Contrast Maintenance of Taps and Trills in Dominican Spanish: Data and Analysis

Erik W. Willis and Travis G. Bradley
Indiana University, Bloomington and University of California, Davis

1. Introduction

Spanish has two contrastive rhotic segments: the *vibrante simple* ‘tap’, and the *vibrante múltiple* ‘trill’. These segments are contrastive only in intervocalic position, as in *caro* [kaRo] ‘expensive’/ *carro* [karo] ‘car’ or *pero* [pero] ‘but’/ *perro* [pero] ‘dog’. The Spanish tap/trill distinction is restricted to intervocalic word-internal position and in practice is limited to less than 30 minimal pairs. Hammond (1999) argues that the trill prescribed by the Real Academia Española does not exist in normal Spanish discourse in the speech of most native speakers. Furthermore, he claims that there is a neutralization of the tap/trill distinction in many Spanish dialects.1

Our goal in this paper is to examine the claim of rhotic neutralization in Spanish by comparing new acoustic data on taps with data on trills reported by Willis (2006, 2007) from an innovative variety with a non-normative trill, Dominican Spanish (hereafter DS). Specifically, we will compare new acoustic measurements of intervocalic taps and previously reported findings on trills from the two main dialects of the Dominican Republic, Santo Domingo and the Cibao in narrative DS.

The organization of the remainder of the paper is as follows. Section 2 reviews reports on taps and trills in Spanish and DS. Section 3 details the methodology for the study. Section 4 contains the findings of the acoustic examination of the DS tap and discusses the claim of neutralization. Section 5 posits a theoretical explanation of the development of pre-breathy voicing in DS trills, based on articulatory and perceptual considerations. The conclusions of the study are presented in Section 6.

2. Review

2.1. General characteristics of taps and trills

Ladefoged describes a tap or flap as “a single contraction of the muscles so that one articulator is thrown against another. It is often just a very rapid articulation of a stop” (2001:150). More detail is provided in Ladefoged and Maddieson (1996:231): “Taps are most typically made by a direct movement of the tongue tip to a contact location in the dental or alveolar region.”

Ladefoged and Maddieson (1996) explain that “the primary characteristic of a trill is the vibration of one speech organ against another, driven by aerodynamic conditions” (217). However, the generally accepted phonological trill has a variety of manifestations. Lindau (1985:161) notes that canonical trills are not produced as often as expected despite their phonological characterization as trills. Due to the precise articulatory requirements of a trill, multiple occlusions are often not realized, and there is considerable variation among the world’s languages. Maddieson (1984) reports that only about 18% of the world’s languages, including Spanish, have contrastive rhotics. Cross-linguistic variations in trill production reviewed by Lindau include the number of closures (one to six or more), a prolonged

* For helpful discussion of some of the ideas appearing in this paper, we acknowledge John Kingston, Tony Lewis, Joaquín Romero, Miquel Simonet, and Maria-Josep Solé, as well as the audience members of the 3rd Conference on Laboratory Approaches to Spanish Phonology. The current paper has also benefited from the comments and questions of two anonymous reviewers. Figures 4 and 5 are used with permission from Cambridge University Press. We are responsible for any shortcomings in the present work.

1 This claim by Hammond goes against multiple laboratory examinations of the trill in Peninsular Spanish such as Recasens (1991), Blecua (2001), and Solé (2002).

opening phase in the case of Bomo, variable place of articulation (alveolar versus dorsal or uvular), the presence or absence of energy during the closure, and variation in spectral peak, to name a few.

2.2. Spanish taps and trills

Recognizing the considerable dialectal variation in Spanish taps and trills (Lipski 1994), Hualde (2005:183) provides a general distribution of Spanish rhotics shown in (1).

