SQUIB

# The phonetics of insertional code-switching

## Suprasegmental analysis and a case for hyper-articulation

Daniel Olson
Purdue University

This study investigates the phonetic production of Spanish-English insertional code-switches, constituents of an embedded language inserted within a discourse of a matrix language. While previous research on the phonetics of code-switching has focused exclusively on the segmental properties, the current study provides a detailed examination of the suprasegmental features of code-switched tokens, including pitch height and duration. Code-switched productions are compared with non-code-switched tokens, and results indicate that insertional code-switched tokens are produced with a degree of hyper-articulation, evidenced by an increase in pitch height and duration. The results are discussed within Hyper- and Hypo-articulation Theory, drawing on the notion of a decreased local probability incurred by insertional code-switches.

**Keywords:** code-switching, bilingualism, phonetics, suprasegmental, prosody, intonation

## 1. Introduction

While bilinguals are capable of operating effectively in either of their languages, perhaps more intriguing is their ability to communicate alternating between languages. Such language switching has proven to be a valuable ground for linguistic study, given that it represents a unique paradigm for understanding language not readily available through the study of monolingual behavior. Although language switching has been more thoroughly examined in other linguistic domains, the area of phonetics, and phonetic production in particular, has remained relatively unexplored. Of the few studies that have addressed the phonetics of language switching, there has been a focus on segmental properties (e.g. Bullock & Toribio, 2009; Grosjean & Miller, 1994), with no attention to the suprasegmental level.

*Linguistic Approaches to Bilingualism* 2:4 (2012), 439–457. DOI 10.1075/lab.2.4.05ols
ISSN 1879-9264 / E-ISSN 1879-9272 

The current study redresses this dearth of research on the phonetic properties of language switches, specifically by contributing an examination of the suprasegmental reflexes of code-switching, as well as by providing a theoretical account of the results.

### 1.1 Code-switching

Language switching, also termed code-switching (CS), has been defined as "the selection by bilinguals or multilinguals of forms of an embedded variety in utterances of a matrix variety during the same conversation" (Myers-Scotton, 1993, p. 3), and the current study focuses on intrasentential CS, with switches within an utterance. Relying on terminology from the Matrix Language Frame (Myers-Scotton, 1993), an intrasentential CS utterance may be divided into the matrix language and the embedded language. The matrix language (ML) is defined as the language in an interaction from which the greatest number of morphemes is drawn, whereas the embedded language (EL) refers to the other language used in the interaction, but with a lower overall frequency. Cross-linguistically, the EL in intrasentential CS has been shown to consist most frequently of, although not limited to, a single N or NP (for review see Chan (2003); for Spanish-English, see Poplack (1980)). In this study, the focus is on the suprasegmental properties of insertional CS productions, specifically instances in which EL material is inserted into an ML context.

### 1.2 Bilingual phonetic production

A substantial body of research on bilingual productions has sought to describe the phonetics of bilingual speakers operating in each of their distinct languages. To that end, Caramazza, Yeni-Komshian, Zurif, and Carbone (1973) demonstrated that early Canadian French-English bilinguals maintain separate voice onset times (VOT) in each of their languages. More recently, these findings have been extended to Spanish-English bilinguals (Bullock & Toribio, 2009; Magloire & Green, 1999). However, as Caramazza et al. (1973) note, the productions of the bilinguals are not identical to those of their monolingual counterparts (see also Flege & Eefting, 1987), with productions subject to effects of age of acquisition and dominance (Hazan & Boulakia, 1993; Thornburgh & Ryalls, 1998, but see also MacLeod & Stoel-Gammon, 2010), and with varying degrees of cross-linguistic influence even within a single speaker (Heselwood & McChrystal, 1999; Khattab, 2002).

As Queen (2001) notes, there is a general lack of research at the suprasegmental level from the bilingual perspective, and the majority of extant studies have focused on intonation production and transfer in second language learners (for

review see Queen, 2001). A limited number of studies have been conducted on the intonation of balanced bilinguals, including examinations of balanced Spanish-English (Penfield & Ornstein-Galicia, 1985), Canadian French-English (Cichocki & Lepetit, 1986), German-Spanish (Lleó & Rakow, 2011; Lleó, Rakow, & Kehoe, 2004), and Catalan-Spanish bilinguals (Simonet, 2008). Again, these studies have focused only on non-mixed language contexts, examining speakers operating in one of their two languages. Taken as a whole, the studies concerning the suprasegmental level seem to support the findings at the segmental level. Specifically, while speakers distinguish intonation patterns in their two languages, there is evidence of cross-linguistic interference. While it is clear that, on the whole, bilinguals maintain separate phonetic categories for their differing languages in isolation, both at the segmental and suprasegmental level, bilingual speech is often characterized by language switches within continuous speech, yet little is known about the prosodic characteristics of code-switched speech.

## 1.3 Code-switching and phonetic production

Representing one of the first attempts to address the phonetics of CS production (for perception see Bürki-Cohen, Grosjean, Miller, 1989, and Elman, Diehl, Buchwald, 1977), Grosjean and Miller (1994) investigated the phonetic productions of insertional CS, exploiting the VOT difference between French and English. French-English bilinguals, reporting daily use of both languages and frequent code-switching, participated in a reading and re-telling task. In addition to unilingual French and English control stimuli, the code-switched stimuli for the task consisted of French ML utterances with insertional English EL tokens. The results indicated that there was no significant difference between the VOT of the English control stimuli and the English code-switched tokens in the French ML. These results led the authors to conclude that there is no 'phonetic momentum' from the ML that impacts the production of the EL (Grosjean & Miller, 1994, p. 203).

