Duration as a secondary cue in Spanish rhotics: Sliding to understand perceptual categorization among heritage and L2 learners of Spanish

Rachel Sangster Garza
Sul Ross State University
ORCID ID 0000-0002-8933-8096
Rachel.SangsterGarza@sulross.edu

Erik W. Willis
Indiana University
ORCID ID 0000-0002-2421-6899
EWWillis@iu.edu

Abstract

This study investigates language learners' attention to segmental duration when perceiving the Spanish tap-trill contrast. Previous research has demonstrated that native speakers (NSs), heritage language learners (HLLs), and second language (L2) learners maintain distinct segmental durations when maintaining the tap-trill contrast (e.g., Amengual, 2016; Henriksen, 2015; McCandless, 2020; Willis & Bradley, 2008). In this study, 31 HLLs and 104 L2 learners from a Midwest university listened to five minimal pairs in Spanish with modified closure durations (40 experimental stimuli) ranging from /ɾ/-like (tap-like) (22-40ms) to /r/-like (trill-like) (52ms-85ms). Visual analog sliders were used to connect learners' learner characteristics to continuous perceptions. Results of a regression analysis suggest that both L2 learners and HLLs attend to closure duration when perceiving contrast (β = .009, p < .001), though HLLs demonstrate less categorical overlap than L2 learners. In both groups, linguistic

confidence is the strongest predictor of selection certainty for both taps (β = 0.824, p < .001) and trills (β = 1.125, p < .001). The results support a clearer understanding of heritage and developing L2 phonological mapping of the Spanish /ɾ/-/r/ contrast, as well as the utility and limitations of employing continuous measurements in perception research.

Keywords: continuous variables, t, taps, trills, Spanish-English bilinguals, perception, phonology.

1. Introduction

Though the Spanish /ɾ/-/r/ (tap-trill) contrast is prototypically explained as a difference in apical occlusions (1 versus 2 or more) (Blecua 2001; Hualde, 2013), previous empirical evidence demonstrates that native speakers (NSs), heritage language learners (HLLs) and second language (L2) learners often produce /r/ with a single occlusion, and employ segmental duration to maintain the phonemic contrast (e.g., Amengual, 2016; Bradley & Willis, 2012; Cummings Ruiz & Montrul, 2020; Henriksen, 2015; McCandless, 2020; Willis & Bradley, 2008). Perception research of /ɾon ther/ contrast similarly suggests that NSs and advanced L2 learners differentiate between the two phones based primarily on tone (e.g., Melero-García & Cisneros, 2020; Scarpace, 2014). Meanwhile, the perception of the tap/trill distinction by HLL speakers has only been explored in one paper, and finds that heritage speakers make use of closure duration[1] as a perceptual cue (Garza et al., 2024). However, HLLs demonstrate greater perceptual overlap between phonemic categories when compared with NSs, and according to their type of bilingualism – being a simultaneous, early, or late bilingual was also found to predict the degree of this overlap between categories significantly. The present study further explores Spanish-English bilingual learners' perceptions by including Spanish proficiency and linguistic confidence (operationalized as in Gardner et al., 1997, as one's belief in their ability to learn and communicate in Spanish) as two learner characteristics that may confound perceptions of phonemic contrasts. In doing so, it aims to fill gaps in existing research by examining how these factors influence the perception of phonemic contrasts in Spanish, offering insight into the broader complexities of language acquisition and phonological processing in both L2 and heritage learners of Spanish.
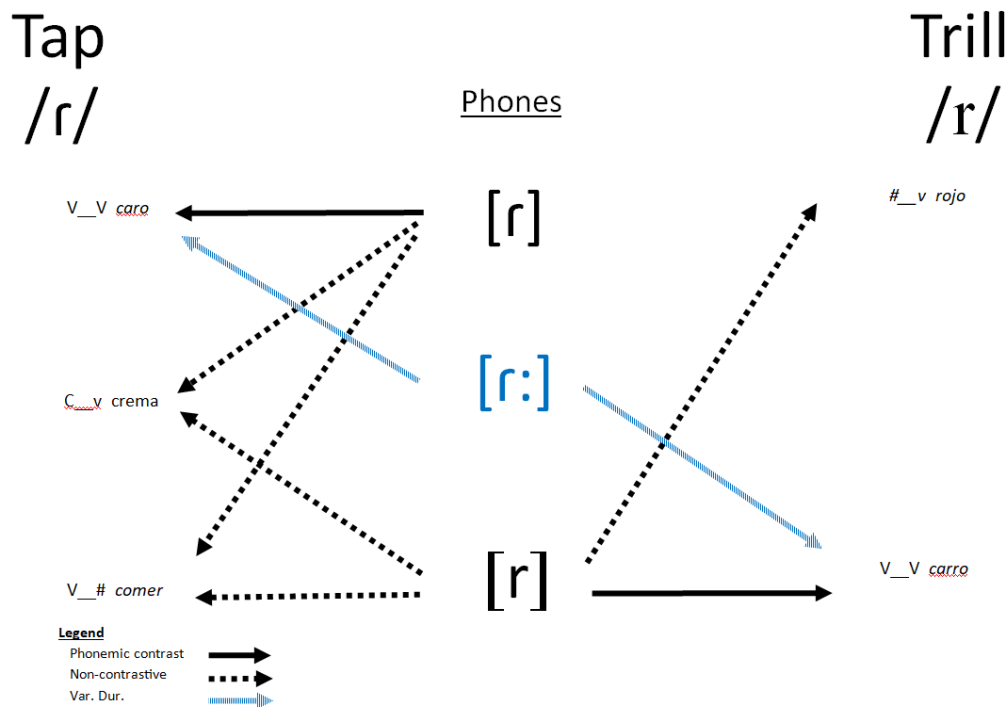
2. Literature Review

2.1. Spanish rhotics

Most reference manuals and handbooks describe the Spanish rhotic system as having two phonemes, the tap and the trill, which are contrastive in intervocalic position, e.g., <vara> [ˈba.ɾa] 'bar/rod small-stick' versus <barra> [ˈba.ra] 'bar/rod large-beam' (Navarro Tomás, 1918; Quilis, 1993; Real Academia Española, 2011). While non-

---

[1]   We employ the terms segmental duration and closure duration as synonyms whose use varies based on context.
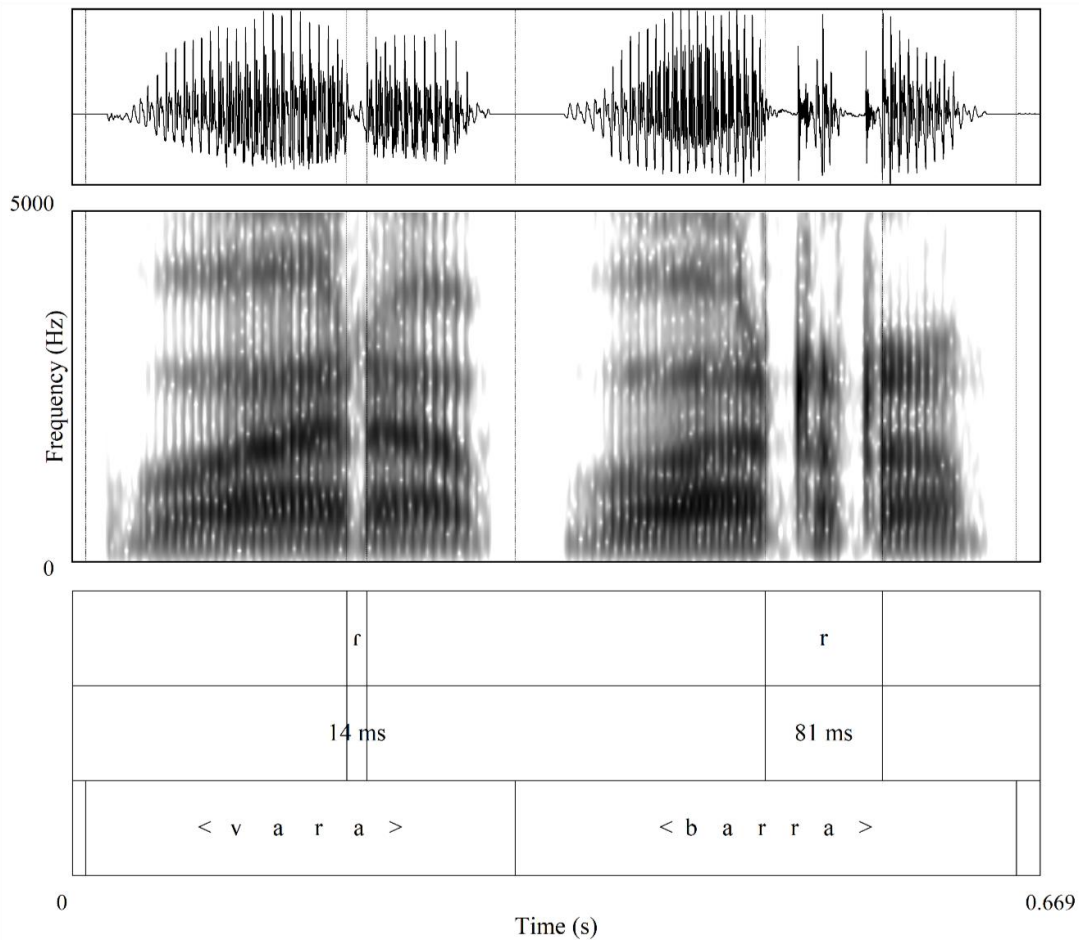
contrastive with taps, word-initial rhotics are typically produced as trills, and in the word-final position, the rhotic distinction is neutralized, with no potential for contrast. Likewise, rhotics that form part of a syllable onset consonant cluster, as in *broma* 'joke', or part of a syllable coda, as in *puerta* 'door', are non-contrastive. Both realizations occur with no change in meaning; however, the tap occurs more frequently with variations by style and dialect. Figure 1 is a schematic of the two phonemes, where they are contrastive (solid lines), where they are non-contrastive (dotted lines), and with an increased single closure duration where they can be interpreted as both (blue striated lines).

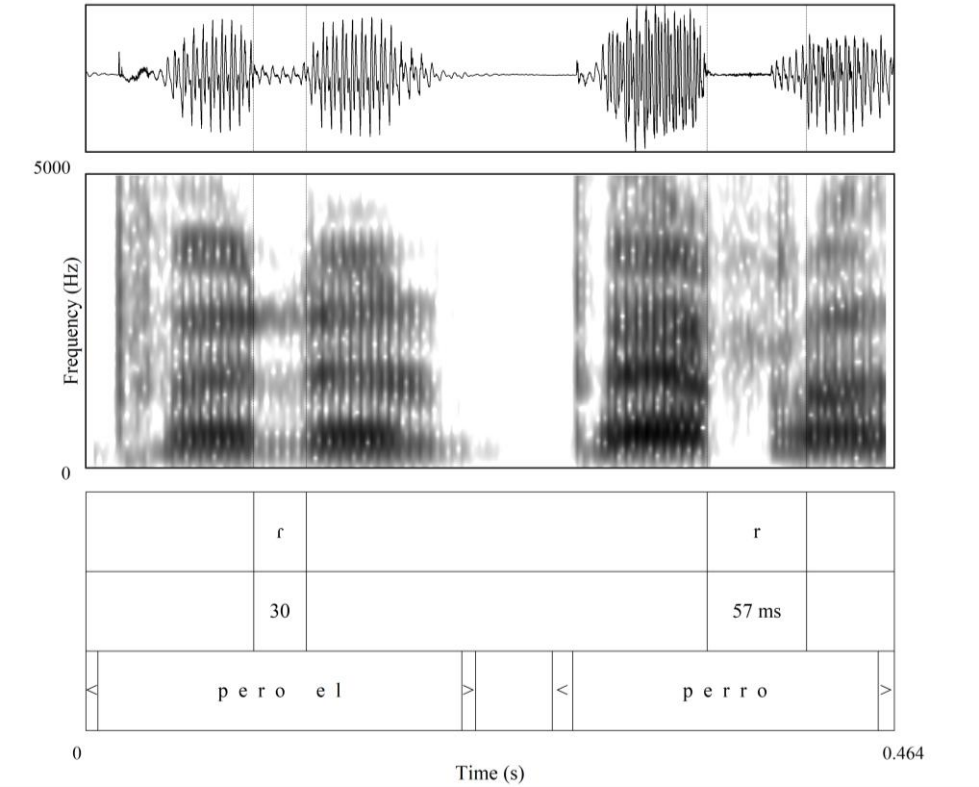**Figure 1.** Contrast mapping of Spanish taps and trills.

