Wendy Herd*, Joan Sereno and Allard Jongman

Cross-modal priming differences between native and nonnative Spanish speakers

Abstract: Training has been shown to improve American English speakers' perception and production of the Spanish /ɾ, r, d/ contrast; however, it is unclear whether successfully trained contrasts are encoded in the lexicon. This study investigates whether learners of Spanish process the /ɾ, r, d/ contrast differently than native speakers and whether training affects processing. Using a cross-modal priming design, thirty-three Spanish learners were compared to ten native Spanish speakers. For native speakers, auditory primes with intervocalic taps (like [koɾo]) resulted in faster reaction times in response to matching visual targets (like *coro*) than to orthographically and phonemically similar targets (like *corro* and *codo*). American English speakers' reaction times were not affected by the relationship between primes and targets before training. After training, trainees responded more quickly to matching targets than to mismatching /ɾ/-/r/ prime-targets (e.g., [koɾo] followed by *corro*) while controls' reaction time patterns did not change. This indicates that native Spanish speakers and Spanish learners process words containing the /ɾ, r, d/ contrast differently and that improvements from training can be encoded in the lexicon.

Keywords: cross-modal priming, high variability training, second language acquisition, Spanish acquisition, American English

DOI 10.1515/shll-2015-0005

1 Introduction

When second language learners begin acquiring a new language, they are often faced with phonemes that do not exist in their native language. Failing to acquire these sounds can lead to being perceived as a nonnative speaker or, even worse, to misunderstandings. While high variability training has been shown to improve the ability of learners to distinguish new phonemes in forced-choice perception tasks and when reading lists of words, it is unclear

***Corresponding author: Wendy Herd,** Mississippi State University, Mississippi State, MS, USA, E-mail: wherd@english.msstate.edu
Joan Sereno, Allard Jongman, University of Kansas, Lawrence, KS, USA

whether training affects the online perception, production, and processing of words in the second language (L2).

1.1 Acquisition of /ɾ, r, d/ by American English learners of Spanish

When native speakers of American English begin learning Spanish, their acquisition of native-like perception and pronunciation can be hampered by the intervocalic /ɾ, r, d/ contrast because they must reanalyze two sounds that exist in their native language and learn a new sound in order to acquire the contrast in Spanish. The tap [ɾ] exists as an allophonic variant of /d/ and /t/ in American English. Native speakers of American English produce [ɾ] when an underlying /d/ or /t/ occurs in intervocalic post-tonic position, so the English word 'rider' is most often produced with a medial [ɾ] and not [d]. Since flapping for intervocalic /d/ and /t/ is such a productive allophonic process in American English, with flapping reported to occur between 94% and 99% of the time in the post-tonic intervocalic position, it is likely that American English learners of Spanish will have difficulties acquiring the /ɾ, d/ contrast (Byrd 1994; Connine 2004; Herd et al. 2010; Patterson and Connine 2001; Zue and Laferriere 1979). With respect to perception, Boomershine et al. (2008) found that monolingual English speakers rated auditorily-presented /ɾ, d/ minimal pairs as more similar than native Spanish speakers and discriminated between the two sounds more slowly than native Spanish speakers, suggesting American English learners of Spanish may confuse the phones. Additionally, Herd et al. (2013) found that American English learners of Spanish before training exhibited more difficulty perceiving the /ɾ, d/ contrast (70% accuracy) than the /ɾ, r/ contrast (82%) or the /r, d/ contrast (96%). When attempting to produce /ɾ/ in Spanish, intermediate and advanced American English learners of Spanish produce the phone correctly only 49% and 79% respectively, significantly less often than native Spanish speakers (Face 2006).

American English learners of Spanish also experience difficulties acquiring the Spanish trill /r/, a phone that does not exist in English. As mentioned above, American English learners of Spanish were not able to distinguish the /ɾ, r/ contrast as accurately as the /r, d/ contrast, indicating American English learners of Spanish experience difficulty perceiving /r/ in some contexts (Herd et al. 2013). American English learners of Spanish experience even more difficulty producing /r/, with intermediate and advanced learners only producing /r/ correctly in about 5% and 27% of cases respectively (Face 2006; Rose 2010).

While the pronunciation of the /r/ improves with experience, even advanced learners in Spanish doctoral programs produce /r/ accurately in only about 80% of cases, significantly less often than native speakers (Johnson 2008; Rose 2010).

Although /d/ exists as a phoneme in both English and Spanish, American English learners of Spanish experience difficulties acquiring /d/ in intervocalic contexts due to spirantization, a highly productive process in Spanish in which intervocalic voiced stops /b, d, g/ are produced as fricatives [β, ð, ɣ] or, more accurately, approximants [β̞, ð̞, ɣ̞]. Spanish intervocalic spirantization occurs in a very similar environment to American English flapping, which could affect American English speakers' acquisition of the /ɾ, d/ contrast. For example, American English speakers might produce Spanish /d/ as [ɾ], both failing to spirantize /d/ and neutralizing the /ɾ, d/ contrast. Previous research has demonstrated that English speakers have more difficulty perceiving a difference between /ɾ, d/ in similarity rating tasks and discrimination tasks (Boomershine et al. 2008) and distinguishing between the two in identification tasks (Herd et al. 2013). Acoustic analyses also indicate that students in their third or fourth semester of university-level Spanish spirantize intervocalic /d/ only about 6–31%, significantly less often than native speakers of Spanish (Face and Menke 2009; González-Bueno 1995; Herd et al. 2013; Waltmunson 2005; Zampini 1993; 1994).

1.2 High variability training of L2 contrasts

High variability perceptual and production training paradigms have been shown to improve learners' acquisition of L2 contrasts. The training involves having learners complete a series of tasks in response to items that contain the target L2 contrast. For example, during high variability perceptual training, learners might hear [koro] via headphones and then have to decide whether they heard visually presented *coro* or *corro* by selecting the correct response. The training is referred to as high variability because learners are presented with training stimuli produced by a variety of speakers, in a variety of contexts, and in a variety of tokens. High variability perceptual training has been shown to improve the perception of the /ɹ/-/l/ contrast by Japanese learners of English (Bradlow et al. 1997; Logan et al. 1991) and the perception of the four tones by English learners of Mandarin (Wang et al. 1999). Additionally, although participants were only trained in perception, they also improved their production of the trained contrasts immediately following training (Bradlow et al. 1997; Wang et al. 2003) and three months after training (Bradlow et al. 1999).