Results from Bullock and Toribio (2009), using a similar reading paradigm, stand out in contrast. Most pertinent to the current study, results for the early bilinguals indicated a degree of convergence at the phonetic level for VOT. Specifically when switching from Spanish to English, bilinguals' English VOT values became more Spanish-like at the point of the CS. Similarly when switching from English to Spanish, their Spanish VOT values became more English-like. Again, such cross-linguistic phonetic influence on CS tokens may be subject to effects of age of acquisition and/or dominance (Antoniou, Best, Tyler, & Kroos, 2011; Kehoe, Lleo, & Rakow, 2004), inherent language internal features (Bullock, Toribio, González, & Dalola, 2006), and even context (Olson, 2010). While the growing body of research has begun to create a nuanced picture of the phonetics of CS, the focus has

been solely on segmental features, most notably VOT, ignoring possible effects of CS at the suprasegmental level.

### 1.4 Current study aims

While progress has been made in the analysis of CS at the segmental level, any description of code-switched tokens, as well as subsequent theories on the nature of bilingual phonetics, must account for potential impacts of switching at the suprasegmental level. In light of the lack of evidence at the suprasegmental level, the current study investigates the nature of insertional code-switched tokens at the suprasegmental level. Specifically, are CS tokens distinct at the suprasegmental level from non-CS tokens?

To address this question, code-switched tokens, produced by highly proficient early bilinguals, are analyzed with respect to suprasegmental features. It is hypothesized that, while the segmental level may evidence convergence or cross-linguistic influence, code-switched tokens may be marked through suprasegmental features, based on notions of local predictability (e.g. Bell, Brenier, Gregory, Girand, & Jurafsky, 2002).

## 2. Methodology

### 2.1 Participants

Six Spanish-English bilingual participants, ranging in ages from 19–37 (M= 26.8), were recruited on a campus at a university in the southwest region of the United States. Participants were early bilinguals, having acquired both languages before the age of 5, with consistent exposure to northern Mexican and contact varieties of Spanish, as well as southern varieties of English. Following the Language Experience and Proficiency Questionnaire (Marian, Blumenfeld, & Kaushanskaya, 2007), participants were asked to rate their abilities and language use in English and Spanish. Self-ratings did not differ significantly between languages in either comprehension ($t(5)= .688$, $p= 0.53$) or production abilities ($t(5)= .408$, $p= 0.70$), and all participants reported daily use of both languages and comfort switching between languages (see Table 1). Previous research has shown that speaker self-ratings are highly consistent with behavioral performance (e.g. Chincotta & Underwood, 1998; Flege, Yeni-Komshian, & Liu, 1999; Jia, Aaronson, & Wu, 2002), and as such, represent a reliable proficiency measure (Flege, MacKay, & Piske, 2002). Based on both language history and self-rating results, all participants are considered to be highly proficient bilinguals, as well as experienced code-switchers.

**Table 1.** Summary of key factors in language background.

| Language Background | M | SD |
|---|---|---|
| Age of Exposure | SPN= 0 | 0.0 |
| | ENG= 1.5 | 2.5 |
| Language Proficiency | SPN= 9.7 | .41 |
| (1 = None; 10 = Native Speaker) | ENG= 9.6 | .83 |
| Typical Daily Use | 4.8 | .83 |
| (1 = English; 10 = Spanish) | | |
| Comfort with Code-switching | 9.7 | .67 |
| (1 = Never Switch; 10 = Very Comfortable) | | |

## 2.2 Stimuli

Stimuli for this production study consisted of a contextualizing discourse followed by a narrow focus utterance with the target token in post-focal position. In reviewing the functions of CS, many authors have noted its use to create narrow focus, or emphasize a particular constituent (e.g. Gumperz, 1982; Zentella, 1997). Given that narrow focus affects phonetic production (e.g. Hualde, 2005), all target tokens for the current study were placed in a post-focal context to ensure that the CS token was not produced as the focused constituent (Olson & Ortega-Llebaria, 2009).[1]

For this experiment there were three distinct language conditions. The only-English condition (1a) consisted of a contextualizing paragraph in English followed by an utterance containing a non-code-switched English target token. Similarly, the only-Spanish condition (1b) consisted of a contextualizing paragraph in Spanish followed by an utterance containing a non-code-switched Spanish target token. Lastly, the code-switched condition (1c) consisted of a Spanish contextualizing discourse followed by an utterance in which Spanish was the ML and a single English NP was the EL.

In the code-switched context, text color alone was used to indicate the intended language of a given constituent, and color-language pairings were counterbalanced across subjects. To control for possible effects of color distinctions, the target constituent was similarly illustrated in a different color than the surrounding constituents in the only-English and only-Spanish contexts. The constituent to be produced with narrow focus was always presented in capital letters.

(1) a. I was at the art fair, and I was looking to buy something special for my girlfriend. She loves jewelry with little pieces of glass, and lots of color. So I asked the artist if he had a necklace like that.
He answered, "No, but this BRACELET has the **beads** that your girlfriend will like."

b. Estábamos hablando de dichos muy comunes, y mi amigo nos contó que su mamá siempre le decía que los perros tienen siete vidas.
Con una sonrisa le dije, "¿Un perro? No, pero un GATO tiene **vidas** múltiples según el dicho."
"We were talking about common sayings, and my friend told us that his mother always told him that dogs have seven lives.
With a smile I told him 'A dog? No, but a CAT has multiple **lives** according to the saying.'"

c. Le dije al dependiente de la tienda que buscaba algo para mi novia. A ella le gustan las cosas con muchos colores brillantes. Le pregunté si tenía una pulsera con muchos colores.
Me contestó: ¿Una pulsera? No, pero el ANILLO tiene **beads** con muchos colores."
"I said to the store clerk that I was looking for something for my girlfriend. She likes things that have lots of brilliant colors. I asked if he had a bracelet with lots of colors.
He answered, 'A bracelet? No but the RING has **beads** with lots of colors.'"