Note: This graphic is not intended to identify all dialectal phenomena.

The principal stated difference between the two phonemes is the number of occlusions; the tap has a single occlusion, and the trill has two or more (Blecua, 2001; Hualde, 2013). Most reports document the phonemic tap with a segmental duration under 30 ms and the phonemic trill with a duration above 50 ms (Amengual, 2016; Blecua, 2001; Bradley & Willis, 2012; Henriksen, 2015; J. Y. Kim & Repiso-Puigdelliura, 2020; Quilis, 1993; Willis & Bradley, 2008). Figure 2 below demonstrates the two phonemes with the segmental duration included in the second tier. However, recent acoustic work has documented multiple varieties of Spanish with a single occlusion that may be either long or have additional secondary cues such as frication or breathiness (Willis, 2007; Willis & Bradley, 2008).
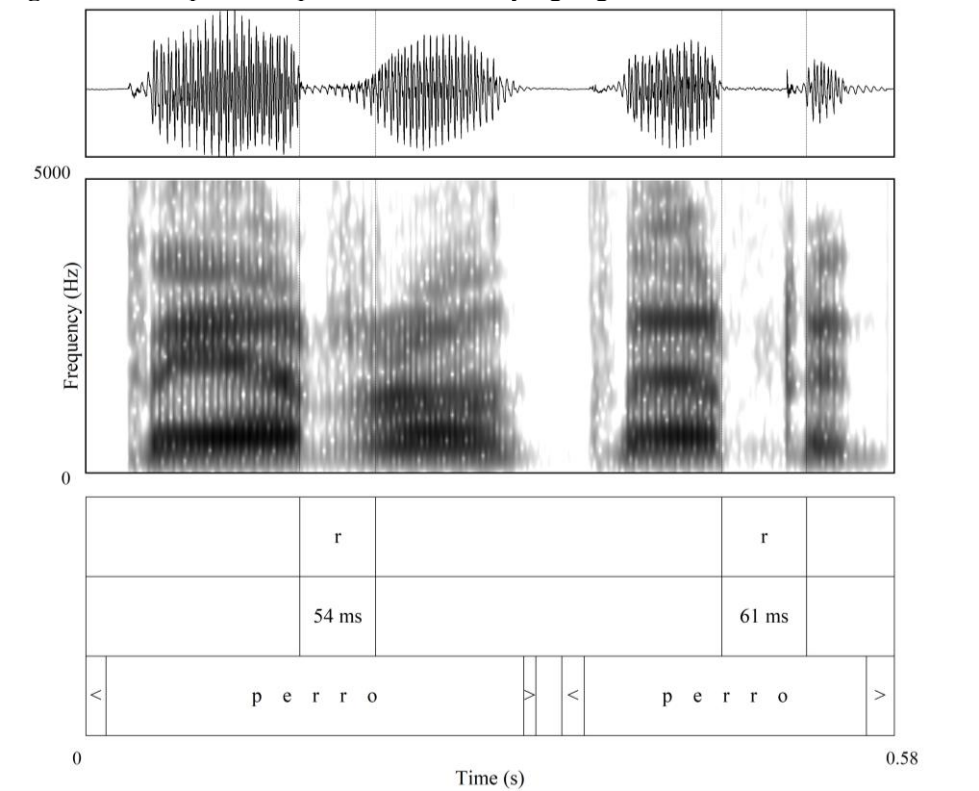
Figure 2. Durational difference between the Spanish tap and trill.

Recently, perception research has indicated that a single occlusion with a duration above the 50 ms range may be interpreted as a phonemic trill by native speakers of Spanish, and that L2 learners may tend to perceive a tap despite the segmental duration (Melero-García & Cisneros, 2020). The association or perceptual identification of a trill based on closure duration without multiple occlusions indicates that there are multiple cues to a phonemic trill, including both occlusions and closure duration. Bradley and Willis (2012) used a narrative retell task of "Frog, Where Are You?" (Mercer Maye 1968), and documented phonemic trills with a single occlusion; however, the average durational values for the two phonemes revealed two distinct duration patterns. Figure 3, taken from the same speaker corpus (Bradley & Willis, 2012), contains a phonemic tap from the word *pero* 'but' and a trill in the word *perro* 'dog' in the phrase *pero el perro* 'but the dog'. Both rhotics are produced with a single closure, but their durations differ considerably. Bradley and Willis (2012) suggest that there is variability in the allophonic productions of the Veracruz phonemic trill. Figure 4, taken from the same audio corpus, illustrates two of these variations, both of which feature a single closure. One production is voiced with light traces of frication, while the second /r/ is voiceless with what appears to be a burst-like VOT.

Figure 3. Example of /r/ produced with varying segmental duration and traces of frication.

Figure 4. Example of /r/ produced with varying segmental duration and burst-like VOT.
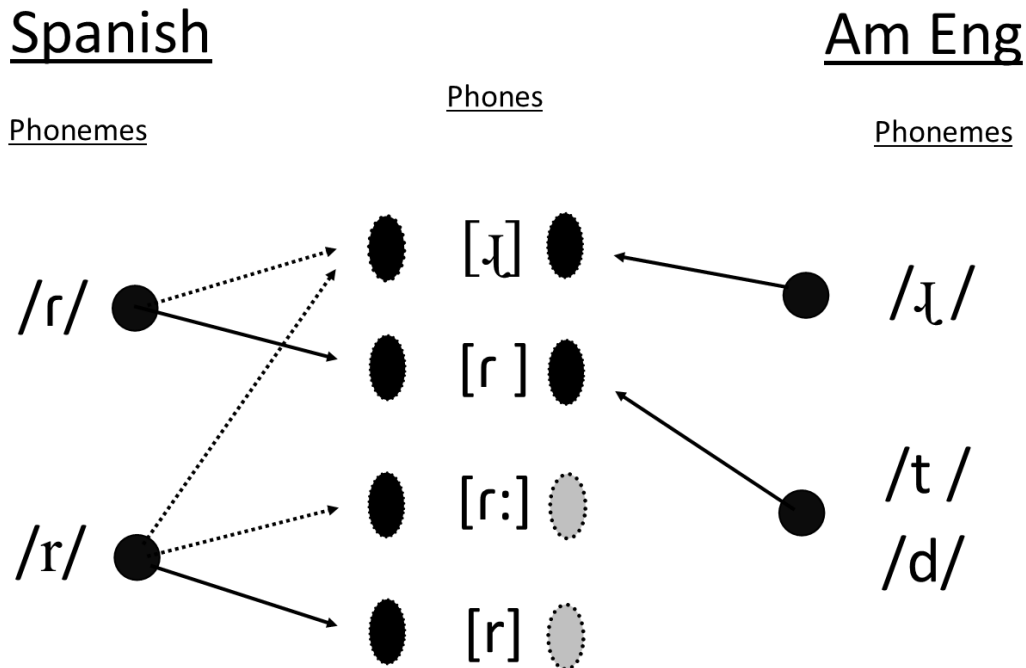
2.2. English-Spanish bilingualism and Spanish rhotics

The realization of Spanish and English rhotics varies considerably, where Spanish rhotics are prescriptively categorized into two phonemes characterized as alveolar taps [ɾ] with single apical occlusions and alveolar trills [r] with two or more apical occlusions (RAE, 2011; Hualde, 2014). However, productions of coda /ɾ/ described as retroflex have been reported in Puerto Rico (Delgado-Díaz et al. 2021), and the Costa Rican trills are also reported to manifest several variants, one of which is described as an approximant (Quesada Pacheco 1996). Meanwhile, American English rhotics are typically characterized as one singular phoneme: the bunched or retroflexed /ɻ/ (Ladefoged & Johnson, 2015; Zhou et al., 2008). Nevertheless, the alveolar allophone [ɾ] exists in American English, though typically as a realization of intervocalic post-stressed /t/ and /d/ as in the word <duty> [ˈdu.ɾi] or when situated between a vowel and syllabic liquid as in <butter> [ˈbɔ.ɾɚ] (Cummings Ruiz & Montrul, 2021; de Jong, 1998; Face, 2006; Hualde, 2014; Ladefoged & Johnson, 2015; Rose, 2010).

Alongside these phonemic differences between English and Spanish rhotics, another important factor affecting the contrast between Spanish taps and trills is variation in Spanish trill production. For example, Amengual (2016) finds that Spanish trills produced by heritage language learners in Northern California are often realized with three categories of taps: true taps, perceptual taps, and approximant taps. This difference in rhotic realization may pose a challenge for English speakers learning Spanish as a second language, due to the need to acquire new pronunciations and reconfigure their existing rhotic system. Figure 5 demonstrates an abstract rendition of what this mapping might look like within the bilingual phonological system of English-Spanish bilinguals. Darkly filled circles indicate phones that are productive in the languages, whereas the greyed circles indicate phones that are not regularly reported as productive in American English, as shown in Figure 5. L2 learners must acquire a novel phonological system with distinct acoustic cues. In the case of Spanish rhotics, many learners may focus on producing a trill with multiple lingual contacts as the primary target for a "correct Spanish <r>", often without considering or worrying about the production of the tap.[2] The focus on the trill leaves open the question of whether the duration of the segment serves as a concomitant or secondary acoustic cue in categorical perception (Melero-García & Cisneros, 2020; McCandless, 2020). This figure visually demonstrates the complexity of acquiring phonological mapping of similar or related phonemes in a new language. Phonological interference in the emerging L2 or less dominant language system may affect both the production and perception of phonemes in the system. This interference may result in difficulty mapping both primary and secondary cues of the new system. The current study focuses on the acquisition of the secondary cue of segmental duration in Spanish rhotics and aims to understand the extent to which heritage and second language learners may utilize this cue to distinguish between Spanish taps and trills.

---

[2]   This claim is anecdotal and reflects comments by students in Spanish phonetics courses.

Figure 5. Phoneme to phone mapping in English-Spanish bilinguals.

Note: Figure modified from Garza et al. (2024). Dotted lines indicate dialectal and learner variants of Spanish that learners may encounter or produce in the process of phonemic acquisition.

2.2.1. Second language acquisition of Spanish rhotics

Previous empirical research examining language learners' production and perception of Spanish rhotics suggests a significant difference in the process and timeframe required to acquire the Spanish /ɾ/ and /r/. The results indicate that learners tend to acquire the perception of contrast earlier than its production, likely due to physical articulatory constraints (e.g., Amengual, 2016; Daidone & Darcy, 2014; Rose, 2010). Some previous research has also considered not only phonemic interference between the Spanish and English rhotic systems but also mapping interference between Spanish taps and English /d/ and /t/ realized as flaps (Daidone & Darcy, 2014; Rose, 2010; Waltmunson, 2005). The results of these studies suggest that beginner or novice learners of Spanish experience difficulty distinguishing between /d/ and/t/, as well as/ɾ/, at the beginning of their phonological acquisition (see Rose, 2010). However, learners were more likely than not to perceive the correct phoneme, regardless of proficiency level.

Regarding production research, multiple acoustic correlates of the tap-trill contrast have been considered, including apical occlusions, overall segmental duration, and other correlates specific to dialectal varieties of Spanish, such as frication (e.g., Amengual, 2016; Reeder, 1998; Rose, 2010). Some literature suggests that L2 learners demonstrate evidence of contrasting via segmental duration (e.g., Amengual, 2016), while other literature suggests that they do not employ this secondary acoustic cue (Rose, 2010). In general, previous studies have found limited evidence of a consistent ability to produce contrast in apical occlusions among beginner or intermediate learners, which is attributed to articulatory constraints (Rose, 2010). In fact, Reeder (1998) suggests that only very advanced speakers with experience living abroad produce predominantly canonical trills (more than 83% of the time). In

contrast, advanced learners without such experience produce these canonical trills only about 37% of the time.

