Prosodic and Segmental Effects on the Durational Variability of Svarabhakti Vowels in Spanish /Cr/ Clusters

Benjamin Schmeiser
Illinois State University, United States
schmeis@ilstu.edu

Abstract

The present study examines the durational variability of svarabhakti vowels in Spanish /Cr/ clusters. Data from twenty-nine speakers from six countries were collected, resulting in a total of 521 /Cr/ clusters that exhibited a svarabhakti vowel whose duration was then measured. Durational measurements were analyzed under the scope of five hypotheses, based on prosodic and segmental factors, and tested by employing single-factor, within-subject ANOVAs, with statistical significance set at a p value of 0.05. Results suggested that segmental factors, rather than prosodic ones, affect the durational variability of the svarabhakti vowel. Using Articulatory Phonology as the theoretical base, segmental effects on svarabhakti vowel duration were discussed.

In short, longer svarabhakti vowel mean duration was evidenced in those consonant clusters whose members shared the same voicing quality; in addition, patterns in svarabhakti vowel duration variability emerged for both manner and place of articulation. Based on the findings, a continuum was posited for future research.

Index Terms: svarabhakti vowels, tautosyllabic consonant clusters, Articulatory Phonology, gestures, Spanish, prosody, segmental features.

1. Introduction

Currently, a significant challenge in linguistics is to determine what factors affect svarabhakti vowel (henceforth, SV) duration. An SV is a vowel-like fragment that exists to transition from one consonant to another. SVs are unique in that they are not part of the syllabic makeup of a word and merit special investigation in that they are a specific subcategory of vowel intrusion and behave independently from vowels and vowel epenthesis (Hall 2003, 2006).

Researchers around the turn of the 20th century (Lenz, 1892; Araujo, 1894; Josselyn, 1907; Colton, 1909) began to seek the environments in which SVs occurred in Spanish and found that an SV is present between the voiced alveolar tap and its adjacent obstruct, as in:

(1) a. the tautosyllabic environment, /Cr/, as in: cuadras [kodras] ‘stables’
   b. the heterosyllabic environment, /rC/, as in: parges [rgrs] ‘parts’

In the examples in (1) and throughout the paper, I follow previous phonetic representation of the SV in Spanish (Malmberg, 1965; Bradley, 2004, 2005, 2006, 2007; Bradley and Schmeiser, 2003; Kilpatrick, Kirby and McGee, 2006; Schmeiser, 2007) by denoting the SV as the superscripted schwa, [˚], in phonetic transcription.

The current study is confined to SVs found in Spanish complex onset clusters (i.e. the tautosyllabic environment) comprised solely of the obstruent and the Spanish tap (for an overview of consonant clusters, see Recasens, 2018; and of the Spanish tap, see Schmeiser, 2019; and Bradley, forthcoming).

The obstruent + tap complex onset cluster is an environment in which SVs occur in the majority of cases (Colantoni and Steele, 2005, 2007; Ramírez, 2002, 2006; Schmeiser, 2007, 2009). The obstruent may be a voiceless labiodental fricative, voiceless stop, a voiced stop or its (unfaithful) allophone, a voiced approximant, as seen in (2):

(2) a. fres.co [fr˚] ‘fresh’
   b. a. tro [tr˚] ‘other’
   c. grean.de [g˚r] ‘big’
   d. ne.go [g˚r] ‘black’

Notice that the cluster may occur either word-initially (2a,c) or word-externally (2b,d). Finally, I note that the voiced approximant in Spanish (2d) occurs between two [+continuant] segments or between a [+continuant] segment and a lateral, in the case of /b, g/.

Though SV occurrence in this environment is well-documented, what causes durational variability of the SV is indeed not well-understood. A growing body of recent work on Spanish (Colantoni and Steele, 2005, 2007, 2011; Kilpatrick, Kirby and McGee, 2006; Ramírez, 2002, 2006; Schmeiser, 2007) has focused on the prosodic and segmental factors of the /Cr/ cluster to determine possible effects on SV duration. With the noted exception from this recent work of significantly longer mean SV duration after a voiced consonant, the testing of other hypotheses has offered conflicting results, which leaves our understanding of the matter still relatively unclear.

The present study fills the current gap by testing five hypotheses based on the findings from previous studies regarding prosodic and segmental factors. I then discuss the results in theoretical terms, using Articulatory Phonology, which differs from conventional phonological frameworks in that it models the articulations that are responsible for variation observed in the acoustic data and is particularly suitable to research in SV durational variability because it illustrates changes in SV duration in terms of the timing relationship between the two consonants in the cluster.

