2022-05-01

The perception of tap and trill in Spanish-English bilinguals

Lesley Vanessa Rodriguez
The University of Texas at El Paso

THE PERCEPTION OF TAP AND TRILL IN SPANISH-ENGLISH BILINGUALS

LESLEY RODRIGUEZ

Master's Program in Linguistics

APPROVED:

Natalia Mazzaro, Ph.D., Chair

Carla Contemori, Ph.D.

Iva M Ivanova, Ph.D.

Stephen L. Crites, Jr., Ph.D.
Dean of the Graduate School

Dedication

This text is dedicated to my family, and friends who have supported me along the way.

THE PERCEPTION OF TAP AND TRILL IN SPANISH-ENGLISH BILINGUALS

by

LESLEY RODRIGUEZ, B.A.

THESIS

Presented to the faculty of the graduate school of

The University of Texas at El Paso

in Partial Fulfillment

of the Requirements

for the Degree of

MASTER OF ARTS

Department of Languages and Linguistics

THE UNIVERSITY OF TEXAS AT EL PASO

May 2022

Acknowledgments

Firstly, I would like to thank my parents, Jose and Veronica, my two brothers, Jose and Sergio, my friends, family, and the professors and classmates I have met throughout my education. Especially my mom, brother Jose, and my friends Monika, Chloe, and Eli. These individuals have believed in me and helped me develop into the person I am. They have supported me throughout my academic career. I would also like to thank Raquel Gonzalez de Anda who guided me academically from my undergraduate to graduate years.

I want to thank my thesis advisor, Dr. Natalia Mazzaro, who had patience and guided me from the beginning to the end of this thesis. Her expertise, feedback, and kindness were valuable during this research project. I would also like to thank Dr. Carla Contemori and Dr. Iva Ivanova for giving me the honor of being part of my thesis committee. They provided me with insightful and meaningful comments when I was revising my thesis. I appreciate every single professor and classmate who has supported me.

I am extremely thankful for the support system that exists in the Linguistics graduate program here at UTEP. My classmates have become wonderful friends who I'll always cherish. I would like to thank Monika Granados Wall, Tracy Huhn, and Juan Davila for their constant support throughout the program.

Abstract

The Spanish [ɾ] tap and [r] trill are two different phonemes that only contrast in the intervocalic position. In Spanish, the tap and trill contrast can only be found in thirty words, for example in <*pero*> (but) and <*perro*> (dog). Studies, such as Bradley and Willis (2012), demonstrate that, in production, taps and trills are variably neutralized among monolingual Spanish speakers. Neutralization of taps and trills creates an ambiguous signal that could be difficult to decode by heritage Spanish speakers. In fact, previous studies found that heritage speakers, such as monolingual Spanish speakers, neutralize these phonemes (Amengual, 2016). The purpose of this study is to determine if variable neutralization in production is caused by difficulty in the perception of the contrast in heritage speakers. To test this hypothesis, a perception study was constructed to compare the discrimination of taps and trills in intervocalic position. Monolingual Spanish speakers and heritage speakers participated in performing a word categorization task where they choose the word they heard from minimal pairs containing the tap and trill shown on the screen. The results of this study showed that heritage speakers can distinguish the sounds but are not on par to the monolingual Spanish participants. Education was a significant factor in being able to distinguish the tap and trill. The errors by all participants showed there is bidirectionality in confusing the trill with the tap and the tap with the trill. The results of this study can shed light onto the dynamics of language variation in contact situations and the nature of bilinguals' phonological systems.

Table of Contents

List of Figures

Lists of Tables

1. Introductions

This study investigates Spanish-English bilinguals' perception of a Spanish phonemic contrast /ɾ/ and /r/ as in *pero* /péɾo/ 'but' and *perro* /péro/ 'dog'. Taps and trills are variably neutralized (i.e., the distinction is lost) among native Spanish speakers (Bradley and Willis, 2012). Recent studies (Henriksen, 2015; Amengual, 2016; Cummings Ruiz and Montrul, 2020) found that early bilingual speakers also neutralize taps and trills in speech. Neutralization of taps and trills creates an ambiguous signal that could be difficult to decode by all listeners, but especially bilinguals, because their linguistic exposure is divided between two languages and their frequent use of Spanish is reduced once they start school (Cummins Ruiz and Montrul, 2020).

The Spanish /ɾ/ tap and /r/ trill are two different phonemes that contrast only in intervocalic position. The tap and the trill are both voiced alveolar sounds and the main difference between them is in their manner of articulation; to produce the tap, the tongue hits the alveolar ridge once, and to produce the trill, the tongue hits the alveolar ridge about three times (Hualde, 2005). The number of rapid touches of the tongue with the alveolar ridge is the main articulatory distinction between words such as <pero> and <perro>. In addition, Quilis (1993) states that on average a trill is 85 ms. long. According to Quilis (1997) a canonical tap lasts 20 ms. In Spanish, there are only thirty minimal pairs where the tap and trill contrast is found, for example in <*cara*> [káɾo] 'expensive' and <*carro*> [káro] 'car'.

Regarding its distribution, the tap appears in the intervocalic position (e.g., <cara> [káɾa] 'face'), in complex onset position (e.g., <tren> [tɾen] 'train'), and in coda position (e.g., <mar> [maɾ] 'sea'). The trill appears in the intervocalic position (e.g., <carro> [karo] 'car'), in word initial position (e.g., <rosa> [rosa] 'rose'), and in syllable onset position after a consonant, such as /s/, /n/, and /l/ (e.g., <Israel> [israel]). In post-nuclear position, the rhotic can be variably realized as [ɾ], [r] and [ɹ] (e.g., <amar> [amáɾ] or [amár] or [amáɹ] 'love'; <puérta> [pu̯éɾta] or [pu̯érta] or [pu̯éɹta] 'door') (Quilis, 1997). So, in this position, the phonemes neutralize (i.e., the contrast between tap and trill is lost), resulting in the rhotic archiphoneme /R/.

Besides the variability found in post-nuclear position, studies have found that that there is also variability in contrastive intervocalic positions. For instance, Bradley and Willis' (2012) study provided evidence that /ɾ/ and /r/ neutralize in intervocalic position in Mexico and other parts of Latin America. Particularly, this study found that taps are produced with longer durations (27 ms. on average) than the canonical taps (20 ms.) reported in other studies such as Quilis (1993). If taps are produced with longer durations, that would mean that they become more similar to trills, which are traditionally longer than taps.

The question addressed in this thesis is whether the neutralization of rhotics in heritage speakers (HSs) could be due in part to the variation in the realization of rhotics in monolingual speech that make it difficult for bilinguals to perceive these sounds. Monolingual speakers have continuous exposure to the L1, unlike bilingual speakers, particularly HS, whose linguistic exposure is divided between two languages, and their exposure to Spanish decreases significantly once they begin school (Cummins and Montrul, 2020). Reduced L1 input coupled with ambiguous signal could partly explain the variation seen in HSs, as explained in the literature review below.

2. Literature Review

2.1. Variation in monolingual speech

Bradley and Willis (2012) analyzed the variation in the production of taps and trills in Mexican Spanish among ten monolingual university students in Veracruz. Their data collection consisted of an informal interview and a narrative task describing a picture book. In the narrative, the average duration of taps was 27 ms., while trills were 72-89 ms. As stated, these taps were longer than the canonical taps reported by Quilis (1993). Some of the trills were realized with two visible contacts of the tongue on the alveolar ridge and some trills were realized with one. Their results revealed that, in the intervocalic position, the trill was sometimes realized as a tap, thus showing neutralization in this position.

Another study (Kouznetsov & Bertránn, 2008) compared the production of trills in the monolingual speech of Spanish and Russian. This study revealed that, in about 15% of the Spanish trill realizations, there was only one closure. Kouznetsov & Bertrán (2008) argued that the Spanish trill is not a geminate correlate of the tap. That is, the trill is not made of several successive taps. So, the question that needs to be explored is, in cases where the trill is produced with only one closure, is it still a trill or has it become a tap? Experimental research was carried out to examine spectral properties of Spanish and Russian trills in an effort to better understand the variation. For their study, they compared Russian speakers with Spanish monolinguals from Spain. Their results showed that, in Russian, 90% of the trills were realized with one closure, whereas in Spanish, this percentage was much lower (15%). In Spanish, when the trill was produced intervocalically, more trills with one closure occurred when /r/ was followed by rounded vowels. In Russian, more trills with one closure occurred before non-front vowels (/e/ and /i/). Thus, there was neutralization of the trill in both languages, Spanish and Russian.

However, neutralization was much more common in Russian than Spanish and they were influenced by different vocalic contexts.

