ora* 's/he prays' form a minimal pair but pronunciations of *body* as [bɑdi] and [bɑɾi] do not. Lastly, [ð] and [ɾ] are contrastive in both English and Spanish, e.g. *ladder* and *lather* in English and *cara* 'face' ([kaɾa]) and *cada* 'each' ([kaða]) in Spanish. The investigators conclude that listeners find phones that are contrastive in their native language to be more perceptually distinct, and the Spanish and English listeners found [ð] and [ɾ] to be equally distinct in the two languages.

This difference in a speaker's perceptual discrimination in his/her native language and languages learned later in life has been explained by several recent theoretical models. Kuhl's Native Language Magnet model contends that the native phonological system begins to lose sensitivity to phonological contrasts to which it is not exposed around the age of six months, and experiments show that listeners perceive non-prototypical realizations of a vowel as closer to the nearest phonological category's prototype than another non-prototypical realization, even though the distance between the two is identical (Greiser and Kuhl, 1989; Kuhl *et*

al., 1992, Kuhl *et al.*, 2008). That is, native speakers form categories in their native language, and these categories affect their perception of contrasts in languages learned later in life (Kazanina *et al.*, 2006). Flege's Speech Learning Model (Flege, 1988, 1990, 1995) and Best's Perceptual Assimilation Model (Best, 1994a, 1994b, 1995) also posit that adults' less successful discrimination of non-native phones is due to their native speech system.

Speech perception may also be impacted by the phonotactics and context-dependence of specific allophones in a speaker's first language (Ingram and Park, 1998). For example, Larkey, Wald and Strange (1978) presented American English listeners with a synthetic /n-ŋ/ continuum and found that respondents were more likely to identify /n/ in initial position even when the stimulus was closer to /ŋ/, while the modal response for word-final position was /ŋ/. These results suggest that English phonotactics, which disallow the velar nasal in word-initial position, influenced participants' perception of nasals.

A similar effect of L1 phonotactic constraints was found for L1-Japanese English language learners: the listeners were least successful at perceiving /l/ and /r/ in initial consonant clusters where liquids are disallowed in Japanese (Mochizuki, 1981). However, the presence of similar allophonic categories in listeners' L1 seem to aid perception in the L2. Although both Korean and Japanese have a single liquid phoneme, Korean listeners were more successful than Japanese listeners at identifying intervocalic liquids in English (Ingram and Park, 1998). This success is attributed to the presence of an intervocalic lateral geminate in Korean that is absent in Japanese, allowing Korean listeners to map English /l/ and /r/ onto the singleton and geminate allophones of the single liquid in their first language.

The aforementioned studies indicate that a listener's L1 serves as a phonotactic filter, affecting perception, and this pattern holds true in Romance languages as well. French listeners showed a tendency to perceive phonotactically disallowed word-initial consonant clusters as similar, phonotactically allowed sequences and transcribe them as such (Hallé *et al.*, 1998). In addition to the segmental level, suprasegmental information in the L1 also impacts perception. Suprasegmental information like stress can be contrastive in Spanish, e.g. ***be****be* 's/he drinks' and *be****bé*** 'baby', but not in French, and French and Spanish listeners' perceptual categories are significantly different in stress-based ABX tasks. While French listeners experience more difficulty in stress-oriented classification tasks, Spanish listeners have trouble ignoring irrelevant accentual information (Dupoux *et al.*, 1997), demonstrating the long-term perceptual impact of the L1.

Fewer studies have explored Trubetzkoy's third hypothesis, but there is some support that certain phonological contrasts are more readily perceived than others. For instance, in Mandarin Chinese the low-falling-rising tone and the mid-rising tone are contrastive in most positions but not in all. The tones are neutralized after a low-falling-rising tone (Hume and Johnson, 2003:2), and the partially contrastive status of the two tones does limit their perceptual distinctiveness for native speakers.

The present study serves as an extension of this hypothesis, raising the question as to whether some allophonic variants are more readily perceived than others. Previous work on allophony has suggested that allophones, while produced distinctively, are perceived as a single perceptual object (Derwing *et al*., 1986; Jaeger, 1980). Derwing *et al*. (1986:55) conclude that listeners

> *readily extrapolate from one positional variant of English /k/ to others; they are generally willing to accept one variant of English /t/ as equivalent to most others; and, with little or no instruction, they seem quite able to segment English speech in a way that largely accords with a traditional taxonomic phonemic analysis.*

Whalen *et al*. (1997) investigate the discrimination of aspirated and unaspirated [p] in English, finding that aspirated [p] was preferred over unaspirated [p] regardless of syllable stress in nonce words, and unaspirated [p] is more difficult to identify than aspirated [p] in an AXB task. In real words, participants preferred the appropriate contextual allophone and found the inappropriate allophone difficult to produce, which indicates that allophones may require contact with the lexicon for their production and perception to be affected.

Delving deeper into the factors contributing to differential discrimination, McQueen and Pitt (1996) point to transitional probabilities of phonological units. They predict that listeners' responses will be faster and more accurate when asked if they heard a particular sound if there is a high probability that the sound would occur in that environment. The authors determine that targets were identified more successfully in CVCC sequences when the preceding consonant and vowel made the following consonantal sequence more probable, demonstrating that listeners are aware of the probability that a phone will occur in a particular environment.

Listeners seem to be highly aware of segmental frequency and even search for frequency-based patterns in new data. After approximately 20 minutes of listening to synthesized nonsense syllables without any indication of the boundaries between words, listeners were able to identify the boundaries between nonce words at a better-than-chance level, suggesting that the listeners were tracking the

nonce words' transitional probabilities, noting the likelihood that a syllable would precede or follow others (Saffran *et al*., 1996). Similarly, Dahan *et al.* (2008) demonstrate that listeners exposed to a dialect that produces tense [æ] before [g] and lax [æ] before [k] in a training exercise were faster and more accurate when identifying stimuli containing [k] than listeners who were not exposed to this dialect. Listeners seem to learn transitional probabilities in nonce words and previously unfamiliar dialects rather quickly, and it follows that native speakers recognize transitional probabilities in their language and are aware that certain phones are more likely to appear in a given context than others.

Numerous studies suggest that listeners do monitor transitional probabilities in their native language. For example, highly educated native speakers of Dutch seem to be influenced by verbal frequency effects when asked to produce the past tense, which is formed by adding the allomorph [tə] or [də] to the verb stem depending on the voicing specification of the stem's final obstruent (Ernestus, 2006). The participants had slower reaction times and sometimes produced the non-standard form when presented with verbs whose internal structure made them similar to verbs with the opposite stem-final voicing specification. These results are supported by Ernestus and Mak (2005), who find a similar tendency in corpus-based writing samples, indicative of the fact that distributional factors condition predictions about the likelihood of particular phones.

In addition to the ability to recognize the transitional probabilities of phonemic segments, listeners seem to discern subphonemic cues as well. For instance, Flagg *et al*. (2006) and Fowler and Brown (2000) show that listeners make use of oral and nasal vowels in English to predict whether the following consonant is oral or nasal, and when listeners' expectations are violated, i.e., an oral vowel is presented before a nasal consonant or a nasal vowel is presented before an oral consonant, listeners' responses were delayed.

Beyond intralinguistic discrimination judgments, listeners are able to use nonphonemic production differences as they make social evaluations about speakers (Barnes, 2015; Campbell-Kibler, 2007, 2008, 2009, 2010; Clopper and Pisoni, 2004; Drager, 2010; Fridland *et al*., 2004; Plichta and Preston, 2005; Szakay, 2008; Walker, 2007). Clopper and Pisoni (2004) contend that listeners categorize speakers' dialects using a limited number of acoustic-phonetic properties. In the United States, American English speakers are acutely aware of social properties indexed by salient vowel shifts (Fridland *et al*., 2004). For instance, American listeners associate degree of (ay) monophthongization with degree of southernness (Plichta and Preston, 2005). Mexican and Puerto Rican

Spanish speakers, on the other hand, have been shown to evaluate salient consonantal lenition processes, with evaluations of social status and heteronormativity linked to coda /s/ aspiration (Walker *et al*., 2014). As noted in the introduction, Chappell (2016) shows that intervocalic /s/ voicing in Costa Rica is associated with lower social status for all speakers and higher evaluations of masculinity, confidence, niceness, and localness for male speakers. In other words, in spite of being perceived as a single perceptual object, different allophones are often imbued with social meaning that is adroitly identified by local listeners.

The research presented in the preceding paragraphs demonstrates the importance of phonotactics, predictability, and variants' association with social information. Previous studies show that listeners make use of linguistic information in real words to determine the appropriate allophone (Whalen *et al*., 1997), highlighting the importance of linguistic factors like stress and syllable/word position in allophonic discrimination. Frequency effects also condition listeners' responses, which suggests that listeners make predictions about the likelihood of a particular phone based on frequency (Flagg *et al*., 2006; Fowler and Brown, 2000). Finally, Costa Rican listeners' clearly associate intervocalic [z] with certain social attributes and yet appear to be unable to explicitly produce or comment on the variants (Chappell, 2016). As a result, linguistic factors like syllable/word position and frequency effects may influence the success with which different allophones are discriminated to a greater degree than allophones imbued with social information. First, I hypothesize listeners will be more sensitive to allophonic variants that carry linguistic information about word position, and second, I hypothesize listeners will be more sensitive to less predictable variants, provoking higher evaluations of difference. While I believe listeners will be able to distinguish between allophones associated with social information like speaker status, I hypothesize that these allophones will be discerned less successfully than those carrying information about word position and frequency.

To determine the validity of these hypotheses, two experiments were designed to measure evaluations of difference for three allophonic difference pairs, two phonemic contrast pairs, and seven identity pairs (identical stimuli), as outlined in Section 3. The identity pairs and phonemic contrast pairs serve as points of comparison for the allophonic difference pairs, which will help answer the first research question:

      1.   How successfully are allophonic differences perceived compared to phonemic contrasts and the same stimuli?

Based on previous research (Boomershine *et al*., 2008; Hume and Johnson, 2003), allophonic differences are expected to receive higher evaluations of difference than the same stimuli and lower evaluations of difference than phonemic contrasts. It is less certain, however, how multiple allophonic differences will be perceived. For this reason, three allophonic difference pairs were selected to investigate the effect of the linguistic and extralinguistic information conveyed by the allophones on listeners' evaluations. As discussed in more detail in Section 3, one allophonic difference pair ([n] ~ [ŋ]) conveys information about word position in Costa Rican Spanish, with [ŋ] generally appearing word-finally and [n] in other positions (Lipski, 1994:222). Another allophonic difference pair ([ð] ~ [d]) involves two allophones that, while both possible intervocalically, differ in frequency. More specifically, [ð] occurs nearly categorically between vowels, and [d] is produced only in hyperarticulated or emphatic speech. The final allophonic difference pair ([s] ~ [z]) includes allophones that can both appear word-initially, word-medially, and word-finally between vowels, which limits the linguistic information available about word position, unlike [n] ~ [ŋ]. Additionally, [z] occurs in over 36% of intervocalic /s/ tokens (Chappell and García, 2017), limiting any potential frequency effects that may occur with the [ð] ~ [d] pair. The inclusion of these three allophonic difference pairs allows for an exploration of the second research question:

2.   How does the type of allophonic difference affect perception? In other words, will allophones that convey linguistic information about word position (intervocalic [n] ~ [ŋ]), allophones that differ in frequency (intervocalic [ð] ~ [d]), and allophones that are less marked for linguistic information and frequency differences but do carry social information (intervocalic [s] ~ [z]) result in divergent evaluations of difference?