All English target tokens were monosyllabic CVC syllables, consisting of /hVd/, /bVd/ and /dVd/, and balanced for the three point vowels, /i, a, u/. Given the prevalence of open syllables, Spanish target tokens consisted of bisyllabic CVCV lexical items. All target syllables were stressed, word-initial, and preceded by an unstressed open syllable ending in a mid-vowel. Target tokens occurred approximately 2 syllables (M= 2.13) from the focalized constituent and a minimum of 4 syllables (M= 7.74) from the end of the utterance. There were a total of 486 productions (9 tokens x 3 language conditions x 3 repetitions x 6 speakers= 486 productions) with less than 2% eliminated for disfluencies.

### 2.3 Procedure

Stimuli consisting of both the contextualizing paragraph and the utterance containing the target token, were visually presented and pace was participant-controlled. Participants were instructed to read both the contextualizing paragraph as well as the target utterance out loud as if they "were speaking with a good friend who also speaks English and Spanish."

Stimuli were blocked by language condition (only-English, only-Spanish, and code-switched), randomized within each block, and the order of presentation of blocks was counterbalanced across the subjects. Recordings were made using Audacity 1.2.5 software, through a single channel with a 44.1 kHz sampling rate.

### 2.4 Measurements

To investigate the nature of the suprasegmental features of insertional CS, the productions of the target tokens were analyzed for both pitch and duration using Praat 5.1.07 (Boersma & Weenink, 2009). While *pitch range,* the difference in Hz between the f0 maximum associated with the tonic syllable and the preceding f0 minimum, has been the most frequently used measure of pitch associated with a particular constituent, the de-accented nature of post-focal target tokens (e.g. Hualde, 2005) made this measure unreliable. To address these difficulties, the analysis relies instead on the measure of *pitch height,* defined as the maximum f0 in Hz within the tonic syllable of the target constituent.[2] To compensate for between-speaker differences in f0, statistical analysis was conducted on normalized pitch values, calculated by dividing the maximum f0 of each token by the mean f0 for the target tokens in the only-English condition.

Stressed vowel duration was analyzed in order to facilitate cross-linguistic comparison. In the initial voiced stop contexts, /bVd/ and /dVd/, stressed vowel duration was defined as the temporal difference (ms) between the release of the stop consonant to the onset of the closure of the following consonant. Given the continual presence of voicing for the initial consonant, the release of the stop was used as opposed to the onset of voicing. In the /hVd/ context, vowel duration was measured from the onset of voicing to the onset of the closure of the following consonant. Crucially, given that the same tokens appeared in each of the three language contexts, measurements were performed in a similar manner, facilitating comparison. Statistical analysis was conducted on a normalized stressed vowel duration value, again using the mean vowel duration for each speaker in the only-English condition as a baseline. Specifically, temporal normalization was calculated as target vowel duration divided by the mean vowel duration for each speaker in the only-English condition.

## 3. Results

Statistical analysis, using a multi-variate ANOVA with main factors of Speaker and Language Condition (only-English, only-Spanish, code-switched) showed a significant impact of both factors on pitch height and stressed vowel duration.

### 3.1 Pitch height

Results of the statistical analysis, using a univariate ANOVA with the main factor of Language Condition, showed a significant effect of language condition on pitch height ($F(2,354)= 90.74$, $p< .001$). A subsequent post-hoc analysis (Bonferroni)

revealed significant differences between the only-English condition and the code-switched condition ($p < .001$), as well as the only-Spanish condition and the code-switched condition ($p < .001$). Thus, pitch height in the only-English and only-Spanish conditions differed significantly from pitch height in the code-switched tokens. Figure 1 depicts the normalized maximum pitch height for each of the language conditions, illustrating that code-switched tokens were produced with a significantly greater maximum pitch relative to the only-English and only-Spanish conditions.

Figures 2 and 3 show the f0 contours of utterances from the only-English condition and the code-switched condition produced by speaker M1. While the only-English token was produced with the expected de-accented post-focal f0 contours, the code-switched condition evidenced a clear f0 movement and correspondingly greater pitch height. Importantly, if subjects were simply switching from Spanish to

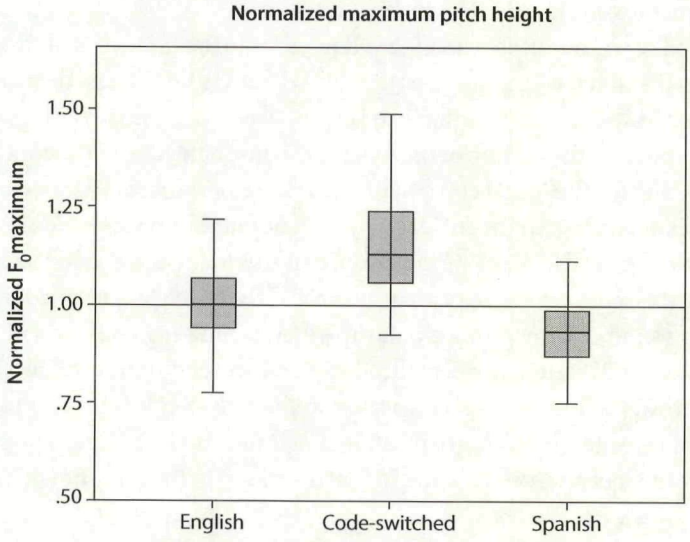


**Figure 1.** Normalized Maximum Pitch Height for all speakers in the only-English, code-switched, and only-Spanish conditions.

