Differences in Discriminating L2 Consonants: A Comparison of Spanish Taps and Trills

Marda Rose
Indiana University

1. Introduction[1]

Models of second- and cross-language speech (e.g., Perceptual Assimilation Model, Best, 1995; Speech Learning Model, Flege, 1995; Native Language Magnet Model, Kuhl, 1993) have shown that the degree of success listeners will have in perceiving second language (L2) sounds depends on the perceived relationship between phonetic contrasts in the first and second language. Adult learners are sensitive to phonetic differences in their L2 that correspond to phonemic contrasts in their first language (L1) (e.g., Best, 1995; Werker & Tees, 1984), but experience difficulty distinguishing between L2 sounds that are not contrastive in their L1 (e.g., Hayes-Harb, 2007; Iverson et al., 2003).

A well-studied example of the influence of the L1 on the acquisition of L2 sounds is the acquisition of English /ɹ/ and /l/ by Japanese L2 learners of English (e.g., Aoyama, Flege, Guion, Akahane-Yamada, & Yamada, 2004; Goto, 1971). These learners have difficulty differentiating between English /ɹ/ and /l/ because these two sounds are associated with one phonemic category in Japanese, the Japanese /ɾ/ (e.g., Aoyama et al., 2004). As a result, it is believed that these learners link both the English /ɹ/ and /l/ in their L2 input to the Japanese /ɾ/, and consequently neutralize this L2 contrast to a single perceptual representation. The neutralization of an L2 contrast can pose persistent difficulty for L2 learners, despite years of experience with an L2 (e.g., Takagi, 2002). However, learners can improve their ability to discriminate nonnative contrasts as they become more proficient (e.g., Flege, Munro, & McKay, 1995) and with training (e.g., Pisoni & Lively, 1995), although they may not reach native-like proficiency.

Nonetheless, questions remain as to why some L2 consonants are persistently more difficult to discriminate and why L2 experience does not affect the discrimination of all L2 sounds equally. Guion, Flege, Akahane-Yamada, and Pruitt (2000) investigated the identification and discrimination of English consonants by Japanese L2 learners of English from three levels of proficiency to determine whether the Perceptual Assimilation Model (PAM) (Best, 1995) or the Speech Learning Model (SLM) (Flege, 1995) would be able to account for differences among learner groups (See details of the PAM in Section 1.1). They found that the PAM was best suited to predict how well L2 contrasts would be discriminated by learners. Only the PAM could be extended to early stages of L2 learning and could account for differences in how nonnative sounds are categorized within the L1.

Therefore, the current study follows this line of research to determine whether the PAM can be used to predict how L2 contrasts in Spanish will be discriminated by L1 speakers of American English. It examines the discrimination of five word-medial intervocalic contrasts in Spanish: /ɾ/-/r/, /ɾ/-[ɾʃ], [r]-[ɾʃ], /ɾ/-/t/, & /ɾ/-/d/. These contrasts specifically focus on the Spanish tap /ɾ/ and trill /r/—two rhotic phonemes in Spanish that contrast in intervocalic, word-medial position where they form minimal pairs (e.g., *caro* 'dear, expensive' vs. *carro* 'car'; e.g., Hualde, 2005; Quilis, 1993). This study also includes both allophonic and phonemic contrasts to determine whether native speakers (NSs) and learners discriminate these contrasts similarly at the word level.

[1] I would like to thank Isabelle Darcy, Laurent Dekydtspotter, and Erik Willis for helpful discussion on earlier versions of this paper, and two anonymous reviewers for their valuable feedback. I am also grateful to the Department of Spanish and Portuguese and the GPSO at Indiana University for travel grants to attend the conference.

The following section provides a description of the PAM. This is followed by an acoustic and articulatory description of taps and trills, a comparison of Spanish phonemes to English, a review of previous research, and a summary of predicted outcomes.

1.1. The Perceptual Assimilation Model

The Perceptual Assimilation Model (Best, 1995) posits the degree of success listeners will have in discriminating nonnative sounds.[2] It proposes that some nonnative contrasts are easier to discriminate than others because of the perceived relationship between L1 and L2 sounds, and outlines three different patterns of perceptual assimilation based on the results of cross-language mapping experiments (Best, 1995, p. 194-195). First, L2 sounds can be assimilated to a native category, in which case they may be heard as a good exemplar, an acceptable but not ideal exemplar, or a poor exemplar of the L1 category. Second, L2 sounds can be assimilated as an uncategorizable speech sound. In this case, the L2 sound is not a clear exemplar of any L1 category and may be perceived as an exemplar of more than one L1 sound. Finally, L2 sounds can be heard as a nonspeech sound and not assimilated into the L1.

Table 1. *Predictions of the Perceptual Assimilation Model (Best, 1995)*

Type of contrast	Type of assimilation	Description	Predicted discrimination of contrast
Two categorizable sounds	Two-category assimilation	Two sounds are assimilated to two different L1 sounds	Great accuracy
	Category goodness	Both L2 sounds assimilate to the same L1 phone, but one is associated more strongly than the other	Moderate to good
	Single category	Both L2 phones are assimilated equally well or equally poorly to the same L1 phone	Poor to moderate
Two uncategorizable sounds	Both uncategorizable	Both L2 phones are associated with more than one L1 sound	Poor to moderate
One categorizable and one uncategorizable sound	Uncategorized vs. categorized	One L2 phone is assimilated to one L1 sound; the other L2 phone is assimilated to more than one L1 sound	Very good
Neither sound categorized in L1	Non-assimilable	Both L2 sounds heard as non-speech (i.e., clicks)	Moderate to good

Each of these assimilation patterns makes specific predictions regarding the discrimination of nonnative contrasts (see Table 1). When both L2 sounds are categorizable, there are three possible outcomes. First, they can be perceptually categorized to two different L1 sounds (*two-category assimilation*). When learners assimilate sounds in this way, they discriminate between them with great accuracy (Best, 1990). Second, both L2 sounds can assimilate to one L1 phone, but one is associated more strongly than the other (*category goodness*). With this type of assimilation, discrimination may be moderate to good. Third, the two L2 phones are assimilated equally well or equally poorly to the

[2] The PAM (Best, 1995) predicts discrimination of nonnative contrasts by naïve listeners (i.e., adult L1 speakers who do not have experience with the L2 being tested). Following Guion et al. (2000), the current study extends the PAM to the perception of L2 contrasts.

same L1 phone (*single category*). When this type of assimilation occurs, discrimination will be poor to moderate. However, multiple outcomes have not been outlined for the other possible types of contrasts. When both L2 phones are uncategorizable (*both uncategorizable*), these phones are associated with more than one L1 sound, and discrimination is poor to moderate. When one L2 phone is categorizable and the other is uncategorizable (*uncategorized versus categorized*), one sound is assimilated to an L1 category while "the other falls in phonetic space, outside native categories" (Best, 1995, p. 195). Discrimination of this contrast type is very good. Finally, the contrast may be *non-assimilable* because both L2 sounds are heard as non-speech. This contrast has good to moderate discrimination and has most often been associated with clicks (Best, McRoberts, & Sithole, 1988).

Guion et al. (2000) collected cross-language mapping data and discrimination data in order to determine how well the PAM could be used to predict learning of English L2 sounds by three groups of Japanese L1 speakers that differed in terms of English proficiency. As the PAM would predict, the learners in their study were unable to discriminate a *both uncategorizable* contrast (/ɹ/-/l/) and did not evidence improvement at higher levels of proficiency. However, discrimination results of two *uncategorized versus categorized* contrasts differed. Discrimination of the /ɹ/-/w/ contrast was good and showed improvement as learners reached higher levels of proficiency, whereas discrimination of the /s/-/θ/ contrast was poor for all participants. These differences in discrimination were attributed to how the L2 phones were mapped to the L1. The /ɹ/ and /w/ were mapped to different phonemes in the L1, while the /s/ and /θ/ were both categorized within the same L1 category, /s/. Therefore, the authors suggested a revision to the PAM "to allow for poor discrimination of this contrast type when the uncategorized sound is close in phonological space to the categorized sound" (p. 2721).