(1) a. Contrast tap /ɾ/ versus trill /ɾ/
   V__V Intervocalic
   /karol/ ‘expensive’ versus /karol/ ‘cart; car’

b. Only trill /ɾ/
   #__ Word-initial
   /roka/ ‘rock’
   C__ After a heterosyllabic consonant
   /alrededor/ ‘around’, /enredo/ ‘mess’, /israelita/ ‘Israeli’

c. Only tap /ɾ/
   C__ After a tautosyllabic consonant (onset cluster)
   /broma/ ‘joke’, /gramo/ ‘gram’

   V__#V Word-final before a vowel
   /ser amigos/ ‘to be friends’

d. Variable rhotic (most commonly [ɾ])
   V__C Before a consonant
   /parte/ [parte] ~ [parte] ‘part’
   V__#C Word-final before a consonant
   /ser poeta/ ‘to be a poet’

   V__## Word-final before a pause
   /ser o no ser/ ‘to be or not to be’

The standard Spanish trill as described by Quilis (1993:329-332), among others, involves two or more brief occlusions between the tip of the tongue and the alveolar ridge (Blecua 2001, Hualde 2005, Ladefoged and Maddieson 1996, Núñez-Cedeño and Morales-Front 1999, Quilis 1993, Real Academia Española 1992). Figure 1 is an example of a Mexican Spanish word-internal trill with three complete closures and an approximation of the fourth that fits the description in (1a). Figure 2 is a word-initial trill with two closures and corresponds to the context described in (1b). Blecua (2001) reports an average duration of 64ms for the Peninsular Spanish trill based on laboratory speech.

In a series of papers based primarily on impressionistic data, Hammond (1999, 2000a,b) claims that the Spanish trill is not produced as a trill or multiple vibrant as described in the literature. He claims that “when one analyzes the real surface manifestations of the intervocalic flap [ɾ] in different Spanish dialects, it becomes clear that a neutralization of the [ɾ] and [ɾ] [sic] has occurred in many dialects in intervocalic environments” (Hammond 1999:147). He also asserts that “the trilled phone [ɾ] occurs systematically in normal Spanish discourse among a very small number of monolingual native speakers and that among all other monolingual speakers, [ɾ] occurs only in highly affected discourse” (1999:136).

---

2 Heterosyllabic /sr/ clusters are typically pronounced in many Spanish dialects with deletion of the fricative, assimilation of the trill, or both (see Bradley 2006 on the realization of /sr/ clusters in Latin American Spanish).

3 As stated, there is considerable variation with syllable-final rhotics in Spanish. An anonymous reviewer suggested that the coda rhotic is more likely a reduced trill than a tap. In the Dominican dialects, they are commonly resolved as a vocoid in the Cibao dialect or as a lateral in other parts of the country.

4 Hammond (1999) would not consider the production in Figure 2 a “trill” as it does not have the three complete closures prescribed by the Real Academia Española; however, it is clearly not a phonological tap.
Figure 1. Mexican intervocalic trill: *tonces agarr*as ‘then he grabs’.

Figure 2. Mexican initial trill: *que había recolectado* ‘that he had picked’.

In making his claim, Hammond restricts his focus to phonological trills following a heterosyllabic consonant as in (1b) above, either word-initial or word-internal. In a somewhat confusing claim referring to word-initial rhotics, Hammond states that since “there is no flap/multiple vibrant contrast in this environment, word-initial position may lend itself to greater occurrences of neutralization than intervocalic environments where the flap and trill do contrast in Spanish” (1999:148).

The standard or normative Spanish tap, according to Quilis (1993:330), involves a brief contact between the tongue tip and the alveolar ridge. The average duration of contact is 20ms, and the constriction is seldom realized as a complete closure. We represent variation in the degree of contact by employing the IPA lowering diacritic in the narrow transcription of the approximant tap, as shown in *[paɾa]* in Figure 3 for Mexican Spanish. Reports of Spanish dialectal variation with the intervocalic tap are limited to describing elision in high frequency words such as *para* *[pa]* ‘for’.

---

5 The initial syllable and the vowel of the final syllable of the word *[en]tonces* ‘then’ are elided, as is common in the Central Mexican dialect.
2.3. DS taps and trills

There have been few studies that describe the production and variation of the DS tap. Based on impressionistic data Jiménez Sabater (1975:88) reports sporadic deletion of intervocalic taps, as in para ‘for’, quiero ‘I want’, and parece ‘seems’. As Jiménez Sabater (1975) is an impressionistic study, there are no acoustic measurements or details for comparison.

