



ARTICLE

Duration as an acoustic cue of the Spanish tap-trill contrast according to age and language background

Gemma Repiso-Puigdelliura^{1,*}  and Scott James Perry²

¹Universitat Autònoma de Barcelona, Spain and ²Universitat Pompeu Fabra, Spain

*Corresponding author. Email: gemma.repiso@uab.cat

(Received 1 August 2025; revised 18 February 2026; accepted 12 March 2026)

Abstract

Previous research has shown that US Spanish heritage speakers produce the tap-trill contrast with varying manners of articulation and a non-canonical number of occlusions, while using duration as a more robust cue to the contrast. This contrast is present in Spanish, the heritage language, and absent in English, the majority language. In this paper, we examine the development of duration as an acoustic correlate of the tap-trill contrast by analyzing semi-spontaneous productions from child and adult heritage speakers as well as age-matched Spanish speakers from a monolingual environment. The durations of phonetic taps and trills were fitted to Bayesian hierarchical models to examine how these speaker groups differ with respect to their use of duration as an acoustic correlate. The results of our analysis show that the average child heritage speaker produces a smaller durational contrast with increased variability, and that some child heritage speakers show near-complete overlap in the duration of the two segments. Adult heritage speakers demonstrate a reduced difference in the tap-trill contrast when compared to Spanish speakers raised in a monolingual environment, but maintain a contrast between the two sounds. Developmental analyses within the child groups do not show significant differences in trill duration by age, but child Spanish monolinguals show a non-linear reduction in tap duration by age that does not occur for heritage children. Overall, our results support a scenario of protracted development in heritage children for the tap-trill contrast, as well as a potential sound change in progress for US Spanish.

Keywords: tap-trill contrast; Spanish heritage children; secondary cues; duration; rhotics

1. Introduction

Heritage speakers are bilinguals who are exposed to a family language at home that differs from the majority language in the community. A well-studied case is that of second-generation speakers whose parents arrive in a host country with a different language and decide to transmit their native language to their children (e.g., Benmamoun, Montrul & Polinsky 2013; Valdés 2014; Montrul 2023). Heritage speakers, in their heritage language, are shown to present quantitative and qualitative differences in several aspects of their speech system compared to speakers raised in monolingual environments (e.g., speakers residing in a monolingual environment or recent immigrants in the host culture) (Godson 2004;

© The Author(s), 2026. Published by Cambridge University Press on behalf of The International Phonetic Association. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

Willis 2005; Amengual 2012, 2016; Ronquest 2013; Colantoni et al. 2016; Elias, McKinnon & Milla-Muñoz 2017; Kissling 2018; Repiso-Puigdelliura 2021). This is thought to result from an early shift in language dominance, as heritage speakers become extensively exposed to the majority language and experience decreased exposure to the heritage language. Arguably due to this shift, the phonetic and phonological systems of heritage speakers show a wide range of variability, from monolingual-like performance to performance similar to other types of bilinguals (e.g., Pallier et al. 2003; Lukyanchenko & Gor 2011; Chang & Yao 2016; Repiso-Puigdelliura & Kim 2021; Rato & Machado 2024). The goal of the present study is to add to this body of literature by examining duration as an acoustic correlate of the tap-trill contrast in heritage Spanish.

The shift in language dominance during childhood constitutes a unique trajectory in heritage language acquisition, which demands further investigation to better understand the scenarios of child-to-adult speech system development. Some research has shown that heritage speakers sound more accented in primary school (7-to-9-year-olds) than in preschool (4-to-6-year-olds) (e.g., Kupisch et al. 2021), indicating early attrition of heritage language sounds. Other research has shown instances of protracted development for some sounds (e.g., Menke 2018; Repiso-Puigdelliura & Kim 2021) and attributed such development to limited input, given that exposure to the heritage language is often restricted to the family environment. These two possibilities are not mutually exclusive, as development is not uniform for all sounds and depends on language-specific properties (e.g., Kehoe 2002; Lleó 2002), as well as the similarity of sounds and phonotactic constraints across the languages of the bilinguals (see Bosch 2024 for discussion). In this paper, we examine the production of two sounds in both child and adult heritage speakers, one of which is more articulatorily complex (i.e., [r]) than the other (i.e., [r̥]) and absent in the sound system of the majority language.

1.1 Phonetics-phonology of Spanish tap-trill

Spanish has two rhotic sounds: an alveolar tap [ɾ] and an alveolar trill [r], which are contrastive in intervocalic position (e.g., /pero/ *pero* 'but' vs. /pero/ *perro* 'dog'), and appear in complementary distribution or free variation in other contexts (Catford 1977; Hualde 2005; Proctor 2011). From a production standpoint, the trill demands greater articulatory control than the tap. The tap is produced with a single rapid ballistic movement of the tongue tip against the alveolar ridge and a secondary uvular gesture of the tongue body (Catford 1977; Proctor 2011). The trill involves lingual trilling, realized by an initial pre-dorsal lowering and post-dorsal retraction to release built-up air pressure, allowing the tip of the tongue to vibrate against the alveolar region (Davis 1998; Recasens & Pallarès 1999; Solé 2002; Dhananjaya, Yegnanarayana & Bhaskararao 2012). As a result of articulatory constraints, Spanish trills are longer than Spanish taps. Studies examining rhotic duration have used interviews and semi-spontaneous speech (i.e., wordless picture books) (Willis & Bradley 2008; Bradley & Willis 2012), reading tasks (Amengual 2016), spontaneous speech (Perry, Kelly & Tucker 2024), or a combination of controlled, semi-spontaneous, and spontaneous speech (Kissling 2018). The average length of a tap has generally been found to lie between 20–30 ms (Willis & Bradley 2008; Bradley & Willis 2012; Amengual 2016; Perry et al. 2024), although Kissling (2018) reported an average duration of just over 40 ms. The average length of a trill generally falls within the range of 70–90 ms (Willis & Bradley 2008; Bradley & Willis 2012; Kissling 2018), with Willis & Bradley 2008 reporting means for Dominican Spanish in semi-spontaneous speech that document by-speaker variability in production.

Trills are among the latest developing sounds in Spanish (Bosch 1983; Acevedo 1993; Fabiano-Smith & Barlow 2010), likely due to their articulatory complexity, as they present

constraints in air stream properties, tongue-tip stiffness, coronal aperture, tongue body placement, and inter-gestural timing (Solé 2002; Proctor 2011). Trills have also been documented to show a wide range of phonetic variation across dialects with respect to place and manner of articulation (e.g., pre-breathy variants, assibilated variants, dorsalized variants) (Lastra & Butragueño 2006; Willis 2006; Colantoni 2006; Henriksen et al. 2022).

Compared to trills, alveolar taps exhibit a more limited range of variation, mainly in their degree of articulatory closure, with realizations ranging from taps with continuous formant structure to non-continuant taps (Bradley & Willis 2012; Perry et al. 2024). Based on spectrographic analyses, these degrees of constriction have been coded in the literature as true taps (i.e., stop closure and/or visible burst release), approximant taps (i.e., visible occlusion on the spectrogram), or perceptual taps (i.e., no visible constriction but a drop in intensity) (Bradley & Willis 2012; Amengual 2016; Kim et al. 2020; McKinnon 2023). Recently, a corpus study of Madrid Spanish by Perry et al. (2024) found that variation in the intensity drop during tap production was associated with changes in speech rate, lexical frequency, and phonetic environment.

Pressures of lexical access may also influence the phonetic realization of this contrast, as trills in word-medial position are produced with more occlusions than those in non-contrastive positions (Lastra & Butragueño 2006; Henriksen 2014). It is also possible that cues other than the number of occlusions are involved in the production and perception of the contrast. Duration has been found to be a secondary cue distinguishing taps from trills (Willis 2006; Henriksen & Willis 2010; Bradley & Willis 2012; Henriksen 2015a). In Kaqchikel-Spanish bilingual speakers, McKinnon (2023) showed that, despite variation in the production of the trill (e.g., 28.3% of tokens produced as post-approximantized taps), the duration of the phones was largely non-overlapping. Trills produced in rhotic minimal pairs (e.g., /karo/ *caro* 'expensive' vs. /karo/ *carro* 'car') were more likely to be realized with at least two occlusions. In perception, Melero-García & Cisneros (2020) found that listeners can use duration as a cue to distinguish between taps and trills, with longer segments more likely to be perceived as trills.