The following section introduces the methodology used to address these research questions, the results and discussion are provided in Section 4, and Section 5 proposes future directions and gives concluding remarks.

3. METHODOLOGY

3.1. Experiment design

The methodology employed in this study follows García (2015), Johnson and Babel (2010), and Boomershine *et al*. (2008), who demonstrate that listeners' allophonic perception can be measured in discrimination tasks and similarity ratings. This allows for a categorical determination as to whether listeners do or do not perceive a contrast between word pairs, on the one hand, and a gradient evaluation as to just how differently the phones are perceived.

In order to answer the research questions posed in Section 2, two broad tasks were created in SurveyGizmo (Vanek and McDaniel, 2006): a similarity rating task, in which participants listened to two words produced by a single speaker and ranked the likeness of those words on a scale ranging from 1 *muy similar* 'very similar' to 6 *muy diferente* 'very different,' and an AX discrimination task, in which participants again listened to two words produced by a single speaker and had to identify the word pair as *igual* 'the same' or *diferente* 'different.' Screenshots of what was presented to each listener in the similarity rating task and the AX discrimination task, respectively, are provided below in figures 1 and 2.

Before beginning the experiment, each participant was asked to first agree to the terms of participation and then fill out a brief survey about demographics and exposure to other languages and dialects of Spanish. Then the two evaluation tasks were presented, with the each audio file containing a word pair playing automatically for the listener.

Figure 1. *Screenshot of the similarity rating task.*

¿Cómo suenan? *

○ Igual

○ Diferente

[Atrás] [Siguiente]

Figure 2. *Screenshot of the AX discrimination task.*

3.2. Stimuli

Throughout the experiment, participants listened to one audio file at a time containing words pairs that were (i) identical (identity pairs), (ii) different only in one intervocalic allophone (allophonic difference) or (iii) different only in one intervocalic phoneme (phonemic contrast). Distractors were not used to limit the length of the experiment and decrease participant attrition. The identity pairs, allophonic differences, and phonemic contrasts included in the stimuli are illustrated below in table 1.

The allophone pairs presented in the experiments include [s] ~ [z], e.g. *asa* 's/he grills' as [asa] ~ [aza]; [n] ~ [ŋ], e.g. *con oro* 'with gold' as [kon oɾo] ~ [koŋ oɾo]; and [d] ~ [ð], e.g. *la daba* 's/he gives it (fem.)' as [la ðaβa] and [la daβa], discussed in more detail in Section 3.3. Each of these variants occurs in Costa Rican Spanish, with /s/ variably undergoing voicing between vowels (Lipski, 1994:223), frequent velarization of word-final, prevocalic /n/ (Lipski, 1994:222), and occlusive intervocalic [d] sometimes appearing where [ð] is typically produced (Lipski, 1994:222-224). I have also observed occlusive intervocalic [d] in hyperarticulated speech in the Costa Rican Central Valley when speakers wish to highlight or emphasize their point. Finally, the phonemic contrast between intervocalic /d/ and /ɾ/ exists in all varieties of Spanish.

Identity pairs	Allophonic difference[3]	Phonemic contrast
[s] ~ [s]	[s] ~ [z]	[ð] ~ [ɾ]
[z] ~ [z]	[n] ~ [ŋ]	[d] ~ [ɾ]
[d] ~ [d]	[d] ~ [ð]	
[ð] ~ [ð]		
[ɾ] ~ [ɾ]		
[ŋ] ~ [ŋ]		
[n] ~ [n]		

Table 1. *Identity pairs, allophonic differences, and phonemic contrasts included in the stimuli.*

As the discrimination of intervocalic /s/ is of particular interest in this paper, special attention was paid to the voicing of /s/ in the recorded stimuli. For all voiced tokens of intervocalic /s/ utilized in the perception tasks, the voiced [z] involved close to 100% voicing, while all [s] tokens had less than 15% voicing. The spectrograms below show an example of the [s] and [z] realizations presented to the listeners along with the other allophonic difference pairs: [d] ~ [ð] and [n] ~ [ŋ].

In the waveforms and spectrograms in figures 3-8, [z] is distinguishable by a periodic wave in the waveform and an associated voicing bar in the spectrogram, which is absent in the waveform and spectrogram corresponding to [s]. The duration of [z] is also shorter than [s], typical of the voiced fricative (Hualde and Prieto, 2014; Strycharczuk *et al.*, 2013). The production of [d] features a low amplitude periodic wave in the waveform along with a voicing bar, no formant structure, and a burst associated with the release of the stop in the spectrogram. On the other hand, [ð] is shorter in duration, and the spectrogram shows greater intensity with a formant structure maintained throughout the consonant. Finally, the spectrogram associated with the alveolar nasal features antiformants and relatively stable formants, while the second formant (F2) of the velar nasal rises (see Johnson, 2003; Ladefoged, 2003).

---

[3] The allophonic difference and phonemic contrast columns do not account for order or presentation, i.e. [s] ~ [z] can refer to a [s] ~ [z] or [z] ~ [s] presentation order.
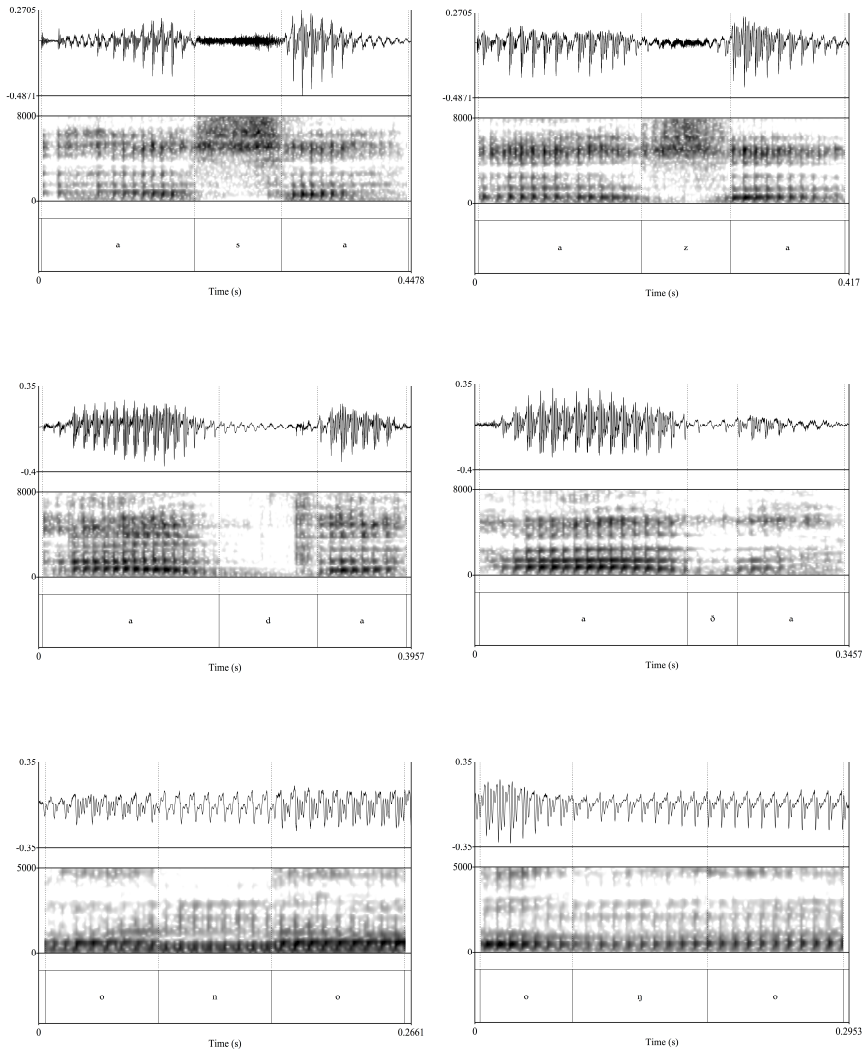
Figures 3-8. *Waveforms and spectrograms for allophonic difference pairs [s] ~ [z], [d] ~ [ð], and [n] ~ [ŋ].*

The stimuli were recorded in a quiet room using a Zoom H2 Handy Recorder. In each recording, the vowels' duration was controlled by trimming longer vowels' duration in Praat (Boersma and Weenink, 2013), and the amplitude of each recording was equated following Boomershine *et al.* (2008:6). The judgments of two professional linguists were employed to ensure that the trimmed and equated recordings sounded naturalistic. They also determined that the difference between allophonic pairs was clear and there were no noticeable differences in the recorded word pairs besides the target phones. The stimuli were also piloted with nine individuals, including three native Spanish speakers who were not linguists, three linguists who were native Spanish speakers, and three linguists who were not native Spanish speakers. Based on the pilot study and professional feedback, the tasks were determined to be easy to comprehend and complete for linguists and non-linguists alike.

3.3. Blocks

The [s] and [z] allophones of intervocalic /s/ may appear word-initially, word-medially, or word-finally in Costa Rican Spanish, and, accordingly, these three positions are represented in the stimuli. While word-final, prevocalic /s/ is postlexically resyllabified to word-initial position at the surface level, e.g. *las uvas* 'the grapes' becomes [la.su.βas], the underlying /s/ continues to be affected by word-final phonological processes (Hualde, 1989, 1991:485-492; Martínez-Gil, 1991:533-5344). The division of /s/ stimuli into word-initial, word-medial, and word-final blocks will establish if the underlying position in the word affects listeners' ability to discriminate between [s] and [z].[4] Rather than using nonce words, which listeners may not successfully process as one or two words as intended by the researcher, real words were used, including [la saka] and [la zaka] for *la saca* 'the sack,' [asa] and [aza] for *asa* 's/he grills,' and [las ata] and [laz ata] for *las ata* 's/he ties them (fem.)'. The use of real words ensures the parsing of distinct lexical items and the interpretation of /s/ in word-initial, word-medial, or word-final position as intended.[5]

---

[4] It should be noted that the same resyllabification takes place with word-final /n/, but intervocalic nasal velarization only takes place in word-final position in Costa Rica (Lipski, 1994:222), rendering the allophonic comparison conducted across word positions for [s] ~ [z] impossible for [n] ~ [ŋ].

[5] No ambiguity regarding the position of the /s/ is expected. While *la saca*, with word-initial /s/, is a lexical sequence in Spanish, *las aca* is unattested. Likewise, while *las ata*, with word-final /s/, is an existing word sequence, *la sata* is not.

Both the similarity rating task (Task 1) and the AX discrimination task (Task 2) were divided into three blocks, respectively. In each block, participants were given a brief training exercise to familiarize themselves with the task and the stimuli they would hear in the test. Following completion of the training exercise in each block, listeners then participated in the test phase. The similarity rating task was presented in Blocks 1-3, with /s/ appearing in a different word position in each block, and the AX discrimination task was presented in Blocks 4-6, again with a different /s/ word position in each block.[6] Following García (2015), the tasks were presented in a fixed order (Task 1, including Blocks 1, 2, and 3, followed by Task 2, including Blocks 4, 5, and 6) as were the stimuli within each block. Sample stimuli from each block are shown below, and for the full list of stimuli in each block, the reader is referred to the appendix.