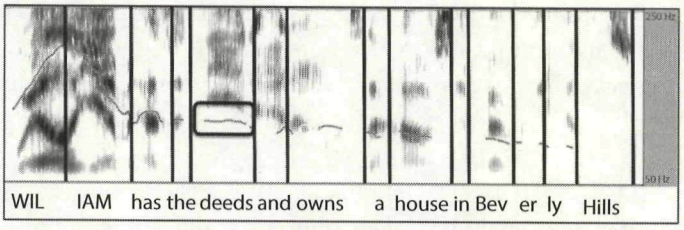


**Figure 2.** Narrow focus utterance in the only-English condition, produced by speaker M1. The constituent WILLIAM receives narrow focus, and the target token *deeds* is produced with a deaccented f0 contour.

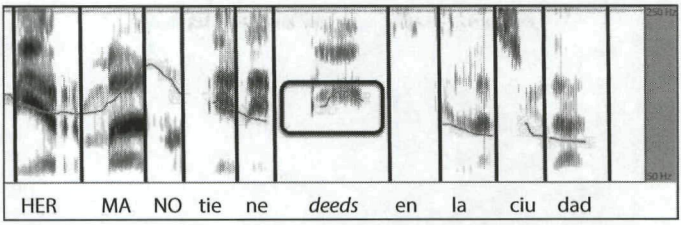


**Figure 3.** Narrow focus utterance in the code-switched condition, produced by the same speaker M1. The constituent HERMANO *brother* receives narrow focus, and the target token *deeds* is produced with a clear pitch movement and increased pitch height.

English intonation patterns in the code-switched condition, we would expect the CS constituent to be realized with a similarly flat or de-accented f0 contour. Figure 3, representative of the majority of CS utterances, instead shows pitch movement that is non-canonical in both monolingual-English and monolingual-Spanish productions, with the post-focal CS constituent being marked with a clear pitch movement.

It should be acknowledged that the results obtained by use of the pitch height measurement, non-relative locally as detailed above, leave open the possibility that the entire post-focal phrase in CS tokens is produced with an elevated, yet flat f0 contour, as opposed to the f0 contour of the CS token itself being raised above that of adjacent constituents. While this does not seem to be the case, as illustrated in Figure 3, such an interpretation would still lead to differentiation of CS and non-CS productions, albeit at the phrasal level as opposed to the constituent level.[3]

### 3.1.1 *Pitch height — creaky voice*

Eliminated from the above pitch height analysis were approximately 8% of the total stimuli productions, due to the presence of *creaky voicing,* defined as a slow vibration of one side of the vocal folds and evidenced by perturbations in pitch tracking. Creaky voice has been shown to correlate with a lower f0 than modal, or normal, voicing, as well as less effortful articulation (Clark & Yallop, 1995, p. 236). Four of 6 subjects produced at least one token with creaky voicing. While not subject to statistical analysis due to the limited number of tokens, an examination of the distribution of creaky-voiced tokens (Figure 4) revealed an unequal distribution across language contexts. Specifically, creaky voicing was much less frequent in the code-switched condition (2.6%) than in the only-English condition (16.3%) and only-Spanish condition (7.2%). This distribution suggests that, overall, code-switched tokens were produced with greater effort and a higher f0 than the non-code-switched tokens.

### 3.1.2 *Pitch Height by Speaker*

Considering individual performance, a multi-variate ANOVA revealed a significant interaction between Stimuli Condition and Subject ($F(1,10)= 6.557$, $p= .009$),

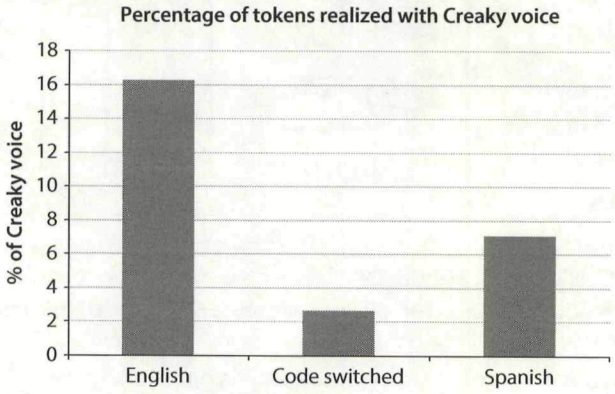


**Figure 4.** Total percentage of stimuli realized with creaky voicing, per stimuli condition.

indicating that the effect of condition was not uniform for all speakers. Subsequent multiple comparisons indicated that all speakers trended in the same direction, with CS tokens being produced with greater pitch than only-English and only-Spanish tokens, and this trend was shown to be statistically significant in 5 of the 6 speakers ($p < .05$).

It should be noted that speakers' Spanish productions tended to be produced with lower overall pitch than their English productions, reaching statistical significance for several speakers, a distinction that is not unprecedented in prosody literature (e.g. Altenberg & Ferrand (2005) for Russian-English bilinguals). Crucial here is that results for pitch height demonstrate a consistent and significant increase in pitch height in the CS condition relative to both the only-English and only-Spanish conditions.

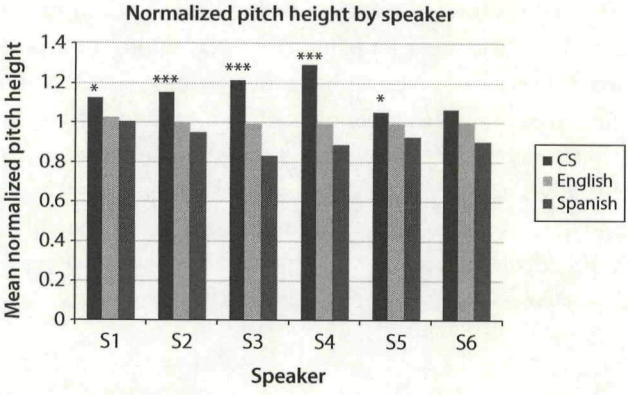


**Figure 5.** Normalized Pitch Height by speaker, in the code-switched, only-English and only-Spanish conditions.

### 3.2 Duration

The results for stressed vowel duration patterned similarly to the results for pitch height. Statistical analysis (ANOVA) with the main factor of Language Condition (only-English, only-Spanish, code-switched) showed a significant effect of condition on stressed vowel duration ($F(2,380)= 73.06$, $p< .001$). Subsequent Bonferroni post-hoc analysis revealed significant differences among all three conditions ($p< .001$). Figure 6 depicts the mean absolute stressed vowel duration for each of the language conditions. Mean stressed vowel durations differ in all three conditions, such that the target tokens in the code-switched conditions were produced with a greater average durations (167.1ms) than the only-Spanish (95.4ms) and only-English (141.2ms) conditions.