Diaz-Campos (2008) investigated rhotics in Venezuelan Spanish. The data consisted of 36 speakers from the corpus of Sociolinguistic study of Caracas' (1987). Diaz-Campos found that 36.2% of the trills were standard and 63.8% were innovative (non-standard) variants. From the non-standard variants, 61% were approximants with zero occlusions. Interestingly, younger participants preferred the use of the non-standard approximant, while older generations preferred the use of the standard trill.

Henriksen and Willis (2010) investigated the acoustic characteristics of taps of trills in urban Andalusian Spanish. Their data consisted of 16 participants who performed a narrative task using pictures from the book *Frog, Where Are You?* In the analysis, they found that the most common phonemic trill was produced with one visible occlusion, but they also found trills with zero occlusions. Similarly to Diaz-Campos' (2008) study, the younger generation had higher rates of the non-standard shorter trills than the older speakers. Their results showed that 44.2% of trills were produced with one-closure, 26.01% with zero-closure, and 29.8% with two or three closures. Overall, their results revealed that, in Andalusian Spanish, the tendency is to produce trills with one closure.

2.2. Research on rhotics in other languages

Ball et al. (2001) investigated the acquisition of rhotics in Welsh-English bilingual children. Welsh contrasts from English in that Welsh has a trilled–r (voiced /r/ and voiceless /r̥/) and English has an approximant-r. The bilinguals (N=100) who participated in their study were divided according to age groups and dominance of the L1 and L2. Ball et al. (2001) recorded a reading list of 50 words containing rhotics in English and Welsh. The results showed that the

Welsh trill had a wide range of variation (around 20 variants) in its realization and the variability decreased over time as subjects gained more experience with the language. The Welsh-dominant group had a 50% correct score for trill production, and this was achieved only by the two oldest subjects in groups 4;0 to 5;0. The English-dominant group had a 30% target realization of trilled-r and this was only in word final position. For most subjects in this study, trilled-r in word initial position was more difficult to produce correctly. In this case, most of the productions were approximant /r/. Rhotic approximant realizations were also the most common in medial and final positions in the word, although trills were also found in final positions in older groups. The authors concluded that the use of [ɹ] for /r/ could be a result of influence from English. However, the use of fricatives for final /r/ has been shown to have an acoustic motivation as the duration of fricatives is fairly close to word- final trills (reported in Ball & Williams, 2001) and the noise component of fricatives may be considered to bear some perceptual resemblance to the short periods of turbulence accompanying the individual strikes of the trill.

Chen and Mok (2019) compared Mandarin-English sequential bilinguals to investigate how bilinguals' two languages interact when the same phoneme in two languages was realized phonetically differently. The authors examined the acoustic and articulatory features of English and Mandarin /ɹ/ produced by highly proficient Mandarin-English bilinguals. Ten Mandarin-English bilingual speakers and ten American English monolingual speakers were recorded with ultrasound imaging technique. Results showed that highly proficient Mandarin-English bilinguals produced English /ɹ/ without L1 influences. They used language-specific phonetic details for English and Mandarin /ɹ/. It implies that language-specific phonetic realizations can be learned for the same phonological category by highly proficient bilingual speakers.

Additionally, it is important to consider the impact of formal language education on distinguishing between different sounds. Studies like Goto (1971) demonstrate that formal education can provide the HSs with the skills to distinguish similar sounds. Goto (1971) investigated how Japanese speakers perceived English /ɹ/ and /l/. In Japanese, these sounds belong to the same phoneme. Goto (1971) collected his data through a perception task where participants listened to words such as <right> and <light> to see if Japanese-English speakers could discriminate between the two phonemes. Results of the perception task demonstrated that advanced Japanese-English speakers are able to produce the different phonemes, but their phonemic perception rated lower with scores of 6/8 to 1/8 respectively. For comparison, monolingual English speakers scored 8/8 in the production and 7/8 in the perception task. Goto's (1971) results showed that Japanese-English speakers find it challenging to discriminate between /ɹ/ and /l/. However, over time, Japanese speakers can be taught to discriminate these sounds. In other words, formal education in a heritage language can be important in being able to distinguish similar sounds. Even though Goto (1971) focused on second language learners of English, formal education for a HS can be important in being able to distinguish sounds that are phonetically similar to each other in the native language.

An influencing factor for a HSs' ability to distinguish between sounds is the setting in which they learn the language. HS generally learn the minority language at home and the majority in a formal setting, such as school. For the first few years of their lives, HSs are mostly exposed to the minority language, and once they begin school, they are exposed to the majority language. As a result, HSs only have a small window to acquire these sounds and learn to distinguish what sounds belong to which language. Monolingual speakers do not face this problem since they are exposed to one language all their life, whether it be an informal/formal setting, so it is easier for them to distinguish between different sounds compared to HSs. In addition, as explained in section 2 that discusses Speech Theories, because HSs generally receive their formal education in the majority language (or L2), they do not have the orthographic representation to clarify their perceptions of ambiguous sounds (Ohala, 1989). On the other hand, monolingual speakers not only have frequent input from native speakers but they are also taught to distinguish contrastive sounds at school through orthography. With no influence from orthography, HSs have no chances of correcting their non-target perception of ambiguous sounds.

2.3. Variation in bilingual speech

As stated earlier, monolingual Spanish speakers are not the only ones neutralizing the distinction between taps and trills; there is evidence of this phenomenon in the speech of Spanish Heritage Speakers (HSSs). Researchers have studied and compared HSSs to second language learners of Spanish to better understand this neutralization phenomenon in bilingual speech.

Daine and Zahler (2021) investigated trill production of L2 learners of Spanish compared to HSSs. The L2 learners produced native-like trills when they followed a central vowel /a/ and front vowels /i, e/, yet native-like trill production was disfavored by back vowels /u, o/. The L2 learners produced the tap more consistently in higher frequency words and were able to produce native-like variants. Even though HSSs grew up hearing both languages, the minority and the majority languages, they showed a phonological preference for the majority language. Cross-linguistic influence between one language to another can occur when the sounds in each language perform different functions, such as phoneme in one language and allophone in another (Daine and Zahler, 2021).

Henriksen (2015) investigated the production of the tap and trill in eight monolingual Spanish speakers who immigrated to the US from different areas of Mexico, and eight HSSs residing in Chicago. They performed a narrative task, where they naturally produced taps and trills. Data analysis demonstrated that monolingual Spanish speakers produced more occlusions and longer-lasting trills compared to HSSs. Trill production with three and four occlusions were uncommon in both groups. Data also showed that HSSs produced some taps with more than one occlusion.

A more recent study on variation in intervocalic rhotics in bilingual speech was conducted by Amengual (2016). Amengual (2016) looked at the phonological contrast of tap and trill in 40 Spanish HSSs and 20 second language (L2) learners of Spanish. The HSSs were further subdivided into Spanish dominant and English dominant. The data was collected in Northern California using a read-aloud task. His results revealed that L2 learners and English-dominant HSs produced non-canonical phonemic trills with one or zero occlusions. On the other hand, Spanish-dominant HSs produced the majority of their trills with two or three brief occlusions. They also maintained the Spanish tap-trill phonological contrast largely by means of segmental duration.

Kissling (2018) investigated rhotic production in 15 Central Mexican and Salvadorian HSSs living in Richmond, Virginia. Besides HSSs, her participants also included 5 long-term immigrants, 8 L2 learners and 10 Spanish monolingual speakers. Participants performed three tasks: a conversation task, a narrative task from *Frog, Where are you?* and a read out aloud task of 22 idioms. Her results showed that proficiency and language dominance were not good predictors of HSSs' rhotic production but, instead, she found a correlation between HSSs' identity as Latino/Hispanic with more target like rhotic production. Those HS who identified as

Latinx/Hispanic produced rhotics that were more similar to those produced by monolingual Spanish speakers than those who did not identify with being Latinx/Hispanic. The HSSs that did not identify as Latino/Hispanic produced rhotics like L2 speakers. Kissling (2018: 50) proposes that "stronger cultural identity could be related to a stronger desire to sound native and motivate a HSSs to pay closer attention to phonetic features". In addition, those HSSs with higher exposure to Spanish from birth had more target-like productions of rhotics than those with less exposure to the language.

Cummings Ruiz and Montrul (2020) investigated the rhotic production of twenty-four adult participants: six sequential bilingual HSs, six simultaneous bilingual HSs, six monolingual Spanish speakers, and six L2 Spanish learners. The heritage groups and the L2 speakers were from the Chicagoland area and the monolingual speakers were from Mexico. Data was collected through a storytelling and a picture naming task. Results showed that the sequential bilinguals, not the simultaneous bilinguals or the L2 learners, patterned similarly to the monolingual Spanish speakers in their production of taps and trills. When producing Spanish words containing a rhotic sound, the sequential bilinguals produced the taps and trills with a similar duration and number of apical occlusions to the productions of the monolingual Spanish speakers. The simultaneous bilinguals patterned differently to the monolingual Spanish speakers. The authors argued that this difference may be due to an incomplete acquisition of the Spanish rhotic system, due to their early exposure to English. In other words, the simultaneous bilinguals may not have received enough Spanish input to fully define L1 categories for the Spanish rhotic system when they already started acquiring English.

