



Faculty Publications

2018

Constructing two phonological systems: A phonetic analysis of /p/, /t/, /k/ among early Spanish–English bilingual speakers

Earl K. Brown

Brigham Young University, ekbrown@byu.edu

Mary T. Copple

Kansas State University

Follow this and additional works at: <https://scholarsarchive.byu.edu/facpub>



Part of the [Linguistics Commons](#)

Original Publication Citation

Brown, Earl K. & Mary T. Copple. (2018). "Constructing two phonological systems: A phonetic analysis of /p/, /t/, /k/ among early Spanish-English bilingual speakers." *International Journal of Bilingualism*, 22(1): 51-68.

BYU ScholarsArchive Citation

Brown, Earl K. and Copple, Mary T., "Constructing two phonological systems: A phonetic analysis of /p/, /t/, /k/ among early Spanish–English bilingual speakers" (2018). *Faculty Publications*. 6552. <https://scholarsarchive.byu.edu/facpub/6552>

This Peer-Reviewed Article is brought to you for free and open access by BYU ScholarsArchive. It has been accepted for inclusion in Faculty Publications by an authorized administrator of BYU ScholarsArchive. For more information, please contact ellen_amatangelo@byu.edu.

Constructing two phonological systems: A phonetic analysis of /p/, /t/, /k/ among early Spanish–English bilingual speakers

International Journal of Bilingualism

2018, Vol. 22(1) 51–68

© The Author(s) 2016

Reprints and permissions:

sagepub.co.uk/journalsPermissions.nav

DOI: 10.1177/1367006916651983

journals.sagepub.com/home/ijb**Earl K. Brown and Mary T. Copple**

Department of Modern Languages, Kansas State University, USA

Abstract

Aims and objectives/purpose/research questions: Many early Spanish-English bilingual speakers in the USA learn Spanish as their first language at home and English in school. This paper seeks to elucidate whether these speakers develop a separate phonological system for English and, if so, the role of primary and secondary cues in the development of the second language (L2) system.

Design/methodology/approach: The phonetic realization of the voiceless stops /p/, /t/, /k/ is analyzed among three groups: early Spanish-English bilinguals; L1 English speakers who are late learners of Spanish; and L1 Spanish speakers who are late learners of English. The participants ($N = 15$) engaged in a reading task and a conversation task in each language during a single recording session.

Data and analysis: 1578 tokens of /p/, /t/, /k/ were extracted and analyzed using acoustic software. Voice onset time in milliseconds and center of gravity in Hertz were analyzed, and monofactorial and multifactorial analyses were performed to determine the role of linguistic background.

Findings/conclusions: Evidence is found of two phonological systems among early bilingual speakers, with varying degrees of assimilation to the phonological systems of the native speakers of each language.

Originality: We argue that early bilinguals construct their L2 system of /p/, /t/, /k/ in English based on the primary cue of voice onset time rather than the secondary cue of center of gravity, as they are accustomed to noticing differences in voice onset time in Spanish and because the center of gravity of /p/, /t/, /k/ in English is more variable than voice onset time, and therefore represents a more variable and less predictable cue for early bilinguals as they construct their L2 system.

Significance/implications: This paper contributes to the literature on the construction of phonological systems and to research detailing the speech of early Spanish-English bilinguals.

Corresponding author:

Earl K. Brown, Department of Modern Languages, Kansas State University, 104 Eisenhower Hall, Manhattan, KS 66506, USA.

Email: ekbrown@ksu.edu

Keywords

Voiceless stops, voice onset time, center of gravity, early bilinguals, second language phonological systems

Introduction*Models of second language sound acquisition*

Within second language (L2) research, discussion has centered on how two languages develop simultaneously in children, when or how two distinct systems emerge, and the effects of the first language (L1), age, and environment on sequential bilingualism (cf. Bhatia & Ritchie, 2012, chapters 4 and 5, for discussion). Usage-based models propose that language use affects the way in which speakers construct their mental representation of a language system. Simply put, language experience, both receptive and productive, shapes language knowledge and speakers' subsequent production; essentially the "substance of linguistic experience (as filtered through human cognition and attention)" shapes which categories are formed and how they evolve over time (Bybee, 2013, p. 68). Within this perspective, speakers' minds actively construct the necessary system based on their interpretation of their experiences with language and that constructed system evolves over time as additional evidence is noticed.