However, some varieties of Spanish in contact with English show overlap in the duration of taps and trills with a single occlusion (Henriksen et al., 2022). Thus, it is not clear to what extent duration is used as a cue to maintain the tap-trill contrast, particularly in cases in which primary canonical cues, such as number of occlusions, are not available to the speakers or found in that specific variety of Spanish.

1.2 Monolingual development of rhotic sounds

When it comes to monolingual phonological development, rhotics are generally acquired later than other consonants. Taps are typically acquired around four or five years of age (Bosch 1983, 2004; Camargo 2006; Torres et al. 2016; Vivar & Alarcón 2019), while trills continue to be acquired until six or seven years of age (Bosch 1983; Jimenez 1987; Acevedo 1993). During acquisition, children use substitution strategies in words containing taps and trills, replacing taps with [l], [n], or [ð], and replacing trills with the liquids [l] or the tap [r], or with the dentoalveolars [d] and [ð]. For instance, Vivar & Alarcón (2019) showed that 2;0-to-3;0-year-olds replace taps at a rate of 42.4% and trills at a rate of 60.7%.

Some studies have reported different stages of rhotic acquisition depending on their position in the word and syllable. Trills, for example, are acquired faster when they appear in intervocalic positions than in word-initial or word-final positions. Acevedo (1993) found slightly higher rates of trill production in intervocalic position (85%) compared to word-initial position (75%) at age 4, with these rates increasing to 100% and 90% respectively at age 5;6. Oropeza Escobar (2017) examined 55 Mexican Spanish-speaking children (2;0 to

6;0 years of age) and found that the acquisition of trills and liquids in general is initially restricted to word-medial contexts, where substitutions occur (e.g., [r] replaced by [l], [d], or [ð]), while word-initial or word-final positions show trill deletion.

Regarding trill duration, Carballo & Mendoza (2000) identified two trends: first, children generally reduce the duration of trills as they age; second, the shortest trills are produced by children who articulate trills with a single occlusion. It is important to note that, when acoustic or articulatory data are not used, the criteria for determining consonant accuracy may vary across studies, which affects the reported age of acquisition. For example, Torres et al. 2016 showed that the age of liquid acquisition can differ by as much as two years depending on the criteria used. For this reason, incorporating phonetic data is crucial in studies of sound development in children.

1.3 Tap and trill development in Spanish heritage speakers

The phonetic and phonological characteristics of the Spanish rhotic sounds regarding their articulatory complexity, limited positions of contrast, and a great deal of dialectal variation may render these sounds an area of vulnerability in heritage language acquisitional trajectories. Seminal research on bilingual development of taps and trills, reporting rates of consonant accuracy, showed that Spanish-English 3-to-4-year-olds demonstrate lower rates of tap and trill acquisition compared to monolingual speakers (Fabiano-Smith & Barlow 2010). More recent studies with Spanish-English bilinguals have included acoustic measures and used cross-sectional designs to examine rhotic development in heritage speakers of different ages (Menke 2018; Repiso-Puigdeliura & Kim 2021). Menke (2018) shows earlier threshold rates of target-like production for the tap than for the trill, with tap duration decreasing as children age, from a mean of 36.8 ms in 6;8-to-7;6-year-olds to 30.8 ms in 12;8-to-13;5-year-olds, and trill duration increasing with age, from a mean of 75.3 ms to 84.65 ms. This trend hints at a potentially different learning trajectory than that found in monolingual speakers (Carballo & Mendoza 2000, see section 1.2). When directly compared to adult heritage speakers, Repiso-Puigdeliura & Kim (2021) show that 9-to-10-year-old heritage speakers produce lower rates of phonetic trills (i.e., with either strong or weak occlusions; $M = 12.84\%$) with fewer occlusions ($M = 0.8$, $SD = 0.7$) and shorter duration ($M = 48.6$ ms, $SD = 25$ ms) compared to adults, whose mean rate of phonetic trill production was 43.69%, with an average of 1.39 occlusions ($SD = 0.98$) and a mean duration of 67.19 ms ($SD = 29.66$ ms).

Studies on the heritage bilingual speech system also find that heritage trills are produced with a wide range of variants, such as segments with periods of frication noise or weakened occlusions (Henriksen 2012; Amengual 2016; Kissling 2018; Kim & Repiso-Puigdeliura 2020; Repiso-Puigdeliura & Kim 2021; Colgiu et al. 2024).

The production of these non-canonical variants has been associated with factors related to bilinguals' relative dominance in the heritage and majority languages. For example, reduced dominance in Spanish has been linked to a preference for strong occlusions in taps (Amengual 2016; Kim & Repiso-Puigdeliura 2020) and a higher number of occlusions in trills (Amengual 2016). Age of acquisition (i.e., simultaneous bilinguals versus Spanish speakers who arrived in the US after the age of 12) predicts both the number of occlusions and the duration of the trill. Finally, self-identification with Latino-Hispano culture has been associated with a higher rate of canonical trill production (Kissling 2018).

Given the observed wide range of phonetic variants for rhotic production, a relevant question is whether heritage bilinguals maintain the tap-trill contrast in intervocalic position with cues other than manner of articulation. Garza, Willis & Melero-García (2024) has recently examined this question with a perceptual design and using duration as a

secondary cue. The study shows that, when presented with word-medial rhotic sounds ranging between 22 ms and 85 ms, Spanish heritage speakers (HSs) are more likely to identify segments above 55 ms as trills. In addition, the category boundary between taps and trills decreased as the age of English acquisition increased. Early bilinguals (AoA: 2–4 years) were more likely to select *tap* in segments below 55 ms relative to late bilinguals (AoA: 5+ years). When compared to Melero-García & Cisneros's (2020) findings on Spanish monolinguals, HSs require a greater duration to identify trills (60+ ms) in the absence of multiple occlusions (Garza et al. 2024), suggesting a less robust tap-trill contrast in the heritage group.

To sum up, the acquisition trajectory of Spanish rhotic sounds appears to differ between heritage bilinguals and monolingual Spanish speakers. Moreover, the mature sound systems of heritage bilinguals continue to exhibit phonetic variability influenced by factors such as language use, dominance, and identity. This suggests, in addition, that the input of the second-generation heritage speakers could be different from that of the speakers raised in monolingual environments, given that they could be exposed to an input with a higher rate of phonetic variants of the trill through the members of their community. Lastly, recent research on the perception of rhotic sounds has explored whether secondary cues help to maintain the tap-trill contrast in intervocalic position. The evidence indicates that age of exposure to the majority language modulates the extent to which duration serves as a cue to this distinction, such that later bilinguals accept segments with shorter durations as trills.

1.4 The present study

In the present study, we investigate how US Spanish heritage speakers employ duration as an acoustic correlate of the tap-trill contrast in intervocalic position, where the sounds are contrastive. Specifically, we are interested in second generation Spanish speakers whose parents migrated from a predominantly Spanish speaking society, and who therefore are likely exposed to Spanish input containing the contrast. The first goal of this paper is to see whether Spanish heritage speakers use duration as a cue to robustly maintain the tap-trill contrast in speech production and compare them to age-matched children raised in monolingual environments. Our second goal is to compare the productions of child and adult heritage speakers to see whether the development of the tap-trill contrast shows protracted development, with child heritage speakers producing a weaker contrast than adult heritage speakers, or attrition, with adult heritage speakers showing a less robust contrast than children. Our third research goal is to examine whether the development of duration as a cue to the tap and trill contrast throughout childhood presents differences between heritage and monolingual children.

2. Methods

2.1 Participants

A total of 102 speakers participated in the semi-spontaneous speech task. The Spanish HSs group consisted of 31 child participants (14 F, 17 M, mean age = 9;1 years, SD = 23 months, range = 5;3–11;11 years) and 16 adult participants (13 F, 3 M, mean age = 20.3 years, SD = 28 months). The Spanish HSs were born in the US or arrived in the US before the age of 3. With regard to their dialectal background, participants had at least one caregiver whose country of origin was Mexico. As for the age of acquisition of the two languages, HSs were exposed to Spanish since birth and to English before the age of five. The adult heritage

speakers were university students at the University of California, Los Angeles. These participants were recruited from courses at the Spanish and Portuguese Department. The child Spanish heritage speakers were sampled from three Los Angeles primary schools that provide Spanish immersion instruction. Child heritage speakers were recruited through flyers distributed to parents.

