Quierro comprar una guitara: Lexical Encoding of the Tap and Trill by L2 Learners of Spanish

Danielle Daidone and Isabelle Darcy
Indiana University

1. Introduction

The Spanish language has two rhotics, the voiced alveolar tap /ɾ/ and the voiced alveolar trill /r/. Depending on context, these phones are either in free variation or complementary distribution; the only exception is in intervocalic position where the presence of the tap or trill signals a difference in meaning, e.g. *pero* /peɾo/ 'but' vs. *perro* /pero/ 'dog' (Hualde, 2005).[1] American English, on the other hand, has a single rhotic which is realized most often as a voiced alveolar approximant and thus is different from either of the Spanish rhotics (Ladefoged & Johnson, 2010). American English-speaking learners of Spanish therefore must learn to produce both the tap and trill in intervocalic position in order to differentiate between minimal pairs. Previous research on the acquisition of these phonemes, however, has found that even advanced speakers often fail to distinguish between the tap and trill in production (Face, 2006; Rose, 2010b).

Initially it may appear possible to attribute this difficulty in production to a difficulty in perception, since Rose (2010a) found that Spanish /ɾ/ and /r/ were assimilated by naïve American English listeners to the same English category /ɹ/ in the majority of cases; 96.2% of Spanish /r/ tokens and 57.7% of /ɾ/ tokens were identified as English /ɹ/, making English /ɹ/ the modal response for both Spanish rhotics. Nevertheless, Rose (2010a) found that learners at all levels were highly accurate at discriminating the /ɾ-r/ contrast, with a mean discrimination accuracy between 86.7% and 94.4% across proficiency levels in an AXB task. Even American English speakers without experience with Spanish discriminated between these phonemes with 80.2% accuracy on average.

Given that learners' difficulties with Spanish rhotics do not appear to be due to a problem in perception, previous researchers have cited a difficulty in articulation as an explanation for learners' failure to distinguish these phonemes in production (Face, 2006; Rose, 2010b). The tap in Spanish "is produced with a single rapid contact of the tip of the tongue against the alveolar ridge," while the Spanish trill "is produced with several such rapid contacts, generally two or three" (Hualde, 2005, p. 181). Although the tongue tip strikes the alveolar ridge in the production of both the tap and the trill, the articulatory gestures required to realize these two segments are quite distinct. The trill is a more complex segment which requires precise control over the positioning of the articulators and the amount of air flow (Solé, 2002). It is among the last segments acquired by native speakers (Jiménez, 1987), and its complexity has led to substantial dialectal variation in its production since even a small change in any of the articulatory gestures of the trill greatly alters the sound produced (Widdison, 1998). Therefore, an inability to reproduce the articulation of the trill could explain the lack of distinction between the tap and the trill in production by many L2 learners. This is substantiated by the type of

* We are grateful for the valuable feedback from three anonymous reviewers on an earlier version of this manuscript.

[1] There is some debate as to whether Spanish has two rhotic phonemes or a single rhotic phoneme that is underlyingly geminate for trills in intervocalic position (see Hualde, 2004). Regardless of the analysis, learners must still learn to differentiate the tap and trill in intervocalic position in their lexical representations.

 Selected Proceedings of the 2012 Second Language Research Forum, ed. Ryan T. Miller et al., 39-50. Somerville, MA: Cascadilla Proceedings Project.

errors that L2 speakers make: at advanced levels learners often use the tap, which is easier to articulate, in the trill environment (Face, 2006; Rose, 2010b).

Other evidence, however, suggests that difficulty articulating the trill may not entirely explain learners' patterns of production. Even when articulation does not play a role, learners often confuse the two Spanish rhotics. In written production, one of the authors who teaches Spanish has repeatedly seen spelling mistakes in which the wrong rhotic is indicated, such as *perro* /pero/ 'dog' for *pero* /peɾo/ 'but' or *caro* /kaɾo/ 'expensive' for *carro* /karo/ 'car.' Even though such anecdotal evidence is not sufficient, similar findings were observed in Rose's (2010b) study on the production of intervocalic tap and trill: twelve out of 21 learners did not differentiate the tap and trill environments in any way, despite /r/ having other acceptable variants that do not require the complex articulatory gestures of an alveolar trill. For example, besides the voiced alveolar trill with two or more occlusions, native speakers in Rose's study exhibited productions such as a tap followed by frication or an assibilated variant. Even supposing that learners could not produce such variants, a longer duration for a tap used in the trill environment would also be a possible production that is characteristic of L1 Spanish speakers (Henriksen & Willis, 2010). Rose (2010b), however, found no evidence for the use of duration by learners to distinguish the two phonemes. These twelve learners did not differentiate between the tap and trill in a nonnative-like way either, e.g. an approximant in the tap environment and a tap in the trill environment, as did several of the other participants. In sum, while it is possible that spelling errors appear because many learners have not mastered the orthographic conventions of Spanish, the fact that more than half of the L2 speakers who participated in Rose's study did not differentiate at all between the tap and trill environments in production suggests that the inability to coordinate the gestures for the trill is not the only contributing factor to this neutralization. The lack of differentiation in production may reflect inaccuracies at a deeper level, in the long-term phonological representations of lexical items.

