



Published in final edited form as:

Clin Linguist Phon. 2012 February ; 26(2): . doi:10.3109/02699206.2011.595526.

Voice onset time of voiceless bilabial and velar stops in 3-year-old bilingual children and their age-matched monolingual peers

LEAH FABIANO-SMITH¹ and FERENC BUNTA²

¹Department of Speech, Language, and Hearing Sciences, University of Arizona, Tucson, AZ, USA

²Department of Communication Sciences and Disorders, Clinical Research Services Center, The University of Houston, Houston, TX, USA

Abstract

This study investigates aspects of voice onset time (VOT) of voiceless bilabial and velar stops in monolingual and bilingual children. VOT poses a special challenge for bilingual Spanish- and English-speaking children because although this VOT distinction exists in both languages, the values differ for the same contrast across Spanish and English. Twenty-four 3-year-olds participated in this study (8 bilingual Spanish–English, 8 monolingual Spanish and 8 monolingual English). The VOT productions of /p/ and /k/ in syllable-initial stressed singleton position were compared across participants. Non-parametric statistical analyses were performed to examine differences (1) between monolinguals and bilinguals and (2) between English and Spanish. The main findings of the study were that monolingual and bilingual children generally differed on VOT in English, but not in Spanish. No statistically significant differences were found between the Spanish and the English VOT of the bilingual children, but the VOT values did differ significantly for monolingual Spanish-versus monolingual English-speaking participants. Our findings were interpreted in terms of Flege’s Speech Learning Model, finding possible evidence for equivalence classification.

Keywords

VOT; bilingual Spanish–English

1. Introduction

The acquisition of voice onset time (VOT) in stop consonants presents a unique challenge to many bilingual individuals. Across languages, small phonetic changes in sound production signal large changes in sound perception (Holt, Lotto, and Diehl, 2004). Small phonetic changes in VOT for stops in languages such as Japanese (Johnson and Wilson, 2002), German (Kehoe, Lleó, and Rakow, 2004), Spanish (Deuchar and Clark, 1996) and English (Macken and Barton, 1980) signal phonemic contrasts important for listener comprehension. Every language that utilizes VOT as an acoustic cue for phoneme differentiation will form a bimodal distribution of VOT values, creating clusters of VOT values that easily identify each phoneme (Maye, Werker, and Gerken, 2002). Bilingual children must form a

© 2012 Informa UK Ltd.

Correspondence: Leah Fabiano-Smith, Department of Speech, Language, and Hearing Sciences, University of Arizona, 1131 E. 2nd St., Tucson, AZ 85721, USA. Tel: (520) 626-9740. Fax: (520) 621-9901. leahfabianosmith@email.arizona.edu.

Declaration of interest: The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

distribution for two languages instead of one, which might lead to clusters of VOT values that are not as clearly defined as they are for monolingual speakers of either language due to cross-language influence. VOT acquisition is especially challenging when a bilingual's two languages differentiate voiced and voiceless stops using diverging patterns. This exploratory investigation focuses specifically on the acquisition of the voicing contrast in bilingual Spanish–English-speaking children and on how cross-language influence might lead to differences in development between monolinguals and bilinguals on this acoustic dimension.

The VOT continuum varies cross-linguistically; however, Lisker and Abramson (1964) found that languages prefer three general timing relations between the stop release and the onset of voicing: (1) voiced (VOT of approximately -90 ms); (2) voiceless unaspirated (VOT of approximately $+10$ ms); and (3) voiceless aspirated (VOT of approximately $+75$ ms). The VOT values of bilingual children who speak a variety of languages have been compared with monolingual speakers in past studies to observe the similarities and differences on this acoustic dimension (Flege and Eefting, 1987; Deuchar and Clark, 1996 (investigating Spanish–English); Kehoe et al., 2004 (investigating Spanish–German); Johnson and Wilson, 2002 (investigating Japanese–English)) (Table I). The current study specifically analyses VOT in bilingual Spanish–English-speaking children using group design.

Lisker and Abramson (1964) found that Spanish initial singleton voiced stops in stressed syllables are pre-voiced (i.e. have lead voicing), while the corresponding English voiced stops, although they may, at times, have optional pre-voicing, typically have a short VOT without lead voicing. The voiceless stops in Spanish and English also differ in that Spanish initial singleton voiceless stops in stressed syllables have a short VOT while their English analogues have a long VOT. Where the two languages overlap is on VOT values of initial voiceless stops in Spanish and initial voiced stops in English (Figure 1). More specifically, the initial, stressed, singleton Spanish voiceless stop VOT corresponds to the English voiced counterpart. This poses a challenge to bilingual Spanish–English speakers, because while analogous voiced–voiceless stop contrasts exist in both languages creating a potential for cross-language transfer (i.e. the use of a language-specific VOT value in the other language context), direct transfer might lead to erroneous phonemic representations and productions. Consequently, it is not surprising that the acquisition of VOT has been investigated in both bilingual children (discussed below) and bilingual adults (e.g. Flege and Eefting, 1987 (investigating adult speakers of Spanish and English); Fowler, Sramko, Ostry, Rowland, and Hallé, 2008; MacLeod and Stoel-Gammon, 2009 (investigating adult speakers of French and English)).

As noted previously, the acquisition of initial, stressed, singleton stop voicing poses a special challenge to bilingual Spanish- and English-speaking children. Cross-language influence could occur because the VOT value for the Spanish *voiceless* stop overlaps with the VOT value for the English *voiced* stop. Conflicting evidence has been found in the literature as to if and how the two languages of bilingual children resolve differences in VOT cross-linguistically. To address issues such as this, Flege (1995) presented the Speech Learning Model (SLM) which stipulates that first language (L1) consonant or vowel categories may trigger assimilation of second language (L2) segmental categories. The formation of novel L2 segmental categories might be blocked if the new segment is identified as an existing, transferable L1 segment. This blockage might occur even though the capacity to form new categories exists. This phenomenon is known as *equivalence classification* (Flege, 1995; MacKay, Flege, and Imai, 2006). The more distant a new vowel or consonant is from the closest existing L1 phoneme, the more likely it is that new category formation will be triggered. Thus, if bilingual children deem, for example, that the English voiced and voiceless stops are analogous and transferable, they might use the same

categories in both languages. Using the same categories in both language contexts could indicate cross-language influence in bilinguals. On the other hand, Flege (1991) also stated that eventually bilingual children, as they grow, might establish distinct phonetic categories for each language. Nonetheless, the conflict between the short-lag voiced English stops and the short-lag Spanish voiceless stops might be a source of confusion for the listener during communication (e.g. the English word ‘pin’ of the bilingual speaker may sound like ‘bin’ to a monolingual speaker of English). However, as Flege (1995) noted, if the L2 segment sounds audibly different from native language segments, modifications to the new sound category are expected.