Previous perception research suggests that both beginner and intermediate learners demonstrate a low but similar ability to perceive a contrast between /ɾ/ and /r/ (e.g., Carballo, 2010; Daidone & Darcy, 2014; Melero-García & Cisneros, 2020; Rose, 2010; Wells, 2013). Daidone & Darcy (2014) present an explanation for the low frequency of /ɾ/ and /r/ contrast in a language learning environment, which could explain why learners do not demonstrate a completely accurate perception of the two sounds, even at advanced proficiency levels. Regarding the tasks previously employed in L2 Spanish rhotic perception, lexical decision tasks are the most prevalent (e.g., Daidone & Darcy, 2014; Rose, 2010, etc.). In these lexical decision tasks, learners hear both real word pairings and nonce word pairings, and in an ABX task, they must determine whether each pairing is the same or different. These studies highlight the phonemes themselves, focusing on prescriptive or prototypical productions of these sounds. In the present study, we incorporate occurrences of descriptive but non-prototypical productions by including single-occlusion trills to understand how learners might handle these productions in real conversation.

2.2.2. The role of listener characteristics

Listener characteristics and learner characteristics have been employed to explain why linguistic structures are readily understood by some learners but not others (Hulstijn, 2019). In a previous iteration of this study, which focused solely on heritage learners of Spanish, we examined how a region's Spanish dominance might predict attention to the closure duration of the trill phoneme (see Garza et al., 2024). Nevertheless, when considering the role of early, simultaneous, and late sequential bilinguals, early sequential bilinguals were more likely to begin perceiving a trill starting at lower closure durations, beginning at 55ms (Garza et al., 2024). However, due to considerable variation, it was difficult to explain the variation in these results, and it was unclear what independent factors may have been related to tap-trill selection. To further understand the distribution of results, within the present study, we have included measurements of proficiency and linguistic confidence (operationalized as in Gardner et al., 1997, as one's belief in their ability to learn and communicate in Spanish) to understand better the effect of learner differences on attention to segmental duration when maintaining a phonological contrast.

Previous research finds that perception can be predicted by Spanish proficiency and linguistic confidence (e.g., Ellis, 1997; Flege & MacKay, 2004; Perea & Swinney, 2021). These studies suggest that higher proficiency in Spanish and greater linguistic confidence lead to improved phonological perception, particularly in vowels, intonational patterns, and overall comprehension. In the present study, it is expected that higher proficiency and linguistic confidence will relate not only to a more robust categorization of tap and trill phonemes as an effect of closure duration, but also that these characteristics will influence certainty, the continuous dependent variable of the study. This prediction is supported by the results of Kapnoula et al. (2017), who examined the perception of the /b/ and /p/ phonemic categories and considered the role that learners' language-processing characteristics play in the gradient, or continuous, perception of these phonemes. Their results suggest that the perceptual categorization of phonemes is influenced by general cognitive processes beyond phonetic processing alone. In summary, the results of these previous empirical studies suggest that learner

characteristics can offer insights into the development of phonemic perception, ultimately contributing to a deeper understanding of how learners navigate phonological contrasts in a second language.

2.2.3. Heritage language production and perception of Spanish rhotics

In general, phonological research on heritage Spanish remains scarce (see Rao & Amengual, 2021), with few studies considering heritage rhotic production and even fewer examining its perception. Previous production research finds a range of results with heritage participants employing non-canonical /r/ with one occlusion and elongated segmental duration or the production of /r/ with two or more occlusions while performing tasks such as storytelling, picture naming, or conversational tasks (Amengual, 2016; Cummings, Ruiz & Montrul, 2021; Henriksen, 2015; Kissling, 2018). These previous studies also found that several independent variables related to identity and language background predicted /r/ production with multiple occlusions: higher Spanish dominance, as well as stronger identification with cultural or ethnolinguistic identity (Amengual, 2016; Cummings, Ruiz, & Montrul, 2021; Henriksen, 2015; Kissling, 2018).

Regarding perception, to our knowledge, Garza et al. (2024) is the only study that considers perceptual contrasts between the Spanish taps and trills in heritage listeners. Each listener heard a word with an intervocalic tap with modified duration (ranging from 22 to 85 milliseconds) and needed to select whether they heard a word containing a tap (e.g., *pera* 'pear') or a trill (e.g., *perra* 'dog'). The results of the study suggest that HLLs do attend to segmental duration length when perceiving a contrast between Spanish taps and trills, but not to the same extent as native Spanish speakers who performed a similar task in Melero-García & Cisneros (2020). Furthermore, Garza et al. (2024) suggest that heritage learners created a category boundary at around 61ms, where they are more likely to perceive words containing taps with 61ms or shorter as a phonemic tap and taps with 61ms or longer as a phonemic trill. In contrast, Melero-García & Cisneros (2020) find that native Spanish speakers formed this perceptual boundary around 50ms. The results of Garza et al. (2024) prompted the researchers to conduct the present study with a sample of L2 participants to understand the process of perceptual rhotic acquisition among another group of bilinguals and to gain new insights regarding this process.

2.3. Methods in phonological perception

Most research on phonological perception in L2 acquisition is based on the AX or ABX paradigm (Flege, 2021; Kapnoula et al., 2017). In this approach, listeners will hear a continuum pair that varies in a particular acoustic measure, and typically, there will be a point of variation in the measure at which the two stimuli are recognized as different. The slope of the distinction between the stimuli, or the duration of the reaction times to identify the different stimuli, is used to characterize the covert confidence of the listener. Flege (2021) suggests that previous work in L2 perception has over-relied on forced-choice identification tests, which, in previous L2 and NS studies, have found that these test results often mimic those of ABX tasks, in which listeners employ the 'X-not-X' strategy (pp. 145-147). Flege's (2021) findings suggest that categorical perception tasks don't strongly support our understanding of listeners'

attention to acoustic correlates. In contrast, continuous perceptual tasks may provide further insight into listeners' perceptual strategies.

Similar to the argument in Flege (2021), Daidone et al. (2023) introduce a novel Free Classification approach modeled after Clopper and Pisoni (2007). A novel approach to phonological perception is the use of a Visual Analog Slider (see Arróniz 2025; Kapnoula et al., 2017; Simonet 2010). In this methodology, there is a continuum between two phonemes, and the listener will slide a cursor toward one of the two extremes, which reflects the binary classification of the two phones based on the certainty of this choice in the listener's mind. The continuous slider approach allows both a binary distinction, which can be analyzed in a strictly categorical manner, and a measure of certainty with respect to the choice of phonemic category. This certainty rating provides an overt measure of the variable certainty associated with a particular choice that is not offered in the 2-alternative forced choice method (2AFC). These certainty measures facilitate a deeper understanding of learners' attention to the acoustic cues of study, as reflected by how changes in these cues interact with learners' certainty. For instance, learners may show higher certainty that they are hearing a tap when the segmental duration is shorter, rather than at longer lengths.

Within the present study, we aim to expand on the findings for L2 learners, utilizing a slightly different methodology than that employed in Melero-García & Cisneros (2020), by controlling the duration of bins. To further understand which factors affect learners' selection choices, we employed a continuous visual analog slider with 5-step durational bins, each with a single closure. This slider methodology has been employed in previous research on VOT and F0 categorical perception in English stops (Kapnoula et al., 2017), as well as in other studies (Kong & Edwards, 2011; Simonet, 2010).

2.4. Research questions

Although previous empirical research has examined the role of secondary acoustic correlates in learner perception (e.g., Garza et al., 2024; Melero-García & Cisneros, 2020), the impact of individual listener characteristics has not been previously considered in this context. The present study aims to fill a gap in the previous literature by considering evidence from HLLs and L2 learners, which incorporates the effects of learner characteristics related to linguistic confidence and Spanish grammar proficiency. The present study also offers insight into a methodology that employs a continuous scale, rather than a strictly categorical or Likert-based method, of understanding gradient perceptual certainty when perceiving stimuli that need to be attributed to different phonemic categories by incorporating the use of Visual Analog Sliders (VAS; see Figure 7; Apfelbaum et al., 2022; Kapnoula et al., 2017; Kong and Edwards, 2016).

Specifically, this research explores the following questions:

(1) Do L2 learners of differing proficiency backgrounds attend to the secondary cue of closure duration as a perceptual cue of tap-trill categorization?

(2) Does the role of duration impact categorization in L2 in the same way it impacts categorization for HLLs?

(3) How do linguistic confidence and grammatical proficiency predict selection certainty using VAS?

Regarding research question (1), it was expected that L2 learners would attend in part to the secondary cue of closure duration. But, responding to research question (2), it was predicted that this attention would not be as robust as the HLLs based on previous results of L2 learners in Melero-García and Cisneros (2020), in which the L2 learners did not attend as robustly to duration as native speakers. For research question (3), it was expected that proficiency and confidence would significantly predict accuracy in categorization, with the reasoning that a learner with higher proficiency and higher confidence in Spanish would likely have more experience perceiving taps and trills and hearing a variety of Spanish productions due to higher exposure to the language.

3. Methods

3.1. Participants

To explore these questions, 31 HLLs (22 female, 9 male) and 104 L2 learners (69 female, 32 male, 3 non-binary/third gender) from a university in the Midwest completed a language background questionnaire and two perception tasks. All learners were over the age of 18, had completed their university's basic language program requirements, and were enrolled in either intermediate or advanced Spanish courses. It was determined that a learner was a heritage or L2 learner by examining responses regarding the age of Spanish and English acquisition, the amount of Spanish spoken at home and outside, as well as the place of birth and languages spoken by parents and grandparents.[3] For example, the typical heritage speaker in this study has parents and grandparents who were born in a Spanish-speaking country; they reported that both parents and grandparents are fluent speakers of Spanish, and they themselves were exposed to Spanish from birth, beginning to learn English around the age of 3-5 years. To better analyze the data, the 104 L2 learners were further divided into groups based on their scores on a grammatical proficiency test (see Linford and Geeslin, 2022; Kanwit and Geeslin, 2014). The grammatical proficiency task consists of a biographical story in which a fictional female student talks about her experience at a dinner party. Throughout the story, there are blank spaces where students must choose one of three possible options to fill in the blank. Choices include selecting between articles of different grammatical genders, choosing between the preterit and imperfect tenses, and selecting between the indicative and subjunctive moods, as well as indirect
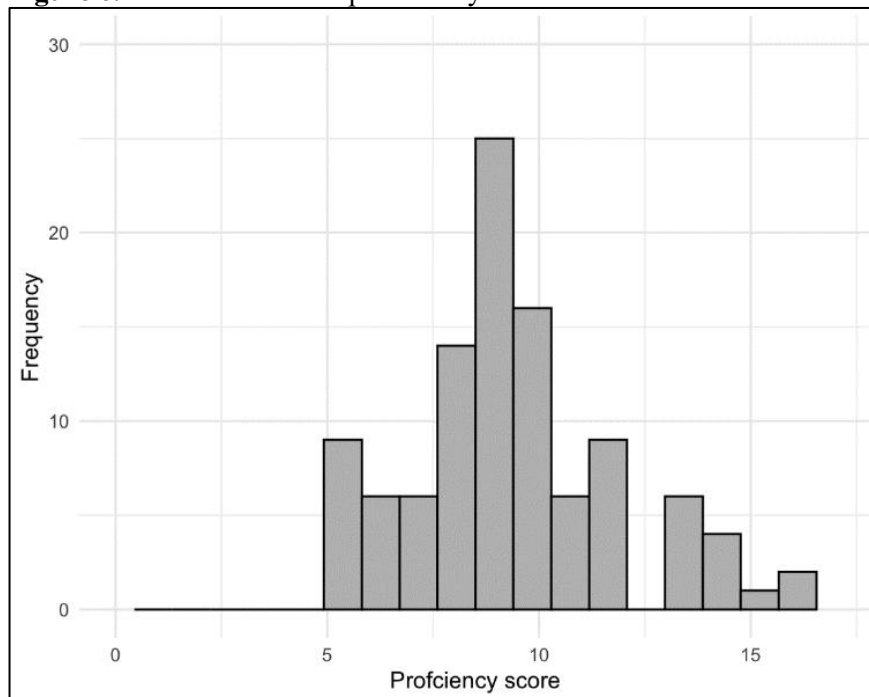
---

[3]   These variables were collectively used to differentiate between heritage speakers and L2 learners. A heritage speaker was identified if they began speaking Spanish from birth, whether or not this was simultaneous with English. Furthermore, a participant was categorized as a heritage speaker if they had family members who spoke Spanish as well (parents and/or grandparents). Other variables, such as place of birth or the amount of Spanish spoken in/outside of the home, gave further support to these categorizations where needed. There were six participants who identified as speaking another language. These additional languages included Russian, Macedonian, Chinese and Portuguese. We did not gather additional data regarding the use of these other languages beyond English and Spanish.

and direct pronouns, among others. Participants also completed a linguistic confidence index based on the English use confidence matrix employed by Gardner et al. (1997). This index asked participants to reflect on their confidence in their ability to communicate in Spanish and learn new vocabulary.