The rest of the paper is designed as follows: §2 discusses what we currently know about SV duration and offers the hypotheses posited in the present study; §3 lays out the experimental design; §4 offers the results of the study; §5 presents a discussion of the results, using Articulatory Phonology as its framework, and §6 concludes.
2. The Prosodic and Segmental Effects on SV Durational Variability

After the above-mentioned initial studies had established the environments in which SVs occurred in Spanish, researchers began to examine SV duration. Navarro Tomás (1918: 386) points out that the SV is highly variable and that its duration often equals or surpasses that of the tap, which evidences a mean duration of 25 ms to 30 ms in his study. Gili Gaya’s (1921) analysis of SVs in Peninsular Spanish is the first to consider the factors that possibly condition SV durational variability. Bradley and Schmeiser (2003) re-categorize Gili Gaya’s (1921) data in terms of prosodic and segmental influences and three trends emerge from the data; namely that SVs tend to be longer in (i) word-initial clusters than in word-internal ones, (ii) stressed syllables than in unstressed ones, and (iii) clusters with a back-to-front order of constriction (i.e. /g/r/ , /k/r/) than clusters with a front-to-back order of constriction (i.e. /p/r/, /b/r/, /f/r/). In recent years, researchers have furthered our understanding of SV durational variability by conducting empirical studies. Ramirez’s (2002, 2006) work evidences significantly longer SV mean duration after a voiced obstruent; his (2006) study also reveals that place of articulation might play a role though the data are not significant. In Kilpatrick, Kirby and McGee’s (2006) study of SVs in Bolivian Spanish, they find only one variable to evidence significant results in the majority of their subjects, namely longer SV mean duration after a voiced obstruent; significant findings are not found for prosodic stress, word position, place of articulation or tongue height of the nucleic vowel.

In Colantoni and Steele’s (2005) study on Argentine Spanish, the findings suggest that word position does play a role in SV duration, with longer SVs evidenced in word-medial clusters. They also find that prosodic stress has a significant effect, with longer SVs found in stressed syllables. In Colantoni and Steele’s (2005, 2007) studies, their results suggest significantly longer mean SV duration after a voiced obstruent as opposed to a voiceless one. Schmeiser’s (2007) study found that SVs were significantly longer following dorsals than labials, and following voiced obstruents than voiceless ones. Word position and stress had no effect. Finally, in Colantoni and Steele’s (2011) study on stop-liquid clusters of Argentine Spanish (and Quebec French), their data suggest that intrinsic duration of the obstruent affects SV durational variability. More specifically, voiceless stops, longer in duration than their voice counterparts, result in shorter SV duration.

2.1 Hypotheses posited in the present study

From this recent body of work on SV durational variability, we draw two conclusions. First, data results consistently suggest that mean SV duration is significantly longer after a voiced obstruent than its voiceless counterpart. Second, data results seem to vary with regard to the other hypotheses posited, namely the prosodic factors of word position and stress, along with place of articulation, which has left the matter unclear as to which factors, outside of obstruent voicing, significantly affect SV duration. To better ascertain what affects SV durational variability, I posit the following five hypotheses, based on prosodic and segmental factors, seen in (3):

1. Word-initial /Cr/ clusters will evidence longer SV mean duration than word-medial ones.
2. Stressed /CrV/ syllables will evidence longer SV mean duration than unstressed ones.
3. /Cr/ clusters with a voiced obstruent will evidence longer SV mean duration than ones with a voiceless one.
4. /Cr/ clusters whose obstruent is an approximant will evidence the longest SV mean duration, followed by a stop and finally, a fricative.
5. /Cr/ clusters whose obstruent is a dorsal will evidence the longest SV duration, followed by a coronal and finally, a labial.

3. Experimental Design

3.1 Subjects

The present study consisted of twenty-nine subjects, nineteen males and ten females, from six countries: Spain, Mexico, Argentina, Guatemala, Colombia, and Ecuador. Subjects were taken from a larger corpus that consisted of subjects from over twenty-five different regions throughout the Spanish-speaking world. Subjects ranged from four to six in number for each country. All subjects were native speakers of Spanish and all produced the Spanish tap (as opposed to an approximant or deletion) in /Cr/ clusters as the canonical realization. SV durational variability in terms of gender was not treated here, given Colantoni and Steele’s (2005) findings that gender plays no role in duration of the SV.

3.2 Data collection

The subjects read a 420-word passage that contained thirty-two /Cr/ clusters and were originally recorded on reel-to-reel tape. The recordings were later digitized and stored on CD in MPEG format at a sample rate of 22,050 Hz and sample size of 16-bit. Spectrographic analyses and waveforms were taken using the software program, Speech Analyzer 2.6. For each cluster, I measured the SV, the consonant, the rhotic, the preceding and following vowels and a burst, if present. The following figure illustrates an SV in a canonical /Cr/ cluster:

Figure 1: A canonical /Cr/ cluster with nucleic vowel in detrás [de´trás] ‘behind’ (Subject 30)

In Figure 1 above, we observe a canonical realization of a /Cr/ cluster, with its nucleic vowel, in the word, detrás.
[de.tˈras] ‘behind’. We detect the voiceless dental stop, [t], followed by the SV, [´], the tap, and the nucleic vowel, [a].