Since the short-lag stop is voiced in English and voiceless in Spanish, a categorical reorganization might be triggered in young bilinguals. However, it is not clear as to when during the course of acquisition such restructuring happens. This problem is compounded by the fact that it is exactly the short-lag stop that is typically acquired earliest in both English (Macken and Barton, 1979) and Spanish (Macken and Barton, 1980) and by bilinguals (Deuchar and Clark, 1996). In other words, the VOT for the initial, stressed, singleton Spanish voiceless stop generally corresponds with the English voiced counterpart, and both of these are acquired relatively early, resulting in a cross-language conflict that young Spanish–English bilingual children must resolve. One possible solution is to deflect these categories away from each other (cf. Bohn and Flege, 1992). Deflecting away from existing categories stems from an effort to clearly separate VOT categories that could cause the structuring of very distinct representations in both languages of the bilingual learner (see Bohn and Flege, 1992). The question in the acquisition of stop voicing by bilingual Spanish- and English-speaking children is whether, as equivalence classification would prompt, there would be categorical assimilation (indicating cross-language influence) or whether separation would override equivalence.

1.1. Previous studies examining bilingual children

Macken and Barton (1979) investigated the acquisition of VOT in monolingual English-speaking children and claimed the existence of three stages of stop voicing acquisition in initial, stressed, singleton position. The first stage was identified as the production of short-lag stops without differentiating voiced and voiceless stops. In the second stage, children begin to differentiate voiced and voiceless stops, but both categories fall within the adult short-lag VOT. Finally, in the third stage, monolingual English-speaking children produce adult-like voiced and voiceless stops around 2 years of age. In a follow-up study, Macken and Barton (1980) compared voiced and voiceless stops produced by monolingual Spanish-speaking children. They found similarities between Spanish-speaking and English-speaking monolinguals in that short-lag stops occurred the earliest. In the second stage, monolingual Spanish-speaking children did differentiate voiced and voiceless stops within the adult short-lag category but did not produce stops with a voicing lead in a consistent fashion; rather, spirantization was used as a mechanism to differentiate voiced from voiceless Spanish stops.

Deuchar and Clark (1996) conducted a study on a bilingual Spanish- and English-speaking child. They found mostly short-lag stops in both Spanish and English at 1;7. At 1;11, differentiation between voiced and voiceless stops showed signs of developing, but the process was incomplete. The first notable VOT contrast appeared at 2;3 in English, but not in Spanish, at which age the study ended.

Johnson and Wilson (2002) conducted a study with two Japanese- and English-speaking bilingual children (Japanese having voiced stops produced with a voicing lead and voiceless stops with a short lag, similar to Spanish) and found similar, albeit not identical, results as Deuchar and Clark (1996). The authors found that while the younger child in the study did not differentiate between the target languages, the older child did show differentiation

between Japanese and English. Another interesting finding was that the bilingual children, when compared to their parents in English, had longer VOT values than both parents in the English voiceless category. Their father was a native English speaker and their mother spoke English as a second language; but nonetheless, the children exhibited differences when compared to both of them. In essence, the children were overcompensating from the existing target categories. This type of overcompensation may be a result of novel segments deflecting away from existing categories, as described by Bohn and Flege (1992).

Kehoe et al. (2004) compared the production of Spanish and German stops by four bilingual children between the ages of 2;0 and 3;0 to their monolingual German- and Spanish-speaking peers. As it was found in other studies, there were few pre-voiced stops even in Spanish; the bilingual children had mostly short-lag or long-lag stops. The authors highlighted three notable patterns in bilingual VOT production: specifically (1) there was delay in the phonetic realization of long-lag voicing, (2) examples of two-way cross-language influence could be found and (3) stop VOT indicated separation between the children's two languages. It is also worth noting that no differences were found between labial and non-labial stop VOT values. This is in spite of what may be expected based on the findings of Lisker and Abramson (1964), who found that labial stops had shorter VOTs than their non-labial counterparts. The authors also acknowledged the existence of notable cross-language influence, as well as the possibility that lead voicing is acquired later due to its linguistic complexity.

Similar findings were presented by Harada (2007), who examined 15 bilingual English–Japanese-speaking children on the production of initial voiceless stops and compared their productions to monolingual Japanese-speaking children. The author predicted that bilingual children might produce Japanese initial voiceless stops with a longer VOT than their monolingual Japanese-speaking peers. The results indicated that the bilingual children did indeed produce initial voiceless stops with a longer VOT value than their monolingual peers; however, they did contrast their English VOT values with their Japanese values. More specifically, bilingual children had established longer VOT values than monolingual speakers of either language, but were not producing initial voiceless stops in the same way across both of their languages. They distinguished their English VOT from their Japanese VOT, indicating that they recognized that initial voiceless stops had to be produced differently in each language context. The results of Harada's (2007) study coincide with the findings of Kehoe et al. (2004) in which bilingual children seem to establish contrasts that are bilingual in nature, that is, not identical to monolingual speakers of either language, but rather a consequence of bilingual acquisition in which both languages are influencing one another.

As suggested by Flege (1991), not all studies found evidence of cross-language influence in young bilinguals. Bond, Edey, and Bermejo (1980) examined two sequential bilingual Spanish–English-speaking siblings, aged 4;0 and 7;0, on VOT values of initial stop consonants. The children were recorded producing initial stops in single words in both English and Spanish and the duration of VOT was measured for each production. The authors found that both children maintained VOT values that were specific to either English or Spanish, indicating a significant voicing distinction for each language and no cross-language influence.

Similarly, Konefal and Fokes (1981) examined three sequential bilingual Spanish–English-speaking children, aged 4;0, 7;0 and 10;0, on VOT in both English and Spanish productions. The two younger children were typically developing and the third older child presented with a language delay (LD). Six words in English and six words in Spanish were analyzed to investigate similarities and differences across languages and across children. The two

younger children exhibited the Spanish short-lag voiceless stops; however, only the 7-year-old had acquired the long-lag voiceless stops for English. Their results indicated separation between the bilingual children's two languages or a lack of cross-language influence.