After running the histogram presented in Figure 6, it was determined that there was a somewhat bimodal distribution of scores with one modality of lower proficiency learners (76 learners) who had scores ranging from 5 to 10 points (out of 19 points) and a second modality of higher proficiency learners (28 learners) with scores ranging from 11 to 16 points (average score of all learners = 9.38). This was evidenced by the dip in frequency of scores at 10-11 points and the subsequent small rise at 13-14 points, although there is a gap in scores where no participant received a score of 12. This method of dividing into smaller groups yields a smaller group of higher proficiency learners and a larger group of lower proficiency learners, with the aim of 'mirroring' the actual proficiency of students enrolled in intermediate and advanced courses in a university, in that there are typically more beginner and intermediate students enrolled in language courses than advanced students.

Figure 6. Distribution of L2 proficiency scores.

All HLLs also completed the proficiency test; however, they were not further divided due to the small sample size. We also recognize that grammatical proficiency exams are not an accurate measurement of HLLs' actual proficiency (e.g., Birdsong, 2014; Shea, 2017), and, as in studies such as Amengual (2016), the role of Spanish dominance is employed instead. The role of Spanish dominance in a previous group of HLLs was considered in our previous study (Garza et al., 2024) and was not found to be significant in that larger sample of participants. Nevertheless, the primary purpose of the present study is to consider how different groups of learners in primarily mixed-classroom environments perceive the tap-trill contrast, especially our L2 learners, given the results of Garza et al. (2024).

3.2. Stimuli

There were 46 stimuli employed in the perception task (40 experimental[4] and 6 non-experimental stimuli). In all, the stimuli were modifications of 7 words that can form a minimal pair in Spanish. Stimuli were modified from natural voice recordings of words containing intervocalic /ɾ/, and all were produced by four speakers of Spanish, whose recordings were made by the researchers in Melero-García & Cisneros (2020); the dialects included Mexican, Puerto Rican, and Peninsular Spanish. Five words were included in the experimental stimuli: *ahora* ('now'), *caro* ('expensive'), *cero* ('zero'), *coral* ('coral'), and *pera* ('pear'). See Table 1 for a listing of each stimulus and its corresponding minimal pair. The non-experimental stimuli were incorporated to include a wider range of productions and were produced by a female monolingual speaker of Costa Rican Spanish who produced trills with frication. The non-experimental Costa Rican stimuli were removed from analysis in the present study.

3.2.1. Stimuli manipulation

The original stimulus words were manipulated in Praat (Boersma & Weenink, 2020) to produce variable closure durations of a single closure. Most naturally produced taps had a segmental closure duration in the range of 22-25 milliseconds (ms), consistent with previous production research (e.g., Amengual, 2016; Blecua, 2001; Bradley & Willis, 2012; Henriksen, 2015; J. Y. Kim & Repiso-Puigdelliura, 2020; Quilis, 1993; Willis & Bradley, 2008). If some did not, the closure duration was first altered by cutting individual pulses within the waveform until the segmental duration fell within the 22-25 ms range. Next, four additional bins of closure duration were created by copying and pasting individual pulses in the waveforms so that increments of approximately 15ms were added to the original starting token. The five bins or ranges of segmental closure duration that resulted were as follows: 22-25ms, 37-40ms, 52-55ms, 67-70ms, and 82-85ms. Employing these five bins of segmental closure durations then allows us to clearly determine at what segmental durational ranges there may be perceptual overlap and evaluation certainty. These ranges were chosen due to the results of Melero-García & Cisneros (2020), in which it was found that native speakers of Spanish created a perceptual category boundary at about 50ms under which they tended to perceive a tap as "good" and above which they tended to perceive a tap as "bad". Notably, L2 learners in this study demonstrated a bias toward perceiving all productions as "good", regardless of segmental duration and status as a tap or trill, suggesting no clear evidence for the development of a category boundary. However, Garza et al. (2024) find that heritage listeners created a perceptual category boundary at around 61ms.

The primary differences between the stimuli creation in the present study (also employed in Garza et al. (2024) and that of Melero-García & Cisneros (2020) involve our employment of controlled bins of segmental duration and the manipulation

---

[4] In all, 4 native Speakers of Spanish recorded the original 5 experimental words with a tap. Then, 5 bins of duration were created per speaker per experimental word (for a total of 100 possible tokens). Then, 40 tokens were pseudo-randomly selected to be included in the survey, taking into account evenly distributing the tokens according to variation of segmental duration, speaker, and word. Forty of the total 100 possible tokens were included to lower the cognitive load on the participant. We aimed to include less tokens per listener and instead opted for more listeners and less tokens per listener.

procedure of lengthening the stimuli. In Melero-García & Cisneros (2020), stimuli were increased in 15 ms intervals, regardless of the starting closure duration of the tap produced in the natural recordings. This approach resulted in the shortest durations produced, ranging from approximately 20 to 45 ms, and the resulting "trill-like" lengths were over 100 ms in length, which is greater than the natural trill lengths found in previous research (e.g., Bradley & Willis, 2012; Willis & Bradley, 2008). Furthermore, rather than copying and pasting individual pulses in the waveforms of the rhotic segment, as done in our study, Melero-García & Cisneros (2020) utilized manipulation of speed in Praat. Participants listened to 40 experimental stimuli of five minimal pairs in Spanish, each with the modified closure durations explained.

Table 1. Minimal pairs employed in the study.

Recorded stimulus	*Corresponding minimal pair**
Experimental	
ahora ('now')	*ahorra* ('save money')
caro ('expensive')	*carro* ('car')
cero ('zero')	*cerro* ('hill')
coral ('coral')	*corral* ('corral')
pera ('pear')	*perra* ('dog')
Non-experimental	
pero ('but')	*perro* ('dog')
moro ('moor')	*morro* ('snout')

*Note: The corresponding minimal pairs containing intervocalic /r/ were not recorded by native or monolingual speakers. Instead, the aim of the study was to create the perception of a word containing a trill by manipulating the segmental closure duration within Praat to sound trill-like (Boersma & Weenink, 2020).

3.3. Perception tasks

Two tasks were administered through an online survey using Qualtrics: a 2AFC and a VAS. Prior to the listening tasks, learners underwent a brief vocabulary training session in which word pairs were paired with photographs, which would be used in the perception tasks rather than orthographic representations of each word. This training was developed based on the results of piloting, which found that listeners required more time with the vocabulary-image pairings before proceeding to the perception task. For example, the image of the diamond is intended to represent an expensive item caro.[5] Graphical representations of the words were employed as previous research suggests that orthography can affect lexical representations in L2 learners (e.g., Hayes-Harb et al., 2010). Specifically, Hayes-Harb et al. (2010) found that orthography significantly affected the remapping of rhotics in developing English-Spanish bilinguals. See Figure 7 for the photographs that learners were presented with
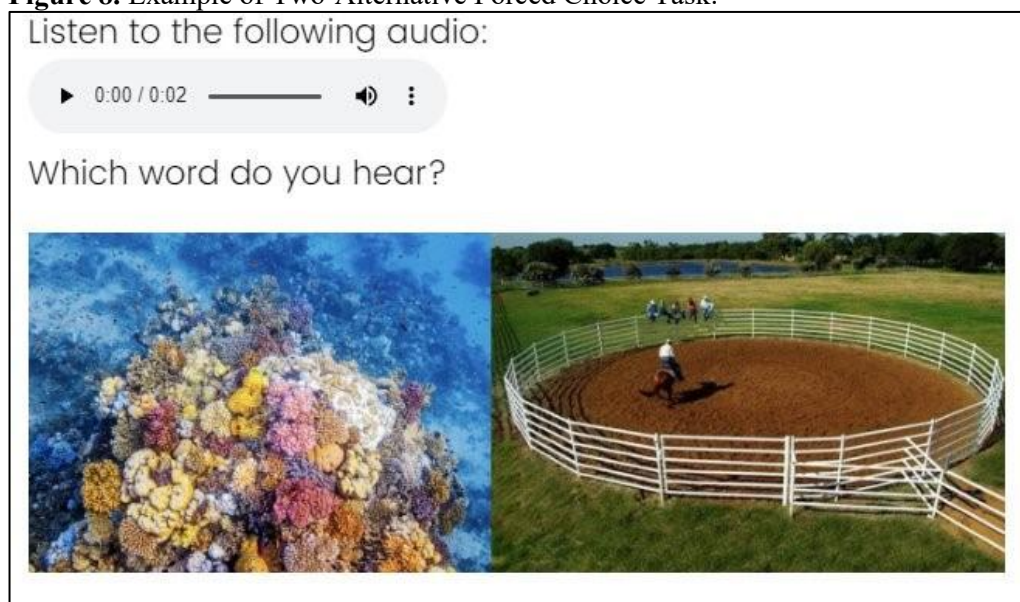
---

[5]   An anonymous reviewer questioned the need for the training component. Several practice items allow the participant to become familiar with the task and reduce confusion (despite the seemingly simplistic nature of the task). It has been our experience through multiple pilots and studies that a simple practice training session helps familiarize the participant with the tasks.

for each word pairing. Two perception tasks were incorporated within the study to elicit both continuous and categorical responses. The first task was a 2 AFC, eliciting categorical responses (see Figure 8). Participants were asked to listen to an audio and indicate which word they heard by clicking on one image or the other.

Figure 7. Vocabulary and photograph pairings.

Figure 8. Example of Two-Alternative Forced Choice Task.

We specifically chose images to facilitate a lexical decision based on understanding the word rather than its spelling. The 2AFC task enabled us to associate the audio with specific visual representations that reflect meaningful differences without requiring additional abstraction. We hoped that this visual connection between the stimuli and phonemic evaluation could then be transferred to the VAS more effectively, encouraging participants to use the slider to reflect their level of certainty in their selection (see Figure 9). The VAS was used to elicit continuous data along a gradient of certainty. Participants were presented with 2-3 questions on each page of their survey, either for the 2AFC or VAS task. The order of these groups of questions was pseudorandomized to avoid repetition of the speaker and lessen task fatigue by presenting speakers with alternating tasks.

Figure 9. Example of VAS perception task.

4. Results

4.1. Listener groups by background, proficiency, and confidence

This section begins with descriptive statistics and distributions of the study's results. Recall that the primary learner characteristics of the study are linguistic confidence (a measurement adapted from the "English use confidence" index in Gardner et al. (1997) and grammatical proficiency (using the abbreviated grammar proficiency task incorporated in Kanwit & Geeslin (2014), Linford & Geeslin (2017). Table 2 presents the distribution of data according to participant group.

Table 2. Proficiency and linguistic confidence scores.