Consequently, the present study analyzes *both uncategorizable* contrasts and *uncategorized versus categorized* contrasts to determine whether differences in discrimination may be attributed to the acoustic distance between L2 sounds once they are assimilated into the L1. It further examines Guion et al.'s (2000) conclusions by including L2 Spanish contrasts that are close in phonological space and assimilate to similar sounds in L1 English to determine whether the combination of contrast type and acoustic distance can be used to consistently predict L2 discrimination.

1.2. Acoustic and articulatory description of the Spanish tap and trill

The Spanish tap "is produced with a single rapid contact of the tip of the tongue against the alveolar region," whereas the trill is produced with two or more similar gestures (Hualde, 2005, p.181). Spanish trills are described as involving two to three occlusions (e.g., Lipski, 1990; Quilis, 1993), although trills can be produced with as few as one occlusion or contain more than three (e.g., Ladefoged & Maddieson, 1996; Lindau, 1985).

Variation in tap and trill production is also possible (e.g., Blecua, 2001; Rose, in press). Approximant productions have been observed in L1 Spanish as variants of taps (Blecua, 2001) and trills (Blecua, 2001; Lewis, 2004), and in some varieties of Spanish, assibilated variants (Quilis, 1993) and pre-breathy voiced trills (Willis, 2006) occur in the phonological trill environment. Of interest to the current study is the tap+ variant [ɾʃ],[3] which has one clear occlusion followed by a period of frication. Both learners and NSs of Spanish produce tap+ variants in spontaneous speech, with NSs producing them almost exclusively in the phonological trill environment (Rose, in press).[4] The tap+ allophone is included in the present study to determine whether it is perceived like a canonical trill /r/, even though it does not have two clear occlusions.

1.3. Comparison of Spanish phonemes to English

The Spanish alveolar trill /r/ is unlike any sound in American English, and consequently its acquisition is not affected by the association of this phone with a similar L1 sound. The acquisition of the Spanish alveolar tap /ɾ/ is slightly more complicated because of its allophonic status in English. Spanish taps are similar to English flaps, which are produced as allophones of /t/ and /d/ between a

[3] The plus sign (+) in *tap+* refers to the period of frication that occurs after the traditional tap realization (Rose, in press).
[4] A similar trill allophone consisting of a tap and additional material (e.g., r-coloring or frication) was also observed in L1 Spanish in Veracruz Mexico (Bradley & Willis, 2008).

stressed and an unstressed vowel (in post-tonic, intervocalic position) in words such as *latter* /læɾəɹ/ and *ladder* /læɾəɹ/. English flaps are produced with a rapid tongue motion characterized by a short closure that quickly moves away toward the configuration of the following vowel (Olive, Greenwood, & Coleman, 1993). Most authors do not differentiate between taps and flaps; however, subtle articulatory differences have been observed in how the tongue strikes the alveolar ridge (Ladefoged & Maddieson, 1996; Monnot & Freeman, 1972). English flaps anticipate the movement and retract before striking the alveolar ridge in passing, like in the production of the words *three* and *throw* (Clark & Yallop, 1990). Taps, on the other hand, involve the direct movement of the tip of the tongue against the alveolar ridge, creating a closure. Despite these articulatory differences, Monnot and Freeman (1972) did not find any criterion to differentiate between English flaps and Spanish taps on a spectogram. Futhermore, they found that English flaps and Spanish taps were cross-identified on a perception task. Listeners from five different L1 backgrounds identified the English flap as either a tap /ɾ/ or as an intervocalic /t/ or /d/, depending on their L1. Similarities between taps and flaps allowed English flaps to be interpreted both ways, leading Monnot and Freeman to conclude that the English flap and the Spanish tap were perceptually similar.

Since the Spanish /ɾ/ is similar to English /t/ and /d/, the current study compares the discrimination of /ɾ/-/t/ and /ɾ/-/d/. The Spanish /t/ and English /t/ are similar in that they are both voiceless stops. However, they differ in terms of place of articulation and presence of aspiration. Spanish /t/ is dental and unaspirated (Hualde, 2005) whereas the English /t/ is alveolar and aspirated (Ladefoged, 2001). The presence of aspiration results in a longer voice onset time (VOT) in the production of English /t/ when compared to the Spanish /t/; consequently, the Spanish /t/ may be perceived as being more similar to the English /d/, which has a shorter VOT (Ladefoged, 2001). On the other hand, when the Spanish /d/ is produced intervocalically, it is realized as a voiced dental fricative [ð] due to spirantization (Hualde, 2005). This sound may be similar to the English voiced interdental fricative /ð/.

1.4. Previous studies on tap and trill acquisition

Previous studies on the acquisition of the Spanish tap and trill have primarily focused on production (e.g., Carballo & Mendoza, 2000; Face, 2006; Major, 1986; Waltmunson, 2005). These studies have shown that the tap is acquired before the trill for both L1 and L2 learners of Spanish (e.g., for L1: Bosch, 1983; Carballo & Mendoza, 2000; Jiménez, 1987; for L2: Face, 2006; Reeder, 1998; Rose, in press; Waltmunson, 2005) and that vocalic environment can affect trill production (Recasens, 1991; Solé, 2002). Reeder (1998) investigated the discrimination of 12 minimal pairs in L2 Spanish by L1 American English speakers from four levels of proficiency. The task focused on 5 vocalic and 6 consonantal features; however, Reeder did not specify which minimal pairs were contrasted, and all results were pooled, making it difficult to draw any conclusions about the perception of specific phonemes. Learners at all proficiency levels showed high levels of accuracy, with learners in their first semester of study discriminating between minimal pairs at 88%. On average, the learners' ability to discriminate between the minimal pairs increased with proficiency, suggesting that the discrimination of some L2 contrasts may improve as learners increase in proficiency level (see also Flege, Munro, & McKay, 1995; Guion et al., 2000).

2. The present study

The current study aims to determine whether the PAM can predict how L2 contrasts in Spanish will be discriminated by L1 speakers of American English. According to the PAM, some contrasts will be discriminated better than others (Best, 1995; Guion et al., 2000). Namely, a *both uncategorizable* contrast is expected to be poor and not to improve as learners increase in proficiency (Guion et al., 2000). However, an *uncategorized versus categorized* contrast was shown to vary, depending on the acoustic distance between L2 sounds once they are assimilated into the L1 (Guion et al., 2000). The present study compares five different word-medial, intervocalic contrasts in Spanish to verify whether the same results are obtained: /ɾ/-/r/, /ɾ/-[ɾʃ], [r]-[ɾʃ], /ɾ/-/t/, & /ɾ/-/d/. Both phonemic and allophonic contrasts are tested to determine whether NSs and learners differ in their discrimination of these contrasts at the word level.

The current study addresses the following research question:

Can the type of contrast as described in the Perceptual Assimilation Model together with acoustic distance in phonological space be used to consistently predict how well a nonnative contrast will be discriminated by learners at different proficiency levels?

2.1. Cross-language categorization of Spanish consonants in L1 English

To predict how well these L2 contrasts would be discriminated within the PAM framework, cross-language mapping data were collected in a pilot study for a forthcoming project (Rose, 2009; Rose & Darcy, 2009). This experiment aimed to determine the perceived relationship between Spanish and English consonants by investigating how the Spanish consonants /ɾ/, /r/, /t/, and /d/ are categorized in English by 13 NSs of American English.[5] The tap+ allophone was not included in the cross-language mapping task due to difficulty in obtaining a sufficient number of tokens. The tap+ is expected to pattern like a trill.