The control group (i.e., SpanMonoS) consisted of Spanish speakers raised in Mexico who resided in Northern or Central Mexico at the time of testing. The group included 34 child speakers (12 F, 22 M; mean age = 8;5, SD = 24.3 months; age range = 5;0–11;8) and 21 adults (15 F, 6 M; mean age = 20.8 years, SD = 1.91 years). All SpanMonoSs were exposed to Spanish at home and to both Spanish and English at school; however, only participants with at least 70% exposure to Spanish were included. These participants were recruited through a snowball sampling approach, whereby existing participants referred to additional individuals. Given that the study was conducted at the outbreak of the COVID-19 pandemic, data collection began in person and later continued online. The in-person recordings were conducted in a quiet space at the primary schools or in a sound-attenuated room for the adult heritage speakers, using an AKG-520 head-mounted microphone and Zoom H4n handy portable digital recorder with a sampling rate of 44.1 kHz and 16-bit quantization. For the online portion of the study, recordings were made with the participants' smartphone microphones using the phone app Shure MOTIVTM, positioned approximately 4 inches from the mouth with a sampling rate of 44.1 kHz and 16-bit quantization. Our duration measurements were not affected by differences in recording quality between the in-person and online sessions.

2.2 Task and segmentation

Semi-spontaneous speech for this study was elicited using the wordless book *Frog Where Are You?* (Mayer 1969). This task elicits taps (e.g., *ahora* 'now', *era* 'was', *vieron* 'they saw') and trills (e.g., *jarra* 'jar', *perro* 'dog') in comparable environments across participants. In addition, the *frog story* has been used in previous studies to elicit naturalistic rhotics in Spanish (e.g., Henriksen 2015a; Repiso-Puigdelliura & Kim 2021). The resulting speech samples were forced aligned using Montreal Forced Aligner (McAuliffe et al. 2017). Boundaries for all instances of intervocalic phonemic taps and trills were manually adjusted.

Based on the manner of articulation as it shows in the spectrogram we classified taps and trills as true tap (i.e., stop closure and burst release), approximant tap (i.e., visible drop in intensity but with a continuation of the formant structure), perceptual tap (i.e., perceptual rhotic sound with intensity drop but no visible presence of a constriction in the spectrogram), true trill (i.e., segment with two or more occlusions), approximant trill (i.e., segment with two or more occlusions presenting continuation of the formant structure), fricativized rhotic (i.e., frication noise in the spectrogram and no visible occlusion), or other (e.g., lateral, retroflex, glides, voiced stops) (see Figure 1). We manually adjusted the boundaries at the onset and offset of the visible occlusion (i.e., single occlusion for taps and multiple occlusions for trills). An occlusion was defined as a disruption in the formant structure between flanking vowels, together with an intensity dip, occurring either between two vowels or between a vowel and an aperture (e.g., see Figure 1a). For tokens that showed a continuous formant structure, the boundaries were aligned at the point where the energy visibly weakened and the intensity decreased between the consonant and the surrounding vowels (e.g., see Figure 1b). For tokens containing frication noise, boundaries were placed at the onset and offset of the aperiodic signal in the spectrogram (e.g., see Figure 1c). Both true and approximant taps and trills were classified as phonetic taps and trills and included in the analysis. Perceptual taps and fricative realizations were excluded.

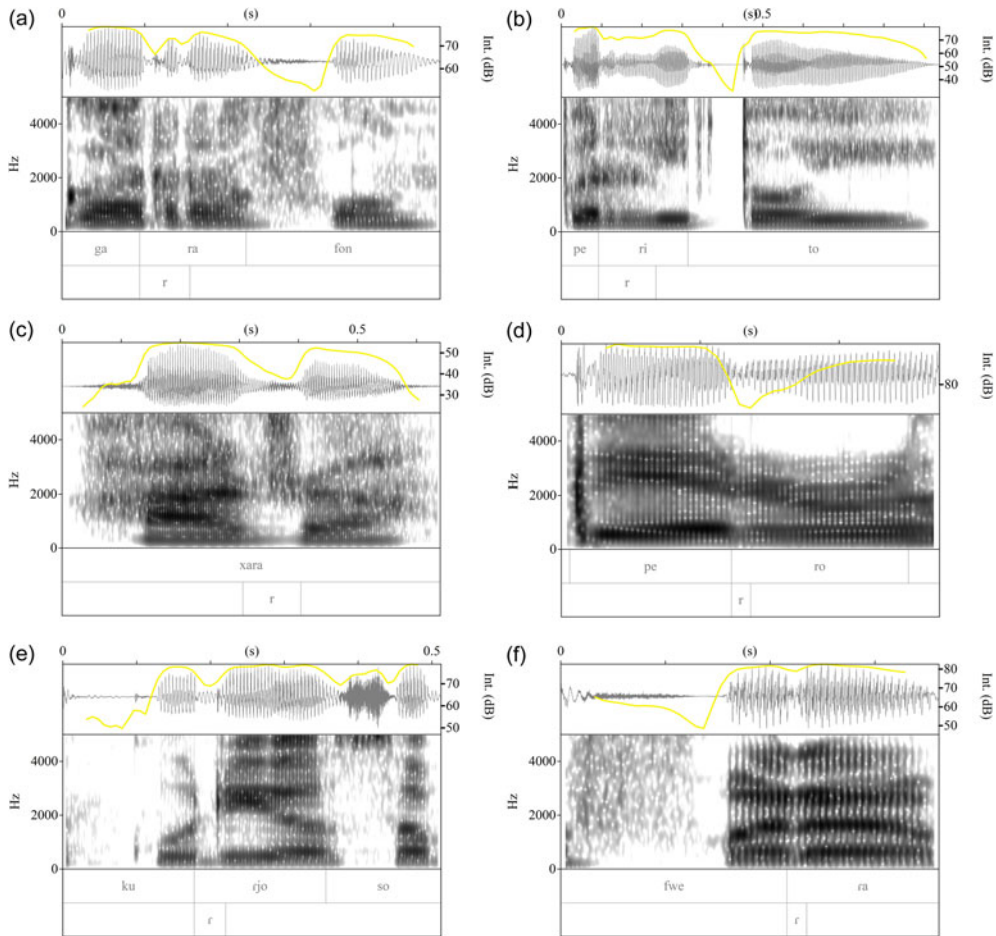


Figure 1. Examples of phonemic trills produced as true trill, approximant trill, trill with friction noise, and tap. The first tier contains an IPA transcription of the word and the second tier indicates the segmentation of the rhotic sounds. (a) True trill in *garrafón* 'jug'. (b) Approximant trill in *perrito* 'doggy'. (c) Fricative in *jarra* 'jar'. (d) Approximant tap (i.e., phonemic trill) in *perra* 'dog'. (e) True tap in *curioso* 'curious'. (f) Approximant tap in *fuera* 'outside'.

2.3 Data analysis

In mapping our first two research objectives to a statistical model, we must define what we mean by 'robustly' maintaining a contrast. In our models, we thought of two distinct ways in which we could define a phonological contrast being maintained. First is the size of the length difference between taps and trills (i.e., how much longer the average trill is compared to the average tap), and second is how variable the durations of taps and trills are in a phonological context where they can be contrastive.

The first model we built was a Bayesian hierarchical model highly similar to a linear mixed-effects model and thus was only able to predict changes in the mean. However, we had evidence from previous research on heritage speakers, as well as general research into bilingualism, that supported the idea that there might also be differences in the variance of duration, not just the means. Our first model assumed a constant variance for all ages, segments, and speaker groups, which is a statistical assumption that does not map well to

the reality of our data. We therefore fit a second model that relaxed this assumption. The model of the mean duration was the same in this second model, but instead of having a constant variance, we modelled the variance by the same three-way interaction that modelled changes in the mean and allowed the variance to vary by speaker. We compared the predictive performance of these two models using the Leave-one-out information criterion. Both of these models were fit using the package *brms* (Bürkner 2017) in R (R Core Team 2021) and used a lognormal likelihood predicting segment duration in milliseconds by the interaction between three sum-coded binary variables: age group (children vs. adults), speaker type (heritage vs. monolingual), and segment (tap vs. trill). For predicting the mean duration, we controlled for speech rate as measured in syllables per second, so that comparisons between groups could be made even though generally adults speak faster than children. We had varying intercepts for participants for both the mean and variance and a varying effect of segment by participant in predicting mean duration. All code and data used to fit these models is detailed in the supplementary materials, along with additional details about our weakly-informative priors which were chosen through prior predictive simulation.