In the first block of the similarity rating task and the AX discrimination task (Blocks 1 and 4), word-medial cases of intervocalic [s] and [z], e.g. [asa] and [aza] for *asa* 's/he grills', were presented alongside other pairs containing word-medial intervocalic [d], [ð], and [ɾ], i.e. [ada] and [aða] for *hada* 'fairy' and [aɾa] for *ara* 's/he plows.' Sibilant identity and difference pairs were presented, e.g. [asa] ~ [asa] and [asa] ~ [aza], as were other identity and difference pairs, e.g. [aða] ~ [aða] and [aða] ~ [aɾa], respectively. The second block of both tasks (Blocks 2 and 5) featured sibilant identity pairs in word-initial position, e.g. *la saca* 'the jacket' as [la saka] ~ [la saka], as well as difference pairs, e.g. [la zaka] ~ [la saka]. These word pairs were presented alongside other identity and difference pairs including [la ðaβa] ~ [la ðaβa] and [la daβa] ~ [la ðaβa] for *la daba* 's/he gave it (fem.)', respectively. Finally, the last block of each task (Blocks 3 and 6) involved sibilant identity and difference pairs in word-final position, e.g. *las ata* 's/he ties them (fem. pl.)' as [las ata] ~ [las ata] and [las ata] ~ [laz ata], and other identity and difference pairs in the same position, e.g. *con oro* as [koŋ oɾo] ~ [koŋ oɾo] and [kon oɾo] ~ [koŋ oɾo], respectively. Tables 2 and 3 show example stimuli from each block.

---

[6] [n] ~ [ŋ] pairs are only presented word finally because they are phonotactically restricted to that environment in Costa Rican Spanish (Lipski, 1994:222). [d] ~ [ð] pairs are included word initially and word medially but not in word-final position because multiple allophones can occur in this position, i.e. [d], [ð], [t], [θ], ø (Quilis, 1993:218-219), which would unnecessarily complicate the task. Finally, because phonemic contrasts are provided as a control against which allophonic difference pairs and identity pairs are compared, their inclusion is restricted to word-medial position to enable a comparison while limiting the length of the experiment.

Task 1	Block 1	Block 2	Block 3
Target phone word position	Medial	Initial	Final
Ex. sibilant identity pairs	[asa] ~ [asa]	[la zaka] ~ [la zaka]	[las ata] ~ [las ata]
Ex. sibilant allophonic difference pairs	[asa] ~ [aza]	[la saka] ~ [la zaka]	[laz ata] ~ [las ata]
Ex. other identity pairs	[aða] ~ [aða]	[la daβa] ~ [la daβa]	[koŋ oɾo] ~ [koŋ oɾo]
Ex. other allophonic difference pairs	[aða] ~ [ada]	[la ðaβa] ~ [la daβa]	[koŋ oɾo] ~ [kon oɾo]
Ex. phonemic contrast pairs	[aɾa] ~ [ada]	--	--
Evaluation options: Scale of 1 (very similar) to 6 (very different)			

Table 2. *Example stimuli from the similarity rating task by block.*

Task 2	Block 4	Block 5	Block 6
Target phone word position	Medial	Initial	Final
Ex. sibilant identity pairs	[asa] ~ [asa]	[la zaka] ~ [la zaka]	[las ata] ~ [las ata]
Ex. sibilant allophonic difference pairs	[asa] ~ [aza]	[la saka] ~ [la zaka]	[laz ata] ~ [las ata]
Ex. other identity pairs	[aða] ~ [aða]	[la daβa] ~ [la daβa]	[koŋ oɾo] ~ [koŋ oɾo]
Ex. other allophonic difference pairs	[aða] ~ [ada]	[la ðaβa] ~ [la daβa]	[koŋ oɾo] ~ [kon oɾo]
Ex. phonemic contrast pairs	[aɾa] ~ [ada]	--	--
Evaluation options: *igual* 'the same' or *diferente* 'different'			

Table 3. *Example stimuli from the AX discrimination task.*

Each listener heard 31 word pairs in Task 1 (the similarity rating task) and 34 word pairs in Task 2 (the discrimination task), for a total of 65 word pairs per participant. In total, 3,286 evaluations were collected in Task 1 and 3,604 evaluations were collected in Task 2. Then, a statistical model was constructed to explore the effect of block (e.g. Training 1, Test 1, Training 2, Test 2, etc.) on listener evaluations. A significant difference was found for evaluations in Training 1 as participants adjusted to the task, but no differences were found for listener evaluations in the other training and test blocks for Task 1 or Task 2. For this reason, the results of Training 1 in the similarity rating task (see the appendix) were excluded from analysis, while the other training and test blocks were combined for analysis in Task 1 and Task 2, respectively. This resulted in 2,862 tokens for analysis in Task 1 and 3,602 in Task 2.

3.4. Speaker

The speaker used to produce these words was a phonetically trained native Spanish speaker from Loja, Ecuador.[7] Although Lojano Spanish represents a dialect other than Costa Rican Spanish, this distinction was not expected to influence listeners' evaluations of the words produced in isolation, as no vowels or consonants in the target stimuli diverged from Costa Rican pronunciations. The intervocalic phonemic contrast produced by the Lojano speaker exists in Lojano and Costa Rican Spanish (/d/ ~ /ɾ/) as do the allophonic variants, including word-final, prevocalic /n/ velarization; intervocalic /s/ voicing; and variable realization of /d/ as [ð] or [d] between vowels (Lipski, 1994:221-224, 247-249). The speaker was trained to produce each variant clearly, using the same speech rate and intonational contour for all recorded stimuli. Multiple productions of the stimuli were recorded, and the clearest recordings were selected for manipulation (see Section 3.2) and inclusion in the perception task.

To determine whether the speaker was judged as local or foreign, an open-ended question was asked of all participants, reading, "*¿Le parece que el hablante que escuchó podría ser un miembro de su comunidad? ¿Por qué sí o por qué no?*" 'Do you think the speaker you heard could be a member of your community? Why or why not?'. This question was also asked of García's (2015) Lojano participants, enabling a comparison between their evaluations of how Lojano the speaker sounded and my participants' evaluations of how Costa Rican the speaker

---

[7] Christina García is responsible for the recording of this Lojano speaker, and I am grateful to her for allowing me to use her stimuli in my experiment.

sounded. The evaluations of the speaker's perceived localness are clearly presented in table 4.

	Lojano evaluations	Costa Rican evaluations
Definitely/probably local	48% (10)	53% (56)
Not sure	14% (3)	3% (3)
Definitely/probably not local	24% (5)	40% (43)
No answer	14% (3)	4% (4)
Total # participants	*n* = 21	*n* = 106

Table 4. *Participants' evaluations of the speaker's localness to their community*.

As predicted, the speaker was evaluated as sounding definitely or probably Costa Rican by most of my participants. He was evaluated as definitely Costa Rican-like by 49% (52/106) of participants, and 4% (4/106) noted he was probably Costa Rican. These results are similar to García (2015), who found that 48% (10/21) of the Ecuadorian listeners thought this speaker, who is in fact from Loja, sounded like he was from Loja, suggesting that the Costa Rican listeners accepted the speaker's membership in their community as much as Lojano listeners did. In other words, while the speaker used in this task was not Costa Rican, he was believed to be a local, native speaker by approximately the same proportion of participants who believed him to be a local, native speaker of his own variety of Spanish (García, 2015). Consequently, his nationality is not expected to influence the Costa Rican participants' phonetic discrimination.

3.5. Participants

A total of 106 Costa Rican participants were recruited for this study, and almost all of these listeners were born and raised in the Central Valley where intervocalic /s/ voicing is commonly observed. As all participants in the experiment were confirmed to be Costa Ricans living in the Central Valley before their participation, the 12 respondents who failed to specify their regional province were not excluded from analysis. The experiment was conducted in the summer of 2015, and participants were recruited in two ways. Most participants were students recruited in person at La Universidad Latina, a private university in Heredia, and others were recruited through emails and social media messages to Costa Rican colleagues, acquaintances, and friends. Table 5 below shows their basic demographic information, collected at the beginning of the experiment.

	Participants' demographics
Median/mean age (years)	19/30
Age range (years)	18-60
Male : Female (*n*)	69 : 37
Regional province (*n*)	Alajuela (22), Cartago (4), Heredia (36), San José (32), Other or unspecified (12)
Total # participants	106

Table 5. *Demographic information of the participants.*

An anonymous reviewer points out that listeners were not subjected to a pure-tone hearing screening, and decreased sensitivity could affect participants' ability to discriminate between allophones, especially for participants above the age of 50. However, in the present study only two participants are above the age of 50: one female participant is 51 years old and another female participant is 60. To determine if their responses affected the results of this study, they were removed from the data and the statistical tests presented in Section 3.6 were run again. Because the results did not change when these two older listeners were removed, their responses are kept in the data discussed in the following sections.

3.6. Statistical analysis

In my analysis of the similarity rating task (Task 1) in Blocks 1-3, mixed effects linear regression models were created with participant as a random effect and participants' evaluations as the continuous dependent variable, which were subjected to a square root transformation to normalize the data distribution. The effect of block (or word position), contrast type (identity pair, allophonic difference, phonemic contrast), specific phones heard (e.g. [s] ~ [z]), and presentation order were tested in the models as independent variables. As each regression model was constructed, independent variables were added one at a time and a comparison of nested models using the analysis of variance function (R Core Team, 2014) established which predictors significantly improved the model. Interactions between the main effects were tested in the model construction, and only the best-fit models are presented in Section 4. In addition to the best-fit mixed effects linear regression models, I follow García (2015) by performing a Welch Two Sample t-test when considering identity pairs, e.g. [s] ~ [s] and [z] ~ [z], and difference pairs, e.g. [s] ~ [z] and [z] ~ [s] (see Section 4.2.1). Because the data showed inflated values at 1 (very similar), I also conducted permutation tests on

each block and across blocks. The permutation tests yielded the same results as the regression models and t-tests, and, as a result, only the results of the regression models and t-tests are presented in the discussion that follows.