The cross-language mapping task used similar stimulus materials to those used in the current study in order to have comparable results. Real words were used to encourage discrimination at the phonological level. The Spanish words selected provide a natural context for the production of the targeted consonants in intervocalic position, allowing the consonants to be contrasted in environments where they form minimal pairs. Two different words were included for each targeted consonant: /ɾ/ (*cara* 'face'; *para* 'for'), /r/ (*carro* 'car'; *perro* 'dog'), /t/ (*rato* 'while'; *pata* 'paw'), and /d/ (*cada* 'each'; *todo* 'all'). All of the targeted consonants were produced between low- and mid-vowels.

Four NSs of Spanish (two from Spain and two from Colombia) produced the stimulus materials by reading the targeted words listed above in the carrier phrase, *Digo __ para ti*, 'I say ___ for you.' (For more details on how the stimuli were recorded, see Section 3.2.1). Both a male and female speaker from each country were included on the cross-language mapping task because the discrimination task in the current study includes both male and female voices to encourage the discrimination of phonological differences (Section 3.2). Only the voices of speakers from Spain are used on the discrimination task, and the words produced by these speakers are the same on both tasks.

Eight NSs of Spanish verified that the items used on the cross-language discrimination task were good examples of the Spanish consonants. They identified the Spanish consonants in terms of Spanish categories and rated them for goodness of fit to the Spanish categories on a five-point scale. The /ɾ/ was identified as a /ɾ/ 100% of the time (goodness-of-fit = 4.8), the /r/ was identified as /r/ 99.2% of the time (goodness-of-fit = 4.8), the /t/ was identified as /t/ 98.4% of the time (goodness-of-fit = 4.8), and the /d/ was identified as /d/ 99.2% of the time (goodness-of-fit = 4.6).

Thirteen NSs of English (mean age = 22.6 years) completed the cross-language mapping task: seven had no prior experience with Spanish and six had studied some Spanish (mean = 3 years in high school). They identified each token as an example of one of the following English consonant categories: /ɹ/, /t/, /d/, /s/, /l/ and /ð/. They then rated the stimuli for goodness of fit on a five-point scale. Each stimulus was repeated three times for a total of 96 tokens on the task (4 consonants x 2 words per consonant x 4 speakers x 3 times).

A one-way ANOVA indicated that for most consonant identification, there was no significant difference between participants who had experience with Spanish and those who did not. The only exception was the classification of the Spanish /d/ in English. Those with some Spanish experience categorized the Spanish /d/ as an English /d/ significantly more than those who did not have experience with Spanish, $F(1, 11) = 7.926$, $p = .017$. Conversely, those without Spanish experience categorized the Spanish /d/ as an English /l/ significantly more often, $F(1, 11) = 5.925$, $p = .033$. Since this was the only significant difference between the two groups of participants, the remainder of the analysis combines the results of all NSs of English.

Table 2 presents the results of the cross-language mapping task. The Spanish trill /r/ was the only consonant categorized within a single category in English (categorizable), being identified as an English /ɹ/ in 96.2% of instances.[6] The other three Spanish consonants (/ɾ/, /d/, and /t/) were

[5] Cross-language mapping tasks determine how nonnative sounds are perceived by having participants categorize nonnative sounds in their L1 using L1 sound categories (e.g., Best, 1995; Guion et al., 2000).

[6] Only L1 sounds identified at 10% or higher are included in Table 2. For further detail, please refer to Rose (2009).

categorized as either two or three different English consonants (uncategorizable). Following the work of Guion et al. (2000), the overall fit of the Spanish stimuli to the English categories was determined through the use of a fit index. The fit index was calculated by multiplying the proportion of identification of each consonant by its corresponding goodness rating.

Table 2. *Cross-language Identification and Fit Indexes Derived for Spanish Consonants in Terms of English Categories*

Spanish consonant	Most common identification in English	Proportion of identifications	Goodness rating	Fit index	
/r/	/ɹ/	96.2%	2.3	2.21	fair /ɹ/
/ɾ/	/ɹ/	57.7%	2.3	1.33	poor /ɹ/
	/d/	30.4%	2.8	0.85	poor /d/
/d/	/l/	54.2%	3.1	1.68	poor /l/
	/d/	32.4%	3.0	0.97	poor /d/
	/ɹ/	11.2%	2.5	0.28	poor /ɹ/
/t/	/t/	65.4%	3.1	2.03	fair /t/
	/d/	21.8%	3.0	0.65	poor /d/
	/ð/	12.5%	2.5	0.31	poor /ð/

Note. The fit index was calculated by multiplying the proportion of identifications by its goodness rating. Only identifications greater than 10% were included.

Based on the results of the cross-language mapping task, predictions were made regarding the discrimination of the two contrast types investigated in the present study: *uncategorized vs. categorized* and *both uncategorizable* (see Table 3).

Table 3. *Expected Discrimination of Five Spanish Contrasts within the Perceptual Assimilation Model*

Contrast	Assimilation in L1	PAM category type	Expected Discrimination	
			Best (1995)	Guion et al. (2000)
/ɾ/-/r/	/ɾ/ → /ɹ/, /d/ /r/ → /ɹ/	uncategorized vs. categorized	very good	poor
/ɾ/-[ɾʃ]	/ɾ/ → /ɹ/, /d/ /r/ → /ɹ/	uncategorized vs. categorized	very good	poor
[r]-[ɾʃ]	[r] → /ɹ/ [ɾʃ] → /ɹ/	NA (see Note)		
/ɾ/-/t/	/ɾ/ → /ɹ/, /d/ /t/ → /t/, /d/, /ð/	both uncategorizable	poor to moderate	good
/ɾ/-/d/	/ɾ/ → /ɹ/, /d/ /d/ → /ɹ/, /d/, /l/	both uncategorizable	poor to moderate	poor

Note. The [r]-[ɾʃ] contrast compares two allophones of the same L2 phoneme. The PAM makes no predictions for this type of contrast.

The /ɾ/-/r/ contrast is an *uncategorized vs. categorized* contrast in which the /ɾ/ is uncategorized, being assimilated to both the English /ɹ/ and /d/, and the /r/ is categorized as a fair exemplar of the English /ɹ/. Previous studies on the production of taps and trills assume that this contrast is discriminated with great accuracy; however, Guion et al. (2000) suggest that this contrast may not be discriminated because both sounds converge on the same L1 category, the English /ɹ/.

The /ɾ/-[ɾʃ] contrast is an *uncategorized vs. categorized* contrast that is expected to assimilate to the L1 in the same way as the /ɾ/-/r/ contrast, with tap+ being categorized as a fair exemplar of the English /ɹ/. Since tap+ is an allophone of the Spanish /r/, discrimination of the /ɾ/-[ɾʃ] contrast is expected to be similar to the discrimination of the /ɾ/-/r/ contrast.

The [r]-[ɾʃ] contrast compares two allophones of the same L2 phoneme that differ in terms of number of occlusions. This is a phonetic contrast that compares a trill with three occlusions to a trill with one occlusion in the word *perro,* ‘dog.’ The PAM does not predict the assimilation of this type of contrast. However, using the model as a guide, it is possible to posit that these two allophones will both be assimilated to the same L1 sound, /ɹ/, to a similar degree and that discrimination at the phonological level will be poor.

The /ɾ/-/t/ contrast is a *both uncategorizable* contrast in which the /ɾ/ is assimilated to both the English /ɹ/ and /d/, and the Spanish /t/ is assimilated to English /t/, /d/, and /ð/. The Spanish /ɾ/ and /t/ primarily are associated with different English phonemes, only sharing a poor association with the English /d/.

The /ɾ/-/d/ contrast is a *both uncategorizable* contrast in which both the Spanish /ɹ/ and /d/ are assimilated to the English /ɹ/ and /d/. Discrimination of this contrast is expected to be poor, and it is expected to be the most difficult contrast for learners to discriminate.