L2 learners have been shown to not accurately encode new contrasts in lexical representations (Darcy et al., 2012; Dupoux, Sebastián-Gallés, Navarrete, & Peperkamp, 2008; Ota, Hartsuiker, & Haywood, 2009; Pallier, Colomé, & Sebastián-Gallés, 2001; Sebastián-Gallés, Rodríguez-Fornells, de Diego-Balaguer, & Díaz, 2006). Darcy et al. (2012) found that although intermediate learners of French could accurately discriminate the /y-u/ contrast, they exhibited positive priming for /y-u/ minimal pairs in a lexical decision task with repetition priming. These results suggest that /y-u/ minimal pairs were encoded as homophones in these learners' L2 lexical representations, despite their ability to discriminate the /y-u/ contrast. Similarly, in Dupoux et al.'s (2008) study on the encoding of stress by French learners of Spanish, learners were unable to contrastively encode stress in long-term representations of words, even though an earlier study found that they could perceive stress differences in an AX discrimination task (Dupoux, Pallier, Sebastián, & Mehler, 1997). These studies provide evidence that a high degree of precision in discrimination of a contrast does not guarantee that learners will fully encode this contrast in their lexical representations (Hayes-Harb & Masuda, 2008). Hayes-Harb & Masuda (2008) argued that L2 learners may initially only be able to use features that are specified in their L1 in the encoding of L2 lexical representations. This possibility would be consistent with the results of Darcy et al. (2012) and Dupoux et al. (1997), since [round] is not used to contrast front vowels in English and stress is not used contrastively in French. Therefore, the unexpected difference between L2 Spanish learners' high accuracy in the perception of the /ɾ-r/ contrast and their lack of differentiation of these phonemes in production may be attributed to problems at the level of lexical encoding. It is possible that since this rhotic distinction is not used in English, L1 English L2 Spanish learners have difficulty maintaining the distinction in long-term memory for the phonological representations of words, independent of their ability to perceive phonetic differences between the tap and trill.

A further complication to the acquisition of the tap and trill by English-speaking learners is the presence of a flap allophone of /t/ and /d/ in the L1. In American English, the phonemes /t/ or /d/ become an alveolar flap when they occur intervocalically after a stressed syllable and before an unstressed syllable. Although they are not articulatorily identical, the English flap and the Spanish tap are very similar (Hualde, 2005), which would explain why 30.4% of the tokens of Spanish /ɾ/ were assimilated to the English /d/ in crosslinguistic perceptual assimilation data (Rose, 2010a). Unlike the tap and trill, which were both most often mapped onto the same L1 category /ɹ/, the L1 categories most often identified with the tap and /d/ differed. The Spanish /ɾ/ was identified with English /ɹ/ at a rate of 57.7%, and Spanish /d/, which is realized as an approximant [ð] intervocalically as well as in other

contexts (Hualde, 2005), was identified with English /l/ at a rate of 54.2%. While this crosslinguistic mapping data might suggest that discrimination of /ɾ-d/ would be more accurate than that of /ɾ-r/, Rose (2010a) found that discrimination by learners of the Spanish contrast /ɾ-d/ was on average between 62.9% and 82.5%, lower than that of /ɾ-r/ (86.7%-94.4%). It is unclear what drove the difference in discrimination accuracy, but orthographic factors or the fact that the tap is an L1 allophonic variant of /d/ might have interfered with discrimination, if the tap was perceived as a realization of English /d/.

If discrimination ability directly predicts how words will be encoded in the L2 lexicon, then minimal pairs containing /r/ and /ɾ/, such as *carro* /karo/ 'car' and *caro* /kaɾo/ 'expensive,' would be more likely to be differentiated in lexical representations than minimal pairs containing /ɾ/ and /d/, such as *miro* /miɾo/ 'I look' and *mido* /mido/ 'I measure,' since discrimination of /r/ and /ɾ/ has been shown to be more accurate than that of /ɾ/ and /d/. However, if discrimination ability is independent of lexical encoding ability, then specific predictions are more difficult to generate. The tap and trill may be less likely to be differentiated in L2 representations than the tap and /d/, given that /ɾ-r/ has a lower functional load than /ɾ-d/ in the Spanish language. In addition, the tap and trill are written with *r* and *rr*, respectively, in Spanish orthography, which both represent /ɹ/ in English orthography, and previous research has shown that orthography can affect lexical representations (e.g., Hayes-Harb, Nicol, & Barker, 2010).

2. Research Questions and Predictions

The present study aims to examine the nature of L2 Spanish learners' lexical representations of words containing /ɾ/, /r/, and /d/ in intervocalic position through a lexical decision task in order to determine if the contrast between the two rhotics and the contrasts between the rhotics and /d/ have been clearly encoded. We also hope to determine if discrimination ability is dissociated from this lexical level by examining learners' ability to discriminate /ɾ/, /r/, and /d/ in an ABX task. We consider the two following research questions:

a. Is the /ɾ-r/ contrast in intervocalic position encoded in the lexical representations of L2 learners of Spanish?

b. Does accuracy in discrimination directly predict accuracy in lexical encoding?

We expect that learners will be able to discriminate between the tap and the trill as Rose (2010a) found in her study. We also hypothesize that the tap and trill are not differentiated in the lexical representations of L2 learners, based on orthographic evidence and the results of Rose (2010b) as previously discussed. If learners can discriminate but not accurately encode the /ɾ-r/ contrast, then the ability to discriminate a contrast does not necessarily entail that it is encoded accurately in the mental lexicon, and therefore we predict that the accuracy order of the lexical encoding of /ɾ-r/, /ɾ-d/, and /r-d/ will differ from the accuracy order in the discrimination of these contrasts.

3. Method

Two main tasks were administered in this experiment: a lexical decision task, to investigate the lexical encoding of words containing /ɾ/, /r/, or /d/, and an ABX task, to test participants' ability to categorically discriminate between these phonemes. A background questionnaire was given at the beginning of the experiment that elicited information about participants' Spanish language experience, as well as information such as age, sex, other languages spoken, etc. Participants then completed the lexical decision task, followed by the ABX task. After finishing the two parts of the experiment, a word familiarity questionnaire was administered that asked participants to rate their familiarity with the words in the lexical decision task.