From the above, the variable realization of taps and trills in HSSs may be due to variability in the input of these sounds by Spanish monolingual speakers and other Spanish-

English bilinguals. Variation in taps and trills can cause difficulties for HSSs, who divide their time between Spanish and English use and input. In addition, HSSs generally have less formal education in Spanish than their monolingual equals and the impact of orthography on the perception of ambiguous acoustic signals has been found to be significant in previous studies reviewed in the section below.

2.4. Speech Theories to explain variation in perception

This section will review two types of theories that relate to the variation in production and perception of sounds. One of the theories was proposed by Ohala (1989; 1993) and was based on sound variation in monolingual speech. This theory explained how variation in the production of certain sounds can lead to non-target perception and eventual sound change. The other theory, the 'Speech Learning Model' proposed by Flege (1995) accounts for age related limits on the ability to perceive sounds in a native like manner in the second language.

Ohala's (1993) theory on variation in monolingual speakers focuses on variation in the speaker. This variation in production can lead to misperception of the intended signal, especially when it comes to sounds that are variably realized and can become difficult to discriminate from each other, such as the Spanish tap and trill. Speakers and listeners are generally not aware of the variation in the signal and listeners may imitate the pronunciation they hear or believe they heard. Inexperienced speakers with low levels of formal education in the language or lack of frequent input from native speakers hear the 'wrong/misheard' pronunciation and take it at face value. Since they do not have orthography to confirm their perceptions, they reproduce the non-target pronunciation. On the other hand, monolingual speakers not only have frequent input from native speakers, but they are also taught to distinguish contrastive sounds at school through orthography, i.e., the tap as <r> in words such as 'pero' (but) and the trill as <rr> in words such as ‘perro’ (dog). There are even rhyming songs that native Spanish parents and teachers teach to children to teach them how to produce the trill.

Research shows that during literacy acquisition, the correspondence between phoneme and grapheme become tightly interconnected such that they could be considered as two faces of the same coin (Frost & Ziegler 2007). Phonological awareness, that is the capacity to consciously manipulate phonemes, is also dependent on literacy (Bassetti 2006). Research on adult monolingual speakers has shown that adults’ performance in phonological awareness tasks is affected by orthographic representations. For instance, learning to read the Roman alphabet seems to improve performance on other types of auditory tasks such as those that require participants to delete certain phonemes from a word or to add them in (Morais et al. 1979).
If the onset of literacy shapes the perceptual system with listeners integrating the orthographic input with the audio input when establishing phonetic/phonemic categories, those speakers who have undergone schooling in L1 Spanish (and have learnt to read and write in Spanish) may achieve a sharper contrast between the tap and the trill than those who have not had schooling in L1 Spanish and, thus, only rely on their auditory perception. Thus, HSSs may exhibit more variability in tap and trill production, because they have not had formal education in Spanish. HSSs usually learn the language from their family, caretakers, and/or friends without much input from school and teachers. It is possible that children learning to read and write are corrected at school when they do not use the correct spelling for the tap and the trill. Learning the grapheme-phoneme correspondence for these sounds may help listeners establish the phonetic/phonemic categories for taps and trills. It may also act as a block for sounds that may be perceptually similar in certain phonetic contexts (Ohala 1993).

The second theory proposed by Flege (1995), called the 'Speech Learning Model' (SLM), relates to the perception and production of L2 sounds by bilingual speakers and the effect of the L1 on the perception and native-like production of sounds in the L2. SLM can be used to explain how the phonological system of two languages interact in the brain of a heritage Spanish speaker. This model is also used to show how the L2 learner's pronunciation is affected by various years of learning their L2 and how the age of L2 acquisition affects the learner' proficiency in the L2. SLM has been applied to determine how learners perceive L2 sounds, whether these sounds are assimilated into the L1 system because they are similar to the L1, or whether new categories are created for perceptually different sounds in the L2.

According to SLM, bilinguals are more likely to acquire a sound in the L2 if it is different to the sound in the L1. On the contrary, they are less likely to acquire an L2 sound if it is similar or identical to a sound that already exists in the L1. Bilinguals are hypothesized to first use L1 sounds as "substitutes" for the L2 sounds, which can potentially interfere or block the formation of new phonetic categories for L2 sounds (Flege, 1995). This is because bilinguals may fail to discern the phonetic differences between pairs of sounds in the L2 and L1, either because phonetically distinct sounds in the L2 are "assimilated" to a single category or because the L1 phonology filters out features (or properties) of L2 sounds that are important phonetically, but not phonologically, or both (Flege, 1995). With increasing exposure to the L2, bilinguals are expected to acquire new categories for the similar, but not identical, categories. With increasing exposure to the L1, bilingual HSSs will develop the categories of their L1, and they are expected to differentiate between the sounds of the L1 and L2, even when these are similar. Those with less exposure to the heritage language, however, may fail to differentiate between similar sounds in their L1 and L2, leading to an "assimilation" of two similar sounds into one phonetic category (Flege, 2007).

If we apply SLM to the acquisition of rhotics in Spanish-English bilinguals, the acquisition path would start with HS acquiring the sounds of their heritage language Spanish, which includes taps and trills. Taps are typically acquired first, and trills are acquired at around 7 years (Cummings and Montrul 2020). Most HSSs would start being exposed to English at school when they are still in the process to acquire the trill. Those who are at the beginning stages of their acquisition of English may use the L1 category (tap) as a substitute of the approximant. With increasing L2 exposure, HSSs predicted to distinguish between taps and approximants and to create different categories for each. Cummings Ruiz and Montrul (2020) mention that even though the tap exists in both languages, the tap can be considered a new sound for L2 learners. This could be due to the fact that in English, the tap is 'registered' as an allophone of a phoneme, and in the L1, they have to 're-register' the tap to be its own phoneme. Then, with more experience in the L1, HS will also create another category for the trill.

The proposed path in the acquisition of rhotics by Spanish-English bilinguals is shown in Table 1:

Table 1. evolution in the acquisition of rhotics by Spanish-English bilinguals.

	HS with little exposure to the L1 and initial exposure to the L2	HS with increased exposure to the L2	HS with increased exposure to the L1 and L2
L1 (Heritage language)	/r/ → [ɾ] tap → [r] (trill)	/r/ → [ɾ] tap → [r] (trill)	/r/ → [ɾ] tap → [r] trill
L2 (English)	/r/ → [ɾ] tap	/r/ → [ɹ] approximant	/r/ → [ɹ] approximant

The trill is shown in parenthesis, meaning that they are still not completely acquired, so learners may not be able to perceive it or produce canonical realizations of this sound. Seeing as the present study is not testing the perception and production of approximants in the L1, the experiment used does not include this sound as an option to listeners. However, this would be an interesting future study to pursue.

Two of the studies reviewed above tested the SLM on HSSs' productions of rhotics. As mentioned previously (Variation in bilingual speech, 2.3), Henriksen (2015) conducted an acoustic analysis of rhotic productions in a semi-informal narrative task. His study included 8 late bilingual speakers who spoke Spanish from birth and acquired English after puberty, and 8 early bilinguals who were born in the US but had Spanish caretakers. Henriksen's results showed that both groups of HSSs maintained the tap-trill contrast throughout adulthood even though their exposure to English had increased. The contrast between the tap and trill was mainly realized via overall duration, not lingual contact, where there was a lot of variation. Most importantly, Henriksen (2015) rejected the hypothesis that phonetic variation in rhotic production results in phonological instability in the heritage grammar. He proposed that sound categories that are acquired early in life are not susceptible to merging in the L1 and L2. Another finding from his study was the high individual variation encountered in the production of trills in monolingual and bilingual speakers. Interestingly, the auditory analysis of the 0-closure variants of the HSSs included approximant-r realizations, while the 0-closure variants of the monolingual Spanish speakers included assibilated rhotics or fricatives.