The current paper, which focuses on the construction of phonological systems in early bilinguals, is aligned with the view that construction occurs through successful processing of contextualized exemplars (i.e. input is converted to intake; cf. Bybee, 2013). Through hearing the articulation of a phone in various contexts, learners accumulate experience with its variants. This processing may be affected by the fact that memory "is necessarily partially abstract insofar as the experience is not recorded completely" (Goldberg, 2013, p. 27) and that during the subsequent analysis of that incomplete memory, selective attention biased toward previously mastered cognitive routines may be employed in distinguishing features (cf. Ellis, 2013; Pulvermüller, Cappelle, & Shtyrov, 2013). Presumably, in organizing their new L2 system, learners initially process input based on their previous L1 experiences. In so doing, their processing of L2 input may be somewhat selective, that is, that which is converted to intake and added to the developing system may be informed by an over-reliance on L1 knowledge and those features of a sound that condition its use in the L1 (cf. Ellis, 2013; Flege & Eefting, 1987). This over-reliance on L1 experience may lead to difficulties in mapping unfamiliar or imperfectly perceived L2 sounds. In attempting to account for such difficulties, Flege and colleagues (cf. Flege, 1995; Flege & MacKay, 2004) address the influence of L1 features in the Speech Learning Model (SLM), which proposes that learners construct new L2 sound categories if they perceive some difference between a new L2 sound and the closest L1 sound. However, if learners are not able to discern differences between these sounds, no new L2 category is formed and learners perceive the L2 sound as equivalent to the closest L1 sound, and would presumably rely on already-familiar cues.

Similarly, Best and colleagues' (cf. Best, 1995; Best & Tyler, 2007) Perceptual Assimilation Model (PAM) focuses on L2 sound contrasts and their relationship with the already established L1 system. It is posited that when two L2 sounds are mapped onto different L1 categories, it is easy for the learners to discriminate between the two L2 sounds. Contrarily, when the two L2 sounds are mapped onto a single L1 category, expectedly, discrimination is inhibited.

Since learners are adept at processing features important in distinguishing L1 variants, it is likely that those same features are privileged in attending to L2 input, thus making the features more accessible or worthy of attention in the L2 and making them more likely to be used as the

basis for initial category formation and classification. This process is aided if the learner's assumptions and deduction of patterns based on L1 perceptions seemingly hold true in the L2. In this case, L2 learners, having identified the use of a recognizable feature, can begin to analyze more carefully how that feature is used in particular contexts.

However, L2 systems are not constructed solely on the expectations held due to the L1, but are also influenced by the ever-evolving nature of the L2 categories themselves as more L2 exemplars are accumulated in a variety of contexts, additional features are noticed, and L2 experiences are more clearly distinguished from previous L1 experiences. As learners accumulate more L2 experience, it presumably becomes easier for them to notice features distinct to the L2, thus redirecting and broadening their attention to include notice of features previously ignored, possibly reshaping their L2 phonological system. As a result of these competing influences, two possibilities are present: (1) more refined categorization may emerge from long-term exposure as systems evolve or (2) initial hypotheses, even if incomplete, may be re-confirmed.

It has been proposed that during this construction process predictability and reliability play a key role in ease of abstraction of patterns of use (cf. Bybee, 2013; Ellis, 2002, 2013). Learners are more easily able to distinguish prototypical exemplars from outliers when sounds are realized in a particular way, thus allowing them to more fully map a feature's use. Features that exhibit a more narrow range of variation are likely easier to identify and classify than features that are more variable. Widdison (1991) discusses the possibility of primary cues outweighing secondary cues in decoding input. These secondary cues would not necessarily be processed as consistently as the primary, but instead would come to the fore when serving to "disambiguate a sound whose principal cues either conflict or have been degraded" (Widdison, 1991, p. 92). In short, successful learning "depends crucially upon the salience of the cue and the importance of the outcome" (Ellis, 2013, p. 371). Presumably, successful learning results from attending to and interpreting cues and, crucially, the need to do so based on whether those cues are necessary to approximate native speech.

Previous research on early bilinguals supports the notion that early bilinguals approximate native articulation. Larsen-Freeman and Long (1991) found that children acquire native or native-like phonological patterns if they begin exposure to an English-speaking environment (e.g. immigration to an English-speaking country) before age seven. Antoniou, Best, Tyler, and Kroos (2010) likewise found that early sequential bilinguals exhibit phonological patterns nearly identical to monolinguals. Similarly, McCarthy, Mahon, Rosen, and Evans (2014) found that while children initially exhibit patterns from their L1, their L2 speech evolves to approximate that of their native-speaking peers.