3.1. ABX Task

ABX tasks and their variants have been used extensively in order to examine the perception of sounds in an second language (e.g., Darcy et al., 2012; Gottfried, 1984; Højen & Flege, 2006; Levy & Strange, 2008; Polka, 1995). In this task, participants heard three stimuli in succession and then had to decide if the last token (X) was more similar to the first token (A) or to the second token (B). Target words were embedded within the sentential context 'Le digo _____ al profe' (I say _____ to the professor) in order to increase cognitive load and be more representative of real-world speech processing, given that speakers rarely hear a series of isolated words (Levy & Strange, 2008). This task type, combined with the high phonetic variety achieved by using different voices and different tokens of each nonce word, ensured that participants could not rely on irrelevant acoustic details to discriminate the stimuli but instead had to process the tokens at the phonological level to categorize them. An example trial is as follows: participants heard the sequence *Le digo* /naɾe/ *al profe; Le digo* /nare/ *al profe; Le digo* /naɾe/ *al profe*, and if they accurately distinguished the tap from the trill, they chose A as the correct answer.

Three native Spanish speakers were recorded for this task, one man and two women; the male speaker was from Costa Rica, one of the female speakers was from Puerto Rico, and the second female speaker was a simultaneous English-Spanish bilingual from the U.S. who spoke Mexican Spanish. The stimuli they recorded consisted of disyllabic nonce words. Nonce words were utilized in order to eliminate any effects of word frequency or differences in lexical knowledge between participants. Stimuli were divided into three test conditions (/ɾ-r/; /ɾ-d/; /r-d/), a control condition (/p/-/f/), and fillers, which are subsequently described in detail below.[2]

1. Test Condition: 5 minimal pairs with /ɾ-r/ in intervocalic position
   e.g. /neɾa/; /nera/

2. Test Condition: 5 minimal pairs with /ɾ-d/ in intervocalic position
   e.g. /faɾe/; /fade/

3. Test Condition: 5 minimal pairs with /r-d/ in intervocalic position
   e.g. /ʧera/; /ʧeda/

4. Control Condition: 5 minimal pairs with /f/-/p/ in intervocalic position, a contrast which also exists in English
   e.g. /bifa/; /bipa/

5. Fillers: 4 minimal pairs, two with a consonant contrast and two with a vowel contrast in various positions
   e.g. /tume/; /dume/

Each of these contrasts was repeated four times, in the sequences AAB, ABA, BAA, and BAB. This yielded a total of 60 test trials, 20 control trials, and 16 filler trials. Nine practice trials with feedback were given at the beginning of the experiment to familiarize participants with the task. None of the practice trials contained the test or control contrasts. In each trial, the first two tokens (A & B) were spoken in the female voices while the last token (X) was spoken in the male voice. These trials were randomized and presented on a PC through headphones with an ISI of 500 milliseconds. The ABX task was administered with the presentation software DMDX (Forster & Forster, 2003).

[2] All except for three of the 60 realizations of /r/ were realized as a voiced alveolar trill with 2-4 occlusions. Of the remaining three tokens of /r/, all were native speaker variants of the trill phoneme: one was realized as tap followed by frication, one was realized as an assibilated variant, and one, in the /r-d/ condition, was realized as a tap.

3.2. Lexical Decision Task

An auditory lexical decision task, adapted from the methodology of Dupoux, Sebastián-Gallés, Navarrete, & Peperkamp (2008), was used to examine the nature of participants' lexical representations of Spanish words. In this task, participants listened to stimuli and had to decide whether each token was a real word or a fake word of Spanish. For example, when learners heard *quiero* /kiɛro/ "I want" they were expected to indicate that this was a word, while when they heard **quierro* /kiero/ they were expected to indicate that this was a nonword. Learners should have correctly rejected the nonword with the trill, **quierro* /kiero/, since [r] is not a legitimate realization of /ɾ/ in *quiero* /kiɛro/. The female Puerto Rican speaker and the male Costa Rican speaker who were recorded for the ABX task were also recorded for this task. Stimuli were created by exchanging /ɾ/ for /r/ or vice versa to create word/nonword pairs. This was also done with /ɾ-d/ and /r-d/.[3] In order to ensure that learners were familiar with the words in the task, an effort was made to use words that are present in the *Beginning Spanish Lexicon*, a database of words from beginner Spanish textbooks (Vitevitch, Stamer, & Kieweg, 2012). Although not all of the stimuli in the task appear in the *Beginning Spanish Lexicon*, a word familiarity questionnaire revealed that participants were generally very familiar with the words, and no participants had to be eliminated based on low word familiarity. A detailed description of the stimuli is presented below:

1. Test Condition: 20 minimal pairs with /ɾ-r/ in intervocalic positions
    a. 10 pairs: word with /ɾ/, nonword with /r/
        e.g. *quiero* /kiɛro/ 'I want'; **quierro* /kiero/
    b. 10 pairs: word with /r/, nonword with /ɾ/
        e.g. *correcto* /korekto/ 'correct'; **corecto* /koɾekto/

2. Test Condition: 20 minimal pairs with /ɾ-d/ in intervocalic position
    a. 10 pairs: word with /ɾ/, nonword with /d/
        e.g. *fuera* /fuɛɾa/ 'outside'; **fueda* /fueda/
    b. 10 pairs: word with /d/, nonword with /ɾ/
        e.g. *abogado* /abogado/ 'lawyer'; **abogaro* /abogaɾo/

3. Test Condition: 20 minimal pairs with /r-d/ in intervocalic position
    a. 10 pairs: word with /r/, nonword with /d/
        e.g. *ocurre* /okure/ 'it occurs'; **ocude* /okude/
    b. 10 pairs: word with /d/, nonword with /r/
        e.g. *estado* /estado/ 'state'; **estarro* /estaro/

4. Control Condition: 20 minimal pairs with /p/-/f/ in intervocalic position, a contrast which also exists in English
    a. 10 pairs: word with /f/, nonword with /p/
        e.g. *difícil* /difisil/ 'difficult'; **dipícil* /dipisil/
    b. 10 pairs: word with /p/, nonword with /f/
        e.g. *guapo* /guapo/ 'handsome'; **guafo* /guafo/

5. Fillers: 24 words and 24 unrelated nonwords
    a. e.g. *cabeza* /kabesa/ 'head'; **leto* /leto/

Stimuli were divided into two lists, so that participants only heard either the word or the nonword of a word/nonword pair. For example, those participants who were tested with List A heard the stimulus *quiero* /kiɛro/ but not its corresponding nonword **quierro* /kiero/, while those participants who were tested with List B heard **quierro* /kiero/ but not *quiero* /kiɛro/. There were 128 stimuli per list, which were randomized, as well as 10 practice trials which consisted of 5 words and 5 unrelated nonwords. The lexical decision task was administered on a PC through headphones with the DMDX software (Forster & Forster, 2003).