The Cummings Ruiz and Montrul (2020) study provided evidence in favor of the SLM in their study of the Spanish rhotic system in bilingual speech. Their study showed that simultaneous and sequential bilinguals did create two separate phonological categories for the tap and the trill. However, the simultaneous bilinguals did not pattern similar to the Spanish monolingual speakers. Their results showed that HSSs were able to produce tap and trill contrast, but differed from the Spanish monolingual speaker and the sequential bilinguals in the number of apical occlusions for the trill and the duration for the tap. Most important to the present study, the authors (Cummings Ruiz and Montrul, 2020) argued that heritage speakers may have gained the ability to perceive but not to produce the alveolar tap and trill in childhood. Thus, the simultaneous bilinguals may be able to produce two distinct sounds but be unaware of the "native" duration of the alveolar trill. This result differs from Henriksen's in that his heritage speakers did resemble the performance of monolingual speakers in rhotic duration, yet not apical occlusions and, according to Henriksen, the HSs have successfully created the tap trill contrast during childhood and maintained it in adulthood. As mentioned, Cummings and Montrul (2020) does support the predictions of SLM, because the simultaneous bilinguals were able to produce target-like Spanish rhotics, which means that the input received throughout childhood was enough to establish the tap and trill categories. The L2 English speakers had approximant realizations for the Spanish trill, which can be seen as an influence from English.

SLM focuses on age of acquisition, and it predicts that learners who are younger will achieve a closer to native-like proficiency in the L2 than those who are older. This is because, as the phonological system matures, it can become more difficult to perceive and differentiate between contrastive sounds. Early learners of an L2 are able to better discern cross-language phonetic differences than individuals who learn an L2 later in life (Flege, 1995). Late L2 learners often have difficulties in forming new phonetic categories if the L2 sound uses phonetic features not used in the learner's L1. The research in Cummings and Montrul (2020) showed that L2 learners of Spanish had non-target realizations of the Spanish rhotics, such as the approximant. Thus, the L2 learners may not have successfully established different categories for Spanish trills and taps. On the contrary, Cummings and Montrul's study did not find evidence of English alveolar approximants in either group of HSSs. In accordance with SLM, these results suggest that the later age of acquisition of Spanish for the L2 learners may negatively affect their Spanish rhotic production. Besides taps and trills, the other variant found in both groups of the HSSs was assibilated rhotics. Assibilated rhotics have been attested in native monolingual Spanish speakers of northern Mexican Spanish (Mazzaro and Gonzalez, 2020), so the use of assibilated rhotics by HSSs provides further evidence of how close their speech is to the monolingual norm.

The studies summarized above focused on the production of rhotics by Spanish-English bilingual speakers of different levels of proficiency in both languages. However, to the best of my knowledge, no study has focused on the perception of Spanish rhotics in bilingual speakers and how their identification and/or discrimination of taps and trills can affect their production of such sounds. This is an important consideration because, although Henriksen (2015) and Cummings Ruiz and Montrul (2020) used SLM to make predictions and explain their results, their studies did not test perception, which makes using SLM problematic. As stated earlier, SLM hypothesizes that the accuracy of perceptual representations for L2 sounds places an upper limit on the accuracy with which the L2 sounds can be produced, thus emphasizing a close connection between perception and production.

In order to explain the production of rhotics in bilingual speech, it is important to test if non-target productions could be due to difficulties in the perception of taps and trills. Thus, the purpose of this study is to determine if there are difficulties with the accurate identification of taps and trills by bilinguals. The following section explains the hypothesis proposed in this study and how this hypothesis will be tested.

3. Hypothesis

I hypothesize that the variation in rhotic production can affect the perception of taps and trills and their accurate categorization. By variation in production, I mean cases where trills are realized with one closure and taps with longer duration than the canonical ones (Henriksen and Willis, 2010; Willis and Bradley, 2012). This variation, or neutralization of taps and trills in production, can lead to difficulties in the identification and categorization of such sounds by bilinguals. The variation in production of taps and trills can be a challenge for all listeners, though particularly for early Spanish-English bilinguals, who have limited input and almost no formal education in the heritage language.

In order to test this hypothesis, the discrimination of taps and trills in HSSs will be compared to that of monolingual speakers. In addition, the study will also compare simultaneous bilinguals against sequential bilinguals to determine if age of acquisition of the L2 can influence the successful creation of two different rhotic categories in the native phonological system.

Monolinguals are expected to correctly identify the tap and trill because, although there is variation in the acoustic signal, they have had formal education in Spanish and ample exposure to Spanish to facilitate the discrimination of the target sounds. Sequential English bilinguals are also expected to behave closer to the monolingual norm because they arrived in the US later (after puberty) and they have had formal education in Spanish during their formative years. In addition, sequential bilinguals had an opportunity to develop and use their L1 before acquiring the L2. However, the variation in the production of the taps and trills is expected to affect the discrimination of taps and trills with greater strength in HSSs, because they had less formal education in Spanish and less input and frequency of use to allow them to disambiguate the variable (confusing) signal. In other words, HSSs are expected to have a higher rate of non-target perception than sequential bilinguals and monolinguals. The latter are expected to have few or zero non-target perceptions. In addition, the sequential Spanish bilinguals are expected to have a higher rate of non-target perception than HSSs, and more than the monolinguals and the sequential English bilinguals.

Another important issue is the directionality of the discrimination in the perception. That is, will taps be discriminated as trills, or will trills be discriminated as taps? Based on the literature reviewed, the type of discrimination is expected to be bidirectional, with taps being identified as trills and trills being identified as taps. Since some studies (Diaz-Campos, 2008; Henriksen and Willis, 2010; Bradley and Willis, 2012) found that trills can be realized with zero or one occlusion, or with shorter duration, trills are expected to be discriminated as taps. Other studies (Bradley and Willis, 2012) showed that taps can be realized with longer duration than canonical taps, so they could be perceived as trills.

Finally, another aspect to consider is whether a strong Hispanic identity and a higher exposure to Spanish from birth can influence the identification of taps and trills. Kissling's study (2012) found that those HSSs who identified as Latinx/Hispanic produced rhotics that were more similar to those produced by monolingual Spanish speakers. In addition, those HSSs with more exposure to the Spanish language had more target-like productions of rhotics than those with less exposure. The author (Kissling, 2018) proposed that a stronger cultural identity could be related to a stronger desire to sound native. Assuming that production is linked to perception, it is hypothesized that HSSs who identify as Latinx/Hispanic will have a higher correct identification of rhotics than those who do not identify with the Latinx culture. Likewise, the HSSs who have more exposure to Spanish will have less difficulties the incorrect identification of taps and trills.

4. Method

4.1. Procedure

Participants were tested on their own time. The QuestionPro link was distributed among family, friends, and students from the University of Texas at El Paso. All participants gave their informed consent for inclusion before they could participate in the study. The study was conducted in accordance with the University of Texas at El Paso Ethics Committee of the Institutional Review Board.

4.2. Participants

There was a total of 116 adult participants, 16 were excluded from analysis due to hearing impairments, incomplete surveys, and selecting random answers. One hundred participants were divided into four groups. Five were Spanish monolinguals who were born and raised in Mexico. The sequential bilinguals were further sub-divided into two groups, because some of them were more English dominant and some of them were more Spanish dominant. There were thirty-two Spanish dominant sequential bilinguals, who learned English after 13 years of age, who were born and/or raised in Mexico, and then immigrated to the US after puberty. There were twenty English dominant sequential bilinguals, who acquired English as a first language and learned Spanish as a second language after puberty. Forty-three participants were simultaneous bilinguals (HSS) that were exposed to Spanish and English since birth. The HSSs were from El Paso, Texas, which is a predominantly Hispanic border town along the US-Mexico border. Approximately 82.9% (717,630) of the 865,65 population in El Paso identifies as Hispanic/Latinx (U.S. Census Bureau, 2020). According to the U.S. Census Bureau (2020), approximately 66.4% (517,796) of the population are exposed to Spanish since they are 5+ years old, thus the majority of the population in El Paso are raised as Spanish-English bilinguals.

Table 2. Participants' demographic information.

	Simultaneous Heritage Speakers (n=43)		Sequential English bilingual (n=32)		Sequential Spanish bilingual (n=20)		Monolingual (n=5)	
Culture Identification scores	Mexican 8.13 American 8.44 Other .62		Mexican 9.9 American 5.15 Other 1.81		Mexican 6.47 American 7.25 Other 0		Mexican 8.6 American 3.4 Other 0	
Years of Education	Mean: 16.6 SD: 1.55		Mean: 16.8 SD: 1.23		Mean: 16.2 SD: 1.15		Mean: 14 SD: 4.58	
	SPAN	ENG	SPAN	ENG	SPAN	ENG	SPAN	ENG
Speaking	7.4	8	9.3	7.2	5.3	8.9	8.8	0
Comprehension	7.8	8	9.3	7.5	6.2	9.1	8.8	0
Reading	7.2	8	9.3	7.6	5	9.1	8.4	0
More comfortable language	40.1%	59.9%	78%	22%	27.8%	72.2%	94%	6%
Daily language exposure	43.6%	56.4%	63.3%	36.7%	32.8%	77.2%	82%	18%

4.3 Materials

The data in this study was collected through an online survey tool, QuestionPro. All participants were over the age of 18. The experiment took an average of eleven minutes to complete. First, through QuestionPro, participants had to complete a linguistic background questionnaire and then they performed a word categorization task. A word categorization task was chosen to verify the discrimination of the tap and the trill. A picture selection task was considered at first, but some of the words fall under different lexical categories and are ambiguous, which makes it difficult to find pictures (e.g. *pero* 'but' and caro 'expensive'). Goto (1971) and Bradlow, Pisoni, Akahane-Yamada, & Tohkura (1997) used a categorization task to test Japanese-English bilinguals' perception of English [l] and [ɹ]. Since the present study is similar in nature to Goto (1971) and Bladlow et al (1997) studies, the categorization task was used to evaluate listeners' discrimination of taps and trills.