During high variability production training, learners might see the word *corro* with a native speaker's waveform and spectrogram that visually represent the target pronunciation of the word. Learners would then record themselves saying *corro*, comparing their waveform and spectrogram to those of the native speaker with the purpose of producing the word in a more native-like manner. Like high variability perceptual training, high variability production training includes training stimuli produced by a variety of speakers, in a variety of contexts, and in a variety of words. Hirata (2004) found that high variability production training improved both the production and perception of pitch and durational contrasts in Japanese (Hirata 2004).

In a study directly related to Spanish, Herd et al. (2013) demonstrated that both high variability perception training and high variability production training significantly improved learners' ability to both perceive and produce the /ɾ, r, d/ contrast. In other words, both of these training paradigms improved participants' abilities in the specific modality being trained and transferred that improvement to the other untrained modality. High variability combination training, a training paradigm in which participants completed half of the sessions in perceptual training and half in production training, also resulted in improved production of the contrasts but not improved perception. While these results support the efficacy of high variability training for the /ɾ, r, d/ contrast, they are based on offline identification and productions tasks and do not reflect whether training affected how the learners processed the L2 contrasts.

1.3 Repetition and minimal pair lexical priming effects

While offline production and identification tasks can be used to show whether learners are able to produce and perceive L2 contrasts when the learners are overtly focused on the contrasts, these task types do not shed light on whether participants process L2 words containing those contrasts in a native-like manner. To that end, lexical priming tasks can be used to measure how American English learners of Spanish and native Spanish speakers process the /ɾ, r, d/ contrast. During a priming task, participants are presented with a series of words. Facilitatory priming occurs when activation due to an earlier presented stimulus, the prime, favorably affects the reaction time and/or accuracy of responses to a subsequent and related stimulus, or target. Inhibition results when the prime adversely affects reaction time and/or accuracy in response to the target. In brief, facilitation occurs when the target of a prime-target pair is responded to more quickly and/or accurately than targets from other prime-target pairs, and inhibition occurs when the target of a prime-target pair is

Table 1: Examples of prime-target pair types.

Prime	Identical target	Rhyming target	Different final phoneme target	Unrelated target
cat	cat	pat	cab	mop

responded to more slowly and/or less accurately than targets from other prime-target pairs.

Previous native language cross-modal priming research has found that auditory primes followed by identical visual targets are responded to significantly more quickly than primes followed by rhyming targets with different onsets (Marslen-Wilson and Zwitserlood 1989), by competitor targets differing in a final phoneme (Marslen-Wilson 1990), and by unrelated targets (Gaskell and Marslen-Wilson 2002; Grainger et al. 2001; Holcomb et al. 2005). Illustrative examples of the prime-target lexical pairs can be found in Table 1. These facilitatory effects are more robust at longer interstimulus intervals (ISIs), the amount of time that elapses between the offset of the prime and the onset of the target (Holcomb et al. 2005); however, effects are still significant when visual targets are displayed at the offset of the auditory primes (Gaskell and Marslen-Wilson 2002; Holcomb et al. 2005). While identical prime-target pairs result in facilitation because identical targets are responded to more quickly than unrelated targets, prime-target pairs that differ in one phoneme result in inhibition because targets are responded to more slowly than unrelated targets. For example, Allen and Badecker (2002) reported inhibitory effects such that auditory primes followed by visual targets differing from the prime by only one phoneme resulted in significantly slower reaction times than those followed by unrelated targets.

Using a within modality auditory priming design, Pallier et al. (2001) reported that fluent Catalan-Spanish and Spanish-Catalan bilinguals performed differently on an auditory lexical decision task. Spanish-dominant bilinguals processed sound contrasts that exist in Catalan but not Spanish (e.g., /e/-/ɛ/) as more similar, with auditory primes containing /ɛ/ resulting in significantly faster reaction times in response to minimal pair auditory targets containing /e/. This facilitatory priming was absent in Catalan-dominant bilinguals. Similar effects have been reported with the /ɛ/-/æ/ contrast for Dutch learners of English (Broersma and Cutler 2011) and the /y/-/u/ contrast for English learners of French (Darcy et al. 2012). The results of Pallier et al. (2001) suggest that, even with fluency or advanced proficiency in the less dominant language, learners might not process new sounds like native speakers. However, while Darcy et al. (2012) reported that /y/ facilitated /u/ and vice versa for intermediate learners of French, the same was not true for advanced learners of French, who processed /y/ and /u/ similarly to native speakers of French. This leaves open the question of whether learners who have perceptually and productively acquired a new L2 contrast also process those new contrasts like native speakers.

1.4 Current study

Herd et al. (2013) demonstrated that perception, production, and combination training improved American English speakers' ability to distinguish the three-way /ɾ, r, d/ contrast in Spanish, but the documented improvements were limited to overtly identifying the contrasts in a two-alternative forced-choice paradigm and pronouncing them when reading a word list. Using a cross-modal priming design, the present study investigates whether improvements due to training also result in lexical processing differences.

Based on previous priming research, one can predict that native speakers of Spanish presented with the auditory stimulus [koɾo] would exhibit the shortest reaction times when responding to the visually identical target *coro* ([koɾo]) due to facilitation while they would exhibit the longest reaction times when responding to a minimal pair visual target differing in one phoneme like *corro* ([koro]) or *codo* ([koð̞o]) due to inhibition. On the other hand, native speakers of American English, who perceive [ɾ] as an allophone of /d/, presented with the same auditory stimulus ([koɾo]) would be expected to exhibit a different pattern, with shorter reaction times for the visually presented *codo* ([koð̞o]) and longer reaction times for visually presented *coro* ([koɾo]) and *corro* ([koro]). A cross-modal priming task administered to participants before and after training could therefore be used as a diagnostic to evaluate whether English learners of Spanish exhibit reaction times in a pattern more similar to native Spanish speakers after successful training.

This study aims to determine if perceptual, production and combination training, which successfully improved learners' perception and production of the /ɾ, r, d/ contrast, can also change the way learners process lexical stimuli containing the contrast. The cross-modal priming data of 33 American English learners of Spanish both before and after training are compared to that of 10 native Spanish speakers to determine how untrained learners of Spanish and native speakers of Spanish process Spanish lexical items and to discover how successful high variability training affects that processing.