In Task 2, which involved the AX discrimination task, binomial logistic regression models with random effects were fitted across Blocks 4-6, with the participants' evaluation of the pairs as *igual* 'the same' or *diferente* 'different' as the dependent variable and participant as a random effect. The same independent variables and interactions were tested here, and the same stepwise procedure described above was utilized in these models. Again, only the best-fit models are presented in Section 4. Following García (2015), I also performed a two-tailed Pearson's chi-squared test with Yates' continuity correction when considering identity pairs and difference pairs. This test determines whether participants ranked difference pairs as *diferente* 'different' significantly more than they ranked identity pairs as different.
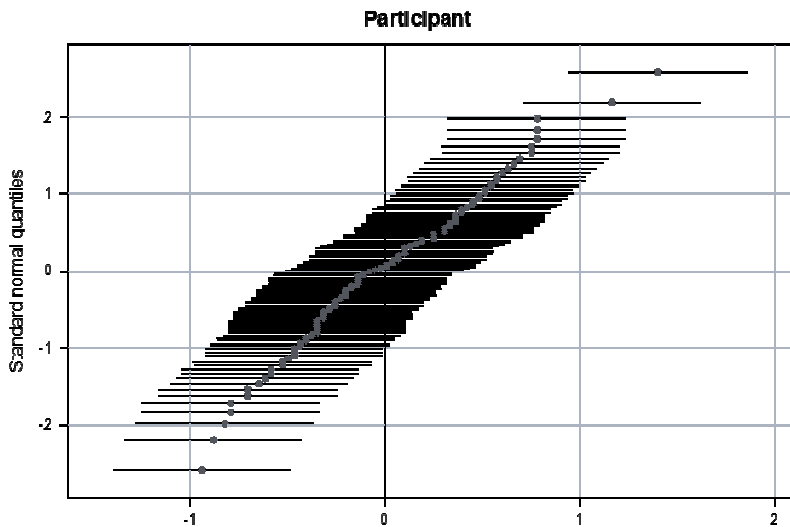

Figure 9. *Caterpillar plot of participant effect in the similarity rating task (Task 1).*
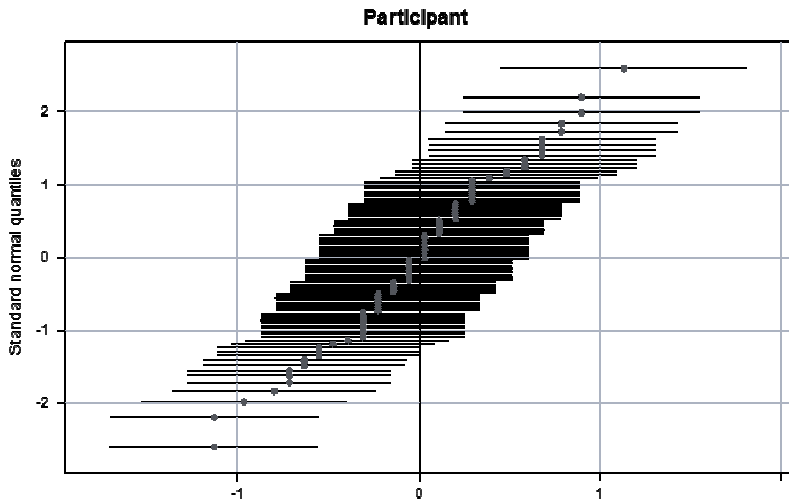
Figure 10. *Caterpillar plot of participant effect in the AX discrimination task (Task 2).*

Gender and age were not found to be significant predictors of listeners' evaluations in the models fitted to Task 1 or Task 2 and, as a result, they are not discussed in the following sections. As expected, evaluations do differ by individual participant, and the overall variability between individuals in Task 1 and Task 2 is captured in the caterpillar plots in figures 9 and 10. These plots include the conditional modes of the random effects' distribution and their corresponding 95% prediction intervals for the random effects across participants (Bates, 2010).

The difference between listeners is unsurprising, as it has long been held that

> *individual participants may have idiosyncratic sensitivities to any experimental manipulation that may have an overall effect, so detecting a 'fixed effect' of some manipulation must be done under the assumption of corresponding participant random effects for that manipulation as well* (Barr *et al*., 2013:2).

For this reason, participant was included as a random effect in all models created to ensure the fixed effects rise above the noise created by individual responses.

4. RESULTS AND DISCUSSION

4.1. Evaluations of phonemic contrast pairs, allophonic difference pairs, and identity pairs

The present section explores the first research question, which asked, "How successfully are allophonic differences perceived compared to phonemic contrasts and the same stimuli?" First, table 6 presents the model fitted to evaluations of difference in the similarity rating task. The independent variable is contrast type (identity pairs, allophonic difference pairs, and phonemic difference pairs), and participant is included as a random effect. Allophonic difference pairs are the reference level, and a negative estimate indicates that the given level received lower evaluations of difference than the reference level, while a positive estimate shows higher evaluations of difference for the given level compared to the reference level. The boxplot in figure 11 clarifies these different evaluations along the six-point similarity-difference scale used. Both table 6 and figure 11 show that phonemic differences are evaluated as very distinct compared to the two other pairs, and allophonic pairs are evaluated as more distinct than identity pairs.

Replicating the data exploration conducted above with the categorical responses provided in the AX discrimination task, table 7 provides the best-fit binomial mixed effects model fitted to evaluations of *diferente* 'different' or *igual* 'the same', including contrast type as an independent variable and participant as a random effect. Allophonic difference pairs serve as the reference level, and a negative estimate indicates lower evaluations of difference, and a positive estimate shows higher evaluations of difference. These evaluations of sameness and difference are graphically visualized in figure 12, demonstrating again that allophone pairs are viewed as more distinct than identity pairs, and phonemes are rated as significantly more different than allophone pairs.

	Estimate	*SE*	*t*-Value	*p*-Value
(Intercept)	2.39353	0.06025	39.72	<0.001
Reference level: Pair = Allophones				
Pair = Identity	-0.48272	0.04558	-10.59	<0.001
Pair = Phonemes	2.20081	0.08551	25.74	<0.001

Table 6. *Best-fit binomial mixed effects model for allophonic differences, phonemic differences, and identity pairs in the similarity rating task (N = 2,862).*
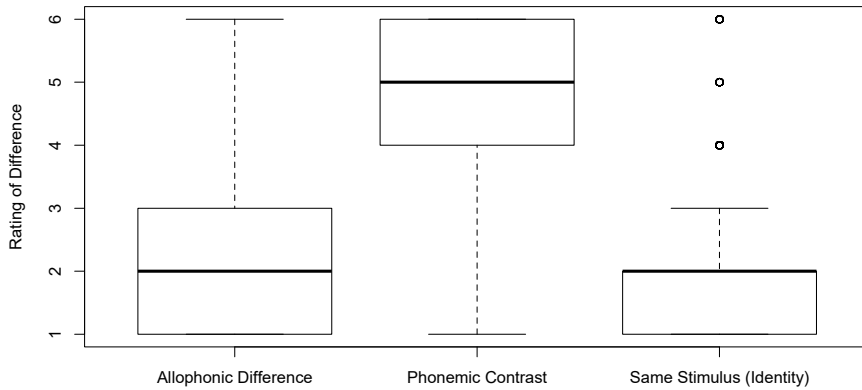
Figure 11. *Evaluations of difference based on the type of contrast in the similarity ratings task.*

	Estimate	*SE*	*t*-Value	*p*-Value
(Intercept)	0.005	0.1	0.055	0.956
Reference level: Pair = Allophones				
Pair = Identity	-1.9	0.09	-21.059	<0.001
Pair = Phonemes	3.64	0.32	11.209	<0.001

Table 7. *Best-fit binomial mixed effects model for allophonic differences, phonemic differences, and identity pairs in the discrimination task (N = 3,604).*

The results presented in this section, consistent in both the similarity rating task and the AX discrimination task, allow us to answer the first research question. This study finds that Costa Rican listeners rate phonemic contrasts as significantly more different than allophone pairs, and they rate allophone pairs as significantly more different than identity pairs. That is, phonemic contrasts in one's native language are perceived more clearly than non-contrastive variants, supporting previous research (Best *et al*., 1998; Boomershine *et al*., 2008; Dupoux *et al*., 1997; Fox *et al*., 1995; Francis and Nusbaum, 2002; Goto, 1971; MacKain *et al*., 1981; Polka and Werker, 1994; Pruitt, 1995; Strange, 1995; Trubetzkoy, 1939).
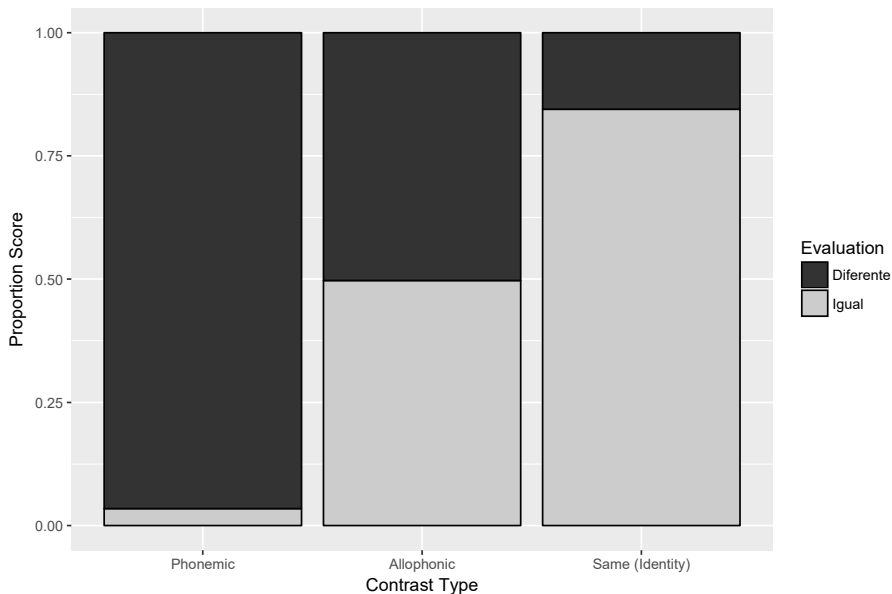
Figure 12. *Evaluations of sameness and difference in the discrimination task based on contrast type.*

4.2. Evaluations of allophonic difference pairs

The results presented in Section 4.1 support previous work on the perception of phonemic contrast and allophony, but little is known about how different noncontrastive phones are evaluated. Section 4.2 seeks to answer research question #2, which asked, "How does the type of allophonic difference affect perception?" Before exploring the evaluations for all allophonic differences, I first focus on intervocalic /s/ evaluations in Section 4.2.1.

### 4.2.1. Evaluations of intervocalic /s/ variants

As noted in Section 1, Costa Rican Spanish speakers struggle to produce or explicitly comment on intervocalic [z], which could suggest that listeners are not able to discriminate between [s] and [z] very successfully. Before exploring all allophonic difference pairs, this section seeks to establish whether Costa Rican listeners can discriminate between /s/ allophones by comparing difference pairs ([s] ~ [z]) to identity pairs consisting of [s] ~ [s] or [z] ~ [z]. To answer this

question, a Welch Two Sample t-test was conducted to find the mean and standard deviation for identity pairs and difference pairs, shown in table 8.

	Difference Pair Evaluations (N = 638)	Identity Pair Evaluations (N = 746)
Mean	2.17	1.87
Standard Deviation	1.27	1.123

Table 8. *Welch Two Sample t-test results for evaluations of /s/ allophones in the similarity ratings task (t = 4.569, df = 279.7, p <0.001).*

The Welch Two Sample t-test demonstrates that the [s] ~ [z] difference pairs are evaluated as significantly more distinct than identity pairs. The juxtaposed histograms in figures 13-14 illustrate the distribution of evaluations for difference pairs and identity pairs, with identity pairs receiving higher evaluations of sameness (1) than difference pairs.
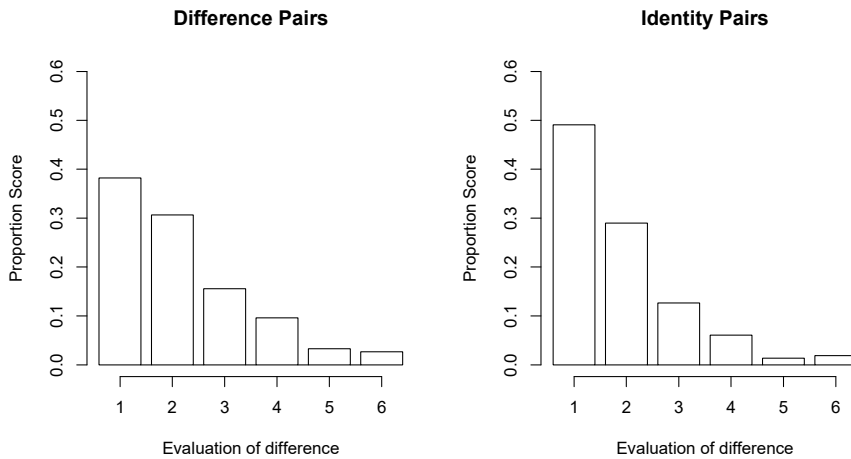
Figures 13-14. *Histograms showing the proportional distribution of difference and identity /s/ pair evaluations in the similarity rating task (N = 1,378).*

Figure 15 presents a boxplot with evaluations of identity pairs ([s] ~ [s] and [z] ~ [z]) and difference pairs ([s] ~ [z] and [z] ~ [s]). In this boxplot and those that

follow, the boxes depict the middle 50% of the evaluations, with the black horizontal line indicative of the median evaluation. If visible, the line extending above the boxes illustrates the top 25% of the evaluations and the line extending below represents the lowest 25% of the evaluations. If no line extends above or below the box, this means that the upper or lower quartile is at the ceiling. In other words, at least 25% of the evaluations are at 6 or 1, respectively. The dots above or below these vertical lines represent outliers. In figure 15, difference pairs clearly receive higher evaluations of difference than identity pairs.