The categorization task was previously recorded in list form by a female L1 Spanish speaker. This task consisted of listeners hearing words with the tap and trill and selecting the word they believed they heard from the screen. For example: the participant heard *pera* [péra] 'female dog' and they saw two words on the screen <perra> and <pera>. Participants had to select the word on the screen that best corresponded with the stimulus they heard. There were two versions of the categorization tasks. The only difference between the two versions was the order in which they were presented. QuestionPro randomly assigned the different versions to participants. Before they performed the main experiment, participants went through six practice trials with six words that did not contain the sound in question. After completing the practice trial, participants completed the real experiment consisting of 15 words containing the tap and trill, with 15 words used as distractors. The lists are included in Appendix A.

4.4 Analysis

First, I conducted the acoustic analysis of the stimuli used for the discrimination study using Praat (Boerma and Weeknik, 2021). According to the literature (Diaz-Campos, 2008; Henriksen and Willis, 2010; Bradley and Willis, 2012), the two acoustic parameters that were found relevant to distinguish between taps and trills are duration and number of occlusions. The stimuli used will be analyzed according to those parameters (see Table 3, Figure 1 and Figure 2). This analysis is meant to verify that the sounds used in the perception study correspond to the target realization for each sound in Mexican Spanish.

Second, I analyzed the performance of listeners in the discrimination study. Rhotic perception was analyzed per group to determine whether the experimental groups (HSs and sequential bilinguals) are significantly different to each other and to the control group (monolinguals). This was done to establish if contact with English affects the discrimination of Spanish rhotics. Then, I analyzed the directionality of the non-target perceptions, to determine whether taps were perceived as trills or trills as taps more often.

Then, cultural identity and exposure to Spanish were analyzed for the experimental groups, to examine whether those bilinguals who identify themselves with the Hispanic culture and that have a higher exposure to Spanish have less difficulties in the correct identification of rhotic categories. In addition, level of formal education was analyzed to determine the influence of orthography on bilinguals' abilities to distinguish the tap and trill. Finally, bilingual dominance will be explored to see if those bilinguals, who are Spanish dominant, have less errors than the ones that are English dominant.

Statistical analysis consisting of Binomial Logistic Regression were used to determine which independent variable better predicts correct discrimination in bilinguals. The dependent variable was whether sound heard matched the stimuli. The independent variables were cultural identity, exposure to Spanish, level of formal education, and bilingual dominance. Participants' performance on distractors and practice runs were not analyzed.

5. Results

5.1. Acoustic analysis of the stimuli used in the discrimination study

As explained, the two parameters used to distinguish the tap and trill were number of occlusions and duration in ms. The literature states that the average duration of taps is 27 ms., and that of trills is 72-89 ms. (Diaz-Campos, 2008; Henriksen and Willis, 2010; Bradley and Willis, 2012). The analysis presented on Table 3 shows that the average duration of taps is 27.8 ms. and trills 85.5 ms. Taps, on average, have one occlusion, and trills have 0-3 occlusions (Diaz-Campos, 2008; Henriksen and Willis, 2010; Bradley and Willis, 2012). As shown in Table A, taps were produced with 1 occlusion and an average duration of 59 ms. In addition, trills were produced with an average of 3 occlusions for 15/15 of the stimuli and an average duration of 85.5 ms. Below, Table 3 provides the acoustic measurements of taps and trills produced by the reader and used as stimuli:

Table 3. Acoustic measurement (number of occlusions and duration in ms.) of taps and trills used in the stimuli.

	Number of occlusions	Duration (ms.)
taps:	1 (15/15)	27.8 ms.
ahora	1	27 ms.
careta	1	29 ms.
caro	1	28 ms.
cero	1	28 ms.
coral	1	26 ms.
coro	1	27 ms.
mira	1	29 ms.
moro	1	29 ms.
para	1	27 ms.
pera	1	29 ms.
amaras	1	27 ms.
pero	1	28 ms.
quería	1	29 ms.
torero	1	29 ms.
vara	1	26 ms.
trills:	3 (1515)	85.5 ms.
ahorra	3	81 ms.
carreta	3	82 ms.

cerro	3	84 ms.
corral	3	82 ms.
carro	3	89 ms.
corro	3	81 ms.
mirra	3	87 ms.
morro	3	88 ms.
parra	3	85 ms.
perra	3	89 ms.
amarras	3	84 ms.
perro	3	85 ms.
querría	3	89 ms.
torrero	3	89 ms.
barra	3	88 ms.

In terms of the duration, the tap is 27.8 ms. on average, and the trill is 85.5 ms. on average. Both the taps and the trills are a bit longer in duration than the ones reported in previous literature, 20 ms. for taps, and 85 ms. for trills, according to Quilis (1993). This could be due to a more careful pronunciation of rhotics by the reader in this study than the realizations found in the speech of participants in previous research. Nevertheless, the average of trills is more than the average of taps and they have more than two occlusions, which would help listeners perceive the difference between categories.

Praat will be used to present the acoustic contrast between the tap and the trill. Figure 1 and Figure 2 show the waveform and spectrogram of the words for <*pero*> and <*perro*>. As shown in Figure 1 the tap [ɾ] in intervocalic position has one contact and a duration of 28 ms. The rhotic in Figure 2 for <*perro*> 'dog' also has the trill [r] in intervocalic position, but instead of one occlusion as in Figure 1, it has three occlusions. The trill in Figure 2 measures 85 ms.

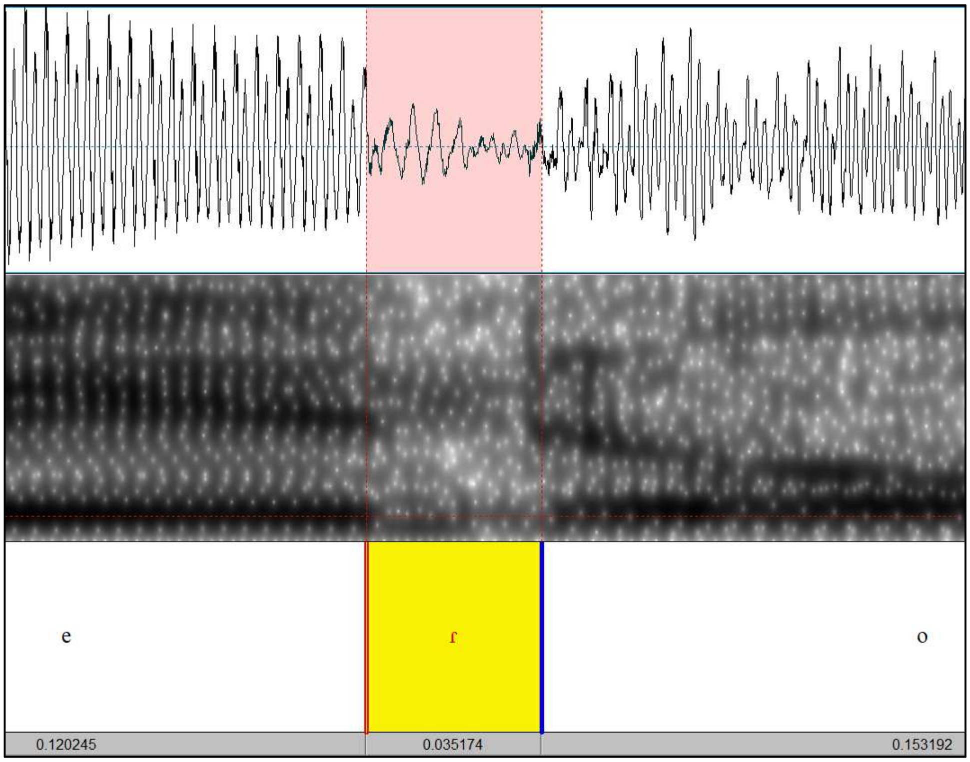

Figure 1. Spectrogram, waveform, and textgrid of the tap in pero /péɾo/ 'but.' The rhotic was realized with one occlusion and measures 35 ms.