The current study seeks to examine the effects of the core issues of predictability and reliability, selective attention to recognized cues, and importance of outcome in the construction of a L2 sound system by examining the production of voiceless stops (/p/, /t/, /k/) in early Spanish-English bilinguals in order to determine if shared primary cues in the L1 and L2 affect the attention paid to secondary cues that differ across languages.

Voice onset time among Spanish-English bilingual speakers

Voice onset time (VOT) is viewed as a primary cue in the production of voiceless stop consonants. It refers to the duration between the release of air from the lungs and the beginning of the vibration of the vocal folds of the following voiced segment, whether a consonant or a vowel. VOT differs among stops and across languages. In general, it has been reported that /p/ has the shortest VOT, with gradual VOT lengthening as the place of articulation moves back in the mouth, with /k/ having the longest VOT (Docherty, 1992; Lisker & Abramson, 1964; Yavas, 2002). Similarly, VOT of /p/,

/t/, /k/ is lengthened when followed by a high vowel. When referring to VOT of /p/, /t/, /k/ in Spanish and in English, the use of “short lag” and “long lag”, respectively, by researchers highlights the large differences between these languages (e.g., Abramson & Lisker, 1973; Flege & Eefting, 1987; Zampini & Green, 2001).

VOT among Spanish-English bilinguals is a well studied area in linguistic variation, likely owing to the fact that the two languages differ greatly in the articulation of these sounds, as noted above. The contrast in VOT has been studied in spontaneous speech (cf. Piccinini & Arvaniti, 2015), in spontaneous code-switches (cf. Balukas & Koops, 2015; Lopez, 2012; Yavas & Byers, 2015), in language switching during a picture-naming task (cf. Olson, 2013), and in child speech (cf. Fabiano-Smith & Bunta, 2012). In addition, it has been shown that phonemic boundaries of consonants, as informed by VOT, differ by language among Spanish-English bilinguals (cf. Garcia-Sierra, Diehl, & Champlin, 2009). Age of acquisition of Spanish and English (cf. Thornburgh & Ryalls, 1998) also affects success in acquiring patterns of VOT, with learners who began learning English before age 12 displaying more native-like English articulation than later bilinguals.

Findings generally support the long- versus short-lag values of English and Spanish, yet success by early bilinguals in establishing separate systems seems to be more mixed. Magloire and Green (1999), as well as Antoniou et al. (2010), found that early bilingual speakers exhibit evidence of separate representation and are nearly identical to monolinguals in each language, while Flege and Eefting (1987) found that early bilinguals establish intermediate values in one or both languages. L2-dominant speakers may still exhibit influence of L1 patterns (Antoniou, Best, Tyler, & Kroos, 2011) or even show loss of L1 VOT patterns as L2 proficiency grows (Major, 1992).

The present study

The present study of voiceless stops in the speech of early bilinguals raises several questions: if one phonetic feature is secondary to another, is it less likely to be incorporated into the developing L2 phonological system? What features make one’s speech native-like? How are these “native or native-like” systems constructed? Does attention to primary cues lead to native-like system construction or must one also attend to secondary cues?

In an effort to delve into the idea of native-like proficiency and to analyze the effects of reliability and predictability, selective attention, and importance of outcome on learners’ responses to primary and secondary cues, this study examines the production of voiceless stops /p/, /t/, /k/ via the analysis of the primary cue of VOT and the secondary cue of the center of gravity (COG). Most previous studies have focused solely on VOT. COG measures the acoustic energy, as measured in Hertz, of the friction between the release of the stop and the start of voicing of the following segment (Boersma & Weenink, 2014). Strong friction returns higher measurements of COG, while weakened aspiration gives lower COG measurements. In measuring both variables, this study seeks to determine to what extent VOT and COG are part of the phonological systems of /p/, /t/, /k/ of the early bilingual speakers when producing both languages.

Specifically, the current paper seeks to answer the following questions:

RQ1: Does the VOT of /p/, /t/, /k/ of early Spanish-English bilinguals approximate that of English speakers in English and that of Spanish speakers in Spanish?

RQ2: Does the COG of /p/, /t/, /k/ of early Spanish-English bilinguals approximate that of English speakers in English, and that of Spanish speakers in Spanish?

RQ3: Do the early bilinguals’ phonetic realizations of /p/, /t/, /k/ provide evidence of two phonological systems, in comparison to control groups?