Characterizations of DS have noted a pre-aspirated trill (Jiménez Sabater 1975, Lipski 1994); however, in an acoustic study of 214 trills by 10 educated speakers of Cibaeño DS, Willis (2007) claims that the CDS “pre-aspirated trill” is a misnomer and that the CDS trill is best characterized as a period of “pre-breathy voice” followed by one or two taps. The most frequently occurring token in the data is a “pre-breathy-voiced tap”, which accounted for 56% of the corpus tokens (see Figure 4). Willis (2007) reports that the pre-breathy-voiced portion of the phonemic trill typically constitutes more than 60% of the overall duration of the segment. The two voiced allophones together, [fir] in Figure 4 and [fr] in Figure 5, account for approximately 75% of the productions. The average number of closures in the trill tokens from CDS (Willis 2007) was 1.2; for SDS (Willis 2006), it was 1.5.

Willis proposes [fr] as an IPA symbol to represent the CDS trill. The second most frequently occurring token type involves a period of pre-breathy voice followed by multiple closures [fr] and accounts for 17% of the total token types. Including normative trill tokens with pre-breathy voiced trills results in only 46 tokens or 21% of the CDS phonological trills. In a later comparative study of the Cibao data with the Santo Domingo dialect, Willis (2006) reports some variation in allophonic token frequency between the two principal dialects of DS, but concludes that the most frequently occurring allophone of the intervocalic and word-initial trill is a pre-breathy voiced tap.

---

6 An anonymous reviewer found the term pre-breathy-voiced tap to be confusing and proposed an alternative such as pre-breathy-voiced one-contact trill. We prefer to continue using the label pre-breathy-voiced tap because the term tap in the already adjectivally heavy name is intended to be phonetically descriptive. This term does not imply the phonological category tap /ɾ/ but rather reflects the appreciable acoustics of the waveform and spectrogram, similar to what would be used to describe the dual articulation of the affricate /tʃ/ without claiming that the affricate realizations are identical to the phonological category of stop and fricative. Additionally, the most accurate IPA symbols that reflect the acoustic data are a portion of pre-breathy voice followed by a tap (see Figure 4).

7 Measurements for Willis (2006, 2007) included the duration of the “trill” segment, the duration of the pre-breathy voice, and the number of occlusions. The pre-breathy-voiced portion was determined by spectrographic and waveform analysis. A reduction and/or cessation of formant structure (typically F2) and a marked reduction in the amplitude of the waveform were the two principal cues used to determine the onset of pre-aspiration as defined more broadly by Helgason (2003).
The present study reports on the production of intervocalic taps compared to previous trill reports by Willis (2006, 2007) from the same elicitation task and speakers in order to determine whether, as suggested by Hammond (1999), the rhotic contrast has been neutralized in DS, a recognized innovative or radical dialect of Spanish.

3. Method

3.1. Speakers

A total of twelve university students from Santiago (Cibao) and the Pontificia Universidad Católica Madre y Maestra, Recinto Santo Tomás de Aquino in Santo Domingo provided the narratives recorded in 2004. The informants were evenly divided by dialect and gender (CDS/SDDS and male/female). They were native speakers of their respective dialects and recorded within the dialectal region.
3.2. Data elicitation

Each speaker was engaged in an informal interview about him or herself to solicit natural conversation prior to performing several linguistic tasks. Topics of the informal conversation included family life, school, a scary experience, or a happy memory from their childhood. One of the linguistic tasks was to narrate the children’s picture book by Mercer Mayer, *Frog, Where Are You?* (1969). The speech from the narrative of this story serves as the basis for the current study. The story is about a little boy, a *perro* [pero] ‘dog’, and a *rana* [rana] ‘frog’, which induced multiple productions of trill segments within a natural context. The tap tokens used for comparison with the trills also came from the same story, and only tap tokens corresponding to word-internal intervocalic position such as *para* [para] ‘for’, *pero* [pero] ‘but’, *parece* [parese] ‘seems’, and *miraba* [miraβa] ‘was looking’ were used for analysis. In the current narrative data each speaker typically produced between 4-24 tap tokens which are then compared to their own trill productions reported in Willis (2006, 2007).

3.3. Acoustic measurements

The current data consist of 154 total taps. The acoustic measurements include the duration of the tap token and whether or not there was a closure (appreciable approximation) defined as a significant reduction in the waveform. For details on the DS trill see Section 2.3 and Willis (2007).