2 Methods

2.1 Participants

The productions of one female native Spanish speaker from Peru (F1) were recorded for the purpose of creating the cross-modal priming task. The stimuli produced by F1 are described in Section 2.2. An additional twelve native Spanish speakers from Peru, Costa Rica, Honduras, Ecuador, and Mexico (4 males, 8 females) and with a mean age of 29 were recruited to participate as native controls for the cross-modal priming task. All of the native Spanish-speaking participants were living in the United States at the time of the study. The group of native controls, henceforth called native Spanish speakers, completed the cross-modal priming task only once and did not participate in any training sessions. It is important to note here that many varieties of Spanish exist; however, the phones investigated in this study (i.e., [ɾ, r, ð̞]) are present in most varieties of Spanish.

Forty-two native speakers of American English (9 male, 33 female) with a mean age of 20 were recruited to participate in the cross-modal priming task. These students completed three to four years of high school Spanish and were enrolled in their second or third semester of college Spanish at the University of Kansas. After completing the cross-modal priming pretest and additional perception and production tasks that are not reported here, participants were randomly assigned to four groups: perception, production, combination, and control. Participants in the perception, production, and combination groups underwent six training sessions, which were scheduled on six different days within a two to three week period. Participants in the control group did not participate in any training sessions. All participants, including those in the control group, then completed the cross-modal priming posttest at the end of the two to three week period. A brief description of the perception, production, and combination training sessions follows. For a detailed description of the training paradigms, the additional perception and production tasks, and the impact training had on perception and production, please see Herd et al. (2013).

Participants in the perception, production, and combination groups practiced thirty different minimal pairs contrasting only one pair of sounds (e.g., /ɾ/-/d/, or /ɾ/-/r/, or /r/-/d/) read by two different speakers (from a total of six native Spanish speakers from Peru) each session. A total of 120 tokens (30 minimal pairs $\times$ 2 contrasts $\times$ 2 speakers) were practiced during each training session. The same sound pair and the same speakers were never repeated in consecutive sessions. Participants in the perception group were trained to perceive the difference between the three critical sounds, but they were not trained to produce the difference. Participants in the perception group completed a different 120-item two-alternative forced-choice identification task with immediate feedback each session. For example, if a participant were practicing the /r/-/d/ pair, the participant might hear [koro], and then the participant would see *codo* and *corro* on the computer screen. If the participant correctly chose *corro,* the message, "Right! That was *corro*, Let's hear *corro* again," would appear. If the participant incorrectly chose *codo,* the message, "Oops! That was *corro*, Let's hear *corro* again," would appear. Conversely, participants in the production group participated in the same amount of training but were only trained to produce the critical sounds. For example, if a participant in the production group were practicing the /r/-/d/ pair, the participant would see the word *corro* presented with the waveform and spectrogram of the word as produced by a native Spanish speaker. The participant would then be prompted to record a version of *corro*, producing a second waveform and spectrogram. The participant would then compare the waveforms and spectrograms of the two version of *corro*. Participants were instructed to attempt different pronunciations until they were able to match the waveforms and spectrograms of the native speakers. Participants in the combination group attended the same total number of sessions but practiced using the perception training method for three sessions and the production training method for three sessions, alternating between perception and production from session to session. Participants in the control group received no training. All three training paradigms resulted in improved production of the trained contrasts, with the largest gains for the production of the /r/. With respect to perception, both perception and production trainees improved their perception of the /ɾ/-/r/ contrast, but only perception trainees improved their perception of the /ɾ/-/d/ contrast. The controls evinced no improvements in production or perception.

All participants filled out a human consent form and a dialect questionnaire before completing any experimental sessions. All participants were paid $10 per hour for their participation, and the learners of Spanish, who were required to visit the lab from two to eight times depending on group, were paid an additional $20 bonus upon the completion of all sessions.

2.2 Stimuli

Participant F1, a female speaker from Peru, was recorded reading a list of 24 words and non-words containing intervocalic taps in the form CVɾV and 6 unrelated CVCV words. These were embedded in a list of 252 other experimental items and fillers to draw attention away from the purpose of the list. The recording took place in an anechoic chamber, using a solid-state recorder

(Marantz PMD671) and Electro-Voice 767a microphone. These stimuli were then used as the auditory primes (and visual targets in one-third of the cases) for the cross-modal priming practice and experimental blocks. The experimental cross-modal priming stimuli were balanced for word frequency based on values from the *Corpus del Español* (Davies 2002) to ensure that prime-target word pairs [t (11) = 0.100, p = 0.922] and different contrasts amongst the visual targets [F (2, 21) = 1.149, p = 0.336] did not differ in frequency. To dissuade participants from guessing that auditory nonwords were always followed by visual nonwords, five of the non-word targets were preceded by lexical items and two were preceded by names. Due to the nature of the cross-modal priming task, it was important that lexical auditory primes and their paired targets be familiar to the American English learners of Spanish as well as the native speakers, who were from different Spanish-speaking regions. To this end, all of the lexical targets and most of the lexical primes, with the exception of *varo*, were high frequency words. Two instructors of Spanish also confirmed that all of the lexical targets and most of the lexical primes, with the exception of *varo* and *moro*, would be familiar to students in the classes from which participants were recruited. Finding familiar primes and targets proved particularly difficult since the intervocalic /ɾ/-/r/ distinction in Spanish is limited to approximately 30 minimal pairs. Table 2 lists the lexical primes and targets included in the task.[1]

2.3 Procedure

The 42 English learners of Spanish were randomly assigned to four groups: perception, production, combination, and controls as described in 2.1 Participants. The four groups of learners and the group of 12 native Spanish speakers completed the cross-modal priming task, which included a block of 6 unrelated practice items and a block of 24 experimental items, immediately after giving informed consent and completing a dialect survey. Participants completed the task presented via Paradigm (Tagliaferri 2011) while individually seated at a partitioned computer station and wearing Sony MDR-V6 headphones. Participants in the perception, production, and combination groups completed the cross-modal priming task both pre- and post-training, and two to three weeks elapsed between the pre- and post-tasks. Likewise, participants in the

1 The following nonword prime-target pairs were used: cere-cere, dera-dera, fara-fara, liro-liro, luro-lurro, rara-rarra, Sara-zarra, ore-horre, flora-floda, paro-pado, fero-fedo, curo-cudo. As noted in Section 2.2, five of the non-word targets were preceded by lexical items and two were preceded by names.