[3] All instances of the phoneme /r/ were realized as a voiced alveolar trill with 2-4 occlusions in the stimuli for this task.

3.3. Participants

Intermediate and advanced American English-speaking learners of Spanish were tested, along with native Spanish speakers. Proficiency level was determined by class level. Intermediate speakers were enrolled in a 5th semester Spanish class at a large Midwestern university, and all but two of the learners had no study abroad experience.[4] Advanced speakers were mostly graduate students in a Hispanic linguistics or literatures program at the same university, although a few were undergraduate students in 300 or 400 level classes (6th semester and above). All advanced speakers but one had 3 or more weeks of study abroad experience. Native Spanish speakers were living in the U.S. and spoke English as a second language. Native speakers and advanced speakers were on average older than intermediate speakers; native speakers had a mean age of 27.2 years (range 18-36), advanced speakers a mean age of 26.3 years (range 18-40), and intermediate speakers a mean age of 18.8 years (range 18-22).

4. Results

4.1. ABX Task

Data from 5 intermediate and 1 native speaker participants were eliminated based on extremely low accuracy in the control condition, which suggests a key reversal. In total, data from 21 intermediate speakers, 20 advanced speakers, and 10 native speakers were analyzed for the ABX task. The overall results for the test, control, and filler conditions are shown in Figure 1.

The average accuracy performance for the intermediate learners was 84% correct. For the advanced, it was 92% correct, and native speakers obtained 95% correct. A mixed model ANOVA was run using SPSS 20, and the alpha level of significance was set at .05 (for all analyses). Because the distribution of accuracy means was positively skewed, we transformed the accuracy means into an arcsine value. The model was run with the arcsine-transformed accuracy means as the dependent variable. We report first the overall effects of group, condition, and the interactions. There was a main effect of group ($F(2, 63.7) = 12.3$, $p < .001$), of condition ($F(2, 198) = 9.7$, $p < .001$), and a marginal interaction ($F(4, 198) = 2.1$, $p = .085$). All three groups performed globally with high accuracy.

Considering the effect of condition within each group, accuracy on the test, control, and filler conditions did not differ significantly for the advanced and native speakers. There was a significant simple main effect of condition for the intermediate learners ($F(2, 198) = 13.6$, $p < .001$), due to lower accuracy on the test condition. Additionally, the intermediate group was significantly less accurate on the test condition than both other groups. The three groups did not differ significantly on the two other conditions.

[4] The two intermediate learners with study abroad experience both had spent 4 weeks in a Spanish-speaking country.

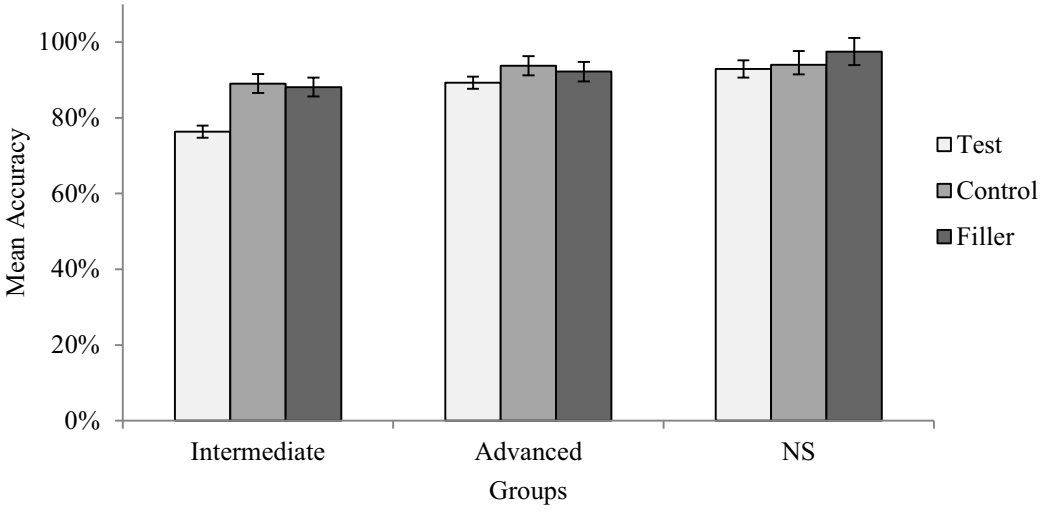

Figure 1. Mean ABX accuracy in each condition for the three participant groups. Error bars enclose +/- 1 SE.