3.4. Equipment and software

A private office within the Department of Languages and Literatures of the Pontificia Universidad Católica Madre y Maestra, Recinto Santo Tomás de Aquino was used for recording the informants. The speech was recorded directly onto a laptop computer using CoolEdit, a USBpre external sound card, and a Shure 512 microphone with a sampling rate of 44.1kHz. The taps were extracted from the recording and analyzed with the acoustic software analysis program PRAAT (Boersma and Weenink 2006).

4. Results

4.1. DS taps

The data revealed a number of distinct realizations of the DS tap, as characterized in (2):

(2) a. [ɾ] a noncontinuant tap (see Figure 6; cf. Figure 3 for Mexican Spanish)
   b. [ɾ] an approximant tap (see Figure 7; cf. Figure 3 for Mexican Spanish)
   c. [ɾ̥] a perceptual tap without salient acoustic cues (see Figure 8)
   d. [Ø] complete reduction or elision (see Figure 9)

Figure 6 is an example of a noncontinuant tap as evidenced by the clear break in the spectrogram and the reduced amplitude of the waveform. The presence of periodic vibrations throughout the constriction indicates that the tap is voiced.

---

8 Cases that involved a phonetic tap due to a following vowel-initial word such as *come[ɾ] un* ‘to eat a(n)’ were not included in the current data set in order to maintain within-lexical-item intervocalic comparisons.
However, in Figure 7, the same word shown in Figure 6 is produced with an approximant variant. We can observe a continuation of the formants, and there is less reduction in the waveform as compared to Figure 6.

A perceptual tap without clearly distinguishable landmarks for delimiting the segment is presented in Figure 8. As suggested by an anonymous reviewer, it is the result of an approximation of the articulators that produces a decreased amplitude of the formant structure. While there is some appreciable reduction in the waveform, the formant structure is maintained throughout the tap gesture.
Figure 8. Perceptual tap [ɾ]: parece ‘it seems’.

Finally, Figure 9 presents a token showing the complete elision of an intervocalic tap. The formant structure of the surrounding vowels is maintained and uninterrupted. In contrast to Figure 8, there is no reduction in the waveform.

Figure 9. Complete elision of the tap [Ø]: parece ‘seems’.9

As just illustrated in Figures 6 through 9, there was considerable variation in the phonetic realization of the tap. A total of 79 tokens, or 51% of the DS taps, were produced with a measurable closure (combining the noncontinuant and approximant allophones illustrated in Figures 6 and 7). The mean closure duration for all speakers was 22ms with a standard deviation of 8, which is in line with characterizations suggested by Quilis (1993). The remaining 75 tokens, or 49% of the DS taps, were produced without a measurable occlusion.10

9 We chose to use a voiced variant of the fricative [z] due to the periodicity in the waveform and energy in the voicing bar compared to the voiceless fricative production seen in Figure 8.

10 An anonymous reviewer suggested separating the perceptual tap and complete elision allophones, but the problem then becomes one of category definition and relevance. We could provide some type of mean energy
4.2. Tap/trill comparison

Willis (2006, 2007) reports an average of 1.2 closures for the Cibao dialect and 1.5 for the Santo Domingo dialect for the trill segment produced by twenty speakers of DS. The claim of rhotic neutralization in Spanish is likely supported if a comparison of the two contrastive rhotics is based exclusively on the number of occlusions. However, as illustrated in Table 1, the duration of the contrastive segments is notably different; on average the mean trill duration is minimally three times as long as the mean tap duration. It is worth noting that the durations of the non-standard pre-breathy voiced DS trill are much longer than the trill average of 64ms reported by Blecua (2001) for Peninsular Spanish. It should also be noted that the means and standard deviations for the two categories do not overlap, which is what should be expected with neutralization.

<table>
<thead>
<tr>
<th>Speakers</th>
<th>Dialect</th>
<th>Gender</th>
<th>Tap durations</th>
<th>Trill durations</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Number</td>
<td>Mean (ms)</td>
</tr>
<tr>
<td>Cibao</td>
<td>f (3)</td>
<td></td>
<td>20</td>
<td>8.1</td>
</tr>
<tr>
<td></td>
<td>m (3)</td>
<td></td>
<td>16</td>
<td>7.8</td>
</tr>
<tr>
<td></td>
<td>f (3)</td>
<td></td>
<td>23</td>
<td>7.9</td>
</tr>
<tr>
<td>Santo Domingo</td>
<td>m (3)</td>
<td></td>
<td>22</td>
<td>8.0</td>
</tr>
</tbody>
</table>

Table 1. DS tap and trill durations.