While the differences between English and Spanish may have been expected due to intrinsic differences between the languages, more pertinent is the comparison between the only-English tokens and the code-switched tokens. In this case, the code-switched English tokens embedded in a Spanish context are an average of 18% longer than the non-code-switched English tokens.

#### 3.2.1 *Duration by speaker*

Again, a multivariate ANOVA for normalized duration values revealed a significant interaction between Stimuli Condition and Subject ($F(1,10)= 6.608$, $p<.001$). Subsequent comparisons revealed that 5 of 6 speakers trended in the same direction, with the stressed vowels in the CS condition being produced with greater duration than in the only-English and only-Spanish conditions. Only 3 of the 6 speakers reached statistically significant levels ($p< .05$). Of particular note is the

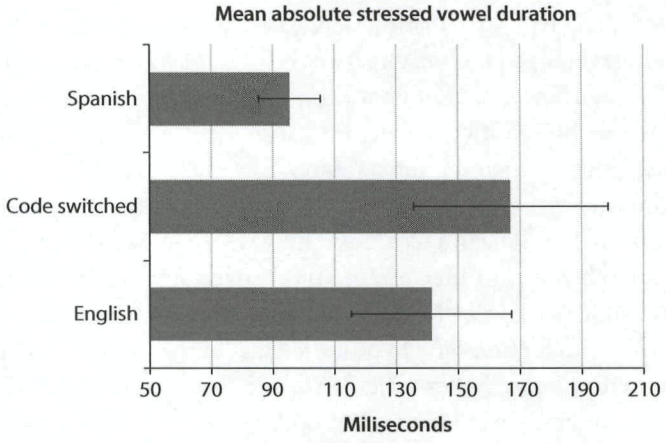


**Figure 6.** Mean stressed vowel duration in ms for all speakers in the only-English, code-switched and only-Spanish conditions.

fact that the only subject who did not produce a statistically significant increase for pitch height in the CS condition did produce a highly significant increase in vowel duration ($p < .001$). Conversely, the subjects who did not produce significantly longer vowel durations did produce significantly greater pitch heights in the CS condition relative to non-CS conditions ($p < .001$). The individual results demonstrate that all speakers display hyper-articulation of CS tokens, employing either an increase in pitch height, an increase in vowel duration, or most commonly both.

## 4. Discussion

The current study, adding to an emerging body of research addressing the phonetics of code-switching, provides a detailed examination of the suprasegmental features of insertional code-switched tokens. While previous research has focused solely on the segmental level, the current results point to a hyper-articulation of the insertional CS token. Addressing the main question posed in this study, results indicate that code-switched tokens were produced with a significantly greater pitch height and greater duration than non-CS tokens. An examination of the data by speaker demonstrated that all speakers hyper-articulated CS tokens, either through an increased pitch, an increased stressed vowel duration, and most commonly a combination of the two strategies. These results may be interpreted by reference to Hyper- and Hypo-articulation theory and local lexical predictability.

### 4.1 H and H Theory

Within Hyper- and Hypo-articulation (H&H) Theory (Lindblom, 1990), speakers constantly adjust their production effort on the basis of constraints in the communicative situation. These constraints can consist of difficulties in the environment (e.g. low signal-to-noise ratio, hearing impairment) or processing demands (e.g. low word frequency, low word predictability). In situations in which communicative constraints are great, speakers tend to hyper-articulate. The phonetic correlates of hyper-articulation can include an increase in duration, increase in pitch, decrease in speech rate and increase in vowel space, among others. Conversely, when faced with lower levels of communicative constraints, speakers economize the effort of their productions and hypo-articulate. In the current study, CS tokens may be analyzed as being hyper-articulated, produced with both an increase in pitch as well as an increase in vowel duration.[4]

Crucial to the current study, research has shown that speakers often hyper-articulate less predictable elements in a discourse relative to more predictable

elements (e.g. Pluymaekers, Ernestus, & Baayen, 2005; Raymond, Dautricourt, & Hume, 2006). Bell et al. (2002), for example, found that function words in English are produced in a "fuller" form, with both fuller vowels and greater word duration, when they occur in less predictable contexts. Such findings have been addressed in the Smooth Signal Redundancy Hypothesis (Aylett & Turk, 2004), which proposes an inverse relationship between predictability, resulting from pragmatic, semantic, and lexical factors, and prosodic prominence, including duration and f0 excursions (53). While code-switches are often high frequency nouns (Chan, 2003), their high frequency does not necessarily imply high predictability. Specifically, when single noun code-switches are inserted into an otherwise monolingual discourse, their local predictability is drastically diminished (e.g. Grosjean, 1988).

It can be concluded that code-switched tokens, by their nature, are less predictable in a discourse than non-code-switched tokens. The act of code-switching, specifically insertional CS in a monolingual-like discourse, may represent a communicative difficulty based on a decreased local predictability. Thus, increased pitch height and increased duration found in the above results may be explained as a response to this communicative difficulty. In the face of lower local predictability, speakers produce more prominent, hyper-articulated code-switched tokens.