In this exploratory investigation, we argue that equivalence classification could yield non-distinct categories across the languages of the bilingual children by providing an analogous phonetic difference in Spanish and English which manifests itself in a phonetically different fashion. We focus on the production of /p/ and /k/ by bilingual Spanish- and English-speaking 3-year-old children and their monolingual age-matched peers acquiring English and Spanish. In order to gain insights into how bilingual 3-year-olds acquire their voiceless labial and velar stops, we compared the VOT of the initial, stressed, singleton stops in English of the bilinguals to the English of their monolingual English-speaking peers, the Spanish VOT of the bilinguals to their monolingual Spanish-speaking peers and the Spanish and English VOT values of the bilinguals. Other relevant analyses were also performed to gain a more complete picture of the various aspects of VOT acquisition.

2. Research Questions

The production of VOT of voiceless stops in singleton onset stressed position poses a unique challenge by providing an analogous phonemic difference in Spanish and English, yet manifesting itself in a phonetically different fashion. Since Flege's SLM posits the possibility of (1) equivalence classification for stop categories in young bilinguals (Flege, 1995) and (2) separation of phonetic categories in young bilinguals (Flege, 1991), the following research questions are raised:

1. *Monolingual versus bilingual participants.* Do bilingual Spanish–English-speaking 3-year-olds produce VOT for voiceless stops similar to their monolingual peers in both of their languages? If bilingual children demonstrate VOT values that are identical to their monolingual peers, it will not provide evidence for *equivalence classification* (Flege, 1995), indicating a lack of cross-language influence. If bilingual children demonstrate VOT values that are significantly different from their monolingual peers, it could provide evidence of *equivalence classification*, indicating cross-language influence.
2. *Cross-language contrasts.* Are bilingual children differentiating between English and Spanish VOT contrasts or are they using the same categorical difference (perhaps an intermediate difference)? If bilingual children utilize the same VOT in both language contexts, or if bilingual children are distinguishing VOT but not in a way that is like monolinguals, it could provide evidence of *equivalence classification* (Flege, 1995) and cross-language influence between Spanish and English. If bilingual children maintain separate VOT values that are comparable to those for monolinguals, it will not provide evidence of *equivalence classification* or cross-language influence.

3. Method

3.1. Participants

Twenty-four 3- to 4-year-old children participated in this study. There were three groups of participants: 8 monolingual English (mean age = 3;3; range 3;0–3;11); 8 monolingual Spanish (mean age = 3;4; range 3;2–4;0); and 8 bilingual Spanish and English speakers (mean age = 3;6; range 3;0–4;0). The method of data collection was consistent with the Bunta, Fabiano-Smith, Goldstein, and Ingram (2009) study because this study re-analysed the data set described there. This particular data set was used to pilot the current methodology for a future study employing a larger and more diverse data set. Table II

displays information regarding the general characteristics of all the study's participants and additional relevant details about the language proficiency, input, output and simultaneous or sequential status of the bilingual participants. The monolingual English, monolingual Spanish and bilingual groups did not differ significantly in age, as measured by a one-way ANOVA ($F(2, 21) = 1.362, p = 0.278$). Children who participated in the study exhibited typically developing language and had no neurological, cognitive, sensory or auditory impairments. Participants were recruited from the metropolitan area of Philadelphia, PA, USA (monolingual English and bilinguals) and from Querétaro, Mexico (monolingual Spanish). The bilingual participants in this study were speakers of Caribbean (notably, Puerto Rican or Dominican) varieties of Spanish. Due to our focus on voiceless stops in stressed, singleton onset position, dialectal variation was not a significant issue for this study.

Bilingual children were being raised in an immigrant community located in northeast Philadelphia. This community was established more than 40 years ago and both English and Spanish are used regularly by its residents. In this study, some of the children's families had recently arrived in the United States while the parents of other bilingual children had grown up in the same community where our data were collected. The mother of one bilingual child was a native speaker of English who had acquired Spanish as an adult. Bilingual children attended a bilingual preschool where the language of the classroom alternated each day and classroom teachers and aides used both languages with all the children. Monolingual English-speaking children were recorded in another neighbourhood of Philadelphia where English was the only language spoken by most of the residents and the only language used in school. Monolingual Spanish-speaking children were recorded in a city in south central Mexico where Spanish is the primary language heard and used both at home and at school. The monolingual children in our study had no exposure to any language other than their indicated language.

This study included bilingual children who could be classified as experiencing bilingual first language acquisition (BFLA)/simultaneous bilinguals (i.e. acquiring both languages from birth) as well as child L2 learners (i.e. acquiring both languages before the age of 5) (see Meisel, 2004). However, all bilingual participants were proficient in both of their languages, as indicated by their parents' report. Furthermore, existing research suggests that simultaneous bilinguals and child L2 learners have commensurate, albeit not identical, speech and language skills (Arnold, Curran, Miccio, and Hammer, 2004; Fabiano-Smith and Goldstein, 2010).

Information about the participants' language background was collected via parent and teacher questionnaires (based on Restrepo, 1998). The questionnaires were designed to collect information that had relevance to the participants' language background, including language input and output, and proficiency ratings in each language for bilingual speakers (see also Table II). The parent and teacher questionnaires used in this study (or very close versions of them) have been validated in a number of studies with both monolingual and bilingual participants (e.g. Gutiérrez-Clellen and Kreiter, 2003; Peña, Bedore, and Rappazzo, 2003). Language proficiency in both Spanish and English of the bilingual participants was assured by (1) following the commonly used minimal 20% use of each language (see Pearson, Fernández, Lewedeg, and Oller, 1997) as well as by (2) native or near-native language proficiency reported by the parents in each of the bilingual participants' languages. Parents were asked to provide proficiency ratings of their children's language abilities in both languages using a scale ranging from 0 (cannot speak the language) to 4 (native proficiency) (Peña et al., 2003). Hence, bilingual participants had native or near-native proficiency reported by their parents in both Spanish and English, and monolingual children had a reported language proficiency score of 3 or 4 in one language

and no input or output reported in any other language. Setting the language proficiency criterion at near-native or native levels of language proficiency provided some assurance that all participants were proficient users of their respective languages.