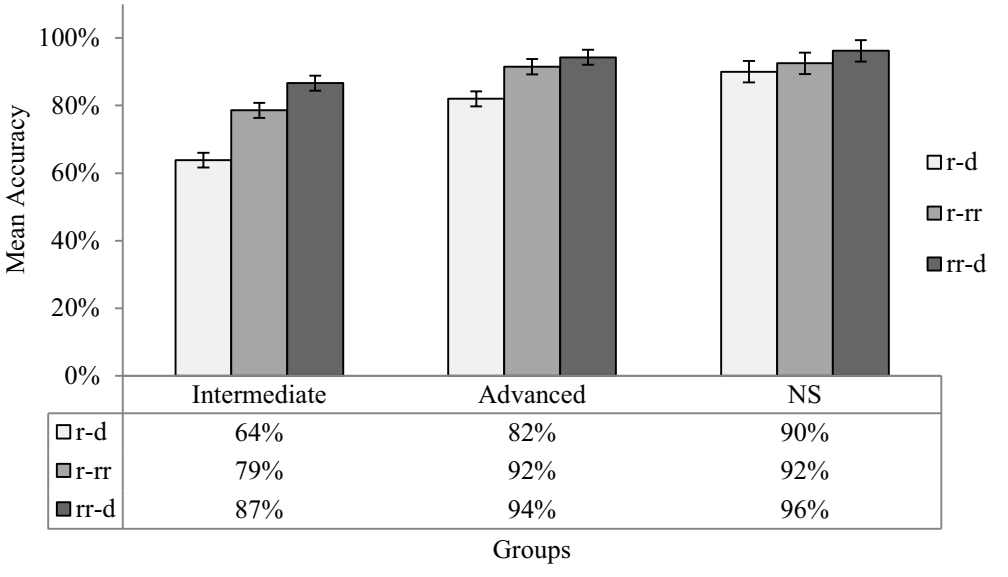

Figure 2. Mean ABX accuracy in the three test contrasts for each group. Error bars enclose +/- 1 SE. The symbol *r* refers to /ɾ/, and *rr* to /r/.

Results for the three contrasts within the test conditions, /ɾ-r/; /ɾ-d/; /r-d/, are presented in Figure 2. A mixed model ANOVA run on the arcsine transformed mean accuracy showed a main effect of group ($F(2, 48) = 24.7$, $p < .001$). This effect was largely due to the intermediate speakers; overall, the advanced speakers did not differ from the native speakers ($p > .6$). The analysis also revealed a main effect of contrast ($F(2, 96) = 21.8$, $p < .001$), but no significant interaction of group and contrast ($F(4, 96) = .7$, $p > .5$). Overall, /ɾ-d/ was discriminated significantly less accurately than the /ɾ-r/ and /r-d/ contrasts, which were not significantly different from each other.

4.2. Lexical Decision Task

Data from one intermediate learner and one native speaker were excluded due to very low accuracy in the control condition. In addition, data from two native speakers were removed because they were early bilinguals. In total, data from 28 intermediate learners, 20 advanced learners, and 8 native speakers were analyzed. The accuracy means for each group on each condition, separated by word and nonword, are presented in Figure 3.

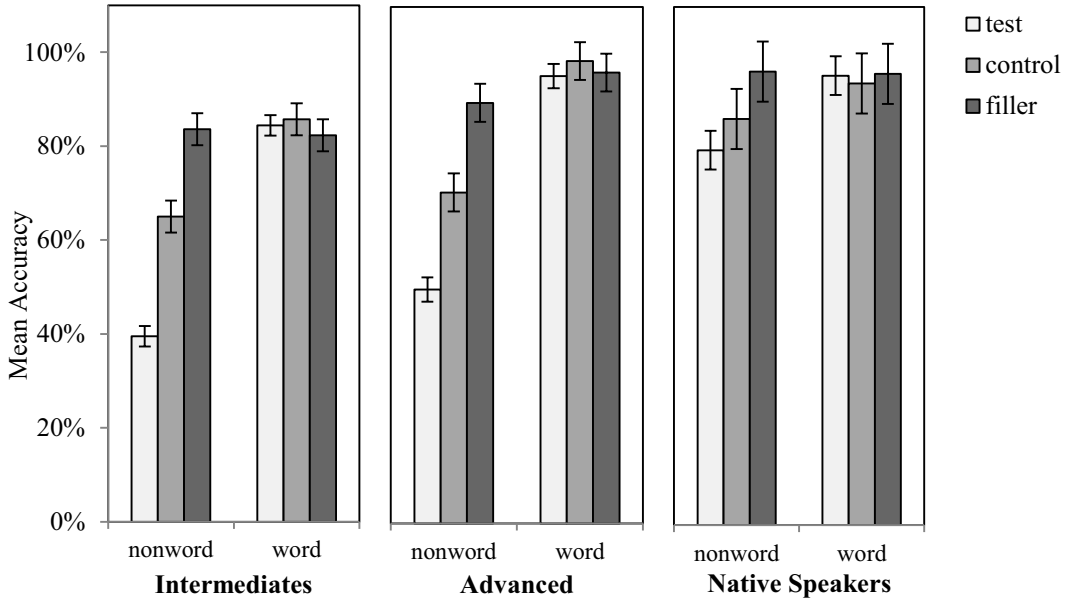

Figure 3. Mean lexical decision accuracy in each condition as a function of lexical status for each group. Error bars enclose +/- 1 SE.

A mixed-model ANOVA was conducted on the mean accuracy for each subject with the factors Group, Condition (test, filler, control), and Lexical Status (word, nonword). Lexical status, condition, and contrast (/ɾ-r/, /ɾ-d/, /r-d/) were declared as repeated effects. There was a large effect of Group ($F(2, 67.4) = 16.0$, $p < .001$), of Condition ($F(2, 489) = 33.6$, $p < .001$), and Lexical Status ($F(1, 489) = 98.2$, $p < .001$). There were two significant interactions (Group x Lexical Status; Lexical Status x Condition). The triple interaction between Group, Lexical Status, and Condition was marginally significant ($F(4, 489) = 2.0$, $p = .09$). The significant interaction between Group and Lexical Status ($F(2, 489) = 6.9$, $p < .01$) further revealed that nonwords were responded to significantly less accurately than words by the intermediate ($F(1, 489) = 85.6$, $p < .001$) and the advanced group ($F(1, 489) = 94.7$, $p < .001$), but not by the native speakers ($p > .05$). The significant interaction between Condition and Lexical Status ($F(2, 489) = 33.7$, $p < .001$) further confirmed that the difference between conditions was only significant for the nonwords ($F(2, 489) = 67.2$, $p < .001$), but not for real words ($F < 1$). Accuracy of lexical decision in L2 Spanish was especially low in the test condition, particularly when it came to rejecting nonwords.