Our third research question asked a question about developmental trajectories within the child groups. We therefore fit a generalized additive mixed model using package *mgcv* (Wood 2020) to allow for the fitting of non-linear developmental trends. This model predicted log-transformed duration by a separate smooth term for age for each of the four possible combinations coming out of speaker group (heritage vs. monolingual) and segment (tap vs. trill) as well as a separate smooth term for speech rate. We also included a random effect of participant. As with the previous models, all data and code used to fit and interpret the model are included in the supplementary file.

3. Results

Here we report a summary of the findings from our modelling of the data. The data and code used to fit the models, additional comparisons and visualizations, the saved model files, and additional information about our prior predictive simulations are located in the supplementary materials. All predictions and comparisons were made using the package *marginalEffects* (Arel-Bundock, Greifer & Heiss 2024). For plotting, we used *ggplot2* (Wickham 2016), *ggdist* (Kay 2024), *patchwork* (Pedersen 2025), and *ggpattern* (FC & Davis 2022).

In our analyses, we only included tokens that were coded as phonetic taps or trills. We also excluded taps without a measurable duration (e.g., excluding perceptual taps and elided segments as in Perry, Kelly & Tucker (2023)). In line with previous research (e.g., Henriksen 2015b; Kehoe and Girardier 2020; Repiso-Puigdelliura and Kim 2021), we also found non-canonical variants, such as fricatives (both heritage and monolingual speakers), lateral and retroflex segments (only child heritage speakers). These productions that were not analyzed constituted less than 7% of tokens in an intervocalic environment. Information regarding these segments is available in the supplementary materials.

3.1 Age group analysis

Here we review the results from the analysis comparing child and adult productions of taps and trills. Before interpreting the model, we report the comparison between the model with a constant variance to the model which allowed the variability of productions to change according to our variables of interest (known as a location-scale model). The predictive power of our location-scale model was an improvement over our model of means (elpd diff = -125.0, se diff = 21.4), indicating a high probability (>99.99%) that dropping

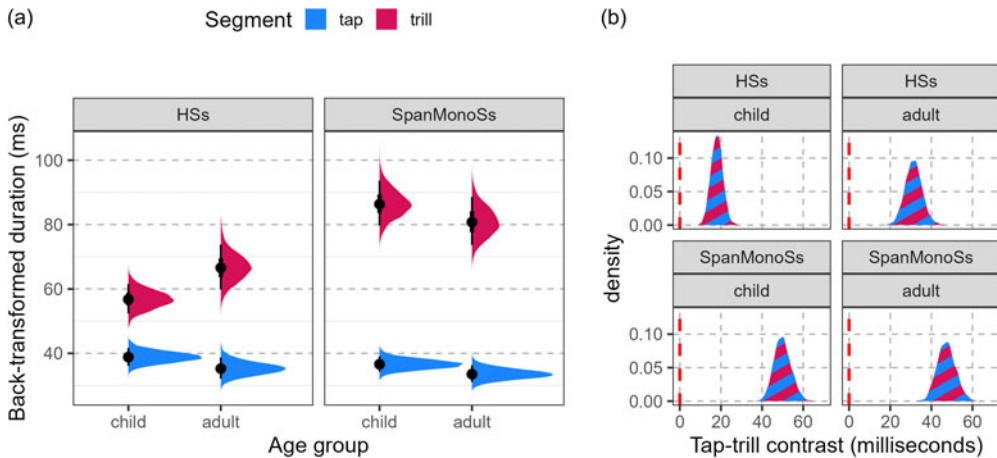


Figure 2. (a) Posterior distributions for the model's predicted duration, in milliseconds, for the two segments according to age and speaker group for a theoretical average participant. The steps on the whiskers indicate the 50% and 89% credible intervals. (b) The posterior distributions for the difference in milliseconds between the tap and trill according to age and speaker groups for a theoretical average participant. The red dashed line at zero indicates the point on the x-axis where taps and trills would be the same length.

the assumption of constant variance improved the out-of-sample predictive power of our model. We therefore report only our location-scale model here, although the reader will find the constant variance model reported in the supplementary file.

We will start by interpreting differences in mean duration according to our three variables of interest. Figure 2a visualizes the predicted durations of taps and trills by age and speaker group. However, our first research question wasn't concerned with specific duration values, but instead with how the duration is employed as an acoustic correlate of the tap-trill contrast. To aid the reader in comparing the size of the tap-trill contrast across groups, Figure 2b plots the posterior distributions for the tap-trill contrast, with the dashed red line indicating the place on the x-axis where the two sounds would be the same length. Our interpretation of the model is that an average monolingual produces trills around 40–60 ms longer than the tap (95% CI is 42.3–58.1 ms for children and 39.1–56.6 ms for adults). This is a longer difference in duration than an average heritage speaker adult (95% CI = 23.6–39.9 ms), who in turn has a larger difference between taps and trills than the average child heritage speaker (95% CI = 12.3–23.2 ms).

In addition to interpreting changes in the mean, our location-scale model allows us to look at changes in how variable productions of the two segments are in our age and speaker groups. Figure 3a visualizes the predicted scale parameter of our model according to our three variables of interest, with higher values meaning a wider range of duration values being modeled for that group. Some general patterns emerge. First, the average heritage speaker child produces trills more variably than taps (95% CI = 0.0135–0.0892), while the two segments are more similar for all other speaker groups, with strong evidence that the contrast is different for heritage children compared to adult monolinguals (95% CI = 0.023–0.119) and somewhat weaker evidence that this pattern is more extreme compared to heritage speaker adults (95% CI = –0.008–0.0937) and monolingual children (95% CI = –0.014–0.0937). The consequences of these differences in variability around the mean can be seen in Figure 3 B, where we output predicted values of taps and trills from the model. These are not predictions of mean values but incorporate the sigma variable to generate the entire distribution of segment durations. So, whereas the tap-trill contrast in

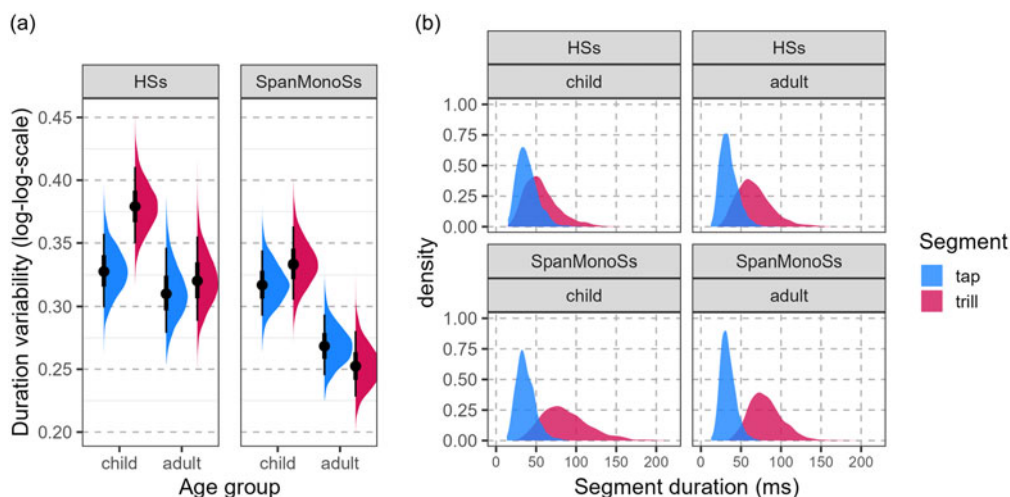


Figure 3. (a) Posterior distributions for the sigma parameter for a theoretical average participant according to segment type, age group, and speaker group. This sigma parameter is how variable durations were around the means shown in Figure 2a. (b) The entire predictive distribution for tap and trill durations according to age and speaker group.

average durations was similar for monolingual children and adults, the increased variability of child productions leads to increased overlap between the tap and trill distributions in the lower left panel compared to the lower right panel. We can also see that an average child heritage speaker has a large degree of overlap in the durations of their taps and trills.