3.2. Materials and procedure

The VOT of voiceless stops (/p/ and /k/) was measured in three English words and two Spanish words (*pencil*, *pants*, *car*, *perro* = ‘dog’ and *cama* = ‘bed’). In English, there were two tokens of initial /p/ and one token of initial /k/. In Spanish, one token of initial /p/ and one token of initial /k/ were available. Bilingual children provided productions of these items in both languages, while monolingual children provided productions in their respective language. For the bilingual children, each session was administered in one language at a time. The person collecting the samples was a fluent speaker of the language and was aware of culturally appropriate ways of interacting with children with diverse language backgrounds. As previously stated, this particular data set was used to pilot the current methodology for a future study employing a larger and more diverse data set. These words were selected because the target phonemes were in word-initial, singleton and stressed position. The target words occurred on the phonology subtest of the *Bilingual English Spanish Assessment* (BESA), a phonological assessment tool designed to gauge children’s speech sound productions in both English and Spanish (Peña, Gutiérrez-Clellen, Iglesias, Goldstein, and Bedore, in press). The words were elicited by asking children to identify items presented via colour photographs in a loose-leaf binder. First, the elicitation involved asking the child to identify the item. If the child did not appropriately identify the object depicted in the picture, the function of the item was provided and the child was asked again to identify it. Finally, if it was necessary, delayed imitation was used to elicit the target word. This is an acceptable form of elicitation that was found not to yield significantly different outcomes than other forms of elicitation when considering phonological information (Goldstein, Fabiano, and Iglesias, 2004). Samples were recorded directly onto a laptop computer equipped with an external Sound Blaster Audigy® 2 ZS Notebook sound card (Creative, Singapore). The microphone system used for the study included a Shure® WL93 wireless microphone (Shure, Chicago, IL, USA) with a wireless transmitter (model T1-CL) and a receiver (model T3-CL). Recordings were saved in an uncompressed PCM wave (.wav) format, sampled at 44.1 kHz, with a 16-bit resolution in mono.

3.3. Voice onset time measurements

In this study, VOT was defined as the duration of the time lag between the burst of the stop consonant and the beginning of voicing for the following vowel (the former constituting part of the VOT). The measurements were done using WaveSurfer (Department of Speech, Music, and Hearing, KTH Royal Institute of Technology, Stockholm, Sweden), a free computer program designed for speech analyses (Sjölander and Beskow, 2010). In order to measure VOT duration, we used a combination of the time waveform and a wide-band spectrogram, as recommended by Ladefoged (2003) who argued for using an expanded waveform for accurate timing information, supplemented by segmental information provided by a wide-band spectrogram. The settings of the spectrogram were the following: analysis bandwidth, 350 Hz; pre-emphasis factor, 0.8; and frequency range, 0–5000 Hz. Measurements were done from left to right, to the nearest millisecond of the respective duration measurement. The VOT measurement began at the beginning of the stop burst and ended at the beginning of the normal voicing for the preceding vowel. Our general criteria for the duration measurements were consistent with established guidelines for measuring VOT (see Peterson and Lehiste, 1960; Shah, 2002). Measuring pre-voicing was not an issue, because this study focused on word-initial voiceless bilabial and velar stops in singleton position.

3.4. Statistical analyses

Because the number of participants included in this study was relatively small ($N = 8$ in each group), non-parametric statistical analyses were employed to control for assumption violations related to sample size. The Mann–Whitney U -test (the non-parametric alternative to the independent samples t -test) was used to examine differences between groups (i.e. monolinguals vs. bilinguals) and the Wilcoxon signed-rank test (the non-parametric alternative to the related samples t -test) was used to examine differences within groups (i.e. English vs. Spanish of bilinguals).

3.5. Measurement reliability

The VOT measurements were verified via inter- and intra-rater reliability measures. Because our data were limited in scope and number, only items with a 100% agreement were included in the analyses. Measurements were considered in agreement if they were within 10 ms of each other (a practice widely accepted for duration measurements), and only two items differed on more than 4 ms. Thus, both inter- and intra-rater reliability were 100% for the items. Five items were discarded due to the lack of agreement, two word tokens were missing at random and one production had the incorrect target segment (glottal stop instead of a velar one), resulting in some unequal group numbers.

4. Results

4.1. Monolingual versus bilingual participants

The first research question inquired about whether bilingual Spanish–English-speaking 3-year-olds produced VOT for voiceless stops similar to their monolingual peers in both of their languages. Average, median and standard deviation (SD) of difference in token measurement can be found in Table III. VOT ranges were as follows: *pencil* 6–117 ms, *pants* 9–230 ms, *car* 27–154 ms, *perro* 4–70 ms and *cama* 7–44 ms. Means and SDs for VOT values for all groups can be found in Table IV.

Results of the Mann–Whitney U -tests indicated that bilinguals differed on VOT from their monolingual peers in English but not in Spanish. Significant differences with moderate effect sizes (i.e. the degree of difference between two means) were found between monolingual English speakers and the English productions of bilinguals on the English words *pencil* ($z = -2.49$, $p = 0.013$, $d = -1.67$) and *pants* ($z = -2.23$, $p = 0.025$, $d = -1.07$). A non-significant difference with a small effect size was found between monolingual English speakers and the English productions of bilinguals on the word *car* ($z = -1.5$, $p = 0.132$, $d = -0.78$). Figure 2 illustrates the VOT values for each lexical item by monolingual and bilingual English speakers.

The Spanish productions of bilinguals and monolingual Spanish speakers did not differ significantly on VOT for the words *perro* ($z = -1.85$, $p = 0.064$, $d = 1.10$) or *cama* ($z = 0.00$, $p = 1.0$, $d = -0.01$), but it is worth noting the moderate effect size for the labial stop /p/. Interestingly, bilingual children seem to be using stop voicing characteristics that are different from monolingual speakers of English but not from monolingual Spanish speakers. Figure 3 illustrates the VOT values for each lexical item by monolingual and bilingual Spanish speakers.

4.2. Cross-language contrasts

The second research question asked how bilingual children were differentiating between the English and the Spanish VOT contrasts. To that end, the English and Spanish productions of /p/ and /k/ were compared for the bilingual participants. In addition, to have a basis of comparison, the productions of /p/ and /k/ were also compared for the monolingual English-

speaking group versus the monolingual Spanish-speaking group. Within the bilingual group, VOT for English and Spanish was compared using Wilcoxon signed-rank tests because they were the same participants. The results of the cross-linguistic comparison for the bilingual participants yielded that VOT for the English and Spanish of bilinguals did not differ for the words *pencil* and *perro* ($z = -0.526, p = 0.599, d = -0.017$); *pants* and *perro* ($z = -0.169, p = 0.866, d = -0.305$); and *car* and *cama* ($z = -1.483, p = 0.138, d = 1.539$). It appears that the Spanish and English VOT values are not different in the two languages of these bilingual participants, but it is worth noting that the effect size was moderate for the cross-linguistic /k/ contrast (i.e. *car* and *cama*).