3.3 Data analysis

The current study limited itself to only those clusters in which both an SV was present and the tap was produced. In light of comparing factors within a hypothesis, I note that the first three hypotheses compare two factors (e.g. word-initial vs. word-medial). For these three hypotheses, I follow Ramirez’s (2006) work in that I compare the SV mean duration of all instances of one category to all instances of the other category, and thus do not compare the two environments for the same phoneme.

In terms of statistical analysis, using a regression framework, durational means were taken across subjects and single-factor, within-subjects ANOVAs were employed to determine significant results (p<0.05). Given the exploratory nature of the study, coupled with the notion that SVs do not necessarily surface in every consonant cluster, an imbalance of tokens was to be expected for any given hypothesis; the analysis of unbalanced data was corrected by the orthogonal sums of squares (also known as ‘Type III sums of squares’). Finally, all data were analyzed using the statistics software program, ‘R.’

4. Results

Of the 812 total /Cr/ clusters examined, 521 /Cr/ clusters contained SVs and a Spanish tap, resulting in an occurrence rate of 64.16%, a percentage in line with Blecu’s study (2001: 63%) on rhotics, and Ramirez (2006: 68%). In what follows, I offer for each hypothesis a graph of the mean SV duration for each factor tested.

4.1 H1: Word position

In light of H1, Figure 2 presents the SV mean duration in word-initial clusters as compared to word-medial ones. Contrary to the hypothesis, shorter mean SV duration in word-initial clusters (22.65 ms) is reported, compared to word-medial clusters (25.35 ms), though the results were not statistically significant (p=0.569):

![Figure 2: H1 - SV mean duration in word-initial vs. word-medial /Cr/ clusters](image)

4.2 H2: Syllable stress

In light of prosodic stress, mean SV duration was slightly longer in stressed syllables (25.07 ms) than in unstressed ones (24.84 ms), though the results were not statistically significant (p=0.523):

![Figure 3: H2 - Stressed vs. unstressed syllable in /Cr/ clusters](image)

4.3 H3: Obstruent voicing

Compared to prosodic factors, however, segmental factors offered statistically significant results. Regarding obstruent voicing, statistically highly significant results (p=<0.001) are reported, with longer mean SV duration found after voiced obstruents (31.09 ms) than after voiceless ones (21.59 ms):

![Figure 4: H3 - Voiced obstruent vs. voiceless obstruent in /Cr/ clusters](image)

4.4 H4: Manner of articulation

A propos manner of articulation of the obstruent, statistically highly significant results (p=<0.001) revealed a continuum of longest mean SV duration after approximants (32.95 ms), followed by stops (23.88 ms) and finally, by fricatives (18.91 ms):

![Figure 5: H4 - Manner of articulation of the obstruent in /Cr/ clusters](image)

4.5 H5: Place of articulation

In light of place of articulation of the obstruent, statistically highly significant results (p=<0.001) revealed a continuum of longest mean SV duration after dorsals (30.08 ms), followed by coronals (23.36 ms), and finally by labials (21.96 ms):

![Figure 6: H5 - Place of articulation of the obstruent in /Cr/ clusters](image)

4.6 Supported hypotheses

In sum, the first two hypotheses (H1-H2), namely word position and syllable stress, were not supported with statistically significant findings. Data results from H3-5, based on segmental factors, all were supported: statistically significant findings were found for obstruent voicing, manner of articulation, and place of articulation. Finally, I note that data for H1 (word position), though not statistically significant, were the only data in the present study that evidenced findings contrary to the hypothesis. In the following section, I discuss these supported hypotheses in theoretical terms.
5. Discussion

5.1 Articulatory Phonology

Articulatory Phonology (Browman and Goldstein, 1989, 1990, 1991, et seq.) is rooted in basic units called ‘gestures’ which are represented on a two-dimensional gestural score. In the case of tautosyllabic /Cř/ clusters, the gestural score consists of two consonantal gestures superimposed onto an underlying vocalic gesture; each consonantal gesture has a timing relationship with the other consonantal gesture and also with the underlying vowel (Gafos, 2002). Crucially, the amount of gestural overlap between the two consonantal gestures is directly proportional to how much of the underlying vocal gesture is perceptible (i.e. the SV). In other words, more overlap between two consonantal gestures entails less SV duration and vice versa.