So far, we have been following the standard practice of reporting the conditional effects from the hierarchical model, i.e., the effect for an average participant. This puts aside the variation taken into account by the group-level (random) effects. However, our model contained a varying slope for the tap-trill contrast by participant, allowing us to model the possibility that not all speakers in our data implement this contrast in the exact same way. Figure 4 visualizes the distributions of individual counterfactual comparisons between the tap and trill for each row in our data, taking into account the nature of how each participant instantiates the contrast. We can see that all of the comparisons for monolingual speakers are positive, meaning that trills are always predicted to be at least 25 ms longer than taps. Much different to this is the predicted contrasts for the heritage group. Predictions for adults are much closer to zero, and many predictions for children fall in on the other side of zero. Thus, while an average child heritage speaker might employ duration as an acoustic correlate of this phonemic contrast in production, there are some who do not. This pattern can also be seen in the raw data, plotted for the readers' convenience in the supplementary material.

3.2 Development analysis

Here we review the main findings of the non-linear regression model that was fit to explore how tap and trill productions changed according to children's ages. The estimated average duration for an average child according to age in months is visualized in Figure 5a. The overall differences between the tap-trill contrast and the length of the trill are similar to the analysis in the previous section, so we will not reiterate that information here. We will discuss the differences (or lack thereof) according to the ages of children within our sample.

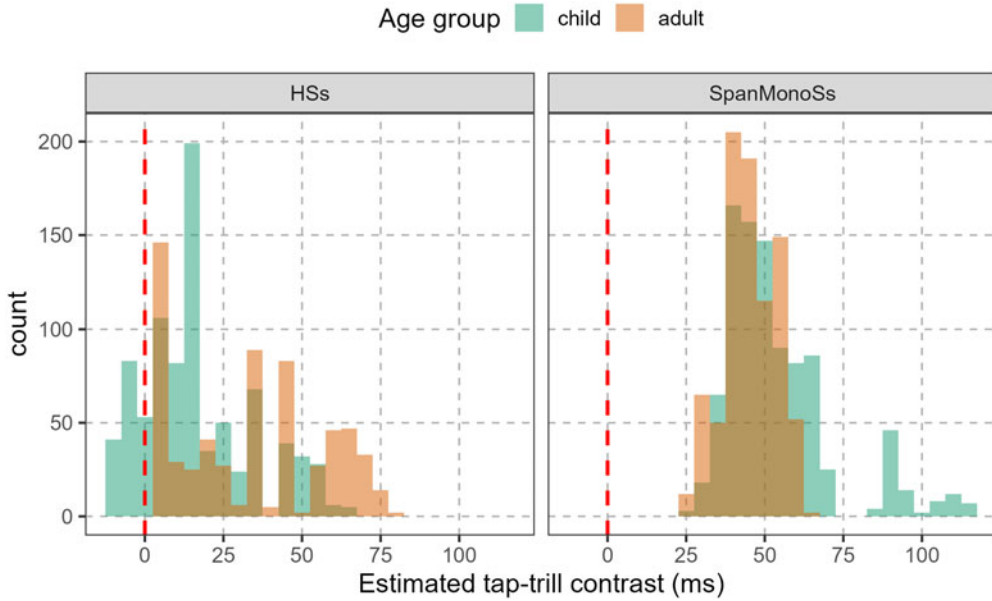


Figure 4. A histogram of the counterfactual comparisons made from our model for each row in our original data frame, split by speaker group and age group. The dashed red line at zero indicates the point at which the model predicts there to be no difference in the duration of the segment if its identity was changed (i.e., if it was a tap instead of a trill or vice versa.)

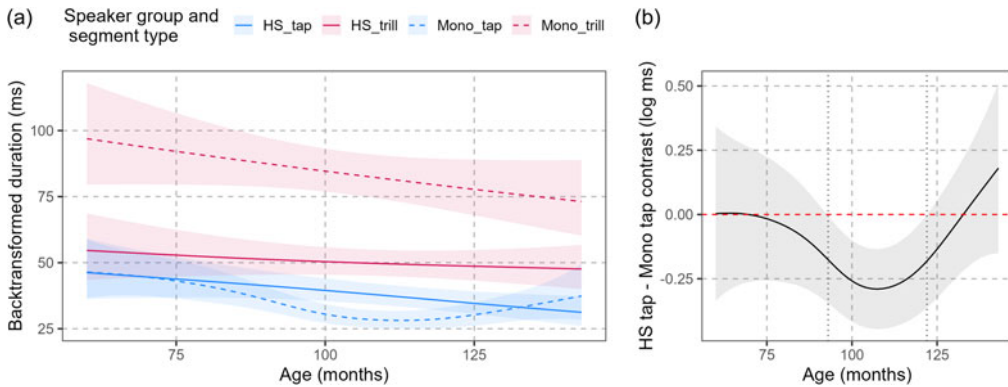


Figure 5. (a) Predicted segment duration in milliseconds for an average participant based on the segment (tap or trill), speaker group (HS or Mono) as well as the child's age in months. (b) The difference in duration (log ms) between a tap produced by an average heritage child vs a tap produced by an average monolingual child. Horizontal dashed line in red is where the two segments are equal in length. Vertical dotted lines in grey at 93 and 122 months indicates approximate start and stop of area where 95% CI does not include zero.

The first thing we point out is that the smooth terms for the trill were not statistically significant for either the heritage group ($p=0.538$) or the monolingual group ($p=0.075$). As this corresponds to an absence of evidence, we will not be interpreting how the length of the trill may vary as a function of age. We do see, however, that the smooth term for the tap is predicted by age for both heritage speakers ($p=0.008$) and monolinguals ($p<0.001$). The pattern predicted for heritage speakers is a linear decrease in tap duration, whereas

monolingual speakers show a non-linear pattern. An average monolingual and heritage speaker child is not significantly different from one another at the youngest or oldest ages observed, but differ between approximately 93 and 122 months as shown in Figure 5b.

4. Discussion

Research on the Spanish rhotic system of heritage speakers has shown that these bilinguals produce trills varying in manner of articulation and number of occlusions (Henriksen 2015b; Amengual 2016; Kissling 2018; Cummings Ruiz & Montrul 2020; Repiso-Puigdelliura & Kim 2021) and taps with varying degrees of constriction (Henriksen 2015b; Amengual 2016; Kim & Repiso-Puigdelliura 2020). The present study investigates how duration, as a secondary cue in the rhotic system, develops in Spanish heritage speakers to support the maintenance of the tap-trill contrast in semi-spontaneous speech. To do so, we examined tap-trill production in a population of child and adult heritage speakers and compared it to that of age-matched Spanish speakers raised in a monolingual environment. Along with comparisons of mean durations, in this study, we modeled variation in segment duration using a location-scale model. We interpret variability in the production of rhotic duration as the degree of articulatory stability of the tap and the trill in the sound system.

We first asked whether adult Spanish heritage speakers and adult Spanish speakers raised in a monolingual environment use duration as an acoustic correlate of the tap-trill contrast to the same extent. We found that, compared to adult monolingual Spanish speakers, adult heritage speakers demonstrate a reduced tap-trill contrast with higher variability, but present largely non-overlapping durations between the two segments. The reduced durational contrast arises from a shorter trill rather than a longer tap. Our results align with previous studies showing shorter trills for heritage bilinguals compared to long-term immigrants or native speakers (Bradley & Willis 2012; Henriksen 2015b; Kissling 2018; Cummings Ruiz & Montrul 2020; Colgiu et al. 2024). We show that, despite differences between groups, heritage speakers maintain the tap-trill contrast using duration as a secondary cue (Amengual 2016; Cummings Ruiz & Montrul 2020). Our study expands on the body of research demonstrating that, even though heritage language learning trajectories affect the specific nature of the contrast, the majority of adult Spanish heritage speakers continue to use duration as an acoustic correlate of the tap-trill contrast.

Our next question asked whether the development of the tap-trill contrast demonstrates a scenario of protracted development, in which child heritage speakers show a weaker contrast than the adult heritage speakers, or a scenario of attrition, where the adult heritage speakers show a weaker tap-trill contrast than the child heritage speakers as a result of shifting language dominance. Our results show child heritage speakers exhibiting a smaller and more variable tap-trill contrast compared to our adult heritage speakers; a pattern not found in our monolingual data. In addition, an average child heritage speaker shows a large overlap between the tap and the trill, and a non-aggregated analysis shows that many child heritage speakers are predicted to produce taps and trills with nearly identical durations. Therefore, similar to previous results (Menke 2018; Repiso-Puigdelliura & Kim 2021), our findings do not support a hypothesis of attrition in which child heritage speakers produce a robust tap-trill contrast and then fail to maintain it as a result of a shift in language dominance during childhood and adolescence. Instead, our results indicate that, in particular, the trill segment in the bilingual population of US heritage bilinguals develops at a later age than that in homeland speakers. This highlights the importance of considering sound-specific properties that may impact development in reduced input situations (e.g., as discussed by Bosch 2024). Tap production was more similar across groups, while trill production was more variable, with the latter being a late-developing sound with unique

articulatory demands (Acevedo 1993; Carballo & Mendoza 2000; Fabiano-Smith & Barlow 2010). Our data indicate that child heritage speakers often produce phonemic trills with a single occlusion, which could be due to failing to trill or to attempting a target tap, as while duration could be employed to distinguish the two sounds in the presence of a single occlusion this does not appear to be the case in our data. Visualizations in the supplementary materials show virtually identical distributions of duration for phonemic taps and trills when both have a single occlusion.