Table 2: Auditory primes and visual targets used during the cross-modal priming task.

Auditory prime	Gloss	Word frequency	Visual target	Gloss	Word frequency	Target type
para	for	237,280	para	for	237,280	/ɾ/
duro	hard-m	1,795	duro	hard-m	1,795	
miro	I look	565	miro	I look	565	
hora	hour/time	10,433	hora	hour/time	10,433	
pero	but	131,191	perro	dog	1,992	/r/
coro	choir	1,391	corro	I run	414	
caro	expensive	849	carro	car	1,102	
varo	peso	34	barro	I sweep	774	
toro	bull	1,068	todo	all	74,189	/d/
cara	face	7,134	cada	each	27,660	
moro	moor	863	modo	style	14,442	
dura	hard-f	1,847	duda	doubt	8,028	

control group completed the cross-modal priming task twice with a two- to three-week break between tasks.

Each trial of the practice and experimental blocks consisted of an auditory prime which was followed by a 500 ms ISI and then a visual target which remained on screen until a lexical decision response was submitted. After completing the six-item practice session, participants were encouraged to ask questions and to repeat the practice if necessary. When participants were ready, they initiated the experimental block by pressing the space bar. During the experimental block, participants heard a series of Spanish words and nonwords that contained an intervocalic tap /ɾ/, like *coro* 'choir' [koɾo]. These were the auditory primes. After the ISI, participants saw the orthographic representation of either the identical item (like *coro* 'choir') or words that differed from the prime by only one phoneme (like *corro* 'I run'). These were the visual targets. The participants' task was to decide whether the visual targets were words. They responded by pressing buttons clearly labeled "WORD" or "NONWORD" on a button box. Participants always pressed the "WORD" button with their dominant hand. Participants were instructed to respond as quickly and accurately as possible.

2.4 Predictions

Reaction time measurements collected using a cross-modal priming design can be used to determine whether trainees have encoded the trained phones [ɾ, r, ð̞] into their lexicons. For example, is *codo* 'elbow' encoded as [koð̞o] rather than [koɾo] and is *coro* 'choir' encoded as [koɾo] after training? If these changes occur as a result of training, trainees will activate *coro,* not *codo,* from their lexicon after hearing the auditory prime [koɾo], resulting in both faster lexical decision reaction times for the visual target *coro* due to facilitation and slower reaction times for *codo* due to inhibition. In other words, posttest trainees will exhibit reaction time patterns more similar to those expected of native Spanish speakers. Likewise, if trainees have categorized the tap [ɾ] and trill [r] as separate phonemes and encoded these separately in words like *coro* 'choir' and *corro* 'I run', reaction times will be slower in response to the visual target *corro* after the auditory prime [koɾo] due to inhibition. However, if trainees do not exhibit reaction time patterns similar to native speakers after training, this will indicate that the gains from training have not yet been encoded in the lexicon and that these learners do not yet process the sounds like native Spanish speakers.

3 Results

The following analyses are based on reaction times and response accuracy to lexical items only. Mean reaction times and accuracy rates in each condition were calculated for each participant. To ensure that reaction times were truly measures of lexical processes, only participants with a mean overall accuracy of 80% were included in further analyses. This resulted in the exclusion of 9 learners of Spanish. Fortunately, this loss of participants was constant across groups, such that the resultant perception, production, and combination groups each included 8 participants and the control group included 9 participants. In addition, two native Spanish speakers were excluded as their mean accuracy was 50%, having completed the lexical decision task as a same-different task, bringing the number of participants in the native speaking group to 10. Correct responses with reaction times beyond two standard deviations of each participant's mean were also excluded. This resulted in a loss of 33 of the 792 Spanish learner data points [2 pre/post × 12 tokens × 33 Spanish learners = 792] and 5 of the 120 native Spanish speaker data points [12 tokens × 10 native Spanish speakers = 120] or about 4.2% of the data. To test for a speed-accuracy tradeoff, bivariate correlations were computed for reaction time and accuracy rates both for all participants and for separate groups. In all cases, reaction time was not significantly correlated with accuracy, indicating there was no speed-accuracy trade-off (all $r < 0.25$, all $p > 0.1$). The results of the perception, production, and combination trainees have been combined and reported as one group (i.e., trainees) in further analyses. This was possible because

Spanish learners were assigned randomly to training groups, because all training groups made gains in the overall perception and production of the /ɾ, r, d/ contrast, and because the three different training groups did not differ significantly in cross-modal priming reaction time or accuracy before or after training.

3.1 Pre-training

Analyses were carried out on pre-training data to determine if participants responded with faster reaction times and better accuracy to matching pairs (e.g., auditory prime [paɾa] followed by visual target *para*) than to mismatching pairs (e.g., auditory prime [koɾo] followed by visual target *corro* or auditory prime [toɾo] followed by visual target *todo*) before training. Specifically, native Spanish speakers were predicted to respond more quickly and accurately to matching prime-target pairs than to mismatching pairs. This pattern was not predicted for learners of Spanish. Instead, the American English-speaking learners of Spanish, for whom [ɾ] is an allophonic variant of /d/, were predicted to respond more quickly and accurately to mismatching /ɾ/-/d/ pairs like auditory [toɾo] followed by visual *todo* than to matching pairs. Repeated measures ANOVAs with three levels of Target Type (/ɾ/, /r/ and /d/) and with Group (controls, trainees, and native Spanish speakers) as a between-subjects factor were conducted separately on reaction time and accuracy.

The repeated measures ANOVA carried out on reaction time revealed a main effect of Target Type [F (2, 80) $=$ 8.161, p $=$ 0.001] and an interaction between Target Type and Group [F (4, 80) $=$ 4.147, p $=$ 0.023] but no main effect of Group. Pairwise comparisons using the Bonferroni correction found that, overall, participants responded more quickly to visual targets containing the orthographic representation of /ɾ/ at 706 ms than to visual targets containing /r/ at 805 ms (p $=$ 0.001) or /d/ at 791 ms (p <0.001).