Comparisons of production data of the two contrastive segments based on both occlusions and duration reveal appreciable differences between the two Spanish rhotics, including allophonic variation, pre-breathy voice, and non-overlapping segmental duration. Future studies could address perceptual robustness of the segmental contrast both within DS as well as for non-DS speakers.

5. On the development of pre-breathy voicing in DS trills

In this section, we explain the innovation of pre-breathy vibrants in aerodynamic terms and then show that contrast maintenance also plays a role in constraining the developments. An interaction of articulatory and perceptual factors is necessary to give a comprehensive account of rhotic behavior in DS varieties and across languages.

In a study of the relationship among oropharyngeal pressure ($P_o$), airflow, and lingual trilling, Solé (2002) finds that voiced trills place a high degree of aerodynamic constraint on $P_o$ in order to maintain simultaneous lingual and glottal vibration. The experimental method involved the use of different sized catheters to vent $P_o$ from behind the back molars while subjects pronounced voiced and voiceless trills. Results indicate that for voiced trills, small changes in pressure may result in the loss of trilling, voicing, or both, as commonly observed in Spanish varieties and other languages. Voiceless trills involve less aerodynamic constraint on $P_o$ because the absence of glottal vibration allows a greater degree of oral airflow, thereby facilitating lingual vibration. Solé does not discuss the relationship between breathy voicing and lingual trilling, but it seems plausible to situate breathy voiced trills between voiced and voiceless trills along a continuum of phonation types, as in Figure 10. In the production of breathy voiced sounds, the vocal cords have little longitudinal tension and are more abducted than for modal voicing (see Gordon and Ladefoged 2001 and studies cited therein).12

11 These numbers reflect the measurable tap tokens only (79 taps or 51% of the tap corpus). The other 49% were lenited or elided forms without consistent landmarks for measurement and, if included, would reflect a zero measurement resulting in a skewed portrayal of the normative taps.

12 Since breathy voicing lies between modal voicing and voicelessness on the phonation continuum, one might expect some of the trill productions in Solé’s experiment to have decayed into breathy voicing instead of complete voicelessness. Maria-Josep Solé (p.c.) explains that “venting the $P_o$ in modally voiced trills resulted in detrilling..."
Relative to modal voicing, breathy voicing offers less resistance at the glottis, which increases the oral pressure necessary to set the tongue tip into vibration. The innovation of breathy voicing in DS trills can be viewed as an articulatory strategy to create favorable aerodynamic conditions for the initiation of lingual trilling. If maximizing oral pressure were the most important factor, then an even better strategy would be to completely devoice the trill via glottal spreading. However, as documented by Willis (2006, 2007), both the breathiness and the oral gesture of the DS tapped [fır] and trilled [fır] are typically voiced, in contrast to previous, impressionistic descriptions of these components as voiceless. Cross-linguistically, voiced trills are statistically preferred to voiceless (and breathy voiced) trills (Maddieson 1984). Aerodynamic factors alone cannot explain the dispreference for voiceless trills.

Solé (2002:680-682) finds that although they are more aerodynamically stable than voiced trills, voiceless trills are perceptually very similar to fricatives. In contrast to voiceless trills, modally voiced trills present an alternation of open and close movements of greater and lesser acoustic energy, respectively, which yields a more distinctive auditory signal. “The preference for voiced trills in phonological systems seems to reflect a trade-off between articulatory stability (i.e., preserving trilling in a narrow range of aerodynamic conditions, as opposed to voiceless trills) and acoustic/auditory saliency (i.e., distinct signal modulation)” (p. 682). We suggest that the innovation of breathy voicing in DS trills can be seen as a compromise between modally voiced and voiceless trills, combining the perceptual distinctiveness of the former and the aerodynamic stability of the latter. Since pre-breathy voiced trills present turbulent airflow, glottal vibration, and lingual trilling, they provide a more distinctive contrast with other fricatives than do voiceless trills — especially in phonological systems that lack voiced fricatives, as in DS and, more generally, modern Spanish. The dispreference for voiceless trills is motivated by the need to maintain perceptually distinctive contrasts, which limits the degree to which glottal spreading can be employed as a strategy to facilitate lingual trilling.