The organization of the remainder of the paper is as follows. Section 3 details the experimental design of the AXB discrimination task and how the stimuli for the task were created. Section 4 reports the results of the AXB discrimination task by NSs of Spanish first, then by NSs of English who have not taken Spanish, and finally by learners of Spanish. Section 5 provides a discussion of the results in light of the research question, and the conclusions of the study are presented in Section 6.

3. Method

3.1. Participants

There were a total of 90 participants: 15 NSs of Spanish, 60 NSs of English who were learners of Spanish, and 15 NSs of English who had never studied Spanish. The 60 learners were enrolled in second- (Level 1), fourth- (Level 2), and eighth-semester Spanish (Level 3), in addition to graduate courses (Level 4) at Indiana University. There were 15 learners in each level. Participants in Levels 1 and 2 (mean age = 21.3 years) were undergraduate students enrolled in language courses that are primarily taken to fulfill a foreign language requirement. The Level 3 participants (mean age = 21.7) were undergraduate students majoring in Spanish or Spanish Education and were enrolled in an advanced literature course. The Level 4 participants (mean age = 27.1) were MA (n = 5) or PhD (n = 10) students of Hispanic Literature or Linguistics in the Department of Spanish and Portuguese.

The 15 NSs of English who had never studied Spanish (henceforth, Level 0) were university students: 12 undergraduate and three graduate (mean age = 21). Eight of them had no prior foreign language experience, while seven had studied one or more of the following languages: Arabic, Chinese, French, German, Greek, Italian, or Russian. Those who had studied Greek, Italian, and Russian (n = 5) had prior experience with trills.

The NSs of Spanish (mean age = 31) were from Colombia (n = 5), Chile (n = 2), Ecuador (n = 1), Peru (n = 1), Spain (n = 4), and Venezuela (n = 2). Speakers from these countries were included because they produce alveolar taps and trills in their speech, as verified by production data collected when they met with the researcher. Eight of the NSs were pursuing a graduate degree at Indiana University, six were studying English in the Intensive English Program at Indiana University, and one was a visiting instructor of Spanish in the Department of Spanish and Portuguese. The graduate students and visiting professor (n = 9) were highly proficient in English whereas the Intensive English students (n = 6) had a high-intermediate level of proficiency.

3.2. Materials
3.2.1. Stimulus materials

Fifteen words were selected to create the contrasts for the discrimination task. These words presented the target sounds in word-medial, intervocalic position so that they could be contrasted in environments where they form minimal pairs. Words were used to encourage discrimination at the phonological level.[7] Two word pairs were selected for each contrast, with the exception of the /ɾ/-[ɾʃ] and /r/-[ɾʃ] contrasts, due to the difficulty of eliciting the production of the [ɾʃ] in the same words by multiple speakers. The words selected for each contrast were: /ɾ/-/r/ (*caro* 'dear,' *carro* 'car'; *quería* 'he wanted,' *querría* 'he would want'), /ɾ/-/t/ (*raro* 'weird,' *rato* 'while'; *para* 'for,' *pata* 'paw'), /ɾ/-/d/ (*cara* 'face,' *cada* 'each'; *toro* 'bull,' *todo* 'all'), /ɾ/-[ɾʃ] (*pero* 'but,' *perro** 'dog'), and /r/-[ɾʃ] (*perro* 'dog,' *perro** 'dog'). The asterisk after the word *perro* indicates it is produced with a tap+ and not a canonical trill.

The stimulus materials were produced by two native Spanish speakers from Spain (mean age = 30), one male and one female, who had been living in the United States for an average of 3 years. The speakers were recorded in a sound proof recording booth using a Shure SM58 microphone, a Shure M267 mixer (digital analog converter) and Peak LE 5.02, and were sampled at 44.1k. To elicit both careful and casual production of the target sounds, the speakers narrated two picture stories (*Frog, where are you?,* Mayer, 1969, and a six-frame picture story from Heaton, 1966) and read word lists. When narrating the stories, speakers were instructed to tell what happened yesterday, using the pictures provided. It was assumed that the use of stories would encourage a less careful speech style. The word lists contained the target words in the carrier phrase, *Digo ___ para ti*, 'I say ___ for you.' Speakers were asked to read the sentences at a speech rate that was appropriate for use with NSs of Spanish and were instructed not to over articulate the target words. The words were presented randomly and allowed for the production of each word to occur at least six times. One production of each word was selected from each speaker to create the task. Since there were multiple tokens of each word per speaker, the words chosen for the task were similar in intonation and duration.

3.2.2. Discrimination task

Each trial of the AXB discrimination task required participants to listen to three words and decide whether the first (A) or the third (B) word was similar to the middle word (X). Following Best and Strange (1992), this discrimination procedure was chosen because it places lower demands on memory and is less sensitive to observer bias than other procedures, such as ABX, oddity, 2IAX and 4IAX (see also Best, McRoberts, & Goodell, 2001). When comparing two words or sounds, the AXB discrimination task has four possible combinations (AAB, ABB, BAA, and BBA). The /ɾ/-/r/ comparison contrasted *caro/carro* and *quería/querría* in three AXB sets for a total of 24 items (2x4x3). These two comparisons differ from each other in terms of stress, number of syllables, and vocalic environment. Since vocalic environment and stress have been shown to affect trill production (e.g., Recasens, 1991; Solé, 2002; Waltmunson, 2005), more items of these two contrasts were included to investigate whether context might affect /ɾ/-/r/ discrimination. Tap+ was compared to a tap and to a trill with three occlusions in two AXB sets (*pero/perro*; perro/perro**) for a total of 16 items (2x4x2), and the /ɾ/-/d/ (*toro/todo; cara/cada*) and /ɾ/-/t/ contrasts (*raro/rato; para/pata*) were compared for an additional 16 items (2x4x2). These 72 items were randomized and divided into six sets of 12 items. These sets comprised Parts B-G of the task. Part A consisted of three control comparisons (*mido/miro* 'I measure/I watch'*; podía/podría* 'he was able to/he could'*; pelo/pero* 'hair/but') of one AXB set each, adding a total of 12 additional items (3x4x1). These 12 items served as training to familiarize the participants with the task. For each AXB comparison, the inter-stimulus interval was 1 second, and the intertrial interval was three seconds. The inter-block interval was 12 seconds.

[7] Discrimination at the phonological level involves distinguishing between minimal pairs such as *para*, 'for' and *pata,* 'paw,' two words that contain different phonemes (/ɾ/ and /t/) and consequently differ in meaning. The use of words and different voices (Section 2.1, Section 3.2) encourages participants to discriminate between two different words as opposed to two different pronunciations of the same word (phonetic level).

The male voice was always in the middle (X) so that his utterance would be compared to two different words pronounced by the female speaker (A and B). The trials were devised in this way to encourage participants to focus on the word itself and not other factors like pitch and intonation. The male voice was also used for the instructions given at the beginning so that his voice might be more familiar to the participants as they completed the task.

Two different versions of the AXB task (Version A and Version B) were created using PowerPoint. Part A was the same on both versions, but the order of Parts B-G varied. In Version B, the order of the blocks was reversed (G-B).

For the learners (Levels 1 to 4) and NSs of Spanish, the entire slide show was presented in Spanish. Level 0 participants completed the same AXB tasks (Version A and Version B), with the only difference being that the instructions were given in English instead of Spanish. The author's voice was used for the English version.

3.3. Procedure

Participants met individually with the researcher outside of class in a language lab. The learners and NSs of Spanish first completed the production task. Then, all participants completed the computer-delivered AXB discrimination task, listening to aural stimuli over individual headsets. Participants heard each item once and recorded their responses on an answer sheet, circling either A or B to indicate which word was similar to the second word that they heard.