To confirm these results, a linear mixed effects model was also fitted to evaluations of identity /s/ pairs ([s] ~ [s] and [z] ~ [z]) and difference pairs ([s] ~ [z] and [z] ~ [s]) in the similarity rating task with pair type as the independent variable and participant as a random effect, shown in table 9. As before, a negative estimate indicates that the given level received a lower rating of difference than the reference level.
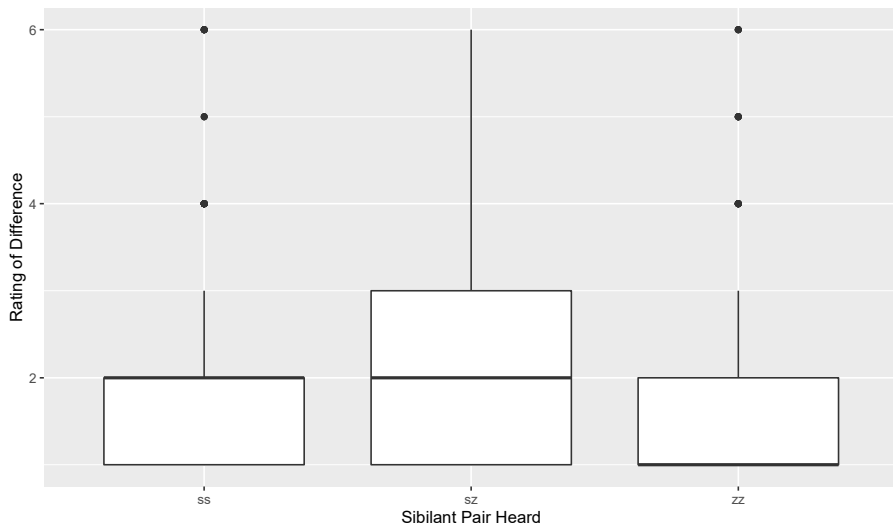

Figure 15. *Evaluations of difference for intervocalic /s/ identity and difference pairs (N = 1,378).*

	Estimate	*SE*	*t*-Value	*p*-Value
(Intercept)	2.17	0.05	45.784	< 0.001
Reference level: Pair = Difference				
Pair = Identity	-0.3	0.06	-4.612	< 0.001

Table 9. *Best-fit mixed effects model for similarity ratings of difference and identity /s/ pairs (N = 1,378).*

Similar results emerged in the AX discrimination task; [s] ~ [z] difference pairs were evaluated as different significantly more than [s] ~ [s] and [z] ~ [z] identity pairs. Table 10 provides the results of a two-tailed Pearson's chi-squared test with Yates' continuity correction, showing a significant difference between identity and difference pairs, and figure 16 visually depicts these evaluations of sameness and difference.

The best-fit binomial mixed effects model presented in table 11, including pair type as an independent variable and participant as a random effect, again shows this significant difference, with a negative estimate indicating lower ratings of difference.

	Difference pair evaluations (N = 636)	Identity pair evaluations (N = 636)
Different	42% (N = 267)	15% (N = 98)
Same	58% (N = 369)	85% (N = 538)

Table 10. *Results of the Pearson chi-squared test for Blocks 4-6 involving AX discrimination of /s/ allophones ($\chi^2 = 108.44$, df = 1, p <0.001).*

	Estimate	*SE*	*t*-Value	*p*-Value
(Intercept)	-0.41	0.14	-2.986	<0.01
Reference level: Pair = Difference				
Pair = Identity	-1.63	0.15	-10.78	< 0.001

Table 11. *Best-fit binomial mixed effects model for difference and identity /s/ pairs in the AX discrimination task in Blocks 4-6 (N = 1,272).*
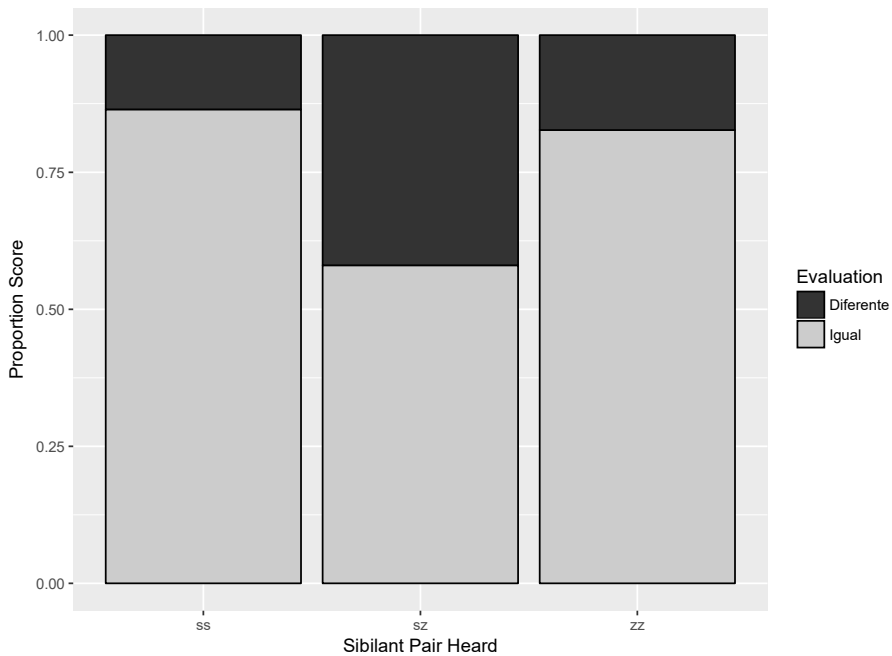
Figure 16. *Evaluations of sameness and difference in the AX discrimination task.*

In spite of the fact that Costa Rican Spanish speakers struggle to produce or comment on intervocalic /s/ variants (Chappell, 2016), Section 4.2.1 has proven that they do discriminate between intervocalic [s] and [z] difference pairs when compared to identity pairs. However, it remains to be seen how differently [s] and [z] are perceived when compared to other allophonic difference pairs. This issue will be investigated in Section 4.2.2.

4.2.2. Evaluations of all allophonic difference pairs compared to phonemic contrasts and identity pairs

To determine how [s] ~ [z] evaluations compared to all other pairs heard in the similarity rating task, table 12 provides the best-fit mixed effects model fitted to evaluations of difference, with the specific pair heard as an independent variable and participant as a random effect. The table again confirms that the difference pair [s] ~ [z] was perceived as significantly more distinct than the identity pairs [s] ~ [s] and [z] ~ [z], but the [s] ~ [z] allophonic difference was perceived as

significantly less distinct than the other allophonic differences; [d] ~ [ð] and [n] ~ [ŋ] were evaluated as significantly more different than [s] ~ [z] pairs. These findings were additionally supported by a Tukey's HSD test. The evaluations for phonemic contrast pairs, allophonic difference pairs, and the same stimuli are visualized in the boxplot in figure 17.

		Estimate	*SE*	*t*-Value	*p*-Value
	(Intercept) Reference level: Pair = [s] ~ [z]	0.78	0.03	27.591	<0.001
Same pairs	Pair = [ð] ~ [ð]	-0.27	0.07	-3.876	<0.001
	Pair = [z] ~ [z]	-0.17	0.04	-4.329	<0.001
	Pair = [s] ~ [s]	-0.13	0.04	-3.71	<0.001
	Pair = [ŋ] ~ [ŋ]	-0.06	0.04	-1.494	0.135
	Pair = [d] ~ [d]	-0.001	0.05	-0.025	0.979
Allophonic difference	Pair = [d] ~ [ð]	0.12	0.029	4.237	<0.001
	Pair = [n] ~ [ŋ]	0.22	0.03	7.0	<0.001
Phonemic contrast	Pair = [ɾ] ~ [ð]	0.61	0.03	17.807	<0.001
	Pair = [d] ~ [ɾ]	0.8	0.03	26.25	<0.001

Table 12. *Best-fit mixed effects model for similarity ratings of all pairs (N = 2,862).*

Then, to establish how [s] ~ [z] evaluations compared to all other pairs heard in the AX discrimination task, a binomial mixed effects model (shown in table 13) was fitted to the data with evaluation (*diferente* 'different' or *igual* 'same') as the dependent variable, pair heard as the independent variable, and participant as a random effect. As before, interactions among individual blocks and pairs were explored and eliminated as they did not reach significance, and the best-fit model with [s] ~ [z] as the reference level is presented below, the results of which were also confirmed by a Tukey's HSD test. Again, a negative estimates signals a lower rating of difference than [s] ~ [z], and a positive estimate should be interpreted as a greater rating of difference than [s] ~ [z]. Figure 18 graphically depicts these evaluations of sameness and difference.
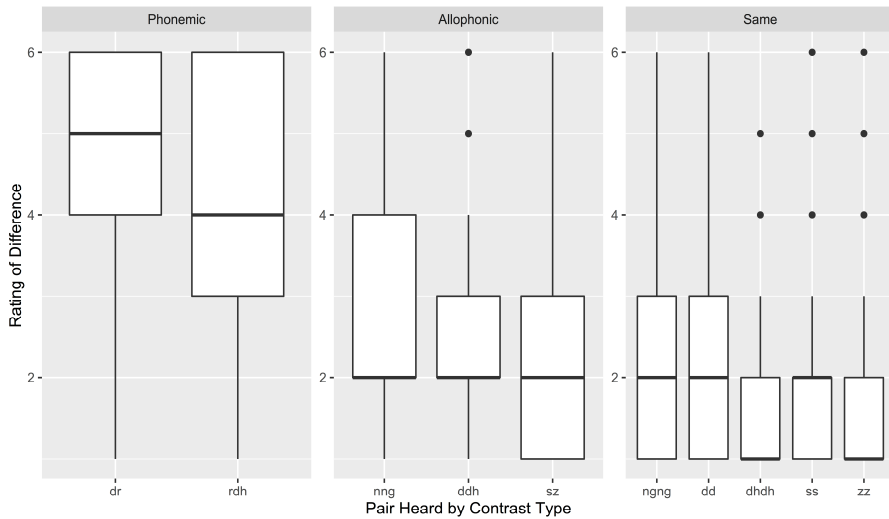
Figure 17. *Listener evaluations of all pairs heard in the similarity rating task.[8]*