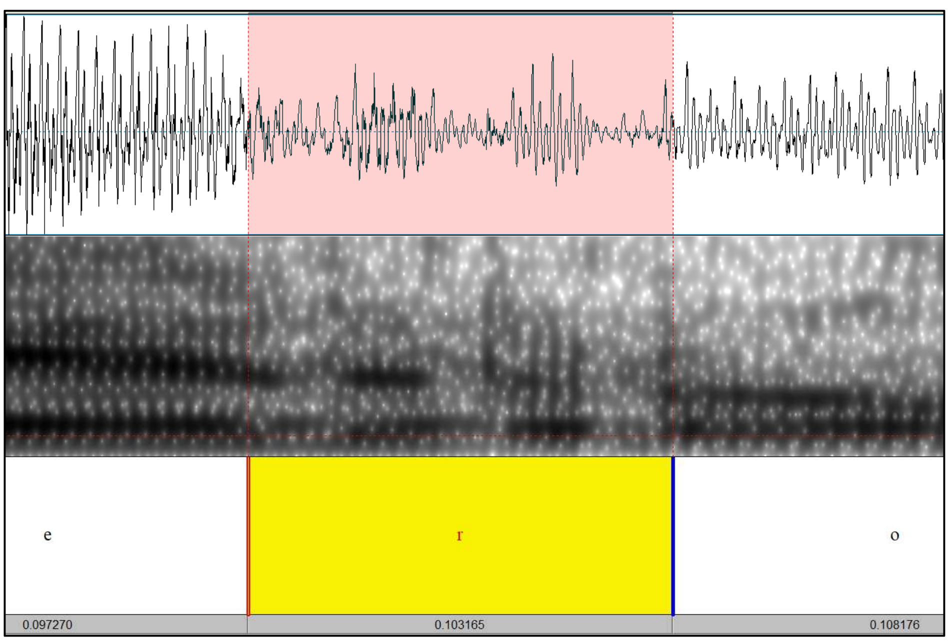

Figure 2. Spectrogram, waveform, and textgrid of the trill in perro /péro/ 'dog.' The rhotic was realized with three occlusions and measures 103 ms.

5.2. Results of the discrimination analysis

A cross tabulation of errors versus group was performed as a first step to analyze the performance of each group (Table 4). Overall, participants had an overall rate of 93.1% target discrimination and 6.9% non-target discrimination. The analysis by group showed that the group with the highest rate of correct discrimination were the sequential English bilinguals (LBE), with 95.1% target discrimination and 4.9% non-target discrimination. The monolinguals (M) had 92% target discrimination and 8% non-target discrimination rates. The simultaneous Spanish-English bilinguals (HSSs) had 94% target discrimination and 6% non-target discrimination. Finally, the sequential Spanish bilinguals (LBS) had 88.8% target discrimination and 11.2% non-target discrimination. The cross tabulation of errors shows that there are significant differences between the groups (p = 0.004). Further analysis is required to see which groups are different.

Table 4. Percentage Errors by Group

Group		Discrimination		Total
		Correct	Incorrect	
Heritage Speakers	N	546	35	581
		94.0 %	6.0 %	100.0 %
Late Bilinguals English	N	448	23	471
		95.1 %	4.9 %	100.0 %
Late Bilinguals Spanish	N	285	36	321
		88.8 %	11.2 %	100.0 %
Monolinguals	N	69	6	75
		92.0 %	8.0 %	100.0 %
Total	N	1348	100	1448
		93.1 %	6.9 %	100.0 %

χ^2 Tests

	Value	df	p
χ^2	13.1	3	0.004
N	1448		

The bar graph below clearly shows the low overall non-target discrimination by all the groups, including how each group compares with one another.

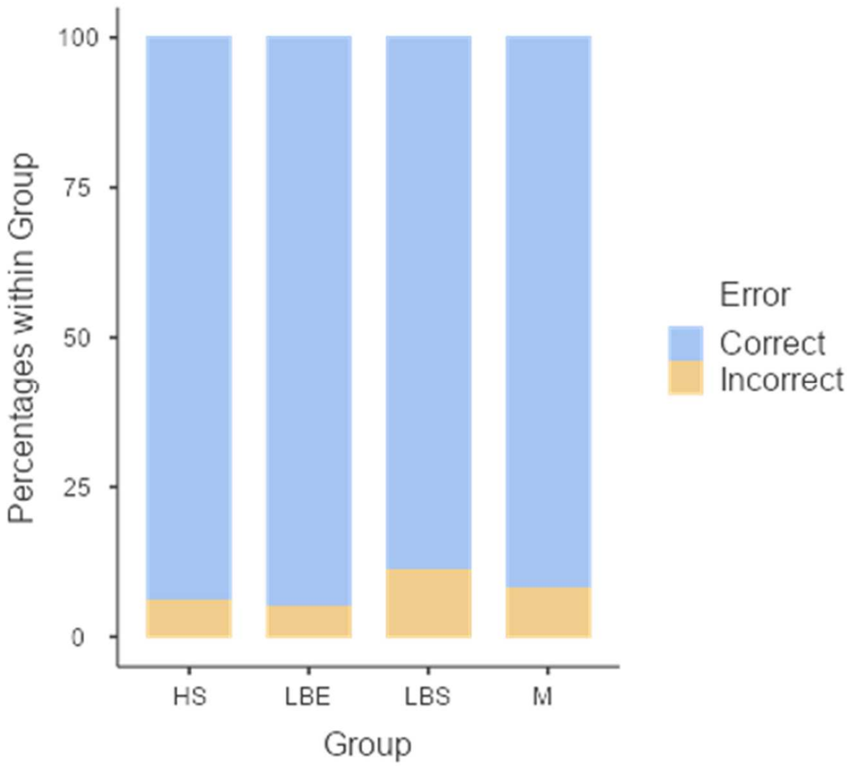

Figure 3. Bar graph showing percentage errors by group

Since there was significant difference in the percent of discrimination errors, the next step was Binary Logistic Regressions to find which groups are different from each other. Binary Logistic Regressions can further show the differences in the error rates between the groups. The first Binary Logistic Regression shows that there was no significant difference between HSS and LBE and M. But there was a significant difference between LBS and HS (p = 0.006). This is shown in Table 5.

Table 5. Binary Logistic Regression to determine if heritage speakers are significantly different from other groups.

Predictor	Estimate	SE	Z	p
Intercept	-2.747	0.174	-15.756	< .001
Group:				
Late Bilingual English – Heritage Speakers	-0.222	0.276	-0.805	0.421
Late Bilingual Spanish – Heritage Speakers	0.678	0.248	2.731	0.006
Monolingual – Heritage Speakers	0.305	0.460	0.663	0.507

The second Binary Logistic Regressions in Table 6 shows that there were no significant differences between LBE, and HS and M, but LBE is significantly different from LBS (p =0.001).

Table 6. Binary Logistic Regression to determine if sequential English bilinguals (LBE) are significantly different from other groups.

Predictor	Estimate	SE	Z	p
Intercept	-2.969	0.214	-13.889	< .001
Group:				
Heritage Speakers – Late Bilingual English	0.222	0.276	0.805	0.421
Late Bilingual Spanish – Late Bilingual English	0.900	0.277	3.245	0.001
Monolingual – Late Bilingual English	0.527	0.476	1.106	0.269

The third Binary Logistic Regression in Table 7 compares LBS with other groups. There are no significant differences between LBS and M, but significant differences were found between LBS, LBE (p = 0.001), and HS (p = 0.006).

Table 7. Binary Logistic Regression to determine if sequential Spanish bilinguals (LBS) are significantly different from other groups.

Predictor	Estimate	SE	Z	p
Intercept	-2.069	0.177	-11.697	< .001
Group:				
Heritage Speakers – Late Bilinguals Spanish	-0.678	0.248	-2.731	0.006
Late Bilinguals English – Late Bilinguals Spanish	-0.900	0.277	-3.245	0.001
Monolingual – Late Bilingual Spanish	-0.373	0.461	-0.810	0.418

The results obtained mostly confirm the hypothesis proposed. First, the LBE had the highest percentage of target discrimination and were significantly different from LBS. The hypothesis predicted a good overall performance of LBE who were brought up in Mexico and completed their primary education there. It was also predicted that LBS, whose L1 was English and learned Spanish after puberty, would have some difficulties in the correct discrimination of the sounds under study. This was confirmed by the binomial regression, showing that LBS were significantly different from LBE and HS. LBS had higher proportion of errors than the other two groups (LBS 88.8% vs. LBE 95.1% vs HS 94%). The results of the monolingual controls were unexpected because they did not perform at ceiling level as predicted. However, this could be due to a low number of participants (N=5) and the fact that they had low formal education. As stated in the section on Speech Perception Theory, through formal education children at school are taught to discriminate contrastive sounds that are difficult to distinguish.