The interpretation of CS as being hyper-articulated also has implications at a pragmatic or functional level. Many authors have claimed that one of the various functions of CS is to create emphasis or focus on a particular constituent (e.g. Gumperz, 1982, among many). As such, it should be noted that the suprasegmental features found to be associated with CS in the current study, namely increase in duration and pitch height, correspond to suprasegmental features of narrow focus in both English and Spanish (e.g. for English, Hedberg & Sosa, 2007, and for Spanish, de la Mota, 1997). While these results were found for CS in a non-focal position, they serve to highlight the correspondence between properties of CS and focus. That is, the tendency to produce hyper-articulated CS tokens, with characteristics of narrow focus such as increased pitch and duration, may serve to facilitate its function as creating narrow focus.

### 4.2 Hyper-articulation and language modes

While Grosjean and Miller (1994) found evidence for a complete switch for insertional CS tokens at the segmental level, Bullock and Toribio (2009), among others, obtained distinct results. Their results, for early Spanish-English bilinguals indicated a degree of convergence at the phonetic level, with English VOT values becoming more Spanish-like at the point of the CS and Spanish VOT values becoming more English-like. The question then remains as to how to reconcile these differing findings.

In contrast to the stimuli employed in both the current study and that of Grosjean and Miller (1994), in which the vast majority of constituents were drawn from one language, Bullock and Toribio (2009) used stimuli that were more balanced in their language distribution (Example 2), a difference akin to insertional and alternational, respectively, code-switching in Muysken's (2000) typology.

(2) a. The typhoon damaged *techos y paredes.*
"The typhoon damaged roofs and walls."
b. *Todos mis amigos* talked Spanish as kids.
"All of my friends talked Spanish as kids."
(Bullock & Toribio, 2009, p. 196)

These stimuli can be distinguished with respect to a bilingual continuum, from monolingual to bilingual speech production, termed *language mode* (Grosjean, 2001). As noted by Soares and Grosjean (1984), the place of a bilingual along such a continuum may impact production patterns, and controlling for this variable is crucial in the study of CS (see also Grosjean, 1997; Khattab, 2002). Language mode can be manipulated by controlling the linguistic expectations of a speaker in a given environment (Grosjean, 2008).

The English code-switched tokens in the current study accounted for roughly 2% of the total constituents, inducing a monolingual-mode and indicating a low probability of CS. Similarly, Grosjean and Miller (1994) employed stimuli that were highly French-dominant, with single NP insertional switches. In contrast, the alternational CS stimuli from the Bullock and Toribio (2009) study consisted of approximately 50% English and 50% Spanish constituents, inducing a more bilingual-mode. Thus, speakers operating towards the monolingual end of the language continuum (insertional CS) produced complete or hyper-articulated switches at the phonetic level, while speakers operating in a more bilingual mode (alternational CS) produced switches with a degree of phonetic convergence.

These findings can be restated in terms of H&H Theory and local predictability. While single CS constituents in a more monolingual-like discourse may incur a lower local predictability, within a language-balanced alternational setting, like that of Bullock and Toribio (2009), CS may be more predictable. If roughly half of the constituents come from each language, both speakers and listeners may be prepared for such switches, such that each language is equally predictable. If, as proposed above, insertional CS results in hyper-articulation due to lower predictability, alternational CS would not incur the same diminished local predictability and consequently, may not evidence such hyper-articulation.[5] Thus, while low-predictability, insertional CS may result in hyper-articulation and a complete phonetic switch, alternational CS may represent a more predictable context and may even beget phonetic convergence.

## 5. Conclusion

While previous research has focused exclusively on the segmental level, the current results highlight the importance of integrating a suprasegmental analysis. This work has demonstrated that insertional CS tokens are produced with a degree of hyper-articulation. Specifically, results indicated a consistent and significant increase in pitch height and duration for CS tokens relative to non-CS tokens. The above discussion provides an account for CS within H&H theory, drawing on the notions of local predictability and language mode. This account not only serves to provide a more comprehensive understanding of insertional CS, but also accounts for differing results found for the phonetic productions of insertional and alternational CS.

The findings of the current study call for further research. As indicated in the previous literature, the segmental phonetic realization of code-switched tokens is subject to a variety of internal and external social factors. While the within-subject design of the current study sought to hold such social factors constant, the nuanced results at the segmental level suggest that the suprasegmental level may also be influenced by such factors. For example, as noted, in a context in which switching is highly predictable, driven by both linguistic and social factors, it is conceivable that such hyper-articulation may not be found. As such, an examination of different paradigms, such as a more naturalistic CS paradigm as opposed to the traditional reading task, as well as different language pairings, may serve to further assess CS phonetics and the above predictability account.

## Acknowledgements

Many thanks to Marta Ortega-Llebaria, Almeida Jacqueline Toribio and Barbara Bullock for their insightful comments and discussion, as well as to three anonymous reviewers for their remarks. Parts of this work were presented at the *Hispanic Linguistics Symposium 2009*, and I am grateful for feedback from those in attendance. All errors remain my own.

## Notes

1. Olson and Ortega-Llebaria (2009) provided experimental evidence that code-switched tokens are more likely to be perceived as narrow focus than non-code-switched tokens. They conclude that, even in the absence of prosodic narrow focus cues, code-switching carries a narrow focus interpretation. Given the effect of narrow focus on production, any investigation of the phonetics of CS production must control for placement of narrow focus within an utterance to disentangle effects of focus from those of code-switching.

2. Pitch range, also termed *pitch excursion*, implies a positive change in f0 associated with a pitch accent. Measurement of pitch range in this case, potentially resulting in either positive or negative values, may be taken as implying presence of a pitch accent, either realized or underlying, a discussion that lies outside the scope of the current paper.

3. As pointed out by an anonymous reviewer, this alternative interpretation could still be explained in terms of hyper-articulation, as described below, but would imply hyper-articulation of the entire phrase containing the CS constituent as opposed to being limited to the CS token itself.