		Estimate	*SE*	*t*-Value	*p*-Value
	(Intercept) Reference level: Pair = [s] ~ [z]	-0.3778	0.118	-3.199	<0.01
Same pair	Pair = [s] ~ [s]	-1.7	0.19	-8.92	<0.001
	Pair = [n] ~ [n]	-1.69	0.22	-7.52	<0.001
	Pair = [ð] ~ [ð]	-1.62	0.19	-8.66	<0.001
	Pair = [ŋ] ~ [ŋ]	-1.53	0.22	-7.09	<0.001
	Pair = [z] ~ [z]	-1.39	0.18	-7.833	<0.001
	Pair = [d] ~ [d]	-1.38	0.16	-8.644	< 0.001
Allophonic difference	Pair = [n] ~ [ŋ]	0.41	0.15	2.804	<0.01
	Pair = [d] ~ [ð]	0.83	0.13	6.434	<0.001
Phonemic contrast	Pair = [ɾ] ~ [ð]	3.7	0.35	10.566	<0.001
	Pair = [d] ~ [ɾ]	5.39	1.03	5.253	<0.001

Table 13. *Best-fit mixed effects model fitted to for AX discrimination of all pairs (N = 3,604).*

---

[8] Because some IPA symbols do not appear correctly in R, <dh> is representative of [ð] and <ng> indicates [ŋ].
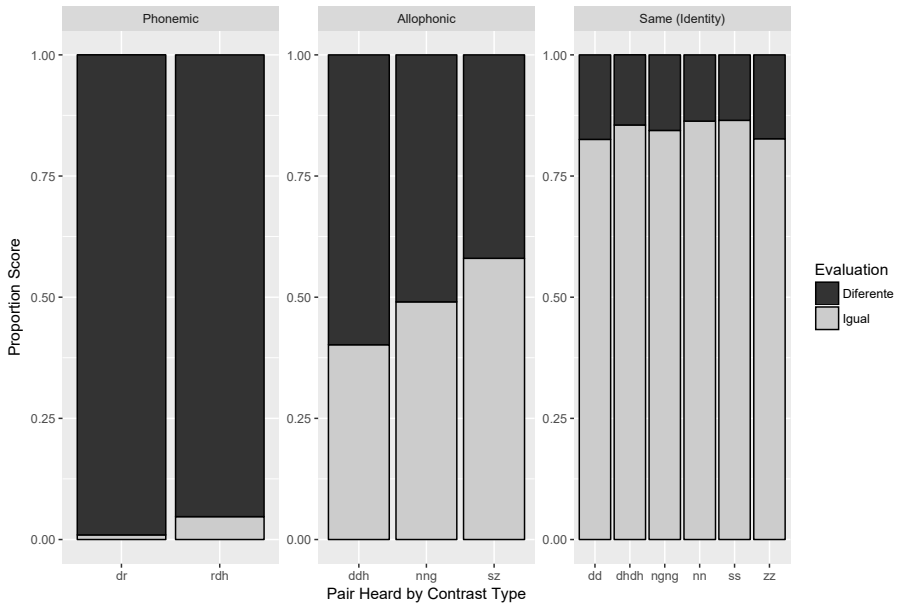
Figure 18. *Evaluations of sameness and difference in the AX discrimination task for all pairs divided by contrast type.[9]*

A similar pattern emerged in table 13 and figure 18 above: the [s] ~ [z] allophones are rated as significantly less different than the other allophonic difference pairs. To simplify the analysis and improve visibility, the evaluations for the three allophonic difference pairs are presented alone in figures 19 and 20.

---

[9] As was the case in figure 17, <dh> is again used for [ð] as is <ng> for [ŋ] in figure 18.
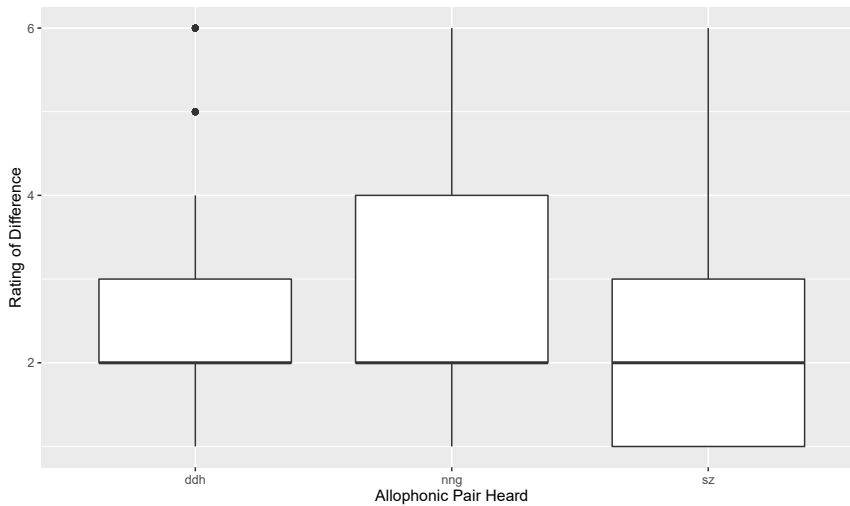
Figure 19. *Perceptual discrimination of the allophonic pairs heard in the similarity rating task.*

The information presented in this section now allows us to address the second research question in its entirety. The question posed was the following: "How does the type of allophonic difference affect perception? In other words, will allophones that convey linguistic information about word position (intervocalic [n] ~ [ŋ]), allophones that differ in frequency (intervocalic [ð] ~ [d]), and allophones that are less marked for linguistic information and frequency differences (intervocalic [s] ~ [z]) result in divergent evaluations of difference?". Interestingly, in both tasks the [s] ~ [z] distinction is heard less successfully than the [n] ~ [ŋ] or the [ð] ~ [d] pairs and, as shown in tables 12 and 13, this difference is statistically significant ($p < 0.01$). In other words, listeners' evaluations of these pairs show that allophonic differences are not categorically perceived in the same way. Rather, these results suggest that listeners attend to allophones differently, but it remains unclear why certain allophones would be more readily perceived than others. Here I propose that listeners utilize linguistic information, i.e. word position, and extralinguistic information, i.e. frequency, to process and distinguish between allophones in Costa Rican Spanish.
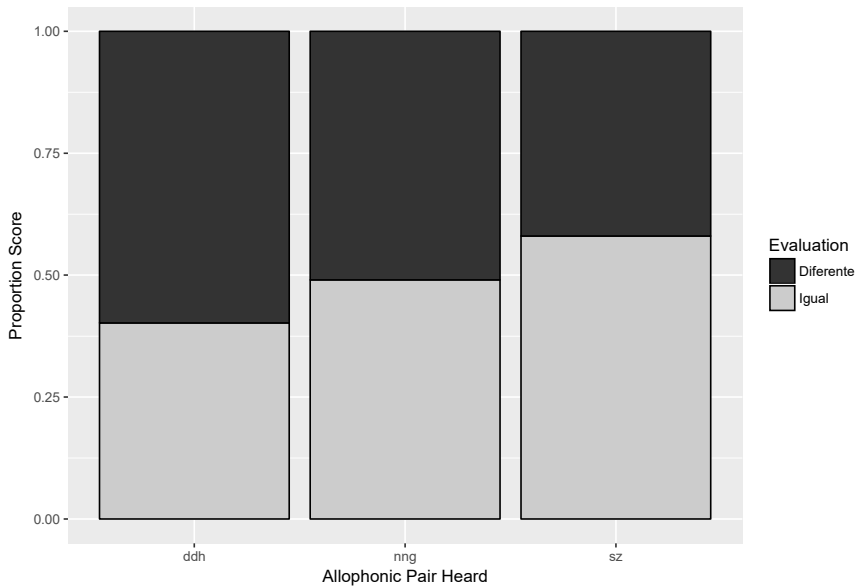
Figure 20. *Evaluations of sameness and difference for the allophonic pairs in the AX discrimination task.*

To explore this proposal, let us consider what linguistic and extralinguistic information is available in each allophone pair. First, the intervocalic alveolar or velar realization of the [n] ~ [ŋ] allophone pair carries linguistic information about word position and provides contextual cues to the listener. The intervocalic velar variant, [ŋ], only occurs in word-final position (Lipski, 1994:222), e.g. *si Nacho* [si nat͡ʃo] 'if Nacho' vs. *sin hacho* [siŋ at͡ʃo] 'without a beacon', which serves to mark the boundary between words. Previous studies have found that listeners attend to linguistic divisions and quickly learn to identify boundaries even in nonce words (Saffran *et al*., 1996). If listeners can identify boundaries in nonce words, it is not surprising that listeners also attend to real word boundaries in their native language, and the transitional and internal acoustic cues available with [n] and [ŋ] facilitate this identification. Consequently, I contend that because [n] and [ŋ] carry linguistic information and can help listeners identify word boundaries, the allophones are perceived as more distinct than some other allophonic difference pairs.

Second, surprisal or predictability may also play an important role in the discrimination of difference. While both [ð] and [d] are allophones of /d/, [ð] is by

far the most common intervocalic realization. Intervocalic [d] generally only appears in hyperarticulated speech, e.g. *de todo* as [de todo] 'all sorts/everything', and such an unexpected realization may be perceived as more distinct because of its lower frequency. This interpretation is supported by McQueen and Pitt's (1996) argument that listeners are sensitive to the probability that a sound will occur in certain environments. In addition to the fact that unexpected phones delay listeners' responses, as has been found in previous studies (Flagg *et al.*, 2006; Fowler and Brown, 2000), I propose that the appearance of less predictable phones also enhances the salience of allophonic differences, making listeners more aware of the subphonemic distinction between [ð] and [d]. That is, violating perceptual expectations about the distribution of [ð] and [d] seems to actually enhance listeners' awareness of allophony.

Finally, let us consider the [s] and [z] allophone pair, which is perceived as significantly less distinct than [n] ~ [ŋ] and [ð] ~ [d]. [s] and [z] both occur frequently between vowels in Costa Rican Spanish, with [z][10] appearing in 36.1% (N = 595/1,647) of all intervocalic /s/ contexts (Chappell and García, 2017). This makes [z] an expected variant in intervocalic position, unlikely to surprise listeners due to its low frequency like intervocalic [d]. Additionally, its production carries no linguistic meaning about word position, as it can occur word-initially, word-medially, and word-finally (Chappell and García, 2017). However, as Chappell (2016) shows, Costa Rican listeners are still adeptly able to evaluate a speaker's social qualities based on allophonic variant, regardless of its lower salience, with [z] evoking lower evaluations of status for all speakers and higher evaluations of niceness, localness, confidence, and masculinity for male speakers. That is, even though listeners do not overtly discriminate between [s] and [z] as successfully as other allophonic pairs, the allophones are still perceived and used to make social distinctions.

Unlike Clopper and Pisoni (2004), who find that naïve listeners tend to categorize unfamiliar dialects using a limited set of acoustic-phonetic cues, the findings presented here suggest that local listeners rely on even the least salient linguistic variables when identifying social information about local speakers. Even though [s] ~ [z] is the least perceptible allophonic difference pair of the three explored in

---

[10] [z] was measured as a continuous rather than a categorical variable in Chappell and García (2017) based on measurements of % voicing in both an interview and reading task. In this case, [z] is defined as a realization with more than 90% voicing following Campos-Astorkiza's (2014) tertiary distinction between unvoiced (0-20%), partially voiced (20-90%), and fully voiced (90-100%) [z].

this study, the variants still evoke social meanings for local listeners, which demonstrates that social properties can be indexed by the least discernable phonetic distinctions.