The next step was to analyze the directionality of the errors by all the groups, i.e., are taps discriminated as trills or the other way around. A Binary Logistic Regression showed no significant differences between directionality of errors, i.e., the discrimination is bidirectional. This can be seen in Table 8.

Table 8. Directionality with all participant groups

Sound		Error		Total
		Correct	Incorrect	
/ɾ/	N	682	55	737
		92.5 %	7.5 %	100.0 %
/r/	N	666	45	711
		93.7 %	6.3 %	100.0 %
Total	N	1348	100	1448
		93.1 %	6.9 %	100.0 %

χ^2 Tests

	Value	df	p
χ^2	0.723	1	0.395
N	1448		

The results on Table 8 confirm the hypothesis that predicted that taps are likely to be discriminated as trills and trills with tap. The directionality of errors was further analyzed by group. The only significant differences were for LBS, with taps being perceived as trills more often than trills being perceived as taps. This can be seen on Table 9. The analysis on Table 10 confirms that the differences were only significant for LBS.

Table 9. Percentage of discrimination errors for taps and trills by participant group.

Group	Sound		Discrimination		Total
			Correct	Incorrect	
Heritage Speakers	/r/	N	274	20	294
			93.2 %	6.8 %	100.0 %
	/r/	N	272	15	287
			94.8 %	5.2 %	100.0 %
	Total	N	546	35	581
			94.0 %	6.0 %	100.0 %
Late Bilinguals English	/r/	N	232	9	241
			96.3 %	3.7 %	100.0 %
	/r/	N	216	14	230
			93.9 %	6.1 %	100.0 %
	Total	N	448	23	471
			95.1 %	4.9 %	100.0 %
Late Bilinguals Spanish	/r/	N	138	24	162
			85.2 %	14.8 %	100.0 %
	/r/	N	147	12	159
			92.5 %	7.5 %	100.0 %
	Total	N	285	36	321
			88.8 %	11.2 %	100.0 %
Monolinguals	/r/	N	38	2	40
			95.0 %	5.0 %	100.0 %
	/r/	N	31	4	35
			88.6 %	11.4 %	100.0 %
	Total	N	69	6	75
			92.0 %	8.0 %	100.0 %
Total	/r/	N	682	55	737
			92.5 %	7.5 %	100.0 %
	/r/	N	666	45	711
			93.7 %	6.3 %	100.0 %
	Total	N	1348	100	1448
			93.1 %	6.9 %	100.0 %

Table 10. χ^2 Tests for the directionality of errors per participant group

Group		Value	df	p
Heritage Speakers	χ^2	0.637	1	0.425
	N	581		
Late Bilinguals English	χ^2	1.402	1	0.236
	N	471		
Late Bilinguals Spanish	χ^2	4.257	1	0.039
	N	321		
Monolinguals	χ^2	1.048	1	0.306
	N	75		
Total	χ^2	0.723	1	0.395
	N	1448		

The next analysis explored which factors contributed more at predicting the discrimination of rhotics. A Binomial Logistic Regressions was conducted and the results are shown in Table 11.

Table 11. Binary Logistic Regression to determine which factors significantly predict the number of errors

Predictor	Estimate	SE	Z	p
Intercept	-1.48998	1.37264	-1.0855	0.278
LangExpSpanish	0.00252	0.00775	0.3254	0.745
CultureMexican	-0.01296	0.04189	-0.309	0.757
Education	-0.30784	0.13099	-2.350	0.019
Dominance:				
Spanish – English	-2.21428	1.03868	-2.131	0.033
Group:				
Late Bilinguals English – Heritage Speakers	1.48347	1.05973	1.3999	0.162
Late Bilinguals Spanish – Heritage Speakers	0.47671	0.32177	1.4815	0.138
Monolinguals – Heritage Speakers	1.36980	1.56275	0.8765	0.381
SpanReading	0.06119	0.06858	0.8923	0.372
EngReading	-0.01284	0.12868	-0.099	0.921
LangPrefSpanish	0.00269	0.00816	0.3298	0.742

Note. Estimates represent the log odds of "Error = Incorrect" vs. "Error = Correct"

Education and Dominance were selected as significant factors in predicting the number of errors, with higher education reducing the number of discrimination errors and higher Spanish dominance having the same effect. This corroborates the hypotheses that education (or orthography) plays a role in the discrimination of sounds that are difficult to perceive. Although there were no specific hypotheses to test the influence of dominance, the information was available from the linguistic profile questionnaires, so it was included in the analysis.

Language exposure was not selected as significant, which fails to prove the hypothesis that more exposure to Spanish will decrease the difficulties to correctly discriminate between taps and trills. Another hypothesis was failed to confirm was regarding the influence of Mexican cultural identity on the correct discrimination of rhotics.

Contrary to what was predicted, grouping was not selected as a significant predictor, and this could be due to a correlation between Group and Education where some groups had more formal education than others. A cross-tabulation between group and education (Appendix B) was done to determine if there could be a correlation between these two factor groups (i.e., some groups being more educated than others). The education levels were categorized as 1 = Less than secondary school, 2 = Secondary/Preparatory School, 3 = Professional Training, 4 = Something of University, 5 = University, 6 = Something of Post-Graduate, 7 = Master's, and 8 = Ph.D. or equivalent.

Figure 4 shows this correlation: HS have education ranging from 4 to 9 (some University to PhD); most of LBE and LBS have university degrees and monolinguals' education range from 1 to 5 (less than secondary school to University). This may explain the poor performance of monolinguals compared to HS.

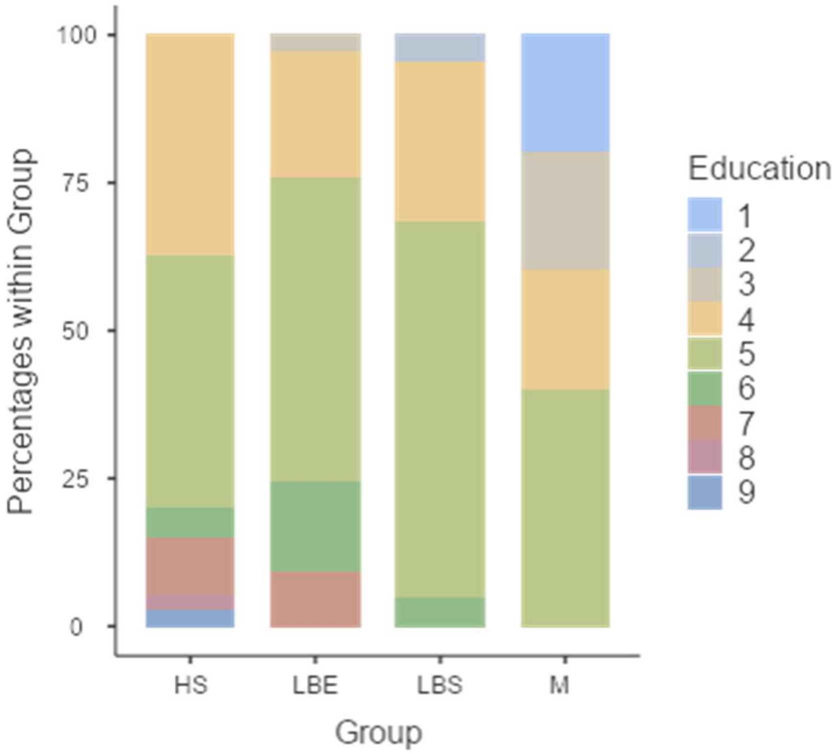

Figure 4. Bar graph showing the education levels of participant groups

6. Discussion and Conclusion

Overall, the results obtained do not confirm the hypothesis that variation in rhotic production can affect the perception of taps and trills and their accurate categorization in early bilinguals. The discrimination performance of HSSs was not significantly different from that of monolingual controls and LBE. The LBE, who were brought up in Mexico and completed their primary education there, had the highest percentage of target discrimination and were significantly different from LBS, whose L1 was English and learned Spanish after puberty. The good performance of HSSs could be related to their higher education, which includes exposure to orthography that may help to clarify ambiguous sounds. In fact, education turned out to be a significant predictor of discrimination errors. Another successful predictor was Spanish dominance, which was not initially predicted to be an important factor in the correct discrimination of taps and trills. Contrary to what was expected, language exposure to Spanish and identity towards the Mexican culture were not selected as significant.

The hypothesis about bidirectionality was confirmed by the analysis. Taps were discriminated as trills, and trills were discriminated as taps in similar proportions. However, when looking at directionality per group, LBS was significantly different from the other groups with higher percentages of errors of tap being perceived as trill. A possible reason for LBS' difference from the other groups is that the tap exists in English and the trill does not. LBS are exposed to the tap since birth and are not exposed to the trill until puberty. Less experience hearing trills could explain why LBS perceive trills instead of taps. As has been reported in the literature, Spanish speakers often realize shorter trills with one or zero occlusions (Kouznetsov and Bertram, 2008; Henriksen and Willis, 2010; Bradley and Willis, 2012; Henriksen 2015). This variability, coupled with less experience hearing this sound, could explain the variation in the discrimination among LBS. On the contrary, the other groups have been exposed to both taps and trills since birth.