4. It should be noted that there is ongoing theoretical debate about the status of de-accented post-focal targets, specifically whether there is an underlying phonological representation of the pitch accent or if there is actually no pitch accent assigned to de-accented tokens. ToBI notation (for Spanish, Beckman, Díaz-Campos, McGory, and Morgan, 2002), for example, traditionally marks de-accented tokens with an "!," followed by the assumed pitch accent structure, such as !LH*, denoting down-stepped or reduced underlying f0 peak (Hualde, 2005). What is clear is that, *ceteris paribus*, greater pitch movement correlates positively with perceived prominence. The notion of "hyper-articulation" adopted here is one of enhanced prominence (e.g. Aylett and Turk's (2004) Smooth Signal Redundancy Hypothesis), rather than strictly enhancement of an existing phonological category or pitch accent, which would hinge upon one's perspective of the underlying nature of de-accented tokens.

5. Preliminary results (Olson, 2010) seem to confirm this hypothesis. By varying the predictability of CS, the study demonstrates that more predictable CS, while prosodically marked, does not evidence the same magnitude of hyper-articulation. Such results point to the gradient nature of this enhanced prominence, with speakers modulating based on the degree of predictability and redundancy.

## References

Altenberg, E., & Ferrand, C. (2005). Fundamental frequency in monolingual English, bilingual English/Russian, and bilingual English/Cantonese young adult women. *Journal of Voice, 20*(1), 89–96.

Antoniou, M., Best, C., Tyler, M., & Kroos, C. (2011). Inter-language interference in VOT production by L2-dominant bilinguals: Asymmetries in phonetic code-switching. *Journal of Phonetics, 39,* 558–570.

Aylett, M., & Turk, A. (2004). The smooth signal redundancy hypothesis: A functional explanation for relationships between redundancy, prosodic prominence, and duration in spontaneous speech. *Language and Speech, 47*(1), 31–56.

Beckman, M.E., Díaz-Campos, M., Tevis McGory, J., & Morgan, T.A. (2002). Intonation across Spanish, in the Tones and Break Indices framework. *Probus, 14,* 9–36.

Bell, A., Brenier, J., Gregory, M., Girand, C., & Jurafsky, D. (2002). Predictability effects on durations of content and function words in conversational English. *Journal of Memory and Language, 60*(1), 92–111.

Boersma, P., & Weenink, D. (2009). Praat: doing phonetics by computer. Version 5.1.07. [computer program], January 10, 2009, http://www.praat.org/.

Bullock, B., Toribio, A.J., González, V., & Dalola, A. (2006). Language dominance and performance outcomes in bilingual pronunciation. In M. O'Brien, C. Shea, & J. Archibald (Eds.), *Proceedings of the 8th Generative Approaches to Second Language Acquisition Conference (GASLA 2006)* (pp. 9–16). Somerville: Cascadilla Proceedings Project.

Bullock, B., & Toribio, A.J. (2009). Trying to hit a moving target: On the sociophonetics of code-switching. In L. Isurin, D. Winford, & K. De Bot (Eds.), *Multidisciplinary approaches to code-switching* [Studies in Bilingualism 41] (pp. 189–206). Amsterdam: John Benjamins.

Bürki-Cohen, J., Grosjean, F., & Miller, J. (1989). Base-language effects on word identification in bilingual speech: Evidence from categorical perception experiments. *Language and Speech, 32*, 355–371.

Caramazza, A., Yeni-Komshian, G.H., Zurif, E.B., & Carbone, E. (1973). The acquisition of a new phonological contrast: The case of stop consonants in French-English bilinguals. *Journal of the Acoustical Society of America, 54*(2), 421–428.

Chan, B.H. (2003). *Aspects of the syntax, the pragmatics and the production of code-switching: Cantonese and English* [Berkeley Insights in Linguistics and Semiotics 51]. New York: Peter Lang Publishing.

Chincotta, D., & Underwood, G. (1998). Non-temporal determinants of bilingual memory capacity: The role of long-term representations and fluency. *Bilingualism, 1*, 117–130.

Cichocki, W., & Lepetit, D. (1986). Intonational variability in language contact: F0 declination in Ontario French. In D. Sankoff (Ed.), *Diversity and diachrony* (pp. 239–247). Amsterdam: John Benjamins.

Clark, J., & Yallop, C. (1995). *An introduction to phonetics and phonology* (2nd ed.). Oxford: Blackwell Publishing.

de la Mota, C. (1997). Prosody of sentences with contrastive new information in Spanish. Proceedings of the ESCA workshop on intonation: Theory, models and applications. Athens, Greece.

Elman, J.L., Diehl, R., & Buchwald, S.E. (1977). Perceptual switching in bilinguals. *Journal of the Acoustical Society of America, 62*, 971–974.

Flege, J.E., & Eefting, W. (1987). Production and perception of English stops by native Spanish speakers. *Journal of Phonetics, 15*, 67–83.

Flege, J., Yeni-Komshian, G., & Liu, S. (1999). Age constraints on second language acquisition. *Journal of Memory and Language, 41*, 78–104.

Flege, J.E., MacKay, I., & Piske, T. (2002). Assessing bilingual dominance. *Applied Psycholinguistics, 23*, 567–598.

Grosjean, F. (1988). Spoken word recognition processes and the gating paradigm. *Perception and Psychophysics, 28*(4), 267–283.

Grosjean, F. (1997). Processing mixed language: Issues, findings, and models. In A. De Groot, & J. Kroll (Eds.), *Tutorials in Bilingualism: Psycholinguistic perspectives*, (pp. 225–254). Mahwah: Lawrence Erlbaum Associates.

Grosjean, F. (2001). The bilingual's language modes. In J. Nicol (Ed.), *One mind, two languages: Bilingual language processing* (pp. 1–22). Malden: Blackwell Publishing.