3.4. Analysis

For each participant, mean accuracy scores for each of the following contrasts were calculated: /ɾ/-/r/ (*carro*), /ɾ/-/r/ (*quería*), /ɾ/-/r/ (all), /ɾ/-[ɾʃ], [r]-[ɾʃ], /ɾ/-/d/, and /ɾ/-/t/. Separate one-way ANOVAs were then performed to identify any significant differences between levels (Levels 0, 1, 2, 3, and 4, and NS). Post hoc tests were conducted using Tukey's HSD procedure. For both of these procedures, the alpha level was set at $p = .05$.

Paired samples (two-tailed) t-tests were performed to compare contrasts within each level, and the alpha level was adjusted according to the number of comparisons made.

4. Results

The task yielded 6480 responses, with 1080 responses per level. NSs discriminated between four of the five contrasts with great accuracy, and learner responses varied. This section is organized in three main sections by participant group. As the results will show, L2 contrasts are more difficult to discriminate when both L2 sounds are assimilated to the same sound(s) in the L1. However, learners can improve their discrimination of these contrasts as they reach higher levels of proficiency.

4.1. Native speakers of Spanish

Table 4 presents the NSs' mean accuracy scores and standard deviations for each of the five contrasts. As expected, the NSs were extremely accurate in discriminating four of the targeted contrasts: /ɾ/-/r/, /ɾ/-/t/, /ɾ/-/d/, and /ɾ/-[ɾʃ]. These results also suggest that the range for native-like discrimination of phonemic contrasts is between 97% and 100%. As for the /ɾ/-[ɾʃ] and [r]-[ɾʃ] contrasts, the tap+ was perceived as being different from a tap, but was not distinguished from the trill at a level significantly above chance: $t(14) = -1.031$, $p = .320$.

Table 4. *Discrimination Accuracy of Contrasts by Native Speakers of Spanish*

Contrast	*M* (%)	*SD*
/ɾ/-/r/	99.7	0.3
/ɾ/-[ɾʃ]	97.5	7.0
[r]-[ɾʃ]	45.0	18.8
/ɾ/-/t/	99.2	3.2
/ɾ/-/d/	97.1	4.6

4.2. Native speakers of English who have not studied Spanish (Level 0)

The results of the Level 0 participants on the /ɾ/-/r/ contrast are presented in Table 5. These results are first presented as an average of all responses and then separated by contrast (*caro/carro* and *quería/querría*). The *quería/querría* contrast, which presents the /ɾ/-/r/ contrast in the tonic syllable, was discriminated better than the *caro/carro* contrast. However, this difference was not statistically significant ($t(14) = -2.057$, $p = .059$). The rest of the analysis focused on the combined /ɾ/-/r/ contrast results.

Table 5. *Discrimination Accuracy of /ɾ/-/r/ by Level 0 Participants*

Contrast	*M* (%)	*SD*
/ɾ/-/r/ (all)	80.2	11.9
/ɾ/-/r/ (carro)	75.6	17.9
/ɾ/-/r/ (querría)	85.0	11.0

Table 6 presents the Level 0 participants' results on the five targeted contrasts. They discriminated four of the contrasts (/ɾ/-/r/, /ɾ/-[ɾʃ], /ɾ/-/t/, /ɾ/-/d/) at levels significantly above chance ($p < .001$), and they discriminated between the /ɾ/-/r/ and the /ɾ/-/t/ contrasts slightly better than the others. Paired samples t-tests revealed that none of these four contrasts were perceived significantly better than each other (*p*s > 0.10), suggesting that without training in Spanish, NSs of American English discriminate these contrasts to a similar degree. The only contrast Level 0 participants did not discriminate at a level significantly above chance was the [r]-[ɾʃ] contrast, $t(14) = -1.288$, $p = .219$.

Table 6. *Discrimination Accuracy of Contrasts by Level 0 Participants*

Contrast	*M* (%)	*SD*
/ɾ/-/r/ (all)	80.2	11.9
/ɾ/-[ɾʃ]	75.8	20.3
[r]-[ɾʃ]	42.5	22.6
/ɾ/-/t/	79.5	17.3
/ɾ/-/d/	72.9	16.8

4.3. Learners of Spanish

The learners' discrimination of the /ɾ/-/r/ contrast is presented in Table 7. Learners at all levels discriminated between taps and trills at a level significantly above chance ($p < .001$), and demonstrated very good discrimination from the lowest levels of proficiency, leaving little room for improvement. The discrimination of the *caro/carro* contrast and the *quería/querría* contrast varied by level, with Levels 1 and 2 discriminating the *quería/querría* contrast better and Levels 3 and 4 discriminating the *caro/carro* contrast better. The difference in the discrimination of these two contrasts was only significant for Level 1, $t(14) = -3.416$, $p = .004$. When discrimination on both /ɾ/-/r/ contrasts is combined (column 4), accuracy scores range from 86.7% to 94.4%, and there is no significant difference among the learner groups ($F(3, 56) = 1.652$, $p = 0.188$). Consequently, these combined accuracy scores suggest that learners do not improve significantly in their ability to discriminate between taps and trills as they become more proficient.

Learner discrimination of the /ɾ/-/r/ contrast was less accurate than the NSs, who discriminated all contrasts at a level of 97% or higher. Overall this difference was significant ($F(5, 84) = 6.803$, $p < .001$). The post hoc reveals that the NSs' discrimination of this contrast was significantly better than Level 2 ($p = .008$) and Level 3 ($p = .036$).

Table 7. *Discrimination Accuracy of /ɾ/-/r/ by Learners of Spanish*

	Contrast					
	/ɾ/-/r/ (*carro*)		/ɾ/-/r/ (*quería*)		/ɾ/-/r/ (all)	
Level	*M* (%)	*SD*	*M* (%)	*SD*	*M* (%)	*SD*
1	88.3	10.4	96.7	5.2	92.5	6.7
2	83.9	17.9	89.4	15.6	86.7	14.7
3	93.3	11.0	83.9	17.7	88.6	10.9
4	96.1	7.6	92.7	12.9	94.4	8.8

Table 8 compares the learners' discrimination of the /ɾ/-/r/ contrast to their discrimination of the /ɾ/-[ɾʃ] and [r]-[ɾʃ] contrasts. All of the learner groups differentiated the the /ɾ/-[ɾʃ] contrast contrast at levels significantly above chance ($p < .001$). However, with the exception of Level 3, learners perceived the /ɾ/-/r/ contrast better than the /ɾ/-[ɾʃ] contrast, and this difference was significant for Level 1 ($t(14) = 3.189$, $p = .007$) and Level 4 ($t(14) = 2.987$, $p = .010$). This suggests that even though learners perceive a tap+ as being different from a tap, a trill with three occlusions is a better exemplar of a trill. When the tap+ was compared to a trill, learners in the first three levels did not discriminate the contrast at a level significantly above chance (Level 1: $t(14) = 1.331$, $p = .205$; Level 2: $t(14) = -.564$, $p = .582$; Level 3: $t(14) = 1.749$, $p = .102$). Learners in Level 4 were the only participant group who discriminated the [r]-[ɾʃ] contrast at a level significantly above chance ($t(14) = 4.006$, $p = .001$).