In order to compare the specific contrasts /ɾ-d/, /ɾ-r/, and /r-d/, we restricted the analysis to the test condition only. The results from the three test contrasts (collapsed across words and non-words) are illustrated in Figure 4. As is immediately visible, for both learner groups, the /ɾ-d/ contrast (white bar) yielded slightly more accurate lexical decision responses than the /ɾ-r/ contrast (grey bar).

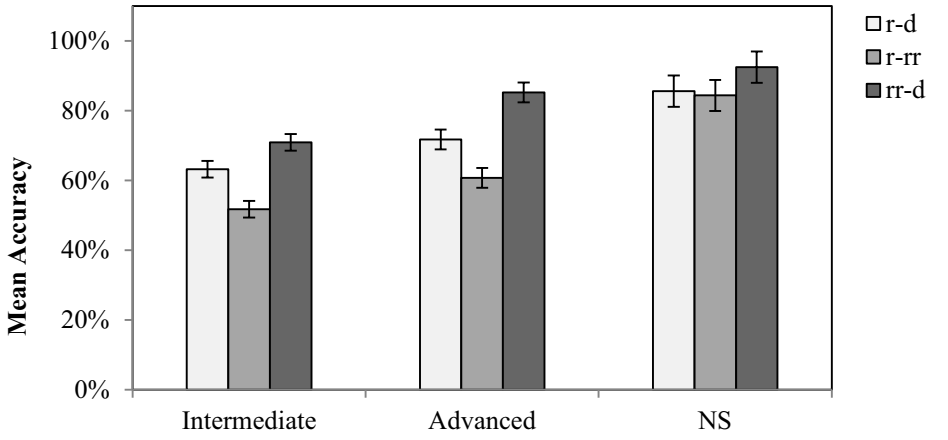

Figure 4. Mean accuracy of the three groups for each contrast in the test condition. Error bars enclose +/- 1 SE. The symbol *r* refers to /ɾ/ and *rr* to /r/.

A mixed model ANOVA declaring the factors Group and Contrast conducted on mean accuracy in this condition revealed a main effect of Group ($F(2, 53) = 24.2$, $p < .001$), of Contrast ($F(2, 610) = 12.8$, $p < .001$), but no interaction between both factors ($F(4, 610) = .9$, $p > .5$). The lack of interaction mostly indicates that the groups, in particular the two learner groups, behaved similarly in the three conditions: learners' accuracy was lowest on the /ɾ-r/ contrast (52% and 60% respectively), intermediate on the /ɾ-d/ condition (63% and 72%) and highest on the /r-d/ condition (71% and 85%).

Table 1 below presents the mean accuracy for each contrast, now separated as a function of lexical status. As shown, accuracy on nonwords for the /ɾ-r/ contrast was extremely low for both intermediate and advanced learners, indicating a limited capacity to distinguish between words and nonwords when the only difference was one of the two rhotics.

Table 1. Mean accuracy for each contrast as a function of lexical status in each group.

Group	Contrast	Nonword	Word
Intermediate	/ɾ-d/	.35	.92
	/ɾ-r/	.23	.81
	/r-d/	.61	.81
Advanced	/ɾ-d/	.46	.97
	/ɾ-r/	.28	.93
	/r-d/	.75	.95
Native Speaker	/ɾ-d/	.71	1.00
	/ɾ-r/	.72	.96
	/r-d/	.95	.90

All groups were less accurate with nonwords than with words. Notably, learners had difficulties correctly rejecting nonwords, and did so at levels below chance. For the /ɾ-r/ contrast specifically, Bonferroni-corrected pairwise comparisons indicate that the difference between intermediate and advanced speakers was significant, despite approaching the .05 significance threshold ($p = .044$); both learner groups also clearly differed from the native speakers ($p < .001$).

5. Discussion

The results of the ABX task show that learners had the most difficulty discriminating /ɾ-d/, less difficulty with /ɾ-r/, and minimal difficulty with the /r-d/ contrast, although all contrasts were discriminated at above chance levels. Even in an ABX task that was made more challenging through the use of a sentential context, three different voices, and all physically different tokens, learners were able to discriminate these phonemes. In fact, advanced learners were not significantly different overall from native speakers in discrimination accuracy. These data closely resemble the findings of Rose (2010a), in which /ɾ-d/ was discriminated less accurately than /ɾ-r/. According to the Perceptual Assimilation Model adapted to second language learning (PAM-L2) (Best & Tyler, 2007), the contrast /ɾ-r/ represents a category goodness assimilation contrast. This is because both the tap and trill are most often mapped onto the same L1 category /ɹ/, as determined by cross-linguistic mapping data, but the trill is a better exemplar of the American English rhotic than the tap, with a fit index rating[5] of 2.21 out of 5 vs. 1.33 out of 5, respectively (Rose, 2010a). In this situation, discrimination is predicted to be good, and this is indeed what has been found for the /ɾ-r/ contrast.

Interestingly, the results of the lexical decision task do not mirror the results of the ABX task. Although learners were significantly less accurate at discriminating the /ɾ-d/ contrast than the /ɾ-r/ contrast, they had the most difficulty correctly distinguishing between words and nonwords with the /ɾ-r/ contrast and less difficulty with the /ɾ-d/ contrast. The PAM-L2 predicts that in the case of a category goodness assimilation contrast, "the perceiver should be able to fairly easily recognize the lexical-functional differences between these L2 phones in minimal lexical contrasts" (Best & Tyler, 2007, p. 27). In contrast to this prediction, learners do not appear to be differentiating between the tap and trill in lexical representations, despite their ability to distinguish these phones. Both groups of learners were below chance level at rejecting nonwords, which means that they were accepting substitutions of the tap for the trill and vice versa as legitimate realizations of those phonemes. Overall, these results support our predictions: learners have not accurately encoded the /ɾ-r/ contrast in intervocalic position in their lexical representations, and the order of accuracy in the discrimination task differed from lexical encoding accuracy. An ability to discriminate two L2 sounds does not guarantee that separate phonological categories will be created for two phonemes, opposed to what has been predicted by PAM-L2.