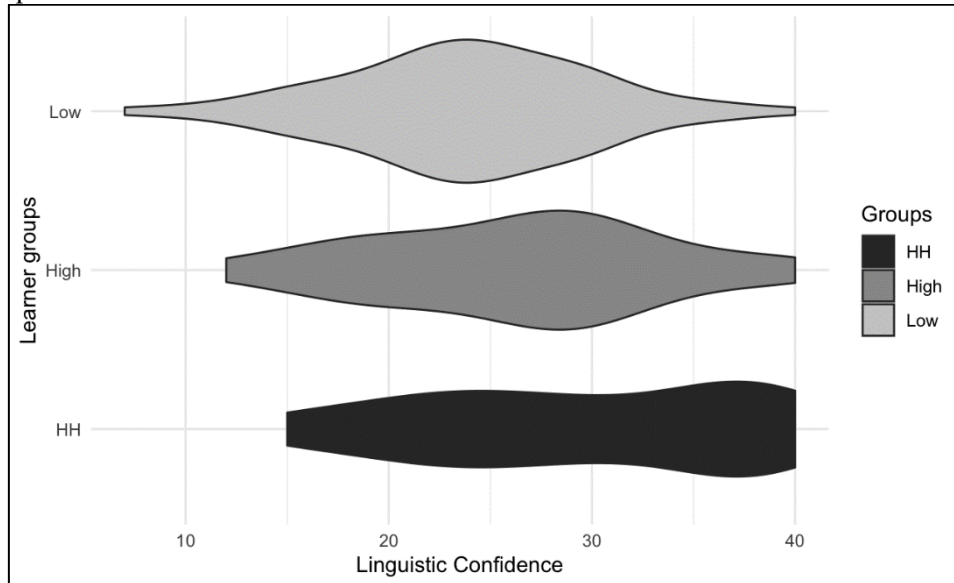| Learner group	Number of learners	Grammatical proficiency scores (out of 20 points)	Linguistic confidence scores (out of 40 points)
Lower proficiency L2	76	Range: 5 - 10 Average: 8.16	Range: 7 - 40 Average: 23.83
Higher proficiency L2	28	Range: 11 - 16 Average: 12.70	Range: 12 - 40 Average: 26.25
*Heritage learners	31	Range: 6 - 19 Average: 12.06	Range: 15 - 40 Average: 29.55

*Note: This data is displayed to present all the data collected in the study. Grammatical proficiency was not used to differentiate HLLs from L2 learners in the study.

Figures 10 and 11 use violin plots to display the distribution of these scores across the three groups of listeners in this study: heritage listeners, lower proficiency L2 listeners, and higher proficiency L2 listeners. As mentioned, the L2 group was divided into two subgroups to better understand the potential effect of grammatical proficiency on phonological acquisition. Figure 11 displays the individual listener groups that were formed. Learners with a score of 10 or below formed the lower
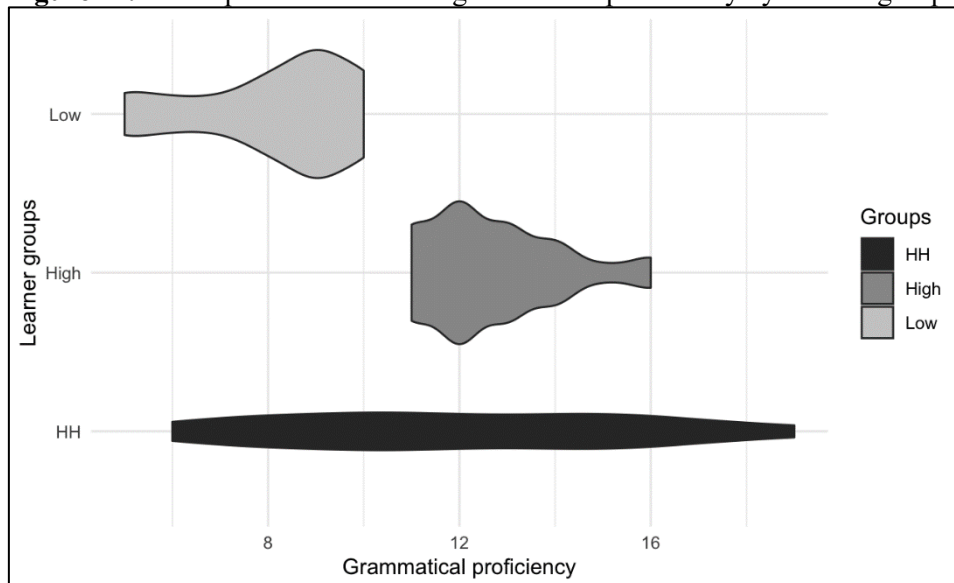
proficiency group, and learners with a score of 11 or above formed the higher proficiency group.

Figure 10. Violin plot of linguistic confidence distribution by listener group based on the ten-question self-evaluation index.

Figure 11. Violin plot distribution of grammatical proficiency by listener group.

All statistical models and figures were conducted in RStudio using various packages, including ggplot2 and dplyr. First, an ANOVA test was conducted to determine whether the listener groupings differed statistically in terms of linguistic confidence and grammatical proficiency. The results of an ANOVA test indicate that the linguistic confidence scores of each group differ significantly from one another $(F(2682) = 7.88, p < .001)$. A Tukey's HSD post-hoc test was conducted to perform pairwise comparisons and provide a better understanding of which variables were significantly different and at what level each differed. This test found that the confidence score of the HLL (mean = 29.55) was significantly higher than that of the

lower proficiency L2 group (mean = 23.83; p < .001). The difference between the higher L2 (mean = 26.25) and lower L2 group (mean = 23.83) was not significant (p = .247), nor was the difference between the higher L2 group (mean = 26.25) and HLL (mean = 29.55; p = .156).

A second ANOVA test was conducted to determine how grammatical proficiency differed significantly between second learner groups. The results suggest that the scores between at least two groups were significantly different from one another (F(2682) = 58.2, p < .001). More specifically, a Tukey's HSD post hoc test for pairwise comparisons finds that the proficiency score of the HLL (mean = 12.06) was significantly higher than that of the lower L2 group (mean = 8.16; p <.001) and that the proficiency score of the higher L2 group (mean = 12.70) was also significantly higher than, the lower L2 group (mean = 8.16; p <.001). The difference between the grammatical proficiency of the heritage (mean = 12.06) and the higher L2 group (mean = 12.70) was not significantly different (p = .550).

The relationship between the scores is presented in Figures 10 and 11, and a scatterplot is shown in Figure 12, illustrating the correlation between these two continuous variables. Figure 12 illustrates that while higher-proficient L2 learners and HLLs demonstrated a positive correlation between proficiency and linguistic confidence, the lower-proficient L2 speakers did not. Additionally, the lower proficiency of L2 speakers reveals an unexpected negative correlation between linguistic confidence and grammatical proficiency.

Figure 12. Plot matrix displaying distributions of and relationships between certainty, linguistic confidence, and grammatical proficiency.
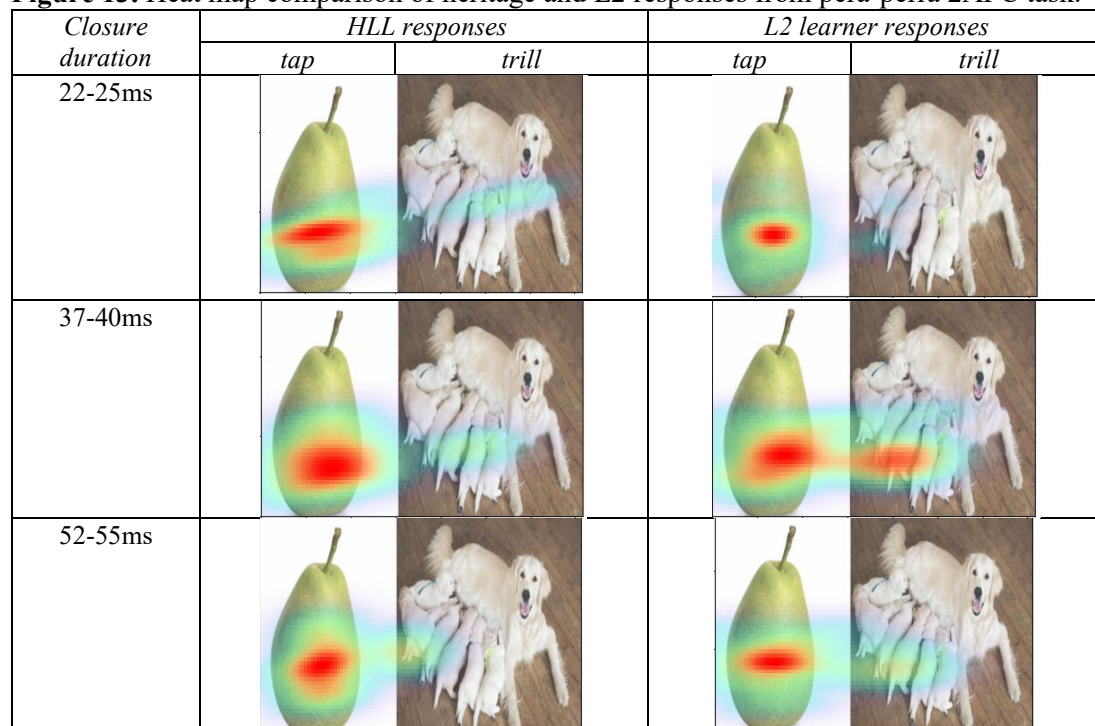

Note: This plot is interpreted as any matrix by comparing the variable labels along the x- and y-axis. For example, to compare certainty and grammatical proficiency, look for the "certainty" label along one axis, and the grammatical proficiency label along the other. Then, it is possible to view the corresponding bubble plot and correlation values. Meanwhile, to view the distribution of only one variable, locate the corresponding variable label on the x- or y-axis to display the density plots.

In sum, comparing linguistic confidence and grammatical proficiency across L2 and heritage learner groups highlights how different learner backgrounds might predict phonemic perception. By examining the correlation between these two variables, these variables demonstrate how learners' self-perception and actual skill relate to their certainty. The interaction between these variables supports the understanding of whether linguistic confidence and proficiency should be treated as overlapping or distinct constructs in data analysis and interpretation.

4.2. 2AFC categorical perception results

This section of results focuses on the perception of taps and trills as a categorical dependent variable. To begin with, a logistic regression was employed with 'trill' as the reference variable and closure duration, grammatical proficiency, linguistic confidence, and word-pair as the independent variables. Participant was included as a random effect that was not found to be significant. The results of a logistic regression suggest that selecting /r/ was predicted by a longer closure duration (β = .009, p < .001), being a HLL (β = .311, p < .001), and lower grammatical proficiency (β = -.063, p < .001). This is exemplified by the results of the forced-choice heat map task displayed in Figure 13. This figure is chosen because it illustrates the photo pairings participants viewed while completing the task, as well as the photo they clicked based on their perception. This figure demonstrates that longer segmental closure duration elicits more selections of the rightward image of <perra> [pe.ra] 'dog' instead of <pera> [pe.ɾa] 'pear'. However, Figure 13 also demonstrates that despite the length of the segmental closure duration of [ɾ], learners of both groups demonstrate a tendency to select <pera> [pe.ɾa] 'pear'. Nevertheless, this tendency is less pronounced among HLLs compared to L2 learners.

Figure 13. Heat map comparison of heritage and L2 responses from pera-perra 2AFC task.
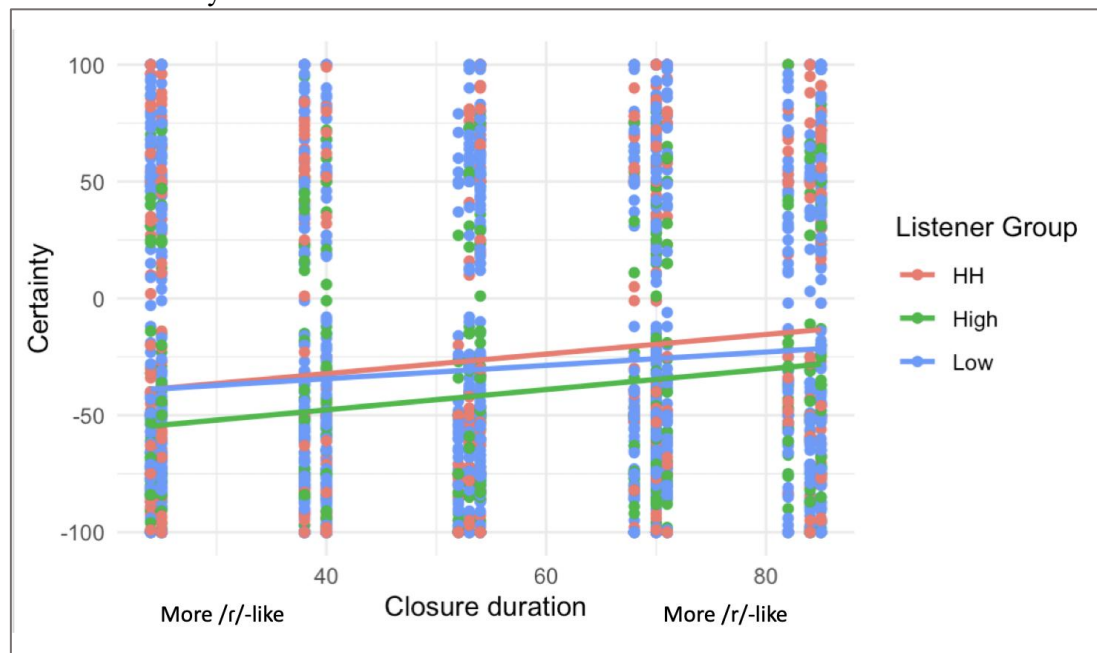
| 67-70ms | | |

4.3. VAS continuous results of perceptual certainty

This section of the results focuses on perception as a continuous dependent variable. To understand how the independent variables of this study predict movement on the variable analog slider (VAS) and its certainty of perception, a linear regression was conducted with the continuous placement of the slider as the dependent variable. The independent variables included in the model were closure duration of the rhotic segment, lexical item, linguistic confidence, and grammatical proficiency. Participant was also included as a random effect. The results of the linear regression suggest that certainty of selecting /r/ was predicted by higher linguistic confidence ($\beta = 1.119$, $p < .001$) and lower grammatical proficiency ($\beta = -1.424$, $p < .05$) but not by closure duration ($\beta = -.009$, $p = .796$) or individual word pairs.