The HSSs showed higher percentage of target discrimination than expected, in other words, they have little difficulty in discriminating rhotics. Besides being exposed to both rhotics since birth, HSS in El Paso have more intense contact with Spanish monolinguals from Chihuahua, LBE and other HSS. Since the Hispanic population in El Paso is among one of the highest in the country (80%) and many of these Hispanic speakers are fluent bilinguals, this increases the chances to hearing Spanish and correcting ambiguous sounds. Even though, HSSs do not have early formal ‘orthographic’ education in Spanish that can help them distinguish the rhotics, they were the group with the higher level of formal education. The higher levels of education implies higher levels of exposure to orthography, which could further help with the discrimination of taps and trills.

Surprisingly, two of the factors hypothesized to be important in the discrimination of taps and trills, which were age of acquisition and frequency of Spanish use, turned out to be not significant. The effect of age of acquisition was tested by putting participants into different groups according to the language they learnt first and second. But, information regarding age of acquisition of the L2 was not requested. It could be possible that exact age of acquisition could be more helpful in explaining differences in rhotic discrimination among bilinguals. As for frequency of the use of Spanish, this information generally refers to present use and not past use. Since frequency of language use can change over time (depending on family, job and other life circumstances) the values obtained in the questionnaire reflect only the present use. A speaker’s history of language use over time could have a higher influence on the discrimination of rhotics, compared to looking only at current use. These two factors should be tested in a future analysis, by requesting more information from participants.

References

Amengual, M. (2016). Acoustic correlates of the Spanish tap-trill contrast: Heritage and L2 Spanish speakers. *Heritage Language Journal*, *13*(2), 88-112.

Ball, M. J., Müller, N., & Munro, S. (2001). The acquisition of the rhotic consonants by Welsh-English bilingual children. *International Journal of Bilingualism*, *5*(1), 71-86.

Bassetti, B. 2006. Orthographic input and phonological representations in learners of Chinese as a foreign language. *Written Language & Literacy* 9:1, 95-114.

Bradley, T. G., & Willis, E. W. (2012). Rhotic variation and contrast in Veracruz Mexican Spanish. *Estudios de fonética experimental*, 43-74.

Bradlow, A. R., Pisoni, D. B., Akahane-Yamada, R., & Tohkura, Y. I. (1997). Training Japanese listeners to identify English/r/and/l: IV. Some effects of perceptual learning on speech production. *The Journal of the Acoustical Society of America*, *101*(4), 2299-2310.

Boersma, Paul, and David Weenink. 2021. Praat: Doing Phonetics by Computer [Computer Program]. Version 6.1.439. Available online: http://www.praat.org/ (accessed on 8 February 2021).

Chen, S., & Mok, P. P. K. (2019). Speech production of rhotics in highly proficient bilinguals: Acoustic and articulatory measures. In *Proceedings of ICPhS*.

Cummings Ruiz, L. D., & Montrul, S. (2020). Assessing Rhotic Production by Bilingual Spanish Speakers. *Languages*, *5*(4), 51.

Daine, D., & Zahler, S. (2021). A Variationist Analysis of Second Language Spanish Trill Production. *Studies in Hispanic and Lusophone Linguistics*, *14*(1), 1-37.

Díaz-Campos, M. (2008). Variable production of the trill in spontaneous speech: Sociolinguistic implications.

Flege, J. E. (1995). Second language speech learning: Theory, findings, and problems. *Speech perception and linguistic experience: Issues in cross-language research*, *92*, 233-277.

Flege, J. E. (2007). Language contact in bilingualism: Phonetic system interactions. *Laboratory phonology*, *9*(353-381).

Frost, R. and J. C. Ziegler. 2007. Speech and spelling interaction: the interdependence of visual and auditory word recognition. In Gaskel, G. (ed.) *The Handbook of Psycholinguistics*. Oxford: OUP, 107-117.

Goto, H. (1971). Auditory perception by normal Japanese adults of the sounds" L" and" R.". *Neuropsychologia*.

Henriksen, N. C. (2015). Acoustic analysis of the rhotic contrast in Chicagoland Spanish: An intergenerational study. *Linguistic Approaches to Bilingualism*, *5*(3), 285-321.

Henriksen, N. C., & Willis, E. W. (2010). Acoustic characterization of phonemic trill production in Jerezano Andalusian Spanish. In *Selected proceedings of the 4th Conference on Laboratory Approaches to Spanish Phonology* (pp. 115-127). Somerville, MA: Cascadilla Proceedings Project.

Hualde, J. I. (2005). *Los Sonidos del Español*. Cambridge University Press.

Jahr, E. H., & Breivik, L. E. (Eds.). (1989). *Language change: Contributions to the study of its causes*. Mouton de Gruyter.

Kissling, E. M. (2018). An exploratory study of heritage Spanish rhotics: Addressing methodological challenges of heritage language phonetics research. *Heritage Language Journal*, *15*(1), 25-70.

Kouznetsov, V., & Pamies, A. (2008). Trill with one closure, still a trill or a tap? Data from Russian and Spanish. *Proceedings of the XX Session of the Russian Acoustical Society*, 672-675

Mazzaro, N., & González de Anda, R. (2020). Men Finally Got It! Rhotic Assibilation in Mexican Spanish in Chihuahua. *Languages*, *5*(4), 38.

McMahon, A. (1993). Charles Jones (ed.), Historical linguistics: problems and perspectives. London: Longman, 1993. Pp. xi+ 405. *Journal of Linguistics*, *29*(2), 522-525.

Morais, J., Cary, L., Alegria, J. and P. Bertelson. 1979. Does awareness of speech as a sequence of phone arise spontaneously? Cognition, 7(4):323-31.

Quilis, A. (1997). *Principios de fonología y fonética españolas* (Vol. 43). Arco libros.

Quilis, A., Cantarero, M., & Esgueva, M. (1993). El grupo fónico y el grupo de entonación en el español hablado. *Revista de Filología española*, *73*(1/2), 55-64.

U.S. Census Bureau quickfacts: El Paso County, Texas. (2020). Retrieved March 7, 2022, from https://www.census.gov/quickfacts/fact/table/elpasocountytexas/POP010220

Appendices:

Appendix A:

The following is the list of the words used in the practice trial:

Practice:

tanto y canto
casa y cana
callo y gallo
foca y boca
pez y vez

The following is the list of distractors used in the experimental session:

Distractors:

tiene y diente
pala y bala
coma y goma
cuando y cuanto
pelo y velo
pasta y vasta
manta y manda
vela y pela
cana y caña
pan y van
dan y tan
cose y goce
sala y pala
vena y pena
beso y peso

The following is the list of stimuli used to test discrimination of the tap and trill:

Stimuli:

ahora y ahorra.
careta y carreta
caro y carro
cero y cerro
coral y corral
coro y corro
mira y mirra
moro y morro
para y parra
pera y perra
amarras y amaras
pero y perro
quería y querría
torero y torrero
vara y barra

Appendix B

Table B1. Correlation table between Education and Group

Group		Education									Total
		1	2	3	4	5	6	7	8	9	
Heritage Speakers	N	0	0	0	225	255	30	60	15	15	600
		0.0	0.0	0.0	37.5	42.5	5.0	10.0	2.5	2.5	100.0
Late Bilinguals English	N	0	0	15	105	255	75	45	0	0	495
		0.0	0.0	3.0	21.2	51.5	15.2	9.1	0.0	0.0	100.0
Late Bilinguals Spanish	N	0	15	0	90	210	15	0	0	0	330
		0.0	4.5	0.0	27.3	63.6	4.5	0.0	0.0	0.0	100.0
Monolinguals	N	15	0	15	15	30	0	0	0	0	75
		20.0	0.0	20.0	20.0	40.0	0.0	0.0	0.0	0.0	100.0
Total	N	15	15	30	435	750	120	105	15	15	1500
		1.0	1.0	2.0	29.0	50.0	8.0	7.0	1.0	1.0	100.0

Vita

Lesley Rodriguez is a Hispanic Linguistics graduate student. She obtained a Bachelor of Arts in Linguistics at UTEP. Lesley has been a teacher's assistant at the University of Texas at El Paso (UTEP). Lesley's current research is in the field of phonology and second language acquisition, particularly on the perception and production of sounds, specifically Heritage Spanish Speakers in the El Paso-Ciudad Juarez border region. She looks forward to seeing language change.

Contact Information: lesleyvrodriguez@gmail.com