Given that the trill is a better exemplar of the English rhotic than the tap but both are poor instances of that category, PAM-L2 predicts that a separate phonological and phonetic category is likely to be created for the more deviant of the L2 phones, the tap, but the less deviant phone, the trill, is likely to be equated phonologically to the L1 category /ɹ/ but with a different phonetic category. If this were the case, then the tap and trill would have separate phonological categories, and if learners did have separate categories, we would expect these categories to be used in the encoding of L2 lexical items. Since the results of the lexical decision task suggest an unstable or possibly absent contrast between /ɾ/ and /r/, it appears that learners have not formed clear, separate phonological categories for these phones.

One possible explanation for the lack of differentiation is the low functional load of the /ɾ-r/ contrast in the Spanish language. In total, the tap and trill differentiate between less than 30 minimal pairs (Willis & Bradley, 2008), and in many of these pairs one word is likely more frequent than the other, such as the preposition *para* /paɾa/ 'for' which a search in a Spanish language corpus revealed to be over 1400 times more frequent than the noun *parra* /para/ 'grapevine' (Davies, 2002). Moreover, the tap and trill only contrast in one context, intervocalic position, while elsewhere this contrast is neutralized. Therefore, the low functional load of the /ɾ-r/ contrast, combined with the fact that these phonemes are contrastive in a single phonetic environment, may prevent L2 speakers of Spanish from creating clearly distinct phonological categories for the tap and trill.

In addition, one common realization of the phoneme /r/ besides the standard voiced alveolar trill is a voiced alveolar tap (Hammond, 2006). Although researchers have found that duration tends to distinguish a tap used in the phonological trill context and a tap used in the phonological tap context (Henriksen & Willis, 2010), this feature may not be salient for L2 learners. If learners were accepting a tap in the lexical decision stimuli as a legitimate realization of the trill phoneme, for example **corecto* /koɾekto/ as a possible realization of *correcto* /korekto/, this could help explain why their accuracy for

[5] A higher fit index rating indicates that the sound is a better exemplar of the chosen L1 category.

nonwords with /ɾ/ was very low (29% for intermediate and 20% for advanced). Nevertheless, this does not explain why their accuracy for nonwords with /r/ was also quite low (16% for intermediate and 36% for advanced), and was in fact lower than the accuracy for /ɾ/ nonwords for the intermediate learners. A voiced alveolar trill is not a variant of the tap phoneme, so variation in native speaker production is not a viable explanation for learners' low accuracy on /r/ nonwords. The existence of a tap as a possible realization of the trill phoneme may be contributing to the results of the learners, in particular for the advanced learners who had a harder time rejecting nonwords with /ɾ/ than with /r/, but phonetic variability alone is insufficient as an explanation for the learners' behavior.

In terms of the native speakers' results, the inherent variability in the pronunciation of the trill, as well as the low functional load of the /ɾ-r/ contrast may explain why native speakers were also less accurate on this specific contrast, and tended to accept nonwords containing a rhotic as being real words on average in 29% of the cases (average accuracy on this contrast for nonwords was 71%).

Overall, our results show that an inability to correctly articulate the voiced alveolar trill is not the sole explanation for the lack of differentiation between the tap and trill that previous studies have found in many L2 Spanish learners' production. This neutralization also originates in the way in which words are encoded in the L2 lexicon, which is not wholly determined by perceptual discrimination ability. L2 Spanish learners appear to have an unstable, or even absent, contrast between the tap and trill in their lexical representations, despite their high accuracy in discrimination.

6. Conclusion

The difference between the results from the ABX task and the lexical decision task suggests that discrimination ability does not directly predict how sounds will be encoded in mental representations of words (see also Darcy et al., 2012). Although learners could discriminate /ɾ-r/, this contrast was not clearly differentiated in lexical representations, even by advanced speakers who were mostly graduate students with study abroad experience. These advanced learners had a similar accuracy in the lexical decision task as intermediate speakers, despite not being significantly different from native speakers in the ABX task. These results speak to the importance of investigating L2 lexical encoding, since the ability to discriminate two L2 sounds may be necessary, but not sufficient, for the creation of well-differentiated L2 phonological categories in lexical representations, a possibility which is not addressed by current models of L2 phonological acquisition such as PAM-L2. When only a discrimination task is used to investigate a contrast, the possibility that separate phonological categories have not actually been formed is left unresolved. Future studies should employ additional methods, such as identification tasks or tasks that incorporate priming, in order to obtain a more complete picture of learners' abilities to distinguish novel L2 phones and recognize and encode L2 words. Additional investigations on the Spanish rhotics could examine phonological representations of minimal pairs, in order to see if there is any difference in the representations of words that have a phonological neighbor that differs only in the presence of the other rhotic and words that do not have said phonological neighbor. By more closely studying the nature of L2 lexical representations, future research can tell us more about how new phonological categories are formed and how precisely different phonemes are encoded in the mental lexicon.

References

Best, Catherine T., & Tyler, Michael D. (2007). Nonnative and second-language speech perception: Commonalities and complementarities. In O.S. Bohn & M.J. Munro (Eds.), *Language experience in second language speech learning: In honor of James Emil Flege* (pp. 13–34). Amsterdam/Philadelphia: John Benjamins.

Darcy, Isabelle, Dekydtspotter, Laurent, Sprouse, Rex, Glover, Justin, Kaden, Christiane, McGuire, Michael, & Scott, John. (2012). Direct Mapping of Acoustics to Phonetics: On the lexical encoding of front rounded vowels in L1 English-L2 French acquisition. *Second Language Research, 28*, 5-40.

Davies, Mark. (2002). Corpus del español: 100 million words, 1200s-1900s.