Though listener type was included in the initial regression and not found to be a significant predictor of certainty, we employed individual linear regressions by listener group to uncover whether any listener group effects were being obscured in the single, larger model. However, even when modeled by individual groups, duration was still not a significant predictor of certainty. This result is unexpected and suggests that the effect of learner characteristics significantly affects performance on a task in such a way that their values are more impactful than the experimental variable of the study, particularly in terms of certainty of selection, as illustrated in Figures 14 and 15.
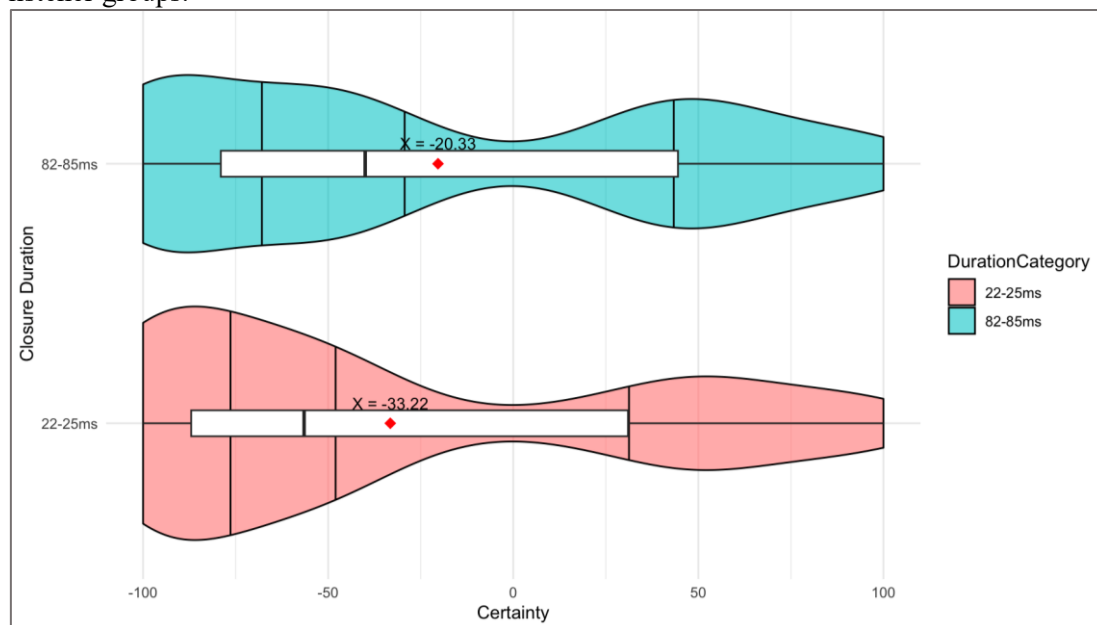
Figure 14 suggests that lower-proficient learners exhibit blurry and non-categorical perception of the difference between the Spanish phonemic tap and trill, based on closure duration. In other words, less proficient learners demonstrate less confidence in closure duration as an acoustic cue, as evidenced by a lack of clear categorical distinctions, such as tap versus trill, when compared to the other listener groups in this study. Meanwhile, higher proficient L2 and heritage listeners demonstrate a similar distinction from one another, as evidenced by their parallel slopes. The inclination of the slope also suggests that highly proficient learners were more likely to select "tap" for all questions, while heritage listeners were somewhat less likely to select "tap". While not statistically significant, the VAS in Figures 14 and 15 also shows a trend of change towards perceiving a longer duration (82-85 ms) as a phonemic trill versus the shorter duration bin (22-25 ms).

Figure 14. Scatterplot displaying the relationship between rhotic closure duration and selection certainty.

Note: On the y-axis, 100 refers to 100% certainty of hearing a word with /r/, and –100 refers to 100% certainty of hearing a word that contains /ɾ/. Meanwhile, the x-axis reflects the segmental duration of the [ɾ].
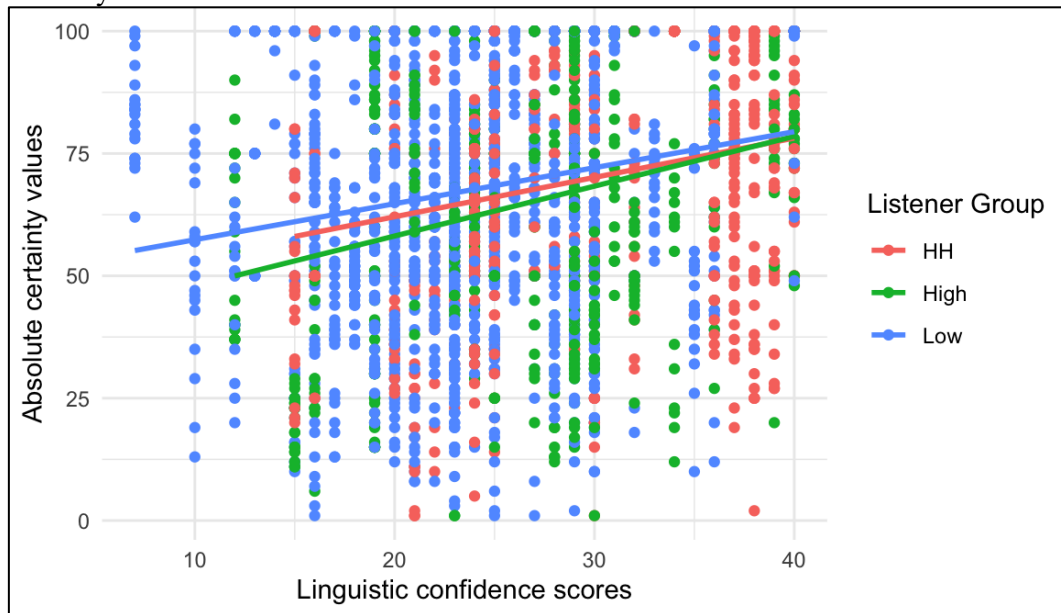
Figure 15. Violin plot with internal boxplot and mean of closure duration extremes among all listener groups.

Note: On the x-axis, 100 refers to 100% certainty of hearing a word with /r/, and –100 refers to 100% certainty of hearing a word that contains /ɾ/.

Figure 16. Scatterplot displaying the relationship between linguistic confidence and selection certainty.

Finally, Figure 16 displays the positive correlation between the certainty of hearing a tap or trill and linguistic confidence scores by listener group. As shown, higher linguistic confidence in each group relates to being more certain of tap or trill selection. When comparing Figures 12 and 16, it is more visually apparent from Figure 16 that the slope is stronger between linguistic confidence and certainty than between segmental duration and certainty. By utilizing absolute certainty values, Figure 16 provides a more precise explanation of overall movements – in either direction, both negative and positive – from zero. This reflects movement toward the extremes, which is used as a metric of higher certainty in the present study.

5. Discussion

5.1. Research questions

5.1.1. RQ (1) Do L2 learners of differing proficiency backgrounds attend to the secondary cue of closure duration as a perceptual cue of tap-trill categorization?

The first research question explored whether L2 learners attend to the closure duration of a single tap as a secondary cue of tap-trill categorization. The results of this study suggest that, while their categorical association based on the perception of an extended closure is not as robust compared to HLLs in this study or Garza et al. (2024) and monolinguals in Melero-García & Cisneros (2020), higher proficient L2 learners demonstrate a small, but non-significant, tendency toward associating the phonemic trill with a longer closure duration and a tap with shorter closure duration. These findings are most pronounced when additional speaker variables are included. For example, L2 listeners with higher grammatical proficiency and a greater sense of linguistic confidence associate the increased duration with a trill more so than listeners with less proficiency and confidence. These differences are not statistically significant

in the results for the segmental duration. Still, the slopes displayed in Figures 14 and 16 suggest a role for learner characteristics, which could be further explored in other ways.

Although closure duration significantly predicted tap-trill perception in this study, the effect is weak (β = 0.009, p < 0.001) and not significant when considering the certainty of selection as predicted by closure duration (β = -0.009, p = 0.796). While the results of this study do not suggest a strong relationship between closure duration and contrastive perception of the phonemic tap and trill in Spanish, they do demonstrate that secondary acoustic cues of duration can be recoverable to these learner groups and that in the absence of occlusions, the tap-trill contrast is not entirely neutralized (see Figures 14 and 15).

5.1.2. RQ (2) Does the role of duration impact categorization in L2 in the same way it impacts categorization for HLLs?

In this study, the listener type – HLL or L2 learner – played a significant role, likely due to the difference in early Spanish phonological exposure experienced by HLLs compared to traditional L2 learners of Spanish. In this study and in the results of Garza et al. (2024), we find that HLL categorization is clearer, likely due to this difference. While previous literature has considered HLL (Garza et al., 2024) and L2 learners' (Melero-García & Cisneros, 2020) perceptions of Spanish rhotics, these studies included learners from different regions of the U.S. who completed different variations of a similar perception task. In the present study, all learners are enrolled in Spanish courses within the same program at a large Midwestern university, and all listeners completed the same task, which serves for a more direct comparison between the two listener groups.
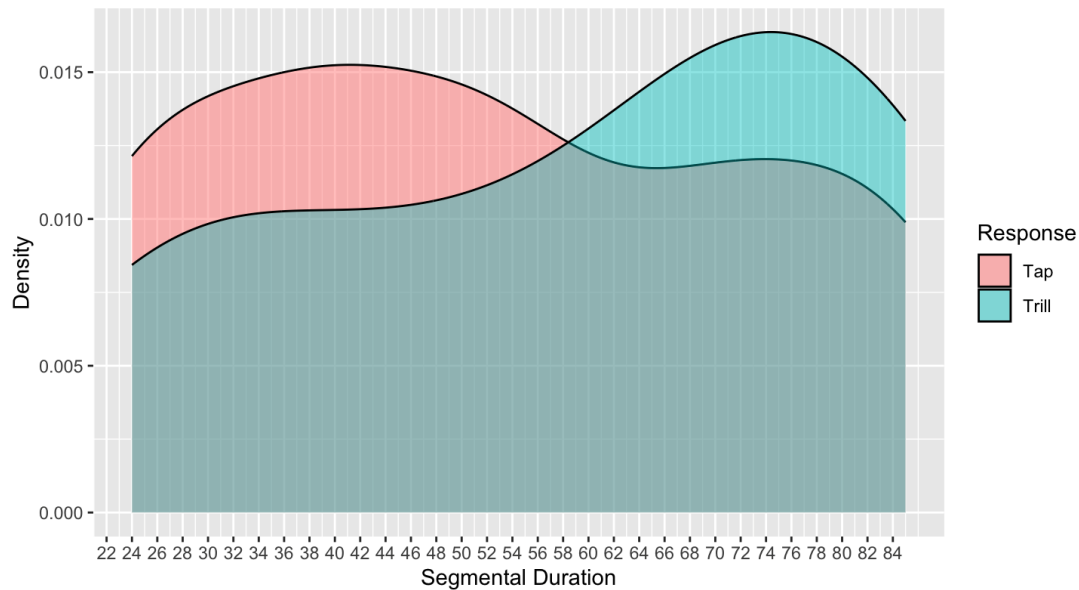
The results suggest that tap-trill categorization based on duration is an emergent and variable acoustic cue across L2 learners. The ability to recognize duration as an acoustic cue appears to reflect a continuum, with a greater role in percentage and a greater impact on the number of listeners who used it as a criterion. Similar to the findings of Melero-García & Cisneros (2020), a majority of NSs perceive a trill when the closure duration exceeds 50 ms. In comparison, HLLs in the present study and in Garza et al. (2024) begin to perceive a trill when the rhotic closure duration reaches about 60ms or more. This also appears to be true for higher-proficiency L2 learners, as indicated by the results of a density plot in Figure 17, which suggests that the category boundary occurs on average at around 59 ms. In this plot, the leftward red wave demonstrates the distribution of tokens perceived as taps, while the rightward blue wave demonstrates the distribution of tokens perceived as trills. Both waves overlap at about 58.5ms.