In order to interpret the Target Type $\times$ Group interaction, a series of 9 paired-sample t-tests were conducted comparing the reaction time of the 3 groups when responding to the 3 target types. Using the Bonferroni correction to control for familywise error, the significant p-value threshold was set at p < 0.0056. When hearing an auditory prime like [paɾa], native speakers responded more quickly to visual targets containing the orthographic representations of /ɾ/ at 756 ms than to visual targets containing /r/ at 952 ms [t (9) $=$ 4.276, p $=$ 0.002] or /d/ at 962 ms [t (9) $=$ 4.770, p $=$ 0.001]. There were no significant differences in the reaction time of the Spanish learners (both trainees and controls) in response to different target types at pre-training. As predicted, native Spanish speakers' responses to matching prime-target pairs were significantly faster than to mismatching

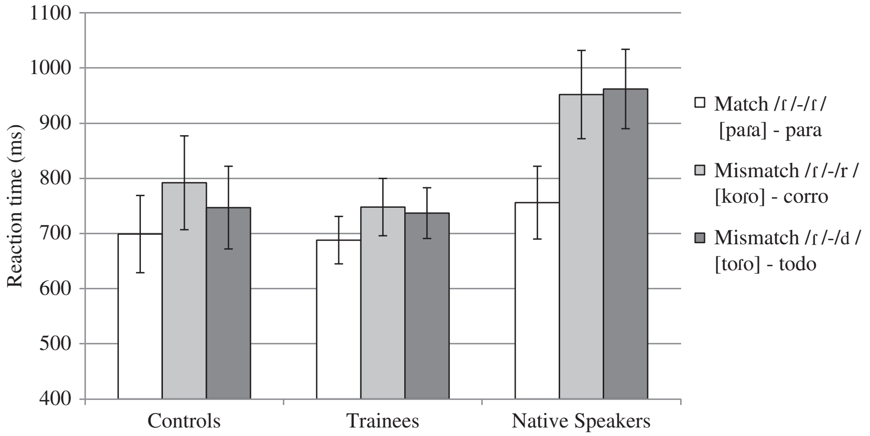

Figure 1: Mean pre-training lexical decision reaction time (ms) organized by Target Type and Group. Error bars indicate standard error.

prime-target pairs. Conversely, learners of Spanish did not exhibit reaction time differences at pre-training in response to matching versus mismatching prime-target pairs. The mean pre-training reaction times of the three groups in response to the three target types are presented in Figure 1.

The repeated measures ANOVA carried out on accuracy revealed a main effect of Target Type [F (2, 80) = 8.901, p <0.001] but no effect of Group and no interaction between Target Type and Group. Pairwise comparisons using the Bonferroni correction found that, overall, participants responded with significantly better accuracy to visual targets containing the orthographic representation of /ɾ/ (e.g., *para*) at 99% than to visual targets containing /r/ (e.g., *corro*) at 88% (p < 0.001) and /d/ (e.g., *todo*) at 91% (p = 0.014). This suggests that response accuracy to words containing /ɾ/ was facilitated by the preceding auditory primes which also contained /ɾ/. While the mean accuracy rates reported in Table 3 suggest the accuracy differences between words containing /ɾ/ and those containing /r/ or /d/ were driven by the learners of Spanish, the Group and Target Type × Group differences did not reach significance in the accuracy data.

3.2 Post-training

Analyses were carried out on the post-training data to determine if Spanish learners responded with faster reaction times and better accuracy to matching pairs (e.g., auditory prime [paɾa] followed by visual target *para*) than to

Table 3: Mean pre-training lexical decision accuracy rate in percentages (and standard deviation) organized by Group and Target Type.

Target Type	/ɾ/		/r/		/d/	
Examples	**prime**	**target**	**prime**	**target**	**prime**	**target**
	[paɾa]	***para***	**[koɾo]**	***corro***	**[toɾo]**	***todo***
Controls	100 (0)		92 (13)		92 (13)	
Trainees	98 (7)		88 (13)		88 (15)	
Native Speakers	100 (0)		88 (18)		98 (8)	

mismatching pairs (e.g., auditory prime [koɾo] followed by visual target *corro* or auditory prime [toɾo] followed by *todo*) after training. Recall that the native Spanish speakers completed the cross-modal priming task only once. However, their results are also included in these post-training analyses in order to compare the post-training learners of Spanish (controls and trainees) to native Spanish speakers. Native Spanish speakers responded more quickly and accurately to matching prime-target pairs than to mismatching pairs. This pattern was predicted for learners of Spanish who underwent training but not for the learners of Spanish in the control condition. Repeated measures ANOVAs with three levels of Target Type (/ɾ/, /r/ and /d/) and with Group (controls, trainees, and native Spanish speakers) as a between-subjects factor were conducted separately on reaction time and accuracy for the post-training data.

The repeated measures ANOVA carried out on reaction time revealed main effects of Target Type [F (2, 80) = 12.389, p <0.001] and Group [F (2, 40) = 6.816, p = 0.003] and an interaction between Target Type and Group [F (4, 80) = 3.504, p = 0.011]. With respect to the main effect of Target Type, pairwise comparisons using the Bonferroni correction found that participants, when hearing an auditory prime like [paɾa], responded more quickly to visual targets containing the orthographic representation /ɾ/ at 652 ms than to visual targets containing /r/ at 780 ms (p < 0.001). Likewise, participants responded more quickly to visual targets containing the orthographic representation /d/ at 703 ms than to those containing /r/ (p = 0.019). Reaction times in response to /ɾ/ and /d/ did not differ significantly. The main effect of Group reflects that trainees responded significantly faster overall at 640 ms than native speakers at 890 ms (p = 0.002); however, controls did not differ significantly from either trainees or native speakers. Recall that Spanish learners' post-training data reflect the second time that they participated in the lexical decision task.