When compared to Spanish speakers raised in a monolingual environment, Spanish heritage speakers are exposed to less input that is more variable and have fewer opportunities to produce speech in Spanish. In addition, a pattern of asymmetrical input-output interactions, in which child heritage speakers are addressed in Spanish but continue the communication in English may impact these opportunities to practice trilling, which, in turn, could impede the refinement of motor control skills associated with trill production. The difference in the development of tap duration between monolingual and heritage children could be an instance of a similar situation in which monolingual children shorten taps at a faster pace due to increased speech output. Due to the relative simplicity of this segment, the differences are minimal for the older children in our sample, but the increased complexity of the trill means that the production of this segment is not mastered until later. This is supported by our preliminary evidence that heritage children show increased variability in trill production compared to the other groups.

In addition to articulatory complexity and reduced Spanish input and output, the tap-trill contrast likely has a low functional load, as it is contrastive only in intervocalic position and creates just 30 minimal pairs in this phonetic context (Willis & Bradley 2008). Neutralizing this contrast is therefore unlikely to lead to communication breakdowns, as top-down information can disambiguate the intended message (e.g., Daidone & Darcy 2014 find large differences in frequency between words in such minimal pairs). A potential scenario can be envisioned in which phonemic trills are initially produced as taps due to articulatory constraints. During this stage, a lack of error signal in communication could delay phonetic category updating and promote neutralization of the contrast for some speakers under specific conditions. In this scenario, phonetic variants of the trill could also gain traction as sociolinguistic variants in US Spanish, aligning with findings of low rates of trilling in US Spanish, such as Albuquerque (Lease & Marchesi 2022), Taos (Vigil 2018), and Los Angeles (Repiso-Puigdeliura & Kim 2021). We suggest that exploring potential interactions between language-level properties, such as functional load, and variation in the quantity and quality of input could help explain some of the heterogeneity of speech production in heritage language populations.

With respect to child to adult monolingual differences, we do not observe the same trend as in the heritage speakers. When comparing our group of child monolinguals to our group of adult monolinguals, we found greater variability in the production of the two sounds in the children than in the adults. However, our results do not show that child Spanish monolinguals produce a reduced tap-trill contrast, as in the case of the child heritage speakers. At the ages we collected in our sample, it appears that child monolingual productions are approximating their adult target on average, but that children are more variable as they refine their motor skills related to speech production.

Given that our design is cross-sectional, we cannot rule out the alternative explanation that the differences between adult and child heritage speakers are due to divergencies in the nature of the input received. A longitudinal design would be necessary to confirm that these differences reflect protracted development. Quality of input is a relevant factor when examining children with dual input, as these speakers are exposed to language communities influenced by the majority language (Stoehr et al. 2019; Brehmer & Treffers-Daller 2020; Bosch 2024). Recent studies have also found a relationship between the type of input and

the acquisition of sounds in bilingual contexts (Stoehr et al. 2019; Ramon-Casas et al. 2023). Thus, an open possibility is that the weakened tap-trill contrast observed in child heritage speakers, relative to adult heritage speakers, reflects an ongoing change in the rhotic system of US Spanish. That is, child heritage speakers may have been exposed to input with a less stable tap-trill contrast. For this reason, recording child-caregiver interactions could provide insight into the types and frequencies of phonetic variants to which children are exposed.

5. Conclusion

Previous research has shown that the production of the tap-trill contrast by Spanish heritage speakers demonstrates variability with respect to the manner of rhotic production and the number of occlusions. In this study, we examined duration as an acoustic correlate of the tap-trill contrast in child and adult US Spanish heritage bilinguals and we compared the productions of heritage speakers to those of age-matched Spanish speakers raised in a monolingual environment. Our study contributes to previous literature by modeling development of duration as an acoustic correlate of the tap-trill contrast in Spanish heritage bilinguals.

Our results show that the child heritage speakers produce the smallest tap-trill contrast among the four groups of speakers, with some speakers showing complete overlap between the tap and the trill. Adult heritage speakers produce a larger contrast than child heritage speakers, with both groups of heritage speakers realizing a smaller contrast than monolinguals. Despite producing shorter trills than the Spanish monolinguals, the majority of adult heritage speakers maintain a tap-trill contrast using duration. These results support a scenario of protracted development by which the child heritage speakers continue to acquire the Spanish trill beyond the age of monolingual trill development. In addition, the child heritage speakers show the greatest variability in trill duration, which we attribute to the articulatory complexity of the trill and decreased language use. As such, this result of protracted development may not be generalizable to sounds that show less complex articulation and decreased variability in the input.

Acknowledgements This work was supported by the US National Science Foundation [BCS-2116801,2021] and partially supported by the Social Sciences and Humanities Research Council of Canada in the form of a postdoctoral fellowship held by the second author.

References

- Acevedo, Mary Ann. 1993. Development of Spanish consonants in preschool children. *Journal of Childhood Communication Disorders* 15(2), 9–15. <https://doi.org/10.1177/152574019301500202>.
- Amengual, Mark. 2012. Interlingual influence in bilingual speech: Cognate status effect in a continuum of bilingualism. *Bilingualism: Language and Cognition* 15(3), 517–530.
- 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>.
- Arel-Bundock, Vincent, Noah Greifer & Andrew Heiss. 2024. How to interpret statistical models using marginal effects for R and Python. *Journal of Statistical Software*, 111(9), 1–32. doi:10.18637/jss.v111.i09.
- Benmamoun, Elabbas, Silvina Montrul & Maria Polinsky. 2013. Defining an “ideal” heritage speaker: Theoretical and methodological challenges Reply to peer commentaries. *Theoretical Linguistics* 39(3–4), 259–294. <https://doi.org/10.1515/tl-2013-0018>
- Bosch, Laura. 1983. El desarrollo fonológico infantil: Una prueba para su evaluación. *Anuario de Psicología* 28, 87–114.
- Bosch, Laura. 2004. *Evaluación fonológica del habla infantil*. Spain: Elsevier Masson.
- Bosch, Laura. 2024. Acquisition of segmental phonology in child bilingualism. In Mark Amengual (ed.), *The Cambridge handbook of bilingual phonetics and phonology*, 339–358. Cambridge University Press. <https://doi.org/10.1017/9781009105767.016>.