Table 8. *Discrimination Accuracy of /ɾ/-/r/, /ɾ/-[ɾʃ], and /r/-[ɾʃ] by Learners of Spanish*

	Contrast					
	tap-trill /ɾ/-/r/		tap-tap+ /ɾ/-[ɾʃ]		trill-tap+ [r]-[ɾʃ]	
Level	*M* (%)	*SD*	*M* (%)	*SD*	*M* (%)	*SD*
1	92.5	6.7	81.7	16.3	56.7	19.4
2	86.7	14.7	76.7	14.1	47.5	17.2
3	88.6	10.9	92.5	9.2	59.2	20.3
4	94.4	8.8	85.0	17.8	72.5	21.8

The learners' discrimination of the /ɾ/-/r/ contrast is compared to their discrimination of /ɾ/-/t/ and /ɾ/-/d/ in Table 9. All learner groups discriminated the /ɾ/-/t/ comparison at levels significantly above chance ($p < .001$). However, unlike the /ɾ/-/r/ comparison, discrimination of the /ɾ/-/t/ contrast improved to a level that is similar to NS discrimination as participants increased in proficiency (Levels 3 and 4). It is the only contrast in this study to improve in this way. The /ɾ/-/d/ contrast also improved as learners reach higher levels of enrollment, and the ANOVA shows a significant difference among all participant groups, $F(5, 84) = 18.296$, $p < .001$. The post hoc indicates that Level 4 participants discriminated this contrast significantly better than all other learner groups (Level 1: $MD = 12.92$, $SE = 4.16$, $p = .030$; Level 2: $MD = 19.58$, $SE = 4.16$, $p < .001$; Level 3: $MD = 15.83$, $SE = 4.16$, $p = .004$).

However, even though learners showed significant improvement in their ability to discriminate the /ɾ/-/d/ contrast at higher levels of enrollment, their ability to perceive this contrast was not as good as their ability to perceive the /ɾ/-/r/ and /ɾ/-/t/ contrasts. This difference was significant for all learners. As Table 10 shows, the /ɾ/-/r/ and the /ɾ/-/t/ contrasts were discriminated significantly better than the /ɾ/-/d/ contrast by learners in all four levels.

Table 9. *Discrimination Accuracy of /ɾ/-/r/, /ɾ/-/t/, and /ɾ/-/d/ by Learners of Spanish*

	Contrast					
	/ɾ/-/r/		/ɾ/-/t/		/ɾ/-/d/	
Level	*M* (%)	*SD*	*M* (%)	*SD*	*M* (%)	*SD*
1	92.5	6.7	90.0	10.8	69.6	12.5
2	86.7	14.7	91.3	8.5	62.9	8.0
3	88.6	10.9	100.0	0.0	66.7	9.6
4	94.4	8.8	99.6	1.6	82.5	12.8

Table 10. *Comparison of /ɾ/-/r/ and /ɾ/-/t/ Discrimination to /ɾ/-/d/ Discrimination by Level*

Level	Pair	Paired samples *t*-test
1	/ɾ/-/r/ and /ɾ/-/d/	$t(14) = 8.496$, $p < .001$
	/ɾ/-/t/ and /ɾ/-/d/	$t(14) = 5.543$, $p < .001$
2	/ɾ/-/r/ and /ɾ/-/d/	$t(14) = 5.109$, $p < .001$
	/ɾ/-/t/ and /ɾ/-/d/	$t(14) = 8.800$, $p < .001$
3	/ɾ/-/r/ and /ɾ/-/d/	$t(14) = 5.538$, $p < .001$
	/ɾ/-/t/ and /ɾ/-/d/	$t(14) = 13.387$, $p < .001$
4	/ɾ/-/r/ and /ɾ/-/d/	$t(14) = 4.119$, $p = .001$
	/ɾ/-/t/ and /ɾ/-/d/	$t(14) = 5.074$, $p < .001$

Note. Results of paired samples (two-tailed) *t*-tests, $\alpha = .025$ (.05 divided by two). Bonferroni adjustment to maintain overall $\alpha = .05$ within each level.

5. Discussion

5.1. The PAM as a predictor of L2 performance

Four of the five word-medial intervocalic contrasts were compared to determine whether the type of contrast as described by the PAM together with acoustic distance in phonological space could predict how well a nonnative contrast would be discriminated. These contrasts represented two main contrasts within the PAM: an *uncategorized vs. categorized* contrast and a *both uncategorizable* contrast. The results of the *uncategorized vs. categorized* contrast were of particular interest as conflicting results have been reported (Best, 1995; Guion et al., 2000).

5.1.1. Uncategorized vs. categorized contrasts

The *uncategorized vs. categorized* contrasts (/ɾ/-/r/ and /ɾ/-[ɾʃ]) both contain an uncategorized phoneme (/ɾ/) that shares a potential mapping with the categorized phoneme (either /r/ or [ɾʃ]). The PAM predicts that all *uncategorized vs. categorized* contrasts will be discriminated very well; however, Guion et al. (2000) found that discrimination was poor when the uncategorized sound was close in phonological space to the categorized sound.

The /ɾ/-/r/ contrast was discriminated well by learners of all levels (ranging from 86.7% to 94.4%), despite the fact that both the /ɾ/ and the /r/ are assimilated to the English /ɹ/. These results differ from those obtained by Guion et al. (2000) and signal a need for further investigation of *uncategorized vs. categorized* contrasts. The results suggest that two sounds that are assimilated to the same L1 sound can be discriminated when they are not assimilated equally well to the same L1 sound. Both the tap and the trill assimilate to the English /ɹ/; however, according to the cross-language mapping data, the /r/ has a higher fit index (2.21) than the /ɾ/ (1.33).

Discrimination of the /ɾ/-/r/ contrast did not improve with proficiency. This suggests that discrimination of an *uncategorized vs. categorized* contrast may not reach a level similar to NS discrimination when the sounds in the contrast are close in phonological space.

The /ɾ/-/r/ contrast was also analyzed to determine whether discrimination was influenced by stress and vocalic environment, factors that have been shown to have an influence in production studies (e.g., Recasens, 1991; Solé, 2002; Waltmunson, 2005). The comparison of the discrimination of *caro/carro* and *quería/querría* contrasts yielded divergent results, with the *quería/querría* contrast being discriminated better by the lower level participants (Levels 0, 1, and 2) and the *caro/carro* contrast being discriminated better by the more advanced participants (Levels 3 and 4). This difference in discrimination of the two contrasts was only significant for Level 1 participants. These results suggest that stress and vocalic environment may influence discrimination by learners, especially at the beginning stages of development. However, further analysis is needed to determine if this is indeed the case because the words used in the current study did not control for vocalic environment or word frequency, two factors that may have influenced the results (see Rose, 2009 and Rose & Darcy, 2009, for further discussion).

The /ɾ/-[ɾʃ] contrast was discriminated at a level significantly above chance, but it was not discriminated as well as the /ɾ/-/r/ contrast by Levels 1, 2, and 4. These results suggest that the tap+ is perceived as being different from a tap, but that a trill with three occlusions is a better exemplar of a trill than the tap+. These results also indicate differences in how learners and NSs are processing the tap+ sound. The NSs were processing all sounds at the word (phonological) level, and consequently discriminated the /ɾ/-/r/ contrast similarly to the /ɾ/-[ɾʃ] contrast. Learners seem to be detecting differences between the tap+ and trill, suggesting that they may also be processing the L2 sounds at an acoustic (phonetic) level.

5.1.2. Both uncategorizable contrasts

The /ɾ/-/t/ contrast was discriminated well by learners at all levels (≥ 90%), despite the fact that both the /ɾ/ and /t/ are uncategorizable and share a mapping as poor exemplars of an English /d/. The fact that this contrast is discriminated well may be because in the majority of cases both phonemes were mapped to two different English phonemes. Spanish /t/ was mapped to English /t/ in 65.4% of instances and Spanish /ɾ/ was mapped to English /ɹ/ in 57.7% of instances. When compared to the other contrasts currently investigated, the /ɾ/-/t/ contrast was also the only contrast to be discriminated at a level similar to NSs as the learners increased in proficiency (Levels 3 and 4). These results are similar to those obtained by Guion et al. (2000).