In order to interpret the Target Type × Group interaction, a series of 6 paired-samples t-tests were conducted comparing the reaction time of trainees and controls when responding to the 3 target types. Native speakers were excluded from this analysis to increase power since their t-test results are reported in Section 3.1. Using the Bonferroni correction to control for familywise error, the significant p-value threshold was set as $p < 0.0083$. Trainees responded significantly faster to visual targets containing the orthographic representation of /ɾ/ at 594 ms than to those containing /r/ at 707 ms ($p = 0.001$). There were no significant differences in the reaction time of the controls in response to different target types. As predicted, trainees exhibited a pattern different from controls and more similar to native Spanish speakers after training, with faster responses to matching prime-target pairs than to mismatching /ɾ/-/r/ prime-target pairs. The mean post-training reaction times of controls, trainees, and native Spanish speakers in response to the three targets are presented in Figure 2.

The repeated measures ANOVA carried out on post-training accuracy revealed a main effect of Target Type [F (2, 80) = 5.635, $p = 0.005$] but no main effect of Group and no interaction between Target Type and Group. Pairwise comparisons using the Bonferroni correction found that participants responded with significantly better accuracy to visual targets containing the orthographic representation of /ɾ/ (e.g., *para*) at 97% than to visual targets containing /r/ (e.g., *corro*) at 83% ($p = 0.009$), but the accuracy rate in response to /d/ did not differ significantly from /ɾ/ or /r/. The mean accuracy rates reported in Table 4 suggest that the accuracy differences in response to /ɾ/ targets and /r/ targets were driven by the native Spanish speakers and the trainees. The low accuracy rate of the native Spanish speakers and the trainees

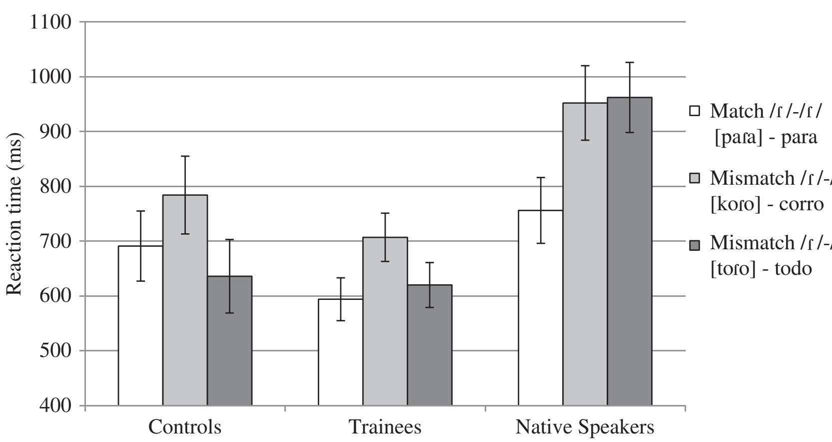

Figure 2: Mean post-training lexical decision reaction time (ms) organized by Target Type and Group. Error bars indicate standard error.

Table 4: Mean post-training lexical decision accuracy rate in percentages (and standard deviation) organized by Target Type and Group.

Target Type	**/ɾ/**		**/r/**		**/d/**	
Examples	**prime**	**target**	**prime**	**target**	**prime**	**target**
	[paɾa]	***para***	**[koɾo]**	***corro***	**[toɾo]**	***todo***
Controls	92 (13)		92 (13)		89 (18)	
Trainees	98 (7)		77 (21)		94 (13)	
Native Speakers	100 (0)		88 (18)		98 (8)	

in response to /r/ targets could indicate that both native Spanish speakers and trainees, but not controls, exhibited the predicted inhibition patterns (e.g., slower reaction time and lower accuracy rate) for the mismatching /ɾ/-/r/ pair; however, the Group and Target Type × Group differences did not reach significance for the accuracy data.

4 Discussion

The main aims of this study were to measure whether native Spanish speakers and nonnative Spanish speakers process the /ɾ, r, d/ contrast differently and to determine if nonnative speakers would process the three-way contrast in a more native-like manner as a result of successful high variability training. As predicted, the reaction time data clearly indicated that native Spanish speakers' reaction times in response to visual targets containing /ɾ/ were facilitated by preceding auditory primes with /ɾ/. On the other hand, Spanish learners responded with similar reaction times regardless of the visual target type that followed the auditory /ɾ/ primes at pre-training. Native Spanish speakers clearly exhibited different reaction time patterns in response to the /ɾ, r, d/ contrast than Spanish learners (both trainees and controls) prior to any participation in the high variability training.

At post-training, trainees responded significantly faster in response to matching prime-target pairs than to mismatching /ɾ/-/r/ pairs, exhibiting a pattern different from pretest and more similar to native Spanish speakers. Controls, on the other hand, exhibited reaction time patterns similar to their pretest results, with reaction times not affected by target type. This suggests that high variability training, which improved trainees' perception and production of the /ɾ, r, d/ contrast, also improved the way trainees processed the lexical items containing these contrasts, with trainees exhibiting a pattern similar to native Spanish speakers.

In addition to matching prime-target pairs resulting in facilitation and faster reaction times, mismatching prime-target pairs, in which the prime and target only differed in one segment (e.g., [koɾo] followed by *corro* or [toɾo] followed by *todo*), resulted in inhibition and slower reaction times (see also Allen and Badecker 2002). In other words, native Spanish speakers' reaction times were slowed for two-thirds of the stimuli. This inhibition for mismatched items in the native speakers coupled with the fact that both groups of learners took the lexical decision task twice (pre-training and post-training) may account for the fact that native Spanish speakers, who only completed the lexical decision task once, responded with slower overall reaction times than the Spanish learners did post-training.

At posttest, trainees' reaction times in response to matching prime-target pairs and mismatching /ɾ/-/r/ prime-target pairs were similar to native Spanish speakers' response patterns, but this was not the case for mismatching /ɾ/-/d/ pairs. Why might this be the case? One possibility is that not all contrasts resulted in similar improvements in perception and production. The most robust improvements in perception and production for all trainees occurred with the /ɾ/-/r/ pair (Herd et al. 2013). This suggests that the /ɾ/-/d/ contrast in particular is resistant to training, which is not surprising considering that [ɾ] is an allophone of /d/ in American English. Future Spanish training studies should examine whether this difficulty is unique to American English learners of Spanish due to the organization of the two phones in English or whether it is a more general learner pattern due to the acoustic similarity of /ɾ/ and /d/ as compared to /r/.