(i.e., no tongue tip vibration) due to the decreased pressure difference across the lingual constriction. No devoicing was observed in these cases; most likely speakers could not maneuver fast enough to compensate for an abrupt decrease in $P_o$. Devoicing was only found when speakers tried to prolong trilling (for about 10 seconds) and there was not sufficient airflow due to decreased $P_s$ [subglottal pressure — EWW and TGB] at the end; then they most likely removed the glottal resistance to increase $P_o$. There might have been a breathy voiced trill in the transition between a voiced and voiceless trill in prolonged trills, but not noticeably so.”

Pre-breathy voicing may also be related to processes of velarization observed in languages such as German, French, and Brazilian Portuguese (see Willis 2007 for discussion).

An anonymous reviewer questioned the functional motivation of contrast maintenance in the development of DS pre-breathy-voiced vibrants and in sound change more generally. If phonetic innovations in lingual trills were constrained by the need to maintain perceptually distinct contrasts, then “the trill-to-fricative change would not occur as often as it does, not only in Spanish, but also in other languages like Swedish, which, like the Spanish dialects where the change has in fact occurred (Andean Spanish among others), also have a tongue-tip /s/-like fricative.” We would argue, however, that contrast preservation need not operate in an “all-or-nothing” fashion across all languages. Rather, different languages and/or dialects have the option of either preserving or neutralizing different phonological contrasts. Furthermore, to claim that contrast plays a role in the DS rhotic innovations does not predict that trill assimilation should be unattested across other dialects and languages because the original contrast between the trill and /s/ or other fricatives may still be maintained via changes in other phonetic parameters such as voicing, periodicity, and center of gravity, among others (see, for example, Colantoni 2006a,b on Argentine Spanish developments involving assimilation in palatals and rhotics). Changes such as these suggest a comparison between contrastive sounds, at least on some level, whether synchronically, within the grammar, or diachronically, outside the grammar.

Figure 10. Trills placed on a continuum of phonation types, defined in terms of the aperture between the arytenoid cartilages (after Gordon and Ladefoged 2001:384).
Another finding reported by Solé (2002:675) is that a higher pressure difference is needed to initiate trilling than to sustain it, which explains why trills are pre- and not post-breathy voiced. Timing the breathy voicing gesture to occur near or before the onset of the oral constriction favors the initiation of trilling, whereas timing it with the oral release would yield no articulatory advantage. Pre-breathy voice timing can also be motivated in terms of perceptual distinctiveness. Cross-linguistically, nonmodal phonation tends to be realized early in sonorants, often on the preceding vowel. “By realizing nonmodal phonation at the beginning of a prevocalic sonorant, the transitions into a following vowel, which are perceptually most valuable, remain modal voiced and thus are clearer from an auditory standpoint” (Gordon and Ladefoged 2001:395).

Consider the timing relationship illustrated in Figure 11, which shows a hypothetical [VfirV] sequence in the form of a gestural score (Browman and Goldstein 1989, 1990, 1991, 1992). In this representation, the activity of each relevant articulator is depicted on a separate tier. Boxes represent articulatory gestures, and the length of a box denotes the period of time during which the articulator is under active control. For purposes of illustration, an acoustic image is time-aligned with the percept to which the gestural score gives rise. The onset of breathy voice occurs before the onset of the alveolar trill, which facilitates the initiation of lingual trilling, as explained above. Early glottal timing colors the preceding vowel with the effects of turbulent airflow, as indicated by the shaded box. The C-V transition from the trill into the following vowel remains modally voiced and auditorily more distinct than the V-C transition preceding the trill.

![Figure 11. Breathy voicing timed to the onset of lingual trilling yields the percept [fir].](image)

Assuming that oral-glottal timing is originally synchronous, the leftward shift of breathy voice in Figure 11 could be interpreted as a change in phasing modes in a coupled oscillator model of intergestural timing (Goldstein 2005). On this view, the glottal gesture shifts from an in-phase mode, coordinated with the following vowel, to an anti-phase mode, coordinated with the preceding vowel, while the oral gesture remains in-phase. By hypothesis, only the in-phase mode is possible in the absence of a preceding vowel. However, pre-breathy voiced vibrants in DS are found in such contexts, namely phrase-initially and word-initially after a consonant (Willis 2006, 2007). This suggests that pre-breathy voice timing is not dependent on the possibility of anti-phase coordination with a preceding vowel. Rather, it seems better to view pre-breathy voicing as a segment-internal coordination strategy motivated by the aerodynamic and perceptual factors discussed above.