In addition to comparing the /p/ and /k/ productions in Spanish and English by bilingual participants, we also compared the production of /p/ and /k/ by monolingual English and monolingual Spanish children using Mann–Whitney *U*-tests. We found differences between *pants* and *perro* ($z = -3.130, p = 0.002, d = 1.91$); *pencil* and *perro* ($z = -3.130, p = 0.002, d = 3.35$); and *car* and *cama* ($z = -3.366, p = 0.001, d = 3.57$). Although bilingual participants did not display significant differences in their Spanish and English productions of /p/ and /k/, there were remarkable differences between the monolingual English and the monolingual Spanish speakers' production of the same stop consonants.

5. Discussion

Summarizing our main findings, bilingual Spanish- and English-speaking 3-year-olds did not differ significantly from their monolingual English-speaking peers on the production of /p/, but a moderate effect size was found for the production of the English /k/. All the VOT values of the bilingual children were shorter, on average, than those of their monolingual peers (see Table IV); however, SDs were large for each token, thus caution must be taken when interpreting these data. This study was performed to determine feasibility for a much larger study, and there were limited opportunities for production of each token. We can, however, examine individual subject data, in addition to group data, to observe how bilinguals might demonstrate cross-language influence on this acoustic property. Bilingual child B03 produced *car* with a VOT of 147 ms, a comparable value to the two longest VOT productions by monolingual English-speaking children. Monolingual English-speaking children E02 and E05 exhibited VOT values of 154 and 128 ms, respectively. The VOT values of the remaining bilingual children ranged from 27 to 90 ms for *car*, while the remaining monolingual English speakers ranged from 71 to 102 ms. For the bilabial stop /p/, bilingual child B05 exhibited the highest VOT value of 125 ms and monolingual English-speaking child E02 exhibited the highest VOT value of 230 ms for *pants*. Both of these children were outliers in their language groups, as the range for the remaining bilingual children was 9–53 ms and for monolingual English speakers was 50–95 ms for the word *pants*. For the production of *pencil*, no outliers were noted for either language group.

A different pattern was observed in the children's Spanish productions. The Spanish VOT values did not differ in the monolinguals and the bilinguals, indicating that there is probably an early Spanish influence on English that is more pronounced on the bilabial than the velar stops. The lowest VOT value for the bilabial stop /p/ in the word *perro* was 6 ms produced by bilingual child B05, and the highest value was 70 ms produced by bilingual child B03 (with the remaining values ranging between 10 and 45 ms). Overall, monolingual and bilingual Spanish-speaking children demonstrated very little variability across subjects for the bilabial stop /p/ in *perro* and for the velar stop /k/ in *cama*. This is perhaps reinforced by a preference for a shorter lag that tends to characterize early VOT development. Therefore, we did find evidence of cross-language influence, which may have been, at least in part, due to the phenomenon known as equivalence classification. These results were inconsistent with the findings of some studies (e.g. Johnson and Wilson, 2002) that did find the long-lag

stops to be acquired earlier in English by bilingual children than by their monolingual English-speaking peers. Moreover, given that the short-lag VOT appears to be the least marked and earliest acquired, it may be that when the markedness values of two languages conflict with one another, the less marked feature will be more easily transferred and persist longer in development.

Our second question inquired about cross-language contrasts, and we found no significant differences between the VOT values for the Spanish and English stops in the bilingual children (just as Kehoe et al. (2004) found no language-based differentiation in two of their bilingual participants). This finding suggests that 3-year-old bilingual children do not reliably differentiate the VOT of their voiceless stops in Spanish versus English. A closer look at our data again reveals that all but one of the bilingual children's mean VOT (the VOT in *car*) values fall in the short category (Table IV). This could be due to the velar place of articulation of /k/, but differences for this item could possibly be due to the lexical item itself or syllable length, two factors that will be examined in future studies. It could also be possible that English /k/ is where the durational differences begin to be realized, so it is conceivable that VOT categories are not established at once; rather, they may begin with a specific contrast and then spread to other stop consonants.

Even though our focus was not on comparing children to adults, it is worth noting that the VOT values we obtained from the English monolingual children were longer than the adult values found by Lisker and Abramson (1964). This finding was shared by Johnson and Wilson (2002), who found longer VOT in the children's productions as compared to their parents. Thus, our findings support previous studies that have found monolingual English-speaking children are still acquiring the adult-like VOT values at this age and have not yet completed the acquisition process. Interestingly, the monolingual English-speaking children's voiceless /p/ and /k/ VOT values were much closer to the adult target than the English productions of their bilingual peers. It is possible to conclude that monolingual English-speaking children's VOT is closer to the adult target than the English productions of bilingual children, indicating that bilingual acquisition of VOT might be slower than that of monolingual acquisition on this construct, at this point in acquisition.

The results of this study could potentially have ramifications for the SLM, but our findings warrant a conservative interpretation, because of the incomplete nature of VOT acquisition in 3-year-old bilingual children and our limited data set. Equivalence classification does appear to affect bilingual children's English VOT values, in spite of the fact that they started acquiring both languages early in life. However, signs of separation do appear involving the VOT of the English voiceless velar stop. There is also evidence for a categorical split being triggered, but the process seems to be gradual rather than abrupt, at least as it manifests itself in production. What does appear to be clear is that the acquisition of voicing and VOT in bilingual children progresses gradually and displays variability, much like that of monolinguals (as existing research has also found; e.g. Kehoe et al., 2004).

This study is not without its limitations. The phonology subtest of the BESA is an assessment tool intended for clinical purposes (i.e. to test each phoneme in initial, medial and final positions of words and to elicit that production at the frequency that it occurs in the target language) and was not designed for research studies that require many types and tokens of each phoneme for English and Spanish. The use of a single word test that provides many more opportunities for voiceless stops in syllable initial position, and the examination of connected speech samples, might allow us to observe the differences in VOT within and across language groups with more clarity. This study was exploratory in nature and performed to determine feasibility for a larger study currently underway. Our next study will expand the number of participants, number of tokens and the analyses, including the full

range of voiceless and voiced stops in both languages. Albeit other studies have found that voicing lead was more marked and acquired later than short or long lag (Kehoe et al., 2004), it would still be beneficial to include them to obtain a more complete picture of stop VOT and voicing acquisition. Ideally, these studies should be performed longitudinally to observe the progression of the acquisition of voicing until it is complete. More lexical items with varied stress patterns and word lengths should be studied to investigate how VOT and lead voicing behave in different environments. In addition, due to the inherent difficulty in matching bilingual children on a number of variables, especially on percent input and output in both languages, a relatively small number of participants were included in this study. The small number of children included in the study required the use of non-parametric statistical analyses, which are not as powerful as parametric statistics. With a larger number of participants and more powerful parametric statistical analyses, we might be able to avoid type II error in statistical analysis and reduce the large amount of variability we found in this particular data set.