Another possibility for the lack of improvement for the mismatching /ɾ/-/d/ prime-target pairs is to consider the process by which second language learners encode new phonemic contrasts in the L2 lexicon. Studies like that of Pallier et al. (2001), where even fluent Spanish-Catalan bilinguals exhibited different reaction time patterns (i.e., having difficulty in perceiving the /e/-/ɛ/ in Catalan) based on their dominant language, suggest that it may not be possible to lexically encode contrasts that are not present in the dominant (or first) language. Similarly, the present study found that trainees did not encode the differences between /ɾ/ and /d/. However, Darcy et al. (2012) showed that advanced second language learners can encode a new contrast in the lexicon, finding that advanced French learners exhibited more native-like reaction time patterns than intermediate French learners in response to /y/-/u/ pairs. Similar to Darcy et al. (2012), the results of the present study also suggest that a novel phoneme (i.e., /r/) can be encoded in the L2 lexicon. While these sets of data seem contradictory at first, the different results may be due to the acoustic salience of the different contrasts. It might be the case that newly acquired novel, acoustically salient contrasts (e.g., /ɾ/-/r/ in Spanish and /y/-/u/ in

French) and newly acquired, less salient contrasts (e.g., /ɾ/-/d/ in Spanish and /e/-/ɛ/ in Catalan) are lexically encoded at different rates.

Hayes-Harb and Masuda (2008) reported that native English speakers were in fact able to lexically encode short versus long consonants, a less salient contrast, for specific lexical entries as a result of training four nonword pairs as brand names. Unlike the present study, however, Hayes-Harb and Masuda (2008) tested whether the short-long consonant distinction was lexically encoded for the specific items trained. It should also be noted that lexical encoding was measured using perception (i.e., auditory word to picture matching) and production (i.e., picture naming) tasks. In the present study, tokens used as primes or targets in the cross-modal priming task were not used during training or any of the other tasks. The cross-modal priming tokens were presented to participants only twice: once during the pre-training cross-modal priming task and once during the post-training cross-modal priming task. The results of the present study and the Hayes-Harb and Masuda (2008) study may differ because less salient contrasts may be encoded in the lexicon gradually, on a word-by-word basis through a process of lexical diffusion (Bybee 2000; 2002; 2012; Chen and Wang 1975). Instead of the new contrast being encoded immediately into all relevant lexical entries, the contrast is instead encoded only when the trainee uses each word, reanalyzing it in light of the newly trained contrast.

In the present study, trainees did not exhibit native-like reaction time patterns in response to /d/ targets after training because the difference between /ɾ/ and /d/ is less salient than the difference between /ɾ/ and /r/ and because participants had not been trained on the specific lexical items used in the cross-modal priming task. The cross-modal priming reaction times thus represent how participants responded to the lexical items the first time they encountered each item after training. Since participants had not been exposed to these specific words during training, they had not reanalyzed their lexical representations. If new contrasts are encoded through lexical diffusion, then Spanish learners who have acquired the new /ɾ, r, d/ contrast should begin encoding the newly acquired contrasts on a word-by-word basis as lexical items are used and reanalyzed. Future training studies should include both trained and untrained items when assessing lexical encoding to see if these predictions are borne out.

5 Conclusion

Using a cross-modal priming design, this study set out to determine if native speakers of Spanish and learners of Spanish exhibit different reaction time patterns, specifically with respect to the /ɾ, r, d/ contrast. This study demonstrates that native Spanish speakers exhibit canonical reaction time patterns, with significantly faster reaction times in response to matching prime-target pairs (e.g., auditory [paɾa] followed by orthographic *para*) than to mismatching prime-target pairs (e.g., auditory [koɾo] followed by orthographic *corro* or auditory [toɾo] followed by *todo*). Conversely, learners of Spanish did not exhibit significantly different reaction times in response to matching and mismatching pairs prior to training. After training, however, learners of Spanish responded with significantly faster reaction times to matching prime-target pairs (e.g., auditory [paɾa] followed by orthographic *para*) than to mismatching /ɾ/-/r/ pairs (e.g., auditory [koɾo] followed by orthographic *corro*), exhibiting a pattern more similar to that of native Spanish speakers. However, reaction times in response to mismatching /ɾ/-/d/ pairs (e.g., auditory [toɾo] followed by *todo*) remained unchanged. Controls, on the other hand, exhibited no change in reaction time patterns across the different primes and targets. Native and non-native Spanish speakers process the /ɾ, r, d/ contrast differently, and learners' reaction time patterns can become more native-like as a result of high variability training. Future research should determine whether the more native-like Spanish learner reaction time pattern is specific to American English learners and whether new contrasts are encoded post-training through a process of lexical diffusion as participants reanalyze relevant lexical entries.

Acknowledgments: This research was part of the first author's dissertation, supervised by Allard Jongman and Joan Sereno and supported by the National Science Foundation (#0843653).

References

Allen, Mark & William Badecker. 2002. Inflectional regularity: Probing the nature of lexical representation in a cross-modal priming task. *Journal of Memory and Language* 46(4). 705–722.

Boomershine, Amanda, Kathleen Currie Hall, Elizabeth Hume & Keith Johnson. 2008. The impact of allophony versus contrast on speech perception. In Peter Avery, B.Elan Dresher & Karen Rice (eds.), *Contrasts in phonology: Theory, perception, acquisition*, 145–171. New York: Mouton de Gruyter.

Bradlow, Ann R., Reiko Akahane-Yamada, David B. Pisoni & Yoh'ichi Tohkura. 1999. Training Japanese listeners to identify English/r/and/l/: Long-term retention of learning in perception and production. *Attention, Perception and Psychophysics* 61(5). 977–985.