In contrast, Figure 18 shows that lower proficiency L2 learners demonstrate more irregular categorization between the tap and trills, lending itself to less clarity. Although the category overlap occurs similarly around 59ms in this group, there is no clear bimodal distribution to support this, as shown in Figure 17. We conclude that, for lower-proficient L2 learners, the majority of the group may not perceive a trill, despite increased closure duration, even though they are partially sensitive to the segmental duration of the rhotic segment.
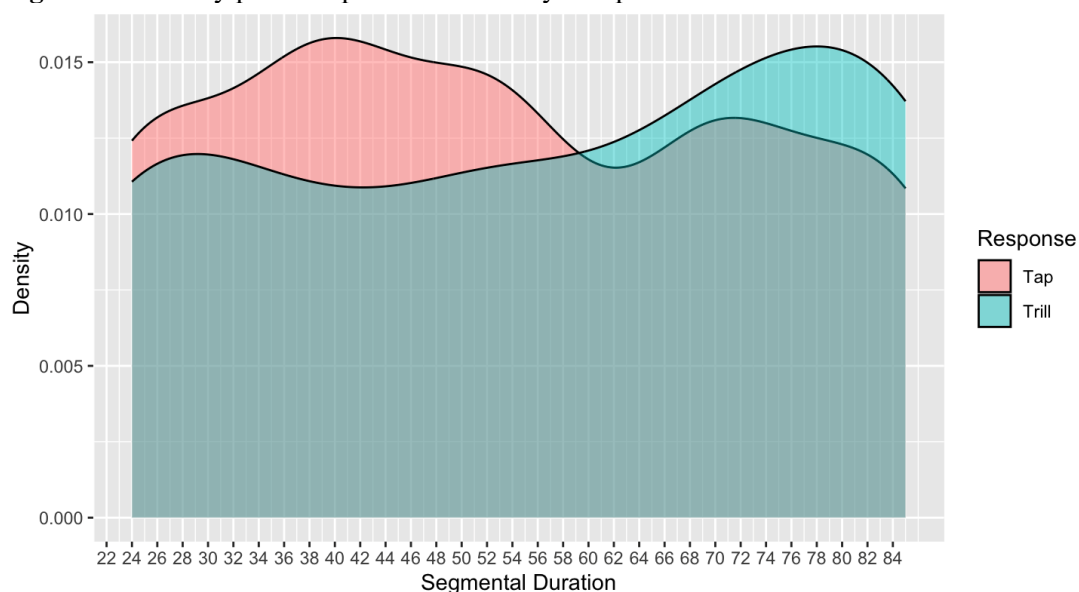
The differences in boundary response may reflect the use of durational bins in the two studies or the nature of the task. Whereas Melero-García & Cisneros used a two-alternative forced-choice task throughout, the current study used a VAS for the

majority of the evaluations. Another difference is that Melero-García & Cisneros modified variable tap/trill stimuli productions by a set percentage. In contrast, the current study employed modified durations of natural speech that fell within set duration bins. Overall, the results of this study suggest that HLLs exhibit somewhat more defined perceptual categories of rhotic phonemes and demonstrate higher certainty in their choices than L2 learners. Meanwhile, L2 learners demonstrate more categorical overlap, suggesting fuzzier categorical boundaries when perceiving taps and trills that differ only by closure duration.

Figure 17. Density plot of tap-trill selection by highly proficient L2 learners.

Figure 18. Density plot of tap-trill selection by low proficient L2 learners.

5.1.3. RQ (3) How do linguistic confidence and grammatical proficiency predict selection certainty using VAS?

The third research question of the study explored the factors that predict selection certainty using VAS. The results suggest that participant listener type (as L2 or HLL), grammatical proficiency, and especially Spanish linguistic confidence predict certainty. The results also suggest that these extralinguistic variables demonstrated a more robust effect on the continuous perception task than the primary independent variable of the study, closure duration. These findings suggest that extralinguistic factors contribute to phonological perception, and it is important to consider the role of these external factors when preparing a perception study. Regarding the present study, we predicted that the continuous variable would relate to certainty in terms of closure duration. Instead, it seems that *categorically,* the direction in which the slider was moved – toward a word with /r/ or a word with /ɾ/, is significantly predicted by closure duration. However, the *continuous* extent to which the slider was dragged – or, rather, the degree of certainty was instead more strongly related to linguistic confidence and proficiency, where higher confidence and lower proficiency were related to higher certainty of selection amongst L2 learners. Conversely, HLL were also more likely to have a higher certainty of selection, despite differences in their linguistic confidence.

Regarding proficiency, as could be expected, the results of this study underscore that grammatical proficiency. In contrast, a convenient measurement of proficiency is not equivalent to phonological encoding or awareness of phonemic categories. For example, we find that low L2 proficient listeners display higher certainty in their choices, while also demonstrating lower attention to closure duration, as evidenced by the more neutral relationship between closure duration and certainty shown in Figure 14. While grammatical proficiency demonstrated a relationship with linguistic confidence in this study, offering a comparison with perception, future phonological research that considers the role of proficiency in predicting learner perception should employ a phonologically-based exam, such as an elicited imitation task (see Kostromitina & Plonsky, 2022; Solon et al., 2019; Solon et al., 2022). Elicited imitation tasks have been employed in both L2 and heritage research as a measure of linguistic proficiency in which listeners hear sentences in the target language of increasing length and repeat what they hear. These tasks may provide a more appropriate measure of proficiency that relates more directly to phonological perception rather than a grammar task.

Finally, linguistic confidence – a measure operationalized as confidence in one's ability to learn and communicate in Spanish – yields the strongest relationship with certainty measurement, suggesting that a learner's feelings regarding their general ability relating to a subject, in this case, listening to Spanish, is more impactful on their continuous responses than perceived linguistic variables. This could be due to linguistic confidence and certainty measuring being two measurements of a similar learner characteristic, in that both are metrics of the learners' belief in their own Spanish abilities, where certainty reflects a belief in their correct perception of a word, and linguistic confidence reflects the belief in their communicative ability and ability to advance in Spanish proficiency.

5.2. Individual variation: linguistic proficiency and confidence

As shown in Figure 12, there is a negative correlation between confidence and proficiency among low-proficiency speakers. This is reminiscent of the Dunning-Kruger effect, which, in the context of this study, supports the idea that the lower-level learners can be unaware of their actual Spanish proficiency (see Dunning, 2011). This could explain why the low-proficiency group reports higher confidence as their proficiency scores become lower. Meanwhile, among higher proficiency learners, higher proficiency is associated with greater confidence.

Previous studies have consistently shown that the trill is one of the most complex segments for L2 learners of Spanish (Amengual, 2016; Daidone & Darcy, 2014; Rose, 2010; Waltmunson, 2005). It has also been found to be among the last segments mastered by native-speaking children (Jiménez, 1987). Additionally, research suggests that lexical associations between the tap and the trill are challenging for learners due to their similarity and the lack of a meaningful distinction in English (Daidone et al., 2014; Menke, 2018; Rose, 2010). Therefore, it is not surprising that learners struggle with perceiving a non-normative variant based on an acoustic feature that is rarely addressed in basic language instruction. One potential interpretation of these findings is that linguistic confidence, as demonstrated by L2 listeners who associated a longer duration with a phonemic trill, may serve as an indirect indicator of sociolinguistic competence in phonemic contrasts. The experiment investigated non-traditional productions of the trill phoneme, a single occlusion tap with variable durations. The variants are produced by native speakers (Bradley & Willis, 2012). In our study, the students most likely to associate longer duration with a dynamic trill were those who reported the highest linguistic confidence—not necessarily those with the highest linguistic proficiency. Future studies should investigate the underlying reasons behind linguistic confidence and explore the connections with willingness to communicate, a similar construct in individual differences research.

6. Conclusion

6.1. Future research, bias, and word pairs

Though word-pair was not found to be a significant predictor in this study, the role of frequency was not fully explored. For example, taking into consideration individual words within each word pair, such as *ahora* 'now' and *ahorra* 'save money', it could be assumed that *ahora* would be much more frequently expected to occur in conversation than *ahorra* 's/he saves', therefore affecting the perception of *ahora* 'now' even with longer, trill-like segmental duration. All participants were familiarized with the word pairs. All word pairs can be considered common and frequent and were found in a corpus of core vocabulary for learners (Davies & Davies, 2017), except for *coral* 'coral' and *corral* 'corral'. Our statistical analysis of word pairs did not reveal a significant difference; however, it is possible that other frequency effects play a role. This frequency effect may reflect the individual speaker's use of the terms. Therefore, future exploration of frequency effects could include individual knowledge, familiarity, and use of particular items. It's possible that a learner's

heightened exposure, due to the frequency of a word within their input, will bias them to hearing one word versus the other, regardless of the acoustic correlates present within the rhotic segment.

Regarding listener characteristics, future research should further consider the role of learner characteristics – such as proficiency and confidence, as considered in the present study – and their resulting biases in perception, especially regarding Likert-scale or continuous perception measurements. It is likely that several learner characteristics influence continuous perception, leading to biases in favor of choosing extremes or remaining neutral in one's certainty of selection (see Kapnoula et al., 2017). Further factors, such as linguistic insecurity, language anxiety, education, experience with survey-taking, motivation, and others, should be considered to gain a more comprehensive understanding of the methodological concerns related to incorporating continuous data in perception research.

Finally, one methodological limitation of this study is that distractor stimuli were not incorporated. Distractors were not included in order to maintain a smaller task size, as initial piloting resulted in evidence of task fatigue. Nevertheless, it is likely that participants were overtly aware of what they were perceiving (i.e., differences in segmental duration) and that this could have resulted in biases. Future replications of this study should consider incorporating phonological distractors to mitigate these possible confounding effects.

6.2. Key takeaways

In summary, the current research suggests that higher ratings of linguistic confidence in Spanish communication, as well as higher Spanish proficiency, are related to higher ratings of perceptual certainty between Spanish taps and trills. While the categorical data from the 2AFC task suggest that segmental duration significantly relates to tap-trill perceptual contrast across listener groups, the results of VAS continuous certainty suggest that linguistic confidence plays a strong role in affecting the certainty of selection. In other words, regardless of segmental duration length, linguistic confidence best predicted certainty on VAS ratings. Nevertheless, the results of this study suggest that even in the absence of occlusions, neutralization of tap-trill phonemes is not guaranteed. The heritage listeners, low-proficiency L2 listeners, and higher-proficiency L2 listeners all demonstrate a tendency to attend to segmental duration in order to contrast between these phonemes. Further research is needed to better understand how listeners process this task and what additional learner characteristics – such as language anxiety or motivation – may contribute to responses in the dependent variable. These results and future suggestions have implications not only within Hispanic linguistic research but also cross-linguistically, in that learner characteristics remain a constant source of variation amongst all types of language acquisition.