Dupoux, Emmanuel, Pallier, Christophe, Sebastián, Núria, & Mehler, Jacques. (1997). A destressing "deafness" in French? *Journal of Memory and Language, 36*, 406-421.

Dupoux, Emmanuel, Sebastián-Gallés, Núria, Navarrete, Eduardo, & Peperkamp, Sharon. (2008). Persistent stress 'deafness': The case of French learners of Spanish. *Cognition, 106*(2), 682-706.

Face, Timothy L. (2006). Intervocalic rhotic pronunciation by adult learners of Spanish as a second language. In Carol A. Klee & T.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 Press.

Forster, Kenneth I., & Forster, Jonathan C. (2003). DMDX: A Windows display program with millisecond accuracy. *Behavior Research Methods, 35*(1), 116-124.

Gottfried, Terry L. (1984). Effects of consonant context on the perception of French vowels. *Journal of Phonetics, 12*(2), 91-114.

Hammond, Robert M. (2006). The status of [r] and [r̃] in Spanish: A functional analysis. In Mercedes Sedano, Adriana Bolívar & Martha Shiro (Eds.), *Haciendo Lingüística: Homenaje a Paola Bentivoglio* (pp. 91-104). Caracas, Venezuela: Universidad Central de Venezuela.

Hayes-Harb, Rachel, & Masuda, Kyoko. (2008). Development of the ability to lexically encode novel second language phonemic contrasts. *Second Language Research, 24*(1), 5.

Hayes-Harb, Rachel, Nicol, Janet, & Barker, Jason. (2010). Learning the phonological forms of new words: Effects of orthographic and auditory input. *Language and Speech, 53*(3), 367-381.

Henriksen, Nick C., & Willis, Erik W. (2010). Acoustic characterization of phonemic trill production in Jerezano Andalusian Spanish. In M. Ortega-Llebaria (Ed.), *Selected Proceedings of the Fourth Conference on Laboratory Approaches to Spanish Phonology* (pp. 115-127). Somerville, MA: Cascadilla Proceedings Project.

Højen, Anders, & Flege, James E. (2006). Early learners' discrimination of second-language vowels. *Journal of the Acoustical Society of America, 119*(5), 3072-3084.

Hualde, José Ignacio (2004). Quasi-phonemic contrasts in Spanish. In V. Chand et al. (eds.), *WCCFL 23: Proceedings of the 23rd West Coast Conference on Formal Linguistics* (pp. 374-98). Somerville, MA: Cascadilla Press.

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

Jiménez, Beatrice C. (1987). Acquisition of Spanish Consonants in Children Aged 3-5 Years, 7 Months. *Language, Speech, and Hearing Services in Schools, 18*(4), 357-363.

Ladefoged, Peter, & Johnson, Keith. (2010). *A course in phonetics* (6th ed.). Boston, MA: Wadsworth Publishing Company.

Levy, Erika S., & Strange, Winifred. (2008). Perception of French vowels by American English adults with and without French language experience. *Journal of Phonetics, 36*(1), 141-157.

Ota, Mitsuhiko, Hartsuiker, Robert J., & Haywood, Sarah L. (2009). The KEY to the ROCK: Near-homophony in nonnative visual word recognition. *Cognition, 111*(2), 263-269.

Pallier, Christophe, Colomé, Angels, & Sebastián-Gallés, Núria. (2001). The influence of native-language phonology on lexical access: Exemplar-based versus abstract lexical entries. *Psychological Science, 12*(6), 445.

Polka, Linda. (1995). Linguistic influences in adult perception of non-native vowel contrasts. *The Journal of the Acoustical Society of America, 97*, 1286.

Rose, Marda. (2010a). Differences in discriminating L2 consonants: A comparison of Spanish taps and trills. In M. T. Prior (Ed.), *Selected Proceedings of the 2008 Second Language Research Forum* (pp. 181-196). Somerville, MA: Cascadilla Proceedings Project.

Rose, Marda. (2010b). Intervocalic tap and trill production in the acquisition of Spanish as a second language. *Studies in Hispanic and Lusophone Linguistics, 3*(2), 379-419.

Sebastián-Gallés, Núria, Rodríguez-Fornells, Antoni, de Diego-Balaguer, Ruth, & Díaz, Begoña. (2006). First-and second-language phonological representations in the mental lexicon. *Journal of Cognitive Neuroscience, 18*(8), 1277-1291.

Solé, Maria-Josep. (2002). Aerodynamic characteristics of trills and phonological patterning. *Journal of Phonetics, 30*(4), 655-688.

Vitevitch, Michael S., Stamer, Melissa K., & Kieweg, Douglas. (2012). Short research note: The beginning Spanish lexicon: A web-based interface to calculate phonological similarity among Spanish words in adults learning Spanish as a foreign language. *Second Language Research, 28*(1), 103-112.

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

Willis, Erik W., & Bradley, Travis G. (2008). *Contrast maintenance of taps and trills in Dominican Spanish: Data and analysis.* Paper presented at the Selected Proceedings of the 3rd Conference on Laboratory Approaches to Spanish Phonology.

Selected Proceedings of the 2012 Second Language Research Forum: Building Bridges between Disciplines

edited by
Ryan T. Miller, Katherine I. Martin, Chelsea M. Eddington, Ashlie Henery, Nausica Marcos Miguel, Alison M. Tseng, Alba Tuninetti, and Daniel Walter

Cascadilla Proceedings Project Somerville, MA 2014

ISBN 978-1-57473-464-5 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, 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:

Daidone, Danielle and Isabelle Darcy. 2014. *Quierro comprar una guitara*: Lexical Encoding of the Tap and Trill by L2 Learners of Spanish. In *Selected Proceedings of the 2012 Second Language Research Forum*, ed. Ryan T. Miller et al., 39-50. Somerville, MA: Cascadilla Proceedings Project. www.lingref.com, document #3084.

http://www.lingref.com/cpp/slrf/2012/index.html