5.2 Obstruent voicing

Results from the current study corroborate all previous work (Ramírez, 2002, 2006; Kilpatrick, Kirby, and McGee, 2006; Colantoni and Steele, 2005, 2007, 2011; Schmeiser, 2007) in that they highly suggest longer mean SV duration after a voiced obstruent. In her study of rhotic duration in Peninsular Spanish, Blecua (2001) observed that voiced consonants are intrinsically shorter than their voiceless counterparts. Crucial to her study is that a consonant shorter in duration (i.e. one that is voiced) implies longer rhotic duration in the cluster. In Colantoni and Steele’s (2011: 20) study, they find statistically significant findings for longer SV and rhotic duration after a voiced consonant.

5.3 Manner of articulation

The current study is novel in that it is the only known study to consider all three categories (fricatives, stops, and approximants) for manner of articulation. Given that voiced approximants have the shortest intrinsic duration, followed by stops, and finally, fricatives (Borzone de Manrique and Signorini, 1983: 121), it is not surprising that data results from the current study suggest the longest mean SV duration after the voiced approximants, followed by the stops, and finally the voiceless labiodental fricative.

5.4 Place of articulation

The findings exhibited statistically significant findings of the shortest SV duration after labials, followed by coronals and finally, dorsals. These data corroborate the tendency noted by Colantoni and Steele (2005) and Ramírez (2006). Labials are in front of a tap in the vocal tract and enjoy an unobstructed release. Data results from the current study support Byrd’s (1996a: 233) findings that front consonants followed by back consonants tend to be more overlapped than an ordering of back consonants to front.

Coronas evidence longer SVs given that the same articulator (i.e. the tongue tip) is required for the two adjacent gestures (i.e. a coronal obstruent and the tap). Clusters with a coronal obstruent require the tongue tip to form a constriction, lower, and then rise back to form the tap constriction, resulting in longer SV, as compared to a labial obstruent.

Dorsals evidenced the longest SV duration and corroborate those data found in Colantoni and Steele (2005) and Ramírez (2006). In this particular case, the back of the tongue moves slower (Hardcastle, 1973: 266) and the contact area is more extended (Cho and Ladefoged, 1999: 211), with the end result of the longest SV duration after dorsals.

5.5 The segmental effects on SV durational variability in Spanish /Cř/ clusters

Thus, there are (at least) two major findings in the current study. First, data results suggest that SV duration varies according to how similar the two consonants in the cluster are in terms of obstruent voicing, manner of articulation, and place of articulation.

Second, results also suggest that articulatory conditions seem to influence SV duration in that (i) labials evidence a full release, (ii) are produced by a different articulator, and (iii) require a front-to-back order of constriction. As such, SV duration following a labial is the shortest. In the case of coronals, they are more similar to the tap in that they (i) require the same articulator as the tap (i.e. the tongue tip) and, (ii) are quite similar in their place of assimilation to the tap. Therefore, they unsurprisingly evidence longer SV duration as compared to their labial counterparts. Finally, dorsals (i) evidence slower tongue movement, and (ii) elicit a back-to-front order of constriction which result in the longest SV duration.

In sum, given the voicing, duration, and articulatory conditions offered above, I posit the following continuum, along which SV duration might be viewed:

\[
\begin{align*}
\text{Fricatives} & \rightarrow \text{Stops} & \rightarrow \text{Approximants} \\
[f] & \rightarrow [p] & \rightarrow [t] & \rightarrow [k] & \rightarrow [b] & \rightarrow [d] & \rightarrow [g] & \rightarrow \{\text{[j]}\} & \rightarrow \{\text{[ɓ]}\} & \rightarrow \{\text{[ɣ]}\} & \rightarrow \{\text{[y]}\}
\end{align*}
\]

shorter SV          longer SV

Figure 7: Proposed continuum of SV duration

In Figure 7, we see a continuum, based on manner of articulation, that suggests SV duration will increase as obstruent voicing and duration decrease (Blecua, 2001; Colantoni and Steele, 2011); within each category of manner of articulation, I hypothesize a ranking of place of articulation that is based on the tendency found in the present study of longer SV duration as the point of constriction moves in a front-to-back order within the vocal tract (i.e. from labials to coronals to dorsals).

6. Conclusions

Present data results suggest that prosodic factors do not play a role in SV durational variability. Of the segmental factors tested, certain tendencies emerged. In terms of voicing, longer SV duration was evidenced after voiced consonants. In light of the two hypotheses with three factors, a continuum developed; as concerns manner of articulation, SV duration increases from fricatives to stops, and from stops to approximants; for place of articulation, SV duration increases from labials to coronals, and from coronals to dorsals.

In sum, this study suggests that when consonants share similar conditions (i.e. voicing, duration, and articulatory factors) with the tap, the cluster is more at risk of being perceptually recovered and thus exhibits an SV that is longer in duration. Then proposed a continuum along which we might view SV durational variability and I called for future research to test the continuum with empirical data.
7. References