6. Conclusion

Monolingual and bilingual children differed on VOT in English, but not in Spanish (at least not significantly), indicating that the phonetic level of speech production, specifically VOT, might be a site of cross-language influence in bilingual children. The findings of this study support previous studies that have found cross-language influence in bilingual phonological acquisition (Paradis and Genesee, 1996; Paradis, 2001; Fabiano-Smith and Barlow, 2009; Fabiano-Smith and Goldstein, 2010). Specifically, these results support studies examining this phenomenon in bilingual German–Spanish-speaking children and English–Japanese-speaking children (e.g. Kehoe et al., 2004; Harada, 2007) and bilingual adults (e.g. MacLeod and Stoel-Gammon, 2009). Future studies should focus on how voicing and other phonological contrasts evolve in young bilingual children. The acquisition of such phenomena holds important clues to how the phonological systems of bilingual and monolingual children are organized and possibly re-organized. Such discoveries not only would have significant theoretical implications, but are also relevant for speech-language pathologists and educators. Knowledge of differences between bilingual and monolingual productions will aid in the accurate differentiation of language difference (i.e. differences in bilingual speech production due to cross-language influence) from language disorder (i.e. an underlying language-learning disability). This differentiation could reduce the number of bilingual children who are overdiagnosed as having a language disorder when they are, in fact, typically developing.

Acknowledgments

The authors thank the children and families who participated in their study. They also thank Donna Jackson Maldonado, Rosa Patricia Bárcenas Acosta and Martha Beatriz Soto Martínez at the Universidad Autónoma de Querétaro for their help with data collection and Brian Goldstein at Temple University for comments on an earlier version of the manuscript.

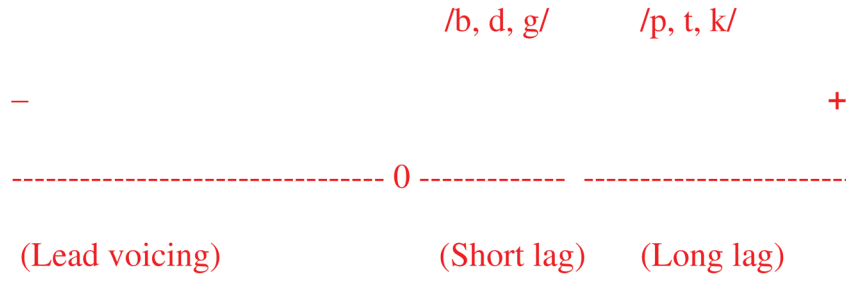
References

- Arnold, E.; Curran, C.; Miccio, A.; Hammer, C. Sequential and simultaneous acquisition of Spanish and English consonants. Poster presented at the Annual Convention of the American Speech, Language, and Hearing Association (ASHA); Philadelphia, PA. 2004.
- Bohn OS, Flege JE. The production of new and similar vowels by adult German learners of English. *Studies in Second Language Acquisition*. 1992; 14:131–158.
- Bond ZS, Eddey JE, Bermejo JJ. VOT del español to English: Comparison of a language-disordered and normal child. *Journal of Phonetics*. 1980; 8:287–290.

- Braunschweiler N. Integrated cues of voicing and vowel length in German: A production study. *Language and Speech*. 1997; 40:353–376.
- Bunta F, Fabiano-Smith L, Ingram D, Goldstein B. Phonological whole-word measures in three-year-old bilingual children and their monolingual peers. *Clinical Linguistics and Phonetics*. 2009; 23(2): 156–175. [PubMed: 19197583]
- Deuchar M, Clark A. Early bilingual acquisition of the voicing contrast in English and Spanish. *Journal of Phonetics*. 1996; 24(3):351–365.
- Fabiano-Smith L, Barlow JA. Interaction in bilingual phonological acquisition: Evidence from phonetic inventories. *International Journal of Bilingual Education and Bilingualism*. 2009; 1:1–17.
- Fabiano-Smith L, Goldstein B. Phonological acquisition in bilingual Spanish-English speaking children. *Journal of Speech, Language, and Hearing Research*. 2010; 53:1–19.
- Flege JE. Age of learning affects the authenticity of voice onset time (VOT) in stop consonants produced in a second language. *Journal of the Acoustical Society of America*. 1991; 89:395–411. [PubMed: 2002177]
- Flege, JE. Second-language speech learning: Theory, findings, and problems. In: Strange, W., editor. *Speech perception and linguistic experience: Issues in cross-language research*. Timonium, MD: York Press; 1995. p. 233-272.
- Flege JE, Eefting W. The production and perception of English stops by Spanish speakers of English. *Journal of Phonetics*. 1987; 15:67–83.
- Fowler C, Sramko V, Ostry D, Rowland S, Hallé P. Cross language phonetic influences on the speech of French–English bilinguals. *Journal of Phonetics*. 2008; 36(4):649–663. [PubMed: 19802325]
- Goldstein B, Fabiano L, Iglesias A. Spontaneous and imitated productions in Spanish-speaking children with phonological disorders. *Language, Speech, and Hearing Services in the Schools*. 2004; 35:5–15.
- Gutiérrez-Clellen V, Kreiter J. Understanding child bilingual acquisition using parent and teacher reports. *Applied Psycholinguistics*. 2003; 24:267–288.
- Harada T. The production of voice onset time (VOT) by English-speaking children in a Japanese immersion program. *International Review of Applied Linguistics in Language Teaching*. 2007; 45(4):353.
- Holt L, Lotto A, Diehl R. Auditory discontinuities interact with categorization: Implications for speech perception. *Journal of the Acoustical Society of America*. 2004; 116(3):1763–1773. [PubMed: 15478443]
- Johnson C, Wilson I. Phonetic evidence for early language differentiation: Research issues and some preliminary data. *International Journal of Bilingualism*. 2002; 6(3):271–289.
- Kehoe M, Lleó C, Rakow M. Voice onset time in bilingual German-Spanish children. *Bilingualism: Language and Cognition*. 2004; 7(1):71–88.
- Konefal JA, Fokes J. Voice onset time: The development of Spanish-English distinction in normal and language disordered children. *Journal of Phonetics*. 1981; 9:437–444.
- Ladefoged, P. *Phonetic data analysis: An introduction to fieldwork and instrumental techniques*. Malden, MA: Blackwell Publishing; 2003.
- Lisker L, Abramson A. A cross-language study of voicing in initial stops: Acoustical measurements. *Word*. 1964; 20:384–422.
- Lisker L, Abramson A. Voice onset time in English stops. *Language and Speech*. 1967; 10(1):1–28. [PubMed: 6044530]
- MacKay I, Flege JE, Imai S. Evaluating the effects of chronological age and sentence duration on degree of perceived foreign accent. *Applied Psycholinguistics*. 2006; 27:157–183.
- Macken M, Barton D. The acquisition of the voicing contrast in English: A study of voice onset time in word-initial stop consonants. *Journal of Child Language*. 1979; 7:41–74. [PubMed: 7372738]
- Macken M, Barton D. The acquisition of the voicing contrast in Spanish: A phonological study of word-initial stop consonants. *Journal of Child Language*. 1980; 7:433–458. [PubMed: 6969264]
- MacLeod A, Stoel-Gammon C. The use of voice onset time by early bilinguals to distinguish homorganic stops in Canadian English and Canadian French. *Applied Psycholinguistics*. 2009; 30(1):53–77.