5. CONCLUSION AND FUTURE DIRECTIONS

This work supports the argument that phonemic contrasts are perceived more successfully than allophonic differences in one's native language (Boomershine *et al.*, 2008), affirming that phonemic status does condition discrimination, as suggested by Trubetzkoy (1939). However, prior to the present study little was known about the discrimination of different allophone pairs, and I have shown that not all allophonic differences are perceived with the same degree of success. More specifically, [s] ~ [z] allophonic difference pairs were evaluated as significantly less distinct than [d] ~ [ð] and [n] ~ [ŋ] allophonic difference pairs. To explain these differential evaluations, I have proposed that non-contrastive phones like [s] ~ [z] that index social information about status, masculinity, localness, confidence, and niceness (Chappell, 2016) are less readily discernible than other variants that convey linguistic information about word position ([n] ~ [ŋ]) or violate frequency expectations ([d] ~ [ð]) in Costa Rican Spanish.

This work has provided a point of departure for future work on allophonic perception, but it should be noted that it is not without its limitations. While I have argued that listeners are clued into the linguistic information conveyed by Costa Rican variants, I have not explicitly proven that linguistic and frequency-based information are responsible for the higher evaluations of difference for [n] ~ [ŋ] and [d] ~ [ð] than [s] ~ [z], and future work is needed to strengthen this proposal. First, other non-contrastive phones in Costa Rican Spanish should be explored to determine how much salience, surprisal, and phonotactic restrictions condition perception. For example, future work could investigate the perception of more salient allophonic differences in Costa Rican Spanish like the retroflex [ɻ] and trilled [r] that appear between vowels, e.g. *carro* 'car' as [kaɻo] or [karo]. The retroflex rhotic is a well-known stereotype of Costa Rican speech, and its salient status as a shibboleth may alter listeners' ability to distinguish between nonphonemic variants.

Additionally, future studies should explore the phonetic discrimination of [s] and [z] in other varieties of Spanish that voice intervocalic /s/, e.g. Loja, Ecuador (García, 2015), Madrid (Hualde and Prieto, 2014; Torreira and Ernestus, 2012),

Buenos Aires (Rohena-Madrazo, 2011), and Mexico (Schmidt and Willis, 2011), among others. Although perception is likely to be similar in other phonetically gradient /s/-voicing dialects like Loja or Madrid, /s/ voicing is more phonemic in nature in Quito and Cuenca. In Quito and Cuenca, word-final, intervocalic /s/ is almost categorically voiced (Chappell, 2011; Robinson, 1979). For example, *los usos* 'the uses' becomes [loz usos], with word-final /s/ undergoing a phonological voicing process while any /s/ in other word positions goes unaffected. As the /s/ voicing in these varieties carries linguistic information, distinguishing *has ido* 'you have gone' from *ha sido* 'it has been', I hypothesize that intervocalic [z] will be more salient and more successfully distinguished from [s].

In sum, this article is one of the first works to investigate the discrimination of different allophones, and a great deal of work remains to be done on the subject. Nonetheless, this paper has contributed to the field of phonetics by expanding our understanding of allophonic discrimination in Costa Rican Spanish. In addition to reaffirming the greater perceptibility of phonemes than allophones, I conclude that (i) subphonemic variants *are* perceived differently and (ii) even the least salient phonetic differences can encode local social meaning and contribute to listeners' evaluations of speakers' social qualities.

6. REFERENCES

BARNES, S. (2015): «Perceptual salience and social categorization of contact features in Asturian Spanish», *Studies in Hispanic and Lusophone Linguistics*, 8(2), pp. 213-241.

BARR, D. J., R. LEVY, C. SCHEEPERS and H. J. TILY (2013): «Random effects structure for confirmatory hypothesis testing: Keep it maximal», *Journal of Memory and Language*, 68(3), pp. 255-278.

BATES, D. M. (2010): *lme4: Mixed-Effects Modeling with R*, New York, Springer.

BENNETT, D. C. (1968): «Spectral form and duration as cues in the recognition of English and German vowels», *Language and Speech*, 11, pp. 65-85.

BEST, C. T. (1994a): «Learning to perceive the sound pattern of English», in C. Rovee-Collier and L. P. Lipsitt (eds.): *Advances in Infancy Research IX,* Hillsdale, Ablex, pp. 217-304.

BEST, C. T. (1994b): «The emergence of native-language phonological influences in infants: A perceptual assimilation model», in H. C. Nusbaum (ed.): *The Development of Speech Perception: The Transition from Speech Sounds to Spoken Words*, Cambridge, MIT Press.

BEST, C. T. (1995): «A direct realist perspective on cross-language speech perception», in W. Strange (ed.): *Speech Perception and Linguistic Experience: Theoretical and Methodological Issues in Cross-language Speech Research*, Timonium, York Press, pp. 171-204.

BEST, C. T. and W. STRANGE (1992): «Effects of phonological and phonetic factors on cross-language perception on approximants», *Journal of Phonetics*, 20, pp. 305–330.

BOERSMA, P. and D. WEENINK (2013): *Praat: Doing Phonetics by Computer*, Version 5.3.56. http://www.fon.hum.uva.nl/praat/ [15/09/2013]

BOOMERSHINE, A., K. C. HALL, E. HUME and K. JOHNSON (2008): «The impact of allophony vs. contrast on speech perception», in P. Avery, E. Dresher and K. Rice (eds): *Phonological Contrast: Perception and Acquisition*, The Hague, Mouton de Gruyter, pp. 146-172.

CAMPBELL-KIBLER, K. (2007): «Accent, (ING) and the social logic of listener perceptions», *American Speech*, 82(1), pp. 32-64.

CAMPBELL-KIBLER, K. (2008): «I'll be the judge of that: Diversity in social perceptions of (ING)», *Language in Society*, 27(5), pp. 637-659.

CAMPBELL-KIBLER, K. (2009): «The nature of sociolinguistic perception», *Language Variation and Change,* 21, pp. 135-156.

CAMPBELL-KIBLER, K. (2010): «Sociolinguistics and perception», *Language and Linguistics Compass*, 4, pp. 377-389.

CAMPOS-ASTORKIZA, R. (2014): «Sibilant voicing assimilation in Peninsular Spanish as gestural blending», in M. H. Côte, E. Mathieu and S. Poplack (eds.): *Variation Within and Across Romance Languages*, Amsterdam, John Benjamins Press, pp. 17-38.

CARRASCO, P. (2008): *An Acoustic Study of Voiced Stop Allophony in Costa Rican Spanish*, PhD dissertation, University of Illinois at Urbana-Champaign.

CARRASCO, P., J. I. HUALDE and M. SIMONET (2012): «Dialectal differences in Spanish voiced obstruent allophony: Costa Rican vs. Iberian Spanish», *Phonetica*, 69, pp. 149-179.

CHAPPELL, W. (2011): «The Intervocalic Voicing of /s/ in Ecuadorian Spanish», in J. Michnowicz and R. Dodsworth (eds.): *Selected Proceedings of the 5th Workshop on Spanish Sociolinguistics*, Somerville, Cascadilla Proceedings Project, pp. 57-64.

CHAPPELL, W. (2016): «On the social perception of intervocalic /s/ voicing in Costa Rican Spanish», *Language Variation and Change*, 28, pp. 357-378.

CHAPPELL, W. and C. GARCÍA (2017): «Variable production and indexical social meaning: On the potential physiological origin of intervocalic /s/ voicing in Costa Rican Spanish», *Studies in Hispanic and Lusophone Linguistics*, 10(1), pp. 1-37.

CLOPPER, C. G. and D. B. PISONI (2004): «Some acoustic cues for the perceptual categorization of American English regional dialects», *Journal of Phonetics*, 32, pp. 111-140.

DAHAN, D., S. J. DRUCKER and R. A. SCARBOROUGH (2008): «Talker adaptation in speech perception: Adjusting the signal or the representations?», *Cognition*, 108, pp. 710-718.

DERWING, B. L., T. M. NEAREY and M. L. DOW (1986): «On the phoneme as the unit of the 'second articulation'», *Phonology Yearbook*, 3, pp. 45-69.

DRAGER, K. (2010): «Sensitivity to grammatical and sociophonetic variability in perception», *Laboratory Phonology*, 1(1), pp. 93-120.

DUPOUX, E., C. PALLIER, N. SEBASTIAN GALLES and J. MEHLER (1997): «A destressing 'deafness' in French? », *Journal of Memory and Language*, 36, pp. 406-421.

ERNESTUS, M. (2006): «Statistically gradient generalizations for contrastive phonological features», *The Linguistic Review*, 23, pp. 217-233.

ERNESTUS, M. and W. M. MAK (2005): «Analogical effects in reading Dutch verb forms», *Memory and Cognition*, 33(7), pp. 1160-1173.

FLAGG, E. J., J. E. ORAM CARDY and T. P. ROBERTS (2006): «MEG detects neural consequences of anomalous nasalization in vowel-consonant pairs», *Neuroscience Letters*, 397, pp. 263-268.

FLEGE, J. E. (1988): «The production and perception of foreign language speech sounds», in H. Winitz (ed.): *Human Communication and its Disorders*: *A Review*, Norwood, Ablex, pp. 224-401.

FLEGE, J. E. (1990): «Perception and production: The relevance of phonetic input to L2 language learning», in C. Ferguson and T. Heubner (eds.): *Crosscurrents in Second Language Acquisition and Linguistic Theories*, Philadelphia, John Benjamins, pp. 249-289.

FLEGE, J. E. (1995): «Second language speech learning: Theory, findings and problems», in W. Strange (ed.): *Speech Perception and Linguistic Experience: Issues in Cross-language Speech Research*, Timonium, York Press, pp. 233-272.

FOWLER, C. A. and J. M. BROWN (2000): «Perceptual parsing of acoustic consequences of velum lowering from information for vowels», *Perception & Psychophysics*, 62(1), pp. 21-32.

FOX, R. A., J. E. FLEGE and M. J. MUNRO (1995): «The perception of English and Spanish vowels by native English and Spanish listeners: A multidimensional scaling analysis», *Journal of the Acoustical Society of America*, 97(4), pp. 2540-2551.

FRANCIS, A. and H. NUSBAUM (2002): «Selective attention and the acquisition of new phonetic categories», *Journal of Experimental Psychology: Human Perception and Performance*, 28(2), pp. 349-366.

FRIDLAND, V., K. BARTLETT and R. KREUZ (2004): «Do you hear what I hear? Experimental measurement of the perceptual salience of acoustically manipulated vowel variants by Southern speakers in Memphis, TN», *Language Variation and Change*, 16, pp. 1-16.

GARCÍA, C. (2015): *Production and Perception of Highland Ecuadorian Spanish*, PhD dissertation, The Ohio State University.

GOTO, H. (1971): «Auditory perception by normal Japanese adults of the sounds 'l' and 'r'», *Neuropsychologia*, 9, pp. 317-323.

GOTTFRIED, T. L. and P. S. BEDDOR (1988): «Perception of temporal and spectral information in French vowels», *Language and Speech*, 31, pp. 57-75.

GRIESER, D. and P. K. KUHL (1989): «Categorization of speech by infants: Support for speech-sound prototypes», *Developmental Psychology*, 25(4), pp. 577-588.