---

15 Thanks to Miquel Simonet (p.c.) for discussion on this point.
As speakers began to favor pre-breathy voice timing in the trill, listeners may have reinterpreted the breathy voice component as a formal part of the segment’s representation. With pre-breathy voicing as a compensatory durational cue, the lingual trilling gesture could then lenite to a pre-breathy voiced one-tap trill without neutralizing the rhotic contrast (Zlotchew 1974). Our findings suggest that the rhotic contrast is maintained in DS on the basis of duration (recall Table 1). As documented by Willis (2006, 2007), [ʃr] and [ʃɪ] are the most frequently occurring allophones of the DS phonological trill. Occasional realizations of the trill as a voiced glottal fricative [ʃ] are also found, which suggests that temporal reduction of lingual trilling can sometimes be complete, leaving behind only the laryngeal component. In fact, the allophonic realizations of trills may be seen as points along a continuum, possibly as a function of articulatory effort. The sequence of phonetic forms [r] > [ʃr] > [ʃɪ] > [ʃ] would constitute progressively less effortful realizations, with the final three forms showing greater degrees of debuccalization.\textsuperscript{16}

The hypothetical gestural score in Figure 12 illustrates the lenition of a pre-breathy voiced trill [ʃr] to a pre-breathy voiced tap [ʃɪ]. Early timing of the breathy voice gesture colors the V-C transition and contributes to the overall duration of the rhotic segment, which is sufficient to maintain a contrast with the phonological single /r/ in word-medial intervocalic contexts (e.g., caʃr]o ‘car’ vs. caʃ]o ‘dear, expensive’). Recall from Section 4 that 51% of the intervocalic /r/ tokens in the present data sample were produced with a measurable occlusion, the mean duration of which was 22ms for all speakers. The mean durations for trill allophones range from 72 to 89ms across different speakers. No measurable occlusion was produced in 49% of the tap tokens, which suggests that the tap itself shows a considerable degree of lenition.

![Figure 12. Temporal lenition of the lingual trilling gesture yields the percept [ʃɪ].](image)

Another issue to consider is why pre-breathy voicing in DS developed with /r/ but never with the phonological tap /ʃ/. The absence of pre-breathy voicing with /ʃ/ is perceptually motivated, as the additional breathy component would have increased the overall duration of the tap, thereby diminishing the contrast with surface realizations of /ʃ/. Furthermore, unlike the aerodynamically driven trill, the tap involves a single ballistic gesture to which pre-breathy voicing offers no particular aerodynamic advantage.

\textsuperscript{16} Assuming that oral-glottal timing is originally synchronous, a pre-breathy-voiced tap could also result from partial debuccalization of the lingual trilling gesture, without the intermediate step involving a leftward shift of the glottal gesture. Thanks to Joaquin Romero and John Kingston (p.c.) for discussion concerning the continuum and debuccalization.
6. Conclusion

An acoustic comparison of taps and trills in DS reveals appreciable differences based on overall duration of the segments. Roughly 50% of the tap tokens were produced with a measurable closure and averaged 22ms. The other 50% of the tap tokens were lenited forms that were reduced or completely elided. While the DS trill does tend to have a reduced number of occlusions per trill token compared to Peninsular Spanish, the overall segmental duration reveals a clear acoustic cue for contrast maintenance. The suggestion of rhotic neutralization in DS is clearly rejected by the data.

We have also sketched a theoretical account of the innovation of pre-breathy voicing in DS trills. We have argued that contrast maintenance plays a role in the development of DS rhotic system in four ways: (i) by disfavoring voiceless trills, which are less perceptually distinct from other fricatives; (ii) by favoring pre-breathy voice timing, which maintains clearer C-V transitions; (iii) by allowing the lingual trilling gesture to lenite, with breathy voice serving as a compensatory durational cue; and (iv) by limiting the innovation of pre-breathy voicing to /r/.

References