On the other hand, /ɾ/-/d/ contrast is a *both uncategorizable* contrast in which both L2 phonemes are assimilated to similar phones in the L1, namely the English /ɹ/ and /d/. Learners at all levels were able to discriminate this contrast significantly above chance, and unlike the results of Guion et al. (2000), discrimination of this contrast improved as learners reached higher levels of proficiency. Learners in Levels 1, 2, and 3 demonstrated poor discrimination (ranging from 62.9% to 69.6% accuracy), whereas Level 4 participants were able to discriminate the contrast at 82.5%. However, despite the fact that discrimination of this contrast improved, the discrimination of the /ɾ/-/r/ and /ɾ/-/t/ contrasts was always significantly better than the discrimination of the /ɾ/-/d/ contrast.

Two factors may have influenced the differences in results between the current study and Guion et al. (2000). First, success in discrimination may depend on the degree to which the two L2 phonemes are mapped to the same sounds in the L1. A *both uncategorizable* contrast may not be discriminated if the two L2 phones are mapped equally well to the same phonemes in the L1. But if one phoneme is mapped better than another, differences may be discriminated. This may signal a potential revision to the PAM.[8] Second, the participants' level of proficiency may differ between the two studies. Level 4 participants in the current study may be more advanced than the "high experience" English learners in Guion et al.'s study. Future studies would have to address both factors to determine why some *both uncategorizable* contrasts may be discriminated better as participants become more proficient.

[8] These results also support Best and Tyler's (2007) more recent model, the PAM-L2, which addresses the perceptual assimilation of L2 phones to L1 categories at both the phonetic and phonological level.

5.2. The [r]-[ɾʃ] contrast

The fifth contrast compared two allophones of the trill /r/ ([r]-[ɾʃ]). As predicted, the NSs and most of the learners were not able to discriminate the [r]-[ɾʃ] contrast at a level significantly above chance, suggesting that a trill with one occlusion is perceived similarly to a trill with three occlusions.

Differences in how learners and NSs discriminate the [r]-[ɾʃ] contrast also suggest that they process this contrast differently. The Spanish NSs did not discriminate the [r]-[ɾʃ] contrast (45% accuracy), suggesting that when these two sounds are presented within words, they are both processed at the phonological level as the Spanish /r/. Level 0 participants, on the other hand, do not know Spanish. Their inability in discriminating this contrast (42.5%) suggests they cannot distinguish between these two allophones at an acoustic (phonetic) level. Conversely, learners appear to process this contrast differently. Learners in Levels 1, 2, and 3 all discriminated the [r]-[ɾʃ] contrast noticeably better than the NSs and Level 0 participants (ranging from 47.5% to 59.2%), and Level 4 participants showed significant improvement in discriminating the contrast (72.5%). These results suggest that learners process this contrast at both a word (phonological) and an acoustic (phonetic) level and that consequently they may be more sensitive to allophonic variations than NSs are. Additionally, these results suggest that this sensitivity increases as learners become more proficient.

6. Conclusion

Differences in L2 discrimination are related to how L2 sounds are assimilated into the L1, as the PAM suggests. However, additional investigations need to clarify how the perceptual assimilation of L2 sounds in the L1 affects L2 discrimination. In the current study, even contrasts that mapped onto similar phonemes in the L1 were discriminated. This occurred with both *uncategorized vs. categorized* contrasts (i.e., /ɾ/-/r/ and /ɾ/-[ɾʃ]) and *both uncategorizable* contrasts (/ɾ/-/t/ and /ɾ/-/d/). The phonemes in the /ɾ/-/r/ contrast both mapped onto the English /ɹ/, and the /ɾ/-/r/ contrast was discriminated significantly better than chance by learners at all levels, contrary to the findings of Guion et al. (2000).

The cross-language mapping data suggested that this could be related to how well these two phonemes were assimilated to the same L1 sound (fit index). Likewise, the phonemes in the /r/-/t/ contrast both mapped onto the English /d/. The /r/-/t/ contrast was discriminated with great accuracy by learners at all levels and improved as learners became more proficient, eventually reaching native-like levels of discrimination. The cross-language mapping data suggested that this might be because the two phonemes primarily mapped to two different phonemes in the L1. Finally, the /ɾ/-/d/ contrast is a *both uncategorizable* contrast in which both phonemes (/ɾ/ and /d/) assimilate to the same phonemes in the L1 (/ɹ/ and /d/). As the PAM would suggest, this was the most difficult contrast for learners to discriminate. It was discriminated significantly worse than the other contrasts (/ɾ/-/t/ and /ɾ/-/r/), and remained challenging for learners over time. However, contrary to the results of Guion et al. (2000), learners improved in their ability to discriminate this contrast as they became more proficient in the L2.

The /ɾ/-/r/ contrast was discriminated at a level significantly above chance by all learners, suggesting the late acquisition of the trill is due to articulatory constraints and not the inability to perceive the contrast (e.g., McGowan, 1992; Widdison, 1998). Results also suggest that vocalic environment and stress may influence the ability to discriminate this contrast, especially at the beginning stages of development. Future studies are needed to determine whether this is the case.

Finally, the tap+ variant was discriminated by all participants when compared to a tap, but was not discriminated by NSs and most learners when compared to a trill, suggesting that trills with one clear occlusion are perceived similarly to trills with three clear occlusions. Differences in the discrimination of the [r]-[ɾʃ] contrast also signaled potential differences in how L2 sounds are processed by learners and NSs. NSs were not able to discriminate this contrast at the word level. However, learners appeared to process this allophonic contrast at both a word and acoustic level. Future investigations should continue to look at allophonic variants as a means of better understanding differences between how learners and NSs process L2 sounds.

References

Aoyama, Katsura, Flege, James E., Guion, Susan G., Akahane-Yamada, Reiko & Yamada, Tsuneo (2004). Perceived phonetic dissimilarity and L2 speech learning: The case of Japanese /r/ and English /l/ and /r/. *Journal of Phonetics, 32*, 233–50.

Best, Catherine T. (1990). Adult perception of nonnative contrasts differing in assimilation to native phonological categories. *Journal of the Acoustical Society of America, 88,* S177.

Best, Catherine T. (1995). A direct realist perspective on cross-language speech perception. In Winifred Strange (Ed.), *Speech perception and linguistic experience: Theoretical and methodological issues in cross-language speech research* (pp.167–200). Timonium, MD: York.

Best, Catherine T., McRoberts, Gerald W. & Goodell, Elizabeth (2001). Discrimination of non-native consonant contrasts varying in perceptual assimilation to the listener's native phonological system. *The Journal of the Acoustical Society of America, 109*(2), 775–794.

Best, Catherine T., McRoberts, Gerald. W., & Sithole, Nomathemba M. (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. & Strange, Winifred (1992). Effects of phonological and phonetic factors on cross-language perception of approximants. *Journal of Phonetics, 20,* 305–330.

Best, Catherine T. & Tyler, Michael D. (2007). Nonnative and second-language speech perception: Commonalities and complementarities. In Ocke-Schwen Bohn & Murray J. Munro (Eds.), *Language Experience in Second Language Speech Learning: In Honor of James Emil Flege* (pp. 13–34). Philadelphia: John Benjamins.

Blecua, Beatriz (2001). *Las vibrantes del español: Manifestaciones acústicas y procesos fonéticos* (Unpublished doctoral dissertation), Universidad Autónoma de Barcelona, Barcelona, Spain.

Bonet, Eulàlia & Mascaró, Joan (1997). On the representation of contrasting rhotics. In Fernando Martínez-Gil & Alfonso Morales-Font (Eds.), *Issues in the phonology and morphology of the major Iberian languages*. (pp. 559–593). Washington, DC: Georgetown University Press.

Bosch, Laura (1983). El desarrollo fonológico infantil: Una prueba para su evaluación. *Anuario de Psicología, 28,* 87–114.