HALLÉ, P. A., J. SEGUI, U. FRAUENFELDER and C. MEUNIER (1998): «Processing of illegal consonant clusters: A case of perceptual assimilation?», *Journal of Experimental Psychology: Human Perception and Performance,* 24, pp. 592-608.

HALLÉ, P. A., C. T. BEST and A. G. LEVITT (1999): «Phonetic vs. phonological influences on French listeners' perception of American English approximants», *Journal of Phonetics*, 27, pp. 281-306.

HARNSBERGER, J. (2001): «The perception of Malayalam nasal consonants by Marathi, Punjabi, Tamil, Oriya, Bengali, and American English listeners: A multidimensional scaling analysis», *Journal of Phonetics*, 29, pp. 303-327.

HUALDE, J. I. (1989): «Silabeo y estructura morfémica en español», *Hispania*, 72(4), pp. 821-831.

HUALDE, J. I. (1991): «On Spanish syllabification», in H. Campos and F. Martínez-Gil (eds.): *Current Studies in Spanish Linguistics*, Washington, Georgetown University Press, pp. 475-494.

HUALDE, J. I. and P. PRIETO (2014): «Lenition of intervocalic alveolar fricatives in Catalan and Spanish», *Phonetica*, 71(2), pp. 109-127.

HUME, E. and K. JOHNSON (2003): «The impact of partial phonological contrast on speech perception», *Proceedings of the Fifteenth International Congress of Phonetic Sciences*, pp. 2385-2388.

INGRAM, J. C. and S. G. PARK (1998): «Language, context, and speaker effects in the identification and discrimination of English /r/ and /l/ by Japanese and Korean listeners», *Journal of the Acoustical Society of America*, 103, pp. 1161-1174.

JAEGER, J. J. (1980): «Testing the psychological reality of phonemes», *Language and Speech*, 23, pp. 233-253.

JOHNSON, K. and M. BABEL (2010): «On the perceptual basis of distinctive features: Evidence from the perception of fricatives by Dutch and English speakers», *Journal of Phonetics*, 38, pp. 127-136.

KAZANINA, N., C. PHILLIPS and W. J. IDSARDI (2006): «The influence of meaning on the perception of speech sounds», *Proceedings of the National Academy of Sciences of the United States of America*, 103(30), pp. 11381-11386.

KUHL, P. K., K. A. WILLIAMS, F. LACERDA, K. N. STEVENS and B. LINDBLOM (1992): «Linguistic experience alters phonetic perception in infants by 6 months of age», *Science*, 255(5044), pp. 606-608.

KUHL, P. K., B. T. CONBOY, S. COFFEY-CORINA, D. PADDEN, M. RIVERA-GAXIOLA and T. 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, pp. 979-1000.

LADEFOGED, P. (2003): *Phonetic Data Analysis: An Introduction to Fieldwork and Instrumental Techniques*, Malden: Blackwell Publishing.

LARKEY, L. S., J. WALD and W. STRANGE (1978): «Perception of synthetic nasal consonants in initial and final syllable position», *Perception & Psychophysics*, 23, pp. 299-312.

LIPSKI, J. M. (1994): *Latin American Spanish*, New York, Longman Group.

MACKAIN, K., C. T. BEST and W. STRANGE (1981): «Categorical perception of English /r/ and /l/ by Japanese bilinguals», *Applied Psycholinguistics*, 2, pp. 369-390.

MARTÍNEZ-GIL, F. (1991): «The insert/delete parameter, redundancy rules, and neutralization processes in Spanish», in H. Campos and F. Martínez-Gil (eds.): *Current Studies in Spanish Linguistics*, Washington, Georgetown University Press, pp. 495-572.

MCQUEEN, J. and M. A. PITT (1996): «Transitional probability and phoneme monitoring», *International Conference on Spoken Language Processing,* 4, pp. 2502-2505.

MOCHIZUKI, M. (1981): «The identification of /r/ and /l/ in natural and synthesized speech», *Journal of Phonetics*, 9, pp. 283-303.

MORGAN, T. A. (2010): *Sonidos en Contexto: Una Introducción a la Fonética del Español con Especial Referencia a la Vida Real*, New Haven, Yale University Press.

PISONI, D. B. and S. E. LIVELY (1995): «Variability and invariance in speech perception: a new look at some old problems in perceptual learning», in W. Strange (ed.): *Speech Perception and Linguistic Experience: Issues in Cross-language Research*, Baltimore, York Press, pp. 433-458.

PLICHTA, B. and D. R. PRESTON (2005): «The /ay/s have it: The perception of /ay/ as a North-South stereotype in US English», *Acta Linguistica Hafniensia*, 37, pp. 107-130.

POLKA, L. and J. F. WERKER (1994): «Developmental changes in perception of non-native vowel contrasts», *Journal of Experimental Psychology: Human Perception and Performance*, 20, pp. 421-435.

PRUITT, J. S. (1995): *The Perception of Hindi Dental and Retroflex Stop Consonants by Native Speakers of Japanese and American English*. PhD dissertation, University of South Florida.

QUILIS, A. (1993): *Tratado de Fonología y Fonética Españolas*, Madrid, Gredos.

ROBINSON, K. (1979): «On the voicing of intervocalic s in the Ecuadorian Highlands», *Romance Philology*, 33(1), 137-43.

ROHENA-MADRAZO, M. (2011): *Sociophonetic Variation in the Production and Perception of Obstruent Voicing in Buenos Aires Spanish*. PhD dissertation, New York University.

SAFFRAN, J. R., R. N. ASLIN and E. L. NEWPORT (1996): «Statistical learning by 8-month-old infants», *Science*, 274(5294), pp. 1926-1928.

SCHMIDT, L. B. and E. W. WILLIS (2011): «Systematic Investigation of Voicing Assimilation of Spanish /s/ in Mexico City», in S. M. Alvord (ed.): *Selected Proceedings of the 5th Conference on Laboratory Approaches to Romance Phonology*, Somerville, Cascadilla Proceedings Project, pp. 1-20.

STRANGE, W. (1995): «Cross-language studies of speech perception: A historical review», in W. Strange (ed.): *Speech Perception and Linguistic Experience*, Baltimore, York Press, pp. 3-45.

STRYCHARCZUK, P., M. VAN'T VEER, M. BRUIL and K. LINKE (2013): «Phonetic evidence on phonology-morphosyntax interactions», *Journal of Linguistics*, 50(2), pp. 403-452.

SZAKAY, A. (2008): «Social networks and the perceptual relevance of rhythm: A New Zealand case study», *University of Pennsylvania Working Papers in Linguistics*, 14(2), pp. 148-156.

TORREIRA, F. and M. ERNESTUS (2012): «Weakening of intervocalic /s/ in the Nijmegen Corpus of Casual Spanish», *Phonetica*, 69(3), pp. 124-48.

TRUBETZKOY, N. S. (1939): *Grundzüge der Phonologie*, Prague, Cercle Linguistique de Copenhague/Ministère de l'Instruccion Publique de la République Tchéco-slovaque.

VANEK, C. and S. McDANIEL (2006): *SurveyGizmo*, Widgix, LLC.

WALKER, A. (2007): «The effect of phonetic detail on perceived speaker age and social class», in J. Trouvain and W. J. Barry (eds.): *Proceedings of the 16th International Congress of Phonetic Sciences (ICPhS),* Dudweiler, Pirrot, pp. 1453-1456.

WALKER, A., C. GARCÍA, Y. CORTÉS and K. CAMPBELL-KIBLER (2014): «Comparing social meanings across listeners and speaker groups: The indexical field of Spanish /s/», *Language Variation and Change*, 26, pp. 169-189.

WHALEN, D. H., C. T. BEST and J. R. IRWIN (1997): «Lexical effects in the perception and production of American English /p/ allophones», *Journal of Phonetics*, 25(4), pp. 501-528.

YONEYAMA, K. and K. JOHNSON (1998): «An instance-based model of Japanese speech recognition by native and non-native listeners», *Journal of the Acoustical Society of America*, 103, p. 3090.

APPENDIX: FULL LIST OF WORD PAIRS HEARD IN EACH BLOCK

TASK 1: SIMILARITY RATING TASK

Block 1 (Word-medial environment)

Training 1
1. [aða] ~ [ada]
2. [aɾa] ~ [ada]
3. [asa] ~ [asa]
4. [aða] ~ [aða]

Test 1
1. [aɾa] ~ [ada]
2. [asa] ~ [asa]
3. [aɾa] ~ [aða]
4. [aza] ~ [asa]
5. [aða] ~ [ada]
6. [asa] ~ [aza]
7. [aza] ~ [aza]

Block 2 (Word-initial environment)

Training 2
1. [lasaka] ~ [lasaka]
2. [laðaba] ~ [ladaβa]
3. [ladaba] ~ [laðaβa]

Test 2
1. [lazaka] ~ [lazaka]
2. [laðaβa] ~ [ladaβa]
3. [lasaka] ~ [lazaka]
4. [lazaka] ~ [lasaka]
5. [ladaβa] ~ [laðaβa]
6. [laðaβa] ~ [laðaβa]
7. [ladaβa] ~ [ladaβa]
8. [lasaka] ~ [lasaka]

Block 3 (Word-final environment)

Training 3
1. [koŋoɾo] ~ [konoɾo]
2. [tanalto] ~ [tanalto]

Test 3
1. [lasata] ~ [lazata]
2. [taŋalto] ~ [tanalto]
3. [lazata] ~ [lazata]
4. [koŋoɾo] ~ [koŋoɾo]
5. [taŋalto] ~ [taŋalto]
6. [lazata] ~ [lasata]
7. [lasata] ~ [lasata]

TASK 2: AX DISCRIMINATION TASK

Block 4 (Word-medial environment)

Training 4
1. [aða] ~ [ada]
2. [aɾa] ~ [ada]
3. [ada] ~ [ada]

Test 4
1. [aða] ~ [ada]
2. [asa] ~ [aza]
3. [asa] ~ [asa]
4. [aða] ~ [aɾa]
5. [aða] ~ [aða]
6. [aza] ~ [aza]
7. [aza] ~ [asa]
8. [aða] ~ [aɾa]

Block 5 (Word-initial environment)

Training 5
1. [laðaβa] ~ [ladaβa]
2. [ladaβa] ~ [ladaβa]
3. [ladaβa] ~ [laðaβa]

Test 5
1. [laðaβa] ~ [laðaβa]
2. [lasaka] ~ [lazaka]
3. [ladaβa] ~ [ladaβa]
4. [lazaka] ~ [lasaka]
5. [laðaβa] ~ [laðaβa]
6. [lasaka] ~ [lasaka]
7. [ladaβa] ~ [laðaβa]
8. [ladaβa] ~ [ladaβa]
9. [lazaka] ~ [lazaka]

Block 6 (Word-final environment)

Training 6
1. [koŋoɾo] ~ [konoɾo]
2. [koŋoɾo] ~ [koŋoɾo]
3. [tanalto] ~ [tanalto]

Test 6
1. [lazata] ~ [lazata]
2. [tanalto] ~ [taŋalto]
3. [lasata] ~ [lasata]
4. [taŋalto] ~ [tanalto]
5. [konoɾo] ~ [konoɾo]
6. [lasata] ~ [lazata]
7. [lazata] ~ [lasata]
8. [koŋoɾo] ~ [koŋoɾo]