- Bradley, Travis G. & Erik Willis. 2012. Rhotic variation and contrast in Veracruz Mexican Spanish. *Estudios de fonética experimental* 21, 43–74. <https://www.raco.cat/index.php/EFE/article/view/260320>.
- Brehmer, Bernhard & Jeanine Treffers-Daller. 2020. Two sides of one coin? In Bernhard Brehmer & Jeanine Treffers-Daller (eds.), *Lost in transmission: The role of attrition and input in heritage language development*, 1–14. John Benjamins. <https://benjamins.com/catalog/sibil.59.int>
- Bürkner, Paul-Christian. 2017. brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software* 80(1), 1–28. <https://doi.org/10.18637/jss.v080.i01>
- Camargo, Maryluz. 2006. Desarrollo fonético-fonológico del español en niños. *Umbral Científico* (9).
- Carballo, Gloria & Elvira Mendoza. 2000. Acoustic characteristics of trill productions by groups of Spanish children. *Clinical Linguistics and Phonetics* 14(8), 587–601. <https://doi.org/10.1080/026992000750048125>
- Catford, John Cunnison. 1977. *Fundamental problems in phonetics*. Bloomington and London: Indiana University Press.
- Chang, Charles B. & Yao Yao. 2016. Toward an understanding of heritage prosody: Acoustic and perceptual properties of tone produced by heritage, native, and second language speakers of Mandarin. *Heritage Language Journal* 13(2), 134–160. <https://doi.org/10.46538/hlj.13.2.4>
- Colantoni, Laura. 2006. Micro and macro sound variation and change in Argentine Spanish. In Núria Sagarra & Almeida Jacqueline Toribio (eds.), *Selected proceedings of the 9th Hispanic Linguistics Symposium*, 91–102. <http://www.lingref.com/cpp/hls/9/paper1369.pdf>
- Colantoni, Laura, Alejandro Cuza & Natalia Mazzaro. 2016. Task-related effects in the prosody of Spanish heritage speakers and long-term immigrants. In Meghan E. Armstrong, Maria del Mar Vanrell & Nicholas Henriksen (eds.), *Intonational grammar in Ibero-Romance: Approaches across linguistic subfields*, 6th edn. 1–24. Amsterdam/Philadelphia: John Benjamins <https://doi.org/10.1075/ihll.6.01col>
- Colgiu, Ioana & Laura Spinu Rajiv Rao & Yasaman Rafat. 2024. Bilingual rhotic production patterns: A generational comparison of Spanish-English bilingual speakers in Canada. *Proc. Interspeech 2024*, 1025–1029. 1 <https://doi.org/10.21437/Interspeech.2024-2287>
- Cummings Ruiz, Laura D. & Silvina Montrul. 2020. Assessing rhotic production by bilingual Spanish speakers. *Languages* 5(4), 51. <https://doi.org/10.3390/languages5040051>.
- Daidone, Danielle & Isabelle Darcy. 2014. Quiero comprar una guitarra: Lexical encoding of the tap and trill by L2 learners of Spanish. In *Selected proceedings of the 2012 Second Language Research Forum: Building bridges between disciplines*, 39–50. Somerville, MA: Cascadia Proceedings Project.
- Davis, Katharine. 1998. Review of Peter Ladefoged & Ian Maddieson, *The sounds of the world's languages*, Blackwell, 1996. *Journal of Linguistics* 34(1). <https://doi.org/10.1017/S002226797276899>
- De Jong, Kenneth. 1998. Stress-related variation in the articulation of coda alveolar stops: Flapping revisited. *Journal of Phonetics* 26(3), 283–310. <https://doi.org/10.1006/jpho.1998.0077>
- Dhananjaya, N., B. Yegnanarayana & Peri Bhaskararao. 2012. Acoustic analysis of trill sounds. *The Journal of the Acoustical Society of America* 131(4), 3141–3152. <https://doi.org/10.1121/1.3688470>
- Eddington, David & Dirk Elzinga. 2008. The phonetic context of American English flapping: Quantitative evidence. *Language and Speech* 51(3), 245–256. <https://doi.org/10.1177/0023830908098542>
- Elias, Vanessa, Sean McKinnon & Ángel Milla-Muñoz. 2017. The effects of code-switching and lexical stress on vowel quality and duration of heritage speakers of Spanish. *Languages* 2(4), 29. <https://doi.org/10.3390/languages2040029>
- Fabiano-Smith, Leah & Jessica Barlow. 2010. Interaction in bilingual phonological acquisition: evidence from phonetic inventories. *International Journal of Bilingual Education and Bilingualism* 13(1), 81–97. <https://www.tandfonline.com/doi/pdf/10.1080/13670050902783528>
- FC, Mike & Trevor L. Davis. 2022. *ggpattern: 'ggplot2' pattern geoms* [Computer software]. <https://cran.r-project.org/web/packages/ggpattern/index.html>.
- Garza, Rachel S., Erik W. Willis & Fernando Melero-García. 2024. More than occlusions: The role of duration in perceiving the Spanish tap-trill contrast by heritage speakers of Spanish. In Michael Gradoville & Sean McKinnon (eds.), *Recent developments in Hispanic linguistics: Studies in structure, variation, and bilingualism*, 122–150. John Benjamins.
- Godson, Linda. 2004. Vowel production in the speech of Western Armenian heritage speakers. *Heritage Language Journal* 2(1), 44–69. <https://www.international.ucla.edu/ccs/article/14648>
- Henriksen, Nicholas. 2012. Chicagoland heritage and native Mexican Spanish intonation: Three contact phenomena. In *Current approaches to Spanish and Portuguese second language phonology*. Columbia, SC: University of South Carolina.
- Henriksen, Nicholas. 2014. Sociophonetic analysis of phonemic trill variation in two sub-varieties of Peninsular Spanish. *Disaster Medicine and Public Health Preparedness* 2(1), 4–24. <https://doi.org/10.1017/jlg.2014.1>

- Henriksen, Nicholas. 2015a. Acoustic analysis of the rhotic contrast in Chicagoland Spanish. *Linguistic Approaches to Bilingualism* 5(3), 285–321. <https://doi.org/10.1075/lab.5.3.01hen>
- Henriksen, Nicholas C. 2015b. 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>
- Henriksen, Nicholas & Shayna Greenley & Amber Galvano. 2022. Sociophonetic investigation of the Spanish alveolar trill /r/ in two canonical-trill varieties. *Language and Speech* 66(4), 836–934. <https://doi.org/10.1177/00238309221137326>
- Henriksen, Nicholas & Erik W. Willis. 2010. Acoustic characterization of phonemic trill production in Jerezano Andalusian Spanish. In *Selected proceedings of the 4th Conference on Laboratory Approaches to Spanish Phonology*, 115–127. Somerville, MA: Cascadia Proceedings.
- Hualde, José Ignacio. 2005. *The sounds of Spanish*. Cambridge: Cambridge University Press.
- Jimenez, Beatrice C. 1987. Acquisition of Spanish consonants in children aged 3–5 years, 7 months. *Language, Speech, and Hearing Services in Schools* 18(4), 357–363. <https://pubs.asha.org/doi/10.1044/0161-1461.1804.357>
- Kay, Matthew. 2024. ggdist: Visualizations of distributions and uncertainty in the grammar of graphics. *IEEE Transactions on Visualization and Computer Graphics* 30(1), 414–424. <https://doi.org/10.1109/TVCG.2023.3327195>
- Kehoe, Margaret. 2002. Developing vowel systems as a window to bilingual phonology. *International Journal of Bilingualism* 6(3), 315–334. <https://doi.org/10.1177/13670069020060030601>
- Kehoe, Margaret & Chloé Girardier. 2020. What factors influence phonological production in French-speaking bilingual children, aged three to six years? *Journal of Child Language* 47(5), 945–981. <https://doi.org/10.1017/S0305000919000874>
- Kim, Ji Young, Matthew Faytak, Gemma Repiso Puigdeliura & Erin Mauffray. 2020. Articulation of non-normative variants of L1 and L2 Spanish trill /r/. *The Journal of the Acoustical Society of America* 148(4). <https://doi.org/10.1121/1.5147165>
- Kim, Ji Young & Gemma Repiso Puigdeliura. 2020. Deconstructing heritage language dominance: Effects of proficiency, use, and input on heritage speakers' production of the Spanish alveolar tap. *Phonetica* 77(1), 55–80. <https://doi.org/10.1159/000501188>
- 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>
- Klein, Harriet B. & Elaine K. Altman. 2002. The acquisition of medial /t, d/ allophones in bisyllabic contexts. *Clinical Linguistics and Phonetics* 16(3). <https://doi.org/10.1080/02699200110112592>
- Kupisch, Tanja, Nadine Kolb, Yulia Rodina & Olga Urek. 2021. Foreign accent in pre- and primary school heritage bilinguals. *Languages* 6(2), 96. <https://doi.org/10.3390/LANGUAGES6020096>
- Lastra, Yolandra & Pedro Martí Butragueño. 2006. Un posible cambio en curso: el caso de las vibrantes en la ciudad de México. In Ana María Cestero Mancera, Isabel Molina Martos & Florentino Paredes García (eds.), *Estudios sociolingüísticos del español de España y América*, 35–68. Madrid: Arco Libros.
- Lease, Sarah & Mariana Marchesi. 2022. A sociophonetic approach to the acquisition of Spanish rhotics in a bilingual community. *Proceedings of the Linguistic Society of America* 7(1), 5231. <https://doi.org/10.3765/plsa.v7i1.5231>
- Lleó, Conxita. 2002. The role of markedness in the acquisition of complex prosodic structures by German-Spanish bilinguals. *International Journal of Bilingualism* 6(3), 291–313. <https://doi.org/10.1177/13670069020060030501>
- Lukyanchenko, Anna & Kira Gor. 2011. Perceptual correlates of phonological representations in heritage speakers and L2 learners. In *Proceedings of the 35th Annual Boston University Conference on Language Development*, vol. 2, 414–426.
- Mayer, Mercer. 1969. *Frog Where Are You?* New York: Dial Press.
- McAuliffe, Michael, Michaela Socolof Sarah Mihuc, Morgan Wagner & Michael Sonderegger. 2017. Montreal Forced Aligner: trainable text-speech alignment using Kaldi. In *Proceedings of the 18th Conference of the International Speech Communication Association*.
- McKinnon, Sean. 2023. Phonological contrast maintenance and language contact: An examination of the Spanish rhotic system in a bilingual Guatemalan speech community. In Brandon Baird, Osmer Balam & M. Carmen Parafita Couto (eds.), *Linguistic advances in Central American Spanish*, 41–71. Brill.
- Melero-García, Fernando & Alejandro Cisneros. 2020. No es tan simple como parece. In Elola, Idoia & Diego Pascual y Cabo (eds.), *Current theoretical and applied perspectives on hispanic and lusophone linguistics*, 295–318. John Benjamins.
- Menke, Mandy R. 2018. Development of Spanish rhotics in Spanish-English bilingual children in the United States. *Journal of Child Language* 45(3), 788–806. <https://doi.org/10.1017/S0305000917000460>
- Montrul, Silvina. 2023. Heritage languages: Language acquired, language lost, language regained. *Annual Review of Linguistics* 9(1), 399–418. <https://doi.org/10.1146/annurevlinguistics-030521-050236>