- Maye J, Werker J, Gerken L. Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition*. 2002; 82:101–111.
- Meisel, J. The bilingual child. In: Bhatia, TK.; Richie, W., editors. *The handbook of bilingualism*. Maiden, MA: Blackwell Publishing; 2004.
- Paradis J. Do bilingual two-year-olds have separate phonological systems? *International Journal of Bilingualism*. 2001; 5:19–38.
- Paradis J, Genesee F. Syntactic acquisition in bilingual children: Autonomous or interdependent? *Studies in Second Language Acquisition*. 1996; 18:1–25.
- Pearson B, Fernández S, Lewedeg V, Oller DK. The relation of input factors to lexical learning by bilingual infants. *Applied Psycholinguistics*. 1997; 18:41–58.
- Peña E, Bedore L, Rappazzo C. Comparison of Spanish, English, and bilingual children's performance across semantic tasks. *Language, Speech, & Hearing Services in the Schools*. 2003; 34:5–16.
- Peña E, Gutiérrez-Clellen V, Iglesias A, Goldstein B, Bedore L. *Bilingual English Spanish Assessment (BESA)*. (in press).
- Peterson GE, Lehiste I. Duration of syllable nuclei in English. *The Journal of the Acoustical Society of America*. 1960; 32(6):693–703.
- Restrepo MA. Identifiers of predominantly Spanish-speaking children with language impairment. *Journal of Speech, Language, and Hearing Research*. 1998; 41:1398–1411.
- Riney T, Takagi N, Ota K, Uchida Y. The intermediate degree of VOT in Japanese initial voiceless stops. *Journal of Phonetics*. 2007; 35(3):439–443.
- Shah, A. Doctoral Dissertation. The City University of New York; New York: 2002. Temporal Characteristics of Spanish-accented English: Acoustic measures and their correlation with accentedness ratings. *Dissertation Abstracts International*, 63-09B, 4311
- Sjölander, K.; Beskow, J. [Accessed 7 December 2010] WaveSurfer (version 1.8.8p2) [computer software]. 2010. from: <http://www.speech.kth.se/wavesurfer/download.html>

English stops



Spanish stops

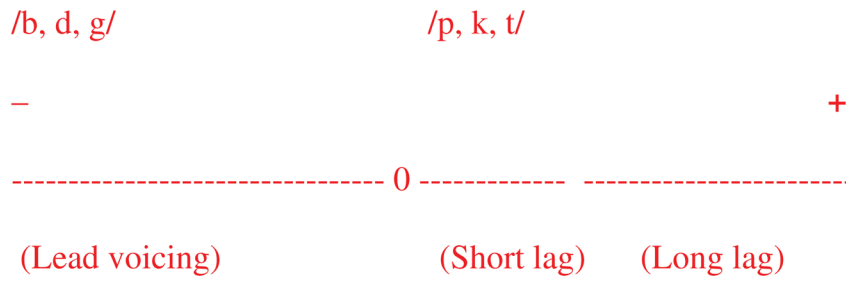


Figure 1.
VOT continuum for English and Spanish from Deuchar and Clark (1996).

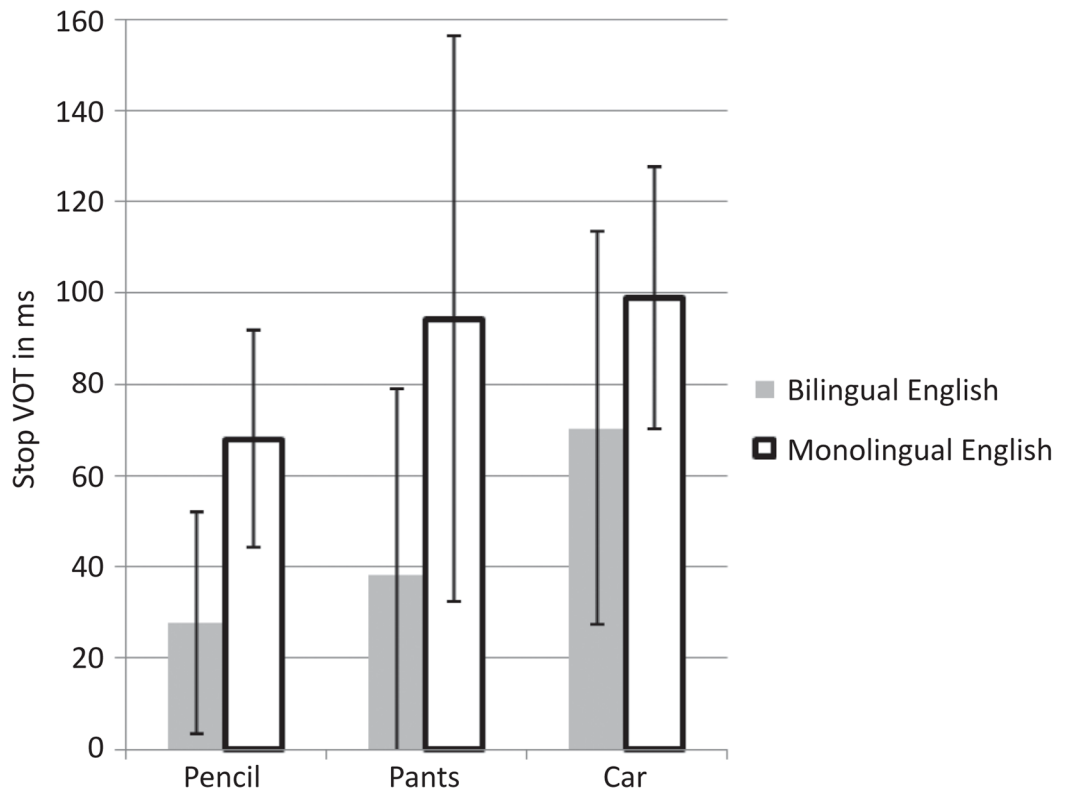


Figure 2.
English VOT values for /p/ and /k/.