This research also contributes to ongoing discussions in social science research at large regarding the merits of incorporating both categorical and continuous dependent variables to tell a more comprehensive story about data trends. In this study, the use of continuous and categorical data complemented one another and helped tell a larger story. The categorical data results demonstrate an overall trend that segmental duration is a salient secondary cue that learners are attuned to when contrasting

segments. Nevertheless, the continuous data employed suggests that certainty about this perception is related to individual learner characteristics, such as linguistic confidence and proficiency, rather than duration alone. These findings suggest that learner characteristics influence the perception of certainty among learners.

Acknowledgments

References

Amengual, Mark. 2016. Acoustic correlates of the Spanish tap-trill contrast: Heritage and L2 Spanish speakers. *Heritage Language Journal* 13(2): 88–112. https://doi.org/10.46538/hlj.13.2.4.

Apfelbaum, Keith S., Kutlu, Ethan, McMurray, Bob, & Efthymia C. Kapnoula. 2022. Don't force it! Gradient speech categorization calls for continuous categorization tasks. *The Journal of the Acoustic Society of America* 152(6): 3728–3745. https://doi.org/10.1121/10.0015201.

Arróniz, Santiago. 2025. *Compensatory cues of /s/ reduction in voiced stop consonants in Western Andalusian Spanish*. Ph.D. thesis. Indiana University.

Blecua, Beatriz. 2001. *Las vibrantes del español: manifestaciones acústicas y procesos fonéticos*. Universitat Autònoma de Barcelona.

Boersma, Paul, & David Weenink. 2020. Praat: doing phonetics by computer [Computer program]. Version 6.1.16, retrieved December 2020 from https://www.praat.org.

Bradley, Travis G., & Erik W. Willis. 2012. Rhotic variation and contrast in Veracruz Mexican Spanish. *Estudios de Fonética Experimental* 21: 43–74.

Clopper, Cynthia G., & David B. Pisoni. 2007. Free classification of regional dialects of American English. *Journal of Phonetics* 35(3): 421–438. https://doi.org/10.1016/j.wocn.2006.05.001.

Cummings Ruiz, Laura. D., & Silvina Montrul. 2020. Assessing rhotic production by bilingual Spanish speakers. *Languages* 5(4): 51. 10.3390/languages5040051.

Daidone, Danielle, & Darcy, Isabel. 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*, 39–50. Somerville, MA: Cascadilla Proceedings Project.

Daidone, Danielle, Lidster, Ryan, & Franziska Kruger. 2023. Free classification as a method for investigating the perception of nonnative sounds. *Studies in Second Language Acquisition* 45(4): 1104–1130. 10.1017/S0272263123000050

Davies, Mark, & Kathy Hayward Davies. 2017. *A Frequency Dictionary of Spanish: Core Vocabulary for Learners* (2nd ed.). Routledge. 10.4324/9781315542638

de Jong, Kenneth. 1998. Stress-related variation in the articulation of coda alveolar stops: flapping revisited. *Journal of Phonetics* 26: 283–310. 10.1006/jpho.1998.0077

Delgado-Díaz, Gibrán, Galarza, Iraida, & Manuel Díaz-Campos. 2021. Complex attitudes towards two sociolinguistic variables and their social meanings: Providing evidence from production and perception data in a speech community. In M. Díaz-Campos, & S. Sessarego (eds) *Aspects of Latin American Spanish Dialectology: In honor of Terrell A. Morgan*, 125–154.

Dunning, David. 2011. The Dunning–Kruger effect: On being ignorant of one's own ignorance. In J.M. Olson, & M.P. Zanna (eds). *Advances in experimental social psychology*, 247–296. Cambridge, MA: Academic Press.

Ellis, Rod. 1997. *The study of second language acquisition.* Oxford / New York: Oxford University Press.

Face, Timothy L. 2006. Intervocalic rhotic pronunciation by adult learners of Spanish as a second language. In C. A. Kleen, & T. L. Face (eds), *Selected Proceedings of the 7th Conference on the Acquisition of Spanish and Portuguese as First and Second Languages*, 47–58. Somerville, MA: Cascadilla Proceedings Project.

Flege, James Emil. 2021. New methods for second language (L2) speech research. In R. Wayland (ed.), *Second Language Speech Learning: Theoretical and empirical progress*, 119–156. Cambridge: Cambridge University Press. https://doi.org/10.1017/9781108886901.004.

Flege, James Emil, & Ian R. A. MacKay. 2004. Perceiving vowels in a second language. *Studies in Second Language Acquisition* 26(1): 1–34. https://doi.org/10.1017/S0272263104261010.

Gardner, Robert C., Tremblay, Paul F., & Anne-Marie Masgoret. 1997. Towards a full model of second language learning: An empirical investigation. *The Modern Language Journal* 81(3): 344–362.

Garza, Rachel S., Willis, Erik W., & Fernando Melero-García. 2024. More than occlusions: The role of duration in perceiving the Spanish tap-trill contrast by heritage

speakers of Spanish. *Recent Developments in Hispanic Linguistics: Studies in structure, variation, and bilingualism* 41: 122.

Hayes-Harb, Rachel, Nicol, Janet, & Jason Barker. 2010. Learning the phonological forms of new words: Effects of orthographic and auditory input. *Language and Speech* 53(3): 367–381. https://doi.org/10.1177/0023830910371460.

Henriksen, Nick. 2015. Acoustic analysis of the rhotic contrast in Chicagoland Spanish: An intergenerational study. *Linguistic Approaches to Bilingualism* 5(3): 285–321. https://doi.org/10.1075/lab.5.3.01hen.

Hualde, José I. 2005. *The Sounds of Spanish with Audio CD*. Cambridge: Cambridge University Press.

Hualde, José I. 2014. *The Sounds of Spanish.* Cambridge: Cambridge University Press.

Hulstijn, Jan H. 2019. An individual-differences framework for comparing nonnative with native speakers: Perspectives from BLC theory. *Language Learning: A Journal of Research in Language Studies* 69(51): 157–183. 10.1111/lang.12317.

Kanwit, Matthew, & Kimberly L. Geeslin. 2014. The interpretation of Spanish subjunctive and indicative forms in adverbial clauses: A cross-sectional study. *Studies in Second Language Acquisition* 36(3): 487–533. 10.1017/S0272263114000021.

Kapnoula, Efthymia C., Winn, Matthew B., Kong, Eun Jong, Edwards, Jan, & Bob McMurray. 2017. Evaluating the sources and functions of gradiency in phoneme categorization: An individual differences approach. *Journal of Experimental Psychology: Human Perception and Performance* 43(9): 1594. https://doi.org/10.1037/xhp0000405.

Kissling, Elizabeth M. 2018. An exploratory study of heritage Spanish rhotics: Addressing methodological challenges of heritage language phonetics research. *Heritage Language Journal* 15(1): 25–70. https://doi.org/10.46538/hlj.15.1.3.

Kong, Eun Jong, & Jan Edwards. 2016. Individual differences in categorical perception of speech: Cue weighting and executive function. *Journal of Phonetics* 59: 40–57. https://doi.org/10.1016/j.wocn.2016.08.006

Kostromitina, Maria, & Plonsky, Luke. 2022. Elicited imitation tasks as a measure of L2 proficiency: A meta-analysis. *Studies in Second Language Acquisition* 44(3): 886–911.

Ladefoged, Peter, & Johnson, Keith. 2015. *A course in phonetics.* Wadsworth Australia.

McCandless, Andrew. 2020. Phonetic training of L2 Spanish rhotics. *Proceedings of the 2020 annual conference of the Canadian Linguistic Association*, 1–16.

Melero-García, Fernando, & Alejandro Cisneros. 2020. No es tan simple como parece: The effect of duration of one closure rhotics on the perception of Spanish /ɾ/ and /r/. In D. Pascual y Cabo, & I. Elola (eds.), *Current Theoretical and Applied Perspectives on Hispanic and Lusophone Linguistics*, 295–318. John Benjamins Publishing Company. https://doi.org/10.1075/ihll.27.14mel.

Menke, Mandy R. 2018. Development of Spanish rhotics in Spanish–English bilingual children in the United States. *Journal of Child Language* 45: 788–806.

Navarro Tomás, Tomás. 1918. *Manual de pronunciación española.* Centro de Estudios Históricos.

Quesada Pacheco, Miguel. 1996. Los fonemas del español de costa rica aproximación dialectológica. *Lexis* 20: 535–562.

Quilis, Antonio. 1993. *Tratado de fonología y fonética españolas.* Gredos.

Rao, Rajiv, & Mark Amengual. 2021. La fonética y fonología del español como lengua de herencia. In D. Pascual y Cabo, & J. Torres (eds), *Aproximaciones al estudio del español como lengua de herencia*, 11–23. Routledge.

Real Academia Española (RAE). 2011. *Nueva gramática de la lengua española: Fonética fonología.* Asociación de academias de la lengua española: Espasa Libros.

Repiso-Puigdelliura, Gemma, & Ji Young Kim. 2021. The missing link in Spanish heritage trill production. *Bilingualism: Language and cognition* 24(3): 454–466.

Kong, Eun Jong, & Jan Edwards. 2011. Individual Differences in Speech Perception: Evidence from Visual Analogue Scaling and Eye-Tracking. In *ICPhS*: 1126–1129.

Rose, Marda. 2010. Differences in discriminating L2 contrasts: A comparison of Spanish taps and trills. In M.T. Prior, Y. Watanabe, & S-K. Lee (eds), *Selected Proceedings of the 2008 Second Language Research Forum*, 181–196. Somerville, MA: Cascadilla Proceedings Projects.

Scarpace, Daniel. 2014. The acquisition of the tap/trill contrast within and across words in Spanish. *Proceedings of the International Symposium on the Acquisition of Second Language: Speech Concordia Working Papers in Applied Linguistics* 5: 580–596.

Shang, Peizhu, Li, Yuejiao, & Yuhao Liang. 2021. Unraveling the contributions of prosodic patterns and individual traits on cross-linguistic perception of Spanish sentence modality. *Plos One* 16(2). https://doi.org/10.1371/journal.pone.0247098

Simonet, Miquel. 2010. Rating accented speech on continua: Nativeness in speech production in highly proficient bilinguals. In M. Ortega-Llebaría (ed.), *4th Conference on Laboratory Approaches to Spanish Phonology*, 37–46. Somerville, MA: Cascadilla Proceedings Project.

Solé, Maria-Josep. 2002. Aerodynamic characteristics of trills and phonological patterning. *Journal of Phonetics* 30(4): 655–688. https://doi.org/10.1006/jpho.2002.0176.

Solon, Megan, Park, Hae In, Henderson, Carly, & Marzieh Dehghan-Chaleshtori. 2019. Revisiting the Spanish elicited imitation task: A tool for assessing advanced language learners? *Studies in Second Language Acquisition* 41(5): 1027–1053.

Solon, Megan, Park, Hae In, Dehghan-Chaleshtori, Marzieh, Carver, Carly, & Avizia Y. Long. 2022. Exploring an elicited imitation task as a measure of heritage language proficiency. *Studies in Second Language Acquisition* 44(4): 1095–1123.

Waltmunson, Jeremy. C. 2005. *The relative degree of difficulty of L2 Spanish /d, t/, trill, and tap by L1 English speakers: Auditory and acoustic methods of defining pronunciation accuracy.* Ph.D. thesis, University of Washington.

Willis, Erik W. 2007. An acoustic study of the 'pre-aspirated trill' in narrative Cibaeño Dominican Spanish. *Journal of the International Phonetic Association* 37(1): 33–49.

Willis, Erik W., & Travis G. Bradley. 2008. Contrast maintenance of taps and trills in Dominican Spanish: Data and analysis. In L. Colantoni, & J. Steele (eds), *Selected proceedings of the 3rd Conference on Laboratory Approaches to Spanish Phonology*, 87–100. Somerville, MA: Cascadilla Proceedings Project.

Zhou, Xin, Espy-Wilson, Carol Y., Boyce, Suzanne, Tiede, Mark, Holland, Christine, & Ai Choe. 2008. A Magnetic Resonance Imaging-Based Articulatory and Acoustic Study of Retroflex and Bunched American English /r/. *The Journal of the Acoustical Society of America* 123: 4455–4481. https://doi.org/10.1121/1.2902168.