- Oropeza Escobar, Minerva. 2017. Adquisición de las consonantes líquidas en el español de México TT [Acquisition of liquid consonants in Mexican Spanish]. *CPU-e. Revista de Investigación Educativa* (25).
- Pallier, Christophe, Stanislas Dehaene, J-B Poline, Denis LeBihan A-M Argenti, Emmanuel Dupoux & Jacques Mehler. 2003. Brain imaging of language plasticity in adopted adults: Can a second language replace the first? *Cerebral Cortex* 13(2), 155–161.
- Pedersen, Thomas Lin. 2025. *patchwork: The composer of plots*. <https://patchwork.data-imaginist.com/>. R package version 1.3.1.9000.
- Perry, Scott James, Matthew C. Kelley & Benjamin V. Tucker. 2023. Measuring and modelling the duration of intervocalic alveolar taps in Peninsular Spanish. In Radek Skarnitzl & Jan Volín (eds.), *Proceedings of the 20th International Congress of Phonetic Sciences*, 699–703. Guarant International.
- Perry, Scott James, Matthew Kelley & Benjamin V. Tucker. 2024. Documenting and modeling the acoustic variability of intervocalic alveolar taps in conversational Peninsular Spanish. *Journal of the Acoustical Society of America* 155(1), 294–305. <https://doi.org/10.1121/10.0024345>
- Proctor, Michael. 2011. Towards a gestural characterization of liquids: Evidence from Spanish and Russian. *Laboratory Phonology* 2(2), 451–485. <https://doi.org/10.1515/labphon.2011.017>
- R Core Team. 2021. The R Project for Statistical Computing. <https://www.r-project.org/>
- Ramon-Casas, Marta, Susana Cortés, Ariadna Benet, Conxita Lleó & Laura Bosch. 2023. Connecting perception and production in early Catalan-Spanish bilingual children: language dominance and quality of input effects. *Journal of Child Language* 50(1), 1–22. <https://doi.org/10.1017/S0305000921000787>
- Rato, Anabela & Vanina Machado. 2024. Perception and production of English and Portuguese voiceless stops by heritage learners. In Rajiv Rao (ed.), *The phonetics and phonology of heritage languages*, 343–360. Cambridge University Press.
- Recasens, Daniel & Maria Dolors Pallarès. 1999. A study of /j/ and /r/ in the light of the “DAC” coarticulation model. *Journal of Phonetics* 27(2), 143–169. <https://doi.org/10.1006/jpho.1999.0092>
- Repiso-Puigdeliura, Gemma. 2021. Empty onset repairs in the semi-spontaneous speech of Spanish child and adult heritage speakers. *International Journal of Bilingualism* 25(5), 1311–1326. <https://doi.org/10.1177/13670069211016547>
- Repiso-Puigdeliura, Gemma & Ji Young Kim. 2021. The missing link in Spanish heritage trill production. *Bilingualism: Language and Cognition* 24(3), 454–466. <https://doi.org/10.1017/S1366728920000668>
- Rimac, Robert & Bruce L. Smith. 1984. Acoustic characteristics of flap productions by American English-speaking children and adults: Implications concerning the development of speech motor control. *Journal of Phonetics* 12(4), 387–396.
- Ronquest, Rebecca. 2013. An acoustic examination of unstressed vowel reduction in heritage Spanish. In Chad Howe, Sarah Blackwell, & Margaret Lubbers Quesada (eds.), *Selected proceedings of the 15th Hispanic Linguistics Symposium*, 151–171. 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.2002.0179>
- Song, Jae Yung, Stefanie Shattuck-Hufnagel & Katherine Demuth. 2015. Development of phonetic variants (allophones) in 2-year-olds learning American English: A study of alveolar stop /t, d/ codas. *Journal of Phonetics* 52, 152–169.
- Stoehr, Antje, Titia Benders, Janet G. van Hell & Paula Fikkert. 2019. Bilingual preschoolers’ speech is associated with non-native maternal language input. *Language Learning and Development* 15(1), 75–100. <https://doi.org/10.1080/15475441.2018.1533473>
- Torres, Felipe, Josué Pino C., Christian Peñaloza C., Daniela Rojas C., María Angélica Fernández G., Patricia Castro A., Mario Bustos R. & Katherine Hormazábal M. 2016. Comparación de criterios para determinar la edad de adquisición de fonemas lateral y róticos. *Revista Chilena de Fonoaudiología* 15, 1–16. <https://doi.org/10.5354/0719-4692.2016.44149>
- Valdés, Guadalupe. 2014. Heritage language students: Profiles and possibilities. In Terrence G. Wiley, Joy Kreeft Peyton, Sarah Catherine K. Moore, Donna Christian & Na Liu (eds.), *Handbook of heritage, community, and Native American languages in the United States: Research, policy, and educational practice*, 41–49. Routledge. <https://doi.org/10.4324/9780203122419>
- Vigil, Donny A. 2018. Rhotics of Taos, New Mexico Spanish: Variation and change. *Studies in Hispanic and Lusophone Linguistics* 11(1), 215–264. <https://doi.org/10.1515/shll-2018-0008>
- Vivar, Pilar Vivar & Leslie Figueroa Alarcón. 2019. Rhotic consonant acquisition in Chilean Spanish in a sample of subjects between 2;0 and 3;11 years. *Estudios Filológicos* 64.
- Warner, Natasha & Benjamin V. Tucker. 2011. Phonetic variability of stops and flaps in spontaneous and careful speech. *Journal of the Acoustical Society of America* 130(3), 1606–1617. <https://doi.org/10.1121/1.3621306>

- Wickham, Hadley. 2016. *ggplot2: Elegant graphics for data analysis*. Springer-Verlag New York. <https://ggplot2.tidyverse.org>.
- Willis, Erik W. 2005. An initial examination of Southwest Spanish Vowels. *Southwest Journal of Linguistics* 24(1–2), 185–198.
- Willis, Erik W. 2006. Trill variation in Dominican Spanish: An acoustic examination and comparative analysis. In Núria Sagarra & Almeida Jacqueline Toribio (eds.), *Selected proceedings of the 9th Hispanic Linguistics Symposium*, 121–131. Somerville, MA: Cascadilla Proceedings Project. <http://www.lingref.com/cpp/hls/9/paper1372.pdf>
- Willis, Erik & Travis Bradley. 2008. Contrast maintenance of taps and trills in Dominican Spanish: Data and analysis. In Laura Colantoni & Jeffrey Steele (eds.), *Selected proceedings of the 3rd Conference on Laboratory Approaches to Spanish Phonology*, 87–100. Somerville, MA: Cascadilla Proceedings Project.
- Wood, Simon. 2020. *mgcv: Mixed GAM Computation Vehicle with Automatic Smoothness Estimation*. <https://CRAN.R-project.org/package=mgcv>