Bradlow, Ann R., David B. Pisoni, Reiko Akahane-Yamada & Yoh'ichi Tohkura. 1997. Training Japanese listeners to identify English/r/and/l/: Some effects of perceptual learning on speech production. *Journal of the Acoustical Society of America* 101(4). 2299–2310.
Broersma, Mirjam & Anne Cutler. 2011. Competition dynamics of second-language listening. *The Quarterly Journal of Experimental Psychology* 64(1). 74–95.
Bybee, Joan L. 2000. The phonology of the lexicon: Evidence from lexical diffusion. In Michael Barlow & Suzanne Kemmer (eds.), *Usage-based models of language*, 65–85. Stanford, CA: CSLI Publications.
Bybee, Joan. 2002. Word frequency and context use in the lexical diffusion of phonetically conditioned sound change. *Language Variation and Change* 14(3). 261–290.
Bybee, Joan. 2012. Patterns of lexical diffusion and articulatory motivation for sound change. In Maria-Josep Solé & Daniel Recasens (eds.), *The initiation of sound change*, 211–234. Amsterdam & Philadelphia, PA: John Benjamins Publishing.
Byrd, Dani. 1994. Relations of sex and dialect to reduction. *Speech Communication* 15(1). 39–54.
Chen, Matthew Y. & William S.-Y. Wang. 1975. Sound change: Actuation and implementation. *Language* 51(2). 255–281.
Connine, Cynthia M. 2004. It's not what you hear but how often you hear it: On the neglected role of phonological variant frequency in auditory word recognition. *Psychonomic Bulletin & Review* 11(6). 1084–1089.
Darcy, Isabelle, Laurent Dekydtspotter, Rex A. Sprouse, Justin Glover, Christiane Kaden, Michael McGuire & John H. G. Scott. 2012. Direct mapping of acoustics to phonology: On the lexical encoding of front rounded vowels in L1 English – L2 French acquisition. *Second Language Research* 28(1). 5–40.
Davies, Mark. 2002. *Corpus del Español*. [Online Corpus]. Retrieved March 15, 2011, from http://www.corpusdelespanol.org/.
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*, 47–58. Somerville, MA: Cascadilla Proceedings Project.
Face, Timothy L. & Mandy R. Menke. 2009. Acquisition of Spanish voiced spirants by second language learners. In Joseph Collentine, Maryellen García, Barbara Lafford & Francisco Marcos Marín (eds.), *Selected Proceedings of the 11th Hispanic Linguistics Symposium*, 39–52. Somerville, MA: Cascadilla Proceedings Project.
Gaskell, M. Gareth & William D. Marslen-Wilson. 2002. Representation and competition in the perception of spoken words. *Cognitive Psychology* 45(2). 220–266.
González-Bueno, Manuela. 1995. Adquisición de los alófonos fricativos de las oclusivas sonoras españolas por aprendices de español como segunda lengua. *Estudios de lingüística aplicada* 13. 64–79.
Grainger, Jonathan, Maryline Nguyen van Kang & Juan Segui. 2001. Cross-modal repetition priming of heterographic homophones. *Memory and Cognition* 29(1). 53–61.
Hayes-Harb, Rachel & Kyoko Masuda. 2008. Development of the ability to lexically encode novel second language phonemic contrasts. *Second Language Research* 24(1). 5–33.
Herd, Wendy, Allard Jongman & Joan Sereno. 2010. An acoustic and perceptual analysis of/t/ and/d/flaps in American English. *Journal of Phonetics* 38(4). 504–516.
Herd, Wendy, Allard Jongman & Joan Sereno. 2013. Perceptual and production training of intervocalic/d, ɾ, r/in American English learners of Spanish. *Journal of the Acoustical Society of America* 133(6). 4247–4255.

Hirata, Yukari. 2004. Computer assisted pronunciation training for native English speakers learning Japanese pitch and durational contrasts. *Computer Assisted Language Learning* 17(3–4). 357–376.
Holcomb, Phillip J., Jane Anderson & Jonathan Grainger. 2005. An electrophysiological study of cross-modal repetition priming. *Psychophysiology* 42(5). 493–507.
Johnson, Keith E. 2008. *Second language acquisition of the Spanish multiple vibrant consonant*. Tucson, AZ: University of Arizona dissertation.
Logan, John S., Scott E. Lively & David B. Pisoni. 1991. Training Japanese listeners to identify English/r/and/l/: A first report. *Journal of the Acoustical Society of America* 89(2). 874–886.
Marslen-Wilson, William. 1990. Activation, competition, and frequency in lexical access. In Gerry T. M. Altmann (ed.), *Cognitive models of speech processing: Psycholinguistics and computational perspectives*, 148–172. Cambridge, MA: MIT Press.
Marslen-Wilson, William & Pienie Zwitserlood. 1989. Accessing spoken words: The importance of word onsets. *Journal of Experimental Psychology: Human Perception and Performance* 15(3). 576–585.
Pallier, Christophe, Angels Colomé & Núria Sebastián-Gallés. 2001. The influence of native-language phonology on lexical access: Exemplar-based versus abstract lexical entries. *Psychological Science* 12(6). 445–449.
Patterson, David & Cynthia M. Connine. 2001. Variant frequency in flap production: A corpus analysis of variant frequency in American English flap production. *Phonetica* 58(4). 254–275.
Rose, Marda. 2010. Intervocalic tap and trill production in the acquisition of Spanish as a second language. *Studies in Hispanic and Lusophone Linguistics* 3(2). 379–422.
Tagliaferri, Bruno. 2011. *Paradigm: Perception Research Systems* [Computer Program]. http://www.perceptionresearchsystems.com/ (accessed 23 March 2008).
Waltmunson, Jeremy C. 2005. *The relative difficulty of L2 Spanish/d, t/, trill, and tap by L1 English speakers: Auditory and acoustic methods of defining pronunciation accuracy*. Seattle, WA: University of Washington dissertation.
Wang, Yue, Allard Jongman & Joan A. Sereno. 2003. Acoustic and perceptual evaluation of Mandarin tone productions before and after perceptual training. *Journal of the Acoustical Society of America* 113(2). 1033–1043.
Wang, Yue, Michelle M. Spence, Allard Jongman, & Joan A. Sereno. (1999). Training American listeners to perceive Mandarin tones. *Journal of the Acoustical Society of America* 106(6). 3649–3659.
Zampini, Mary L. 1993. *Spanish voiced stop phonemes and spirantization: A study in second language acquisition*. Washington, DC: Georgetown University dissertation.
Zampini, Mary L. 1994. The role of native language transfer and task formality in the acquisition of Spanish spirantization. *Hispania* 77(3). 470–481.
Zue, Victor W. & Martha Laferriere. 1979. Acoustic study of medial/t, d/in American English. *Journal of the Acoustical Society of America* 66(4). 1039–1050.

Copyright of Studies in Hispanic & Lusophone Linguistics is the property of De Gruyter and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.