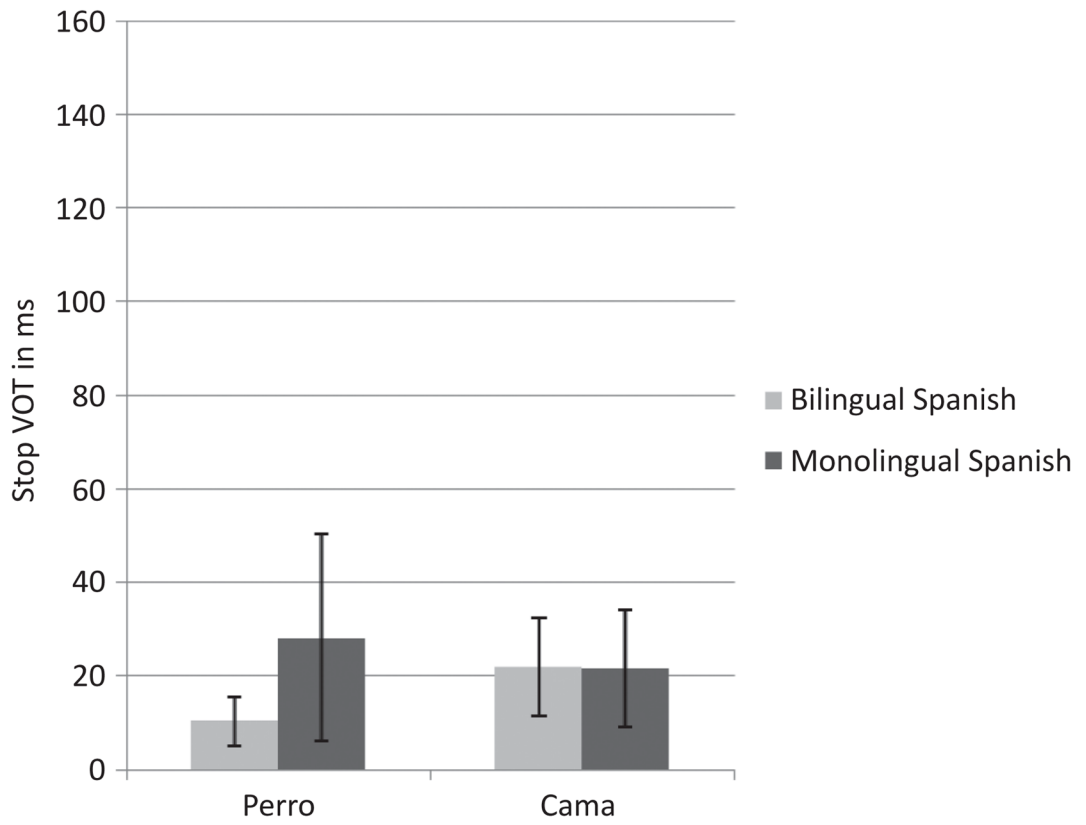


Figure 3.
Spanish VOT values for /p/ and /k/.

Table I

General VOT values for languages discussed in the bilingual acquisition literature.

Language	Voiced stops (ms)	Voiceless stops (ms)
English (Lisker and Abramson, 1967)	-45 to -64	33-48
Spanish (Lisker and Abramson, 1964)	-45 to -235	0-55
German (Braunschweiler, 1997)	16	51
Japanese (Riney, Takagi, Ota, and Uchida, 2007 for voiceless stops; Johnson and Wilson, 2002 for voiced stops)	-15 to -33	30-56

Note: VOT, voice onset time.

Table II

Participant background information.

Participants	Age	Gender	Mother's education	English input (%)	English output (%)	English proficiency ^a	Spanish proficiency ^a	Bilingual status
Bilingual								
B01	3:8	Male	High school	37	37	3	4	Sequential
B02	3:5	Male	High school	25	25	3	4	BFLA
B03	3:8	Male	Some college	42	70	3	3	Sequential
B04	3:5	Male	Undergraduate degree	50	80	4	3	BFLA
B05	3:8	Female	High school	60	80	4	3	BFLA
B06	3:4	Male	Some college	30	80	4	3	BFLA
B07	3:11	Male	High school	60	80	4	3	Sequential
B08	3:5	Male	Some college	20	50	3	3	BFLA
Monolingual English								
E01	3:3	Female	Graduate degree					
E02	3:1	Female	Graduate degree					
E03	3:11	Female	Graduate degree					
E04	3:8	Male	High school					
E05	3:0	Male	Undergraduate degree					
E06	3:0	Male	Undergraduate degree					
E07	3:1	Male	Undergraduate degree					
E08	3:7	Male	Graduate degree					
Monolingual Spanish								
S01	4:0	Male	Some high school					
S02	3:3	Male	Graduate degree					
S03	3:3	Male	High school					
S04	3:10	Male	Some high school					
S05	3:2	Female	Undergraduate degree					
S06	3:6	Female	Graduate degree					
S07	3:4	Female	Some high school					
S08	3:4	Male	Some high school					

Notes: BFLA = bilingual first language acquisition/simultaneous.

^d0 = cannot speak the indicated language; 1 = some expressive skills in the indicated language with frequent morphological, phonological and syntactic errors; 2 = can speak the indicated language with some errors in morphology, phonology and syntax; 3 = near-native-like proficiency in the indicated language with few errors; 4 = native-like proficiency in the indicated language.

Table III

Averages, medians and SDs of token measurement.

	Pencil	Pants	Car	Ferro	Cama
Average difference	1.86	0.93	2.07	1.47	1.33
Median difference	1	0.5	2	1	1
SD of difference	2.35	1.07	1.71	2.53	1.5

Note: SD, standard deviation.

Table IV

VOT values for /p/ and /k/.

	English VOT values mean (SD)			Spanish VOT values mean (SD)		
	<i>Pencil</i>	<i>Pans</i>	<i>Car</i>	<i>Perro</i>	<i>Cama</i>	
Spanish monolingual				10.4 (5.2)	21.9 (10.5)	
<i>n</i>				7	8	
Bilingual	27.7 (24.3)	38.1 (40.8)	70.4 (43.0)	28.1 (22.1)	21.7 (12.6)	
<i>n</i>	7	7	7	8	6	
English monolingual	68.0 (23.7)	94.3 (61.9)	99 (28.7)			
<i>n</i>	7	7	8			

Note: VOT, voice onset time; SD, standard deviation.