Bradley, Travis G. & Willis, Erik W. (2008, April). Rhotic contrast and neutralization in Veracruz Mexican Spanish. Paper presented at the *38th Linguistic Symposium on Romance Languages*, Urbana-Champaign, IL.

Carballo, Gloria & Mendoza, Elvira (2000). Acoustic characteristics of trill productions by groups of Spanish children. *Clinical Linguistics & Phonetics, 14*(8), 587–601.

Clark, John & Yallop, Colin (1990). *An introduction to phonetics and phonology*. Cambridge, MA: Basil Blackwell.

Face, Timothy L. (2006). Intervocalic rhotic pronunciation by adult learners of Spanish as a second language. In Carol A. Klee & Timothy L. Face (Eds.), *Selected proceedings of the 7th conference on the acquisition of Spanish and Portuguese as first and second languages* (pp. 47–58). Somerville, MA: Cascadilla Proceedings Project.

Flege, James E. (1995). Second language speech learning: Theory, Findings, and Problems. In Winifred Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 233–277). Timonium, MD: York Press.

Flege, James E., Munro, Murray J. & McKay, Ian R. A. (1995). Effects of age of second-language learning on the production of English consonants. *Speech Communication*, *16*, 1–26.

Goto, Hiromu (1971). Auditory perception by normal Japanese adults of the sounds 'l' and 'r'. *Neuropsychologia, 9*, 317–323.

Guion, Susan G., Flege, James E., Akahane-Yamada, Reiko & Pruitt, Jesica C. (2000). An investigation of current models of second language speech perception: The case of Japanese adults' perception of English consonants. *Journal of the Acoustical Society of America, 107*(5), 2711-2724.

Hayes-Harb, Rachel (2007). Lexical and statistical evidence in the acquisition of second language phonemes. *Second Language Research, 23*, 65–94.

Heaton, J. B. (1966). *Composition through pictures*. Essex, UK: Longman.

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

Iverson, Paul, Kuhl, Patricia K., & Akahane-Yamada, Reiko (2003). A perceptual interference account of acquisition difficulties for non-native phones. *Cognition*, *87*, B47–B57.

Jimenez, Beatrice C. (1987). Acquisition of Spanish consonants in children aged 3-5 years, 7 months. *Language, Speech and Hearing Services in Schools, 18,* 357–363.

Kuhl, Patricia K. (1993). Innate predispositions and the effects of experience in speech perception: the native language magnet theory. In Bénédicte de Boysson-Bardies, Scania de Schonen, Peter Jusczyk, Peter McNeilage, & John Morton (Eds.), *Developmental neurocognition: Speech and face processing in the first year of life* (pp. 259–274). Dordrecht, Netherlands: Kluwer Academic Publishers.

Ladefoged, Peter (2001). *A course in phonetics* (4th ed.). Boston: Heinle and Heinle.

Ladefoged, Peter & Maddieson, Ian (1996). *The sounds of the world's languages.* Oxford: Blackwell.

Lewis, Julie A. (1999). Perceptual strategies due to trill-vowel coarticulatory effects in Spanish. *Proceedings of the XIVth International Congress of Phonetic Sciences: ICPhP 99 San Francisco, 3,* 2477–2480.

Lewis, Anthony M. (2004). Coarticulatory effects on Spanish trill production. In Augustine Agwuele, Willis Warren, & Sang-Hoon Park (Eds.), *Proceedings of the 2003 Texas Linguistics Society Conference* (pp. 116–127). Somerville, MA: Cascadilla Proeedings Project.

Lindau, Monda (1985). The story of /r/. In Vicki A. Fromkin (Ed.), *Phonetic linguistics: Essays in honor of Peter Ladefoged* (pp. 157–168). Orlando, FL: Academic Press.

Lipski, John M. (1990). Spanish taps and trills: Phonological structure of an isolated opposition. *Folia Linguistica, 24,* 153–174.

Major, Roy C. (1986). The ontogeny model: Evidence from L2 acquisition of Spanish *r*. *Language Learning, 36,* 453–504.

Mayer, Mercer (1969). *Frog, where are you?* New York: Penguin.

McGowan, Richard S. (1992). Tongue-tip trills and vocal tract wall compliance. *Journal of the Acoustical Society of America*, *95*(5), 2903–2910.

Monnot, Michel & Freeman, Michel (1972). A comparison of Spanish single-tap /r/ with American /t/ and /d/ in post-stress intervocalic position. In Albert Valdman (Ed.), *Papers in linguistics and phonetics to the memory of Pierre Delattre* (pp. 409–416). The Hague: Mouton.

Olive, Joseph P., Greenwood, Alice & Coleman, John (1993). *Acoustics of American English speech: A dynamic approach.* New York: Springer-Verlag.

Pisoni, David B. & Lively, Scott E. (1995). Variability and invariance in speech perception. A new look at some old problems in perceptual learning. In Winifred Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 433–459). Timonium, MD: York Press.

Quilis, Antonio (1993). *Tratado de fonética y fonología españolas.* Madrid: Editorial Gredos.

Recasens, Daniel (1991). On the production characteristics of apicoalveolar taps and trills. *Journal of Phonetics, 19,* 267–280.

Reeder, Jeffrey T. (1998). English speakers' acquisition of voiceless stops and trills in L2 Spanish. *Texas Papers in Foreign Language Education*, *3*(3), 101–108.

Rose, Marda (in press). Intervocalic tap and trill production in the acquisition of Spanish as a second language. *Studies in Hispanic and Lusophone Linguistics.*

Rose, Marda (2009). Cross-language identification of Spanish consonants in English. Manuscript submitted for publication.

Rose, Marda & Darcy, Isabelle (2009). The effect of syllable stress and L2 experience on the cross-language identification of L2 Spanish consonants in English. Manuscript in preparation.

Solé, Maria-Josep (2002). Aerodynamic characteristics of trills and phonological patterning. *Journal of Phonetics, 30*, 655–688.

Takagi, Naoyuki (2002). The limits of training Japanese listeners to identify English /r/ and /l/: Eight case studies. *Journal of the Acoustical Society of America, 111*, 2887–2896.

Waltmunson, Jeremy C. (2005). *The relative degree of difficulty of L2 Spanish /d, t/, trill, and tap by L1 English speakers: Auditory and acoustic methods of defining pronunciation accuracy.* Unpublished doctoral dissertation, University of Washington, Seattle.

Werker, Janet F. & Tees, Richard C. (1984). Phonemic and phonetic factors in adult cross-language speech perception. *Journal of the Acoustical Society of America*, *75*, 1866–1878.

Widdison, Kirk A. (1998). Phonetic motivation for variation in Spanish trills. *Orbis: Bulletin international de documentation linguistique*, *40*, 51–61.

Willis, Erik W. (2006). Trill variation in Dominican Spanish: An acoustic examination and comparative analysis. In Nuria Sagarra & Almeida Jacqueline Toribio (Eds.), *Selected proceedings of the 9th Hispanic Linguistics Symposium* (pp. 87–100). Somerville, MA: Cascadilla Proceedings Project.

Selected Proceedings of the 2008 Second Language Research Forum: Exploring SLA Perspectives, Positions, and Practices

edited by Matthew T. Prior, Yukiko Watanabe, and Sang-Ki Lee

Cascadilla Proceedings Project Somerville, MA 2010

Copyright information

ISBN 978-1-57473-439-3 library binding

Ordering information

Orders for the library binding edition are handled by Cascadilla Press.
To place an order, go to 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, e-mail: sales@cascadilla.com

Web access and citation information

This entire proceedings can also be viewed on the web at 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:

Rose, Marda. 2010. Differences in Discriminating L2 Consonants: A Comparison of Spanish Taps and Trills. In *Selected Proceedings of the 2008 Second Language Research Forum*, ed. Matthew T. Prior et al., 181-196. Somerville, MA: Cascadilla Proceedings Project. www.lingref.com, document #2393.