Grosjean, F. (2008). *Studying bilinguals*. Oxford: Oxford University Press.

Grosjean, F., & Miller, J.L. (1994). Going in and out of languages: An example of bilingual flexibility. *Psychological Science, 5*, 201–206.

Gumperz, J.J. (1982). *Discourse strategies*. Cambridge: University Press.

Hazan, V., & Boulakia, G. (1993). Perception and production of a voicing contrast by French-English bilinguals. *Language and Speech, 36*, 17–38.

Hedberg, N., & Sosa, J. (2007). The prosody of topic and focus in spontaneous English dialogue. In C. Lee, M. Gordon, & D. Büring (Eds.), *Topic and focus: Cross-linguistic perspectives on meaning and intonation* (pp. 101–120). Dordrecht: Kluwer Academic Publishers.

Heselwood, B., & McChrystal, L. (1999). The effect of age-group and place of L1 acquisition on the realisation of Punjabi stop consonants in Bradford: An acoustic sociophonetic study. In P. Foulkes (Ed.), *Leeds working papers in linguistics and phonetics, 7* (pp. 49–68). Leeds, UK: University of Leeds.

Hualde, J.I. (2005). *The sounds of Spanish*. Cambridge: University Press.

Jia, G., Aaronson, D., & Wu, Y. (2002). Long-term attainment of bilingual immigrants: Predictive variables and language group differences. *Applied Psycholinguistics, 23*, 599–621.

Kehoe, M., Lleó, C., & Rakow, M. (2004). Voice onset time in bilingual German-Spanish children. *Bilingualism: Language and Cognition, 7*, 71–88.

Khattab, G. (2002). VOT in English and Arabic bilingual and monolingual children. In D. Parkinson, & E. Benmamoun (Eds.), *Perspectives on Arabic Linguistics XIII–XIV* (pp. 1–38). Amsterdam: John Benjamins.

Lindblom, B. (1990). Explaining phonetic variation: A sketch of the H&H theory. In W.J. Hardcastle, & A. Marchal (Eds.), *Speech production and speech modeling* (pp. 403–439). New York: Kluwer Academic Publishers.

Lleó, C., Rakow, M., & Kehoe, M. (2004). Acquisition of language-specific pitch accent by Spanish and German monolingual and bilingual children. In T. Face (Ed.), *Laboratory approaches to Spanish phonology* (pp. 3–27). New York: Mouton.

Lleó, C., & Rakow, M. (2011). Intonation targets of yes/no questions by Spanish German monolingual and bilingual 2;0- and 3;-year-olds. In E. Rinke & T. Kupisch (Eds.), *The development of grammar: Language acquisition and diachronic change* (pp. 263–286). Philadelphia: John Benjamins.

MacLeod, A., & Stoel-Gammon, C. (2010). What is the impact of age of second language acquisition on the production of consonants and vowels among childhood bilinguals? *International Journal of Bilingualism, 14*(4), 400–421.

Magloire, J., & Green, K. (1999). A cross-language comparison of speaking rate effects on the production of voice onset time in English and Spanish. *Phonetica, 56*, 158–185.

Marian, V., Blumenfeld, H., & Kaushanskaya, M. (2007). The language experience and proficiency questionnaire (LEAP-Q): Assessing language profiles in bilinguals and multilinguals. *Journal of Speech, Language and Hearing Research, 50*, 940–967.

Muysken, P. (2000). *Bilingual speech: A typology of codemixing*. Cambridge: Cambridge University Press.

Myers-Scotton, C. (1993). *Dueling languages*. Oxford: Oxford University Press.

Olson, D. (2010). Bilingual language modes and the effect on phonetic production. Presented at the 40th Linguistic Symposium on Romance Languages, Seattle, WA.

Olson, D., & Ortega-Llebaria, M. (2009). *The perceptual relevance of code-switching and intonation in creating narrow focus*. In M. Ortega-Llebaria (Ed.), *Selected Proceedings of the 4th Conference on Laboratory Approaches to Spanish Phonology* (pp. 57–68). Philadelphia: John Benjamins.

Penfield, J., & Ornstein-Galicia, J. (1985). *Chicano English: An ethnic contact dialect*. Philadelphia: John Benjamins.

Pluymaekers, M., Ernestus, M., & Baayen, R.H. (2005). Articulatory planning is continuous and sensitive to informational redundancy. *Phonetica, 62*, 146–159.

Poplack, S. (1980). Sometimes I'll start a sentence in Spanish *y termino en español*: Toward a typology of code-switching. *Linguistics, 18*(7/8), 581–618.

Queen, R.M. (2001). Bilingual intonation patterns: Evidence of language change from Turkish-German bilingual children. *Language in Society, 30*, 55–80.

Raymond, W.D., Dautricourt, R., & Hume, E. (2006). Word-internal /t,d/ deletion in spontaneous speech: Modeling the effects of extra-linguistic, lexical, and phonological factors. *Language Variation and Change, 18*, 55–97.

Simonet, M. (2008). *Language contact in Majorca: An experimental sociophonetic approach*. Doctoral dissertation, University of Illinois at Urbana-Champaign.

Soares, C., & Grosjean, F. (1984). Bilinguals in a monolingual and a bilingual speech mode: The effect on lexical access. *Memory and Cognition, 12*, 380–386.

Thornburgh, D., & Ryalls, J. (1998). Voice onset time in Spanish-English bilinguals: Early versus late learners of English. *Journal of Communication Disorders, 31*(3), 215–228.

Zentella, A.C. (1997). *Growing up bilingual*. Oxford: Blackwell Publishing.

## *Author's address*

Daniel Olson
School of Languages and Cultures
640 Oval Drive
West Lafayette, IN 47907

danielolson@purdue.edu