Data and methods

Participants

The 15 participants in this study were divided into three groups: (1) five early Spanish-English bilinguals, hereafter referred to as “early bilinguals”; (2) a control group of five native speakers of Kansas and Missouri English who are late learners of Spanish as a L2, hereafter referred to as “English speakers”; and (3) a control group of five native speakers of Ecuadorian Spanish who are late learners of English as a L2, hereafter referred to as “Spanish speakers”.¹ The English speakers and their parents were raised in Kansas or Kansas City, Missouri. All were university students studying in an intermediate-level Spanish conversation course at the time of the recording sessions. The Spanish speakers were raised in Ecuador, as were their parents, and were recorded while in the USA during a 7-month stay as visiting scholars of English at the same university. Both of these groups of speakers began their L2 language study in adolescence, and were less proficient in their L2. The data from these two groups served to represent native usage in their respective languages as well as learner usage in their L2, with which the early bilinguals’ production might be compared. The early bilingual speakers were university students, aged 19–23 years, who self-identified as native speakers of Spanish and L2 speakers of English on a survey designed to collect demographic information for the study. Four were born in the USA and one arrived at 19 months of age. All were raised in Kansas in a Spanish-speaking home and were first formally exposed to English in a school setting (e.g. pre-school or kindergarten), although it is likely that they were exposed via other means (e.g. television) prior to those early educational experiences. All had spent varying amounts of time visiting Mexico, the country where their parents were raised.

Data collection

The researchers conducted a single recording session with each informant that included four tasks. The recording session took place in a sound booth, with a desk-mounted Audio-Technica AT2020 microphone connected to a laptop computer, recording directly into the editing software Audacity (Audacity Team, 2013). Firstly, each informant read a text in his L1. After reading the text, which lasted between 2 and 3 minutes, the informant conversed with a researcher in the dominant language for 10–15 minutes about topics related to a university setting, such as the informants’ courses, how their hometown compared to the city in which the university was located, plans for breaks in school, current job, professional plans, etc. This portion of the recording session approximated a sociolinguistic interview (cf. Labov, 1984), as the purpose was to elicit casual and spontaneous speech rather than to discuss any specific topic. The third portion of the recording session consisted of the informant reading a different text in the other language. This portion lasted on average 2–5 minutes, depending on the ease with which the informants read in that language. The final portion of the recording session was a 10–15-minute spontaneous conversation in this language, again employing the techniques of the sociolinguistic interview. The texts for the reading tasks were chosen based on the relatively large number of tokens of /p/, /t/, /k/ and because the topics were familiar to the participants. The Spanish-language text dealt with the creation of new types of bread when the Spaniards invaded what are now Spanish-speaking countries in the Americas (adapted from Blitt, Casas, & Copple, 2015, p. 84). The English-language text dealt with television and its history and importance in modern culture (adapted from “Television” in Wikipedia).

Coding of data

The phonetics software Praat (Boersma & Weenink, 2014) was used to code an average of 105 tokens ($SD = 4$) from each of the 15 participants. While viewing both the waveform and the

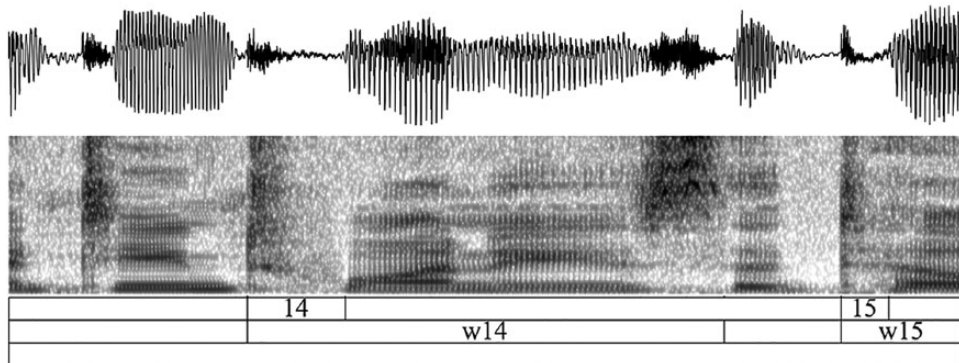


Figure 1. Waveform and spectrogram of English “twenties”, produced by English speaker 1.

spectrogram, the burst of air associated with the release of the stop and the subsequent periodic noise associated with the start of voicing of the following voiced segment were manually delimited in a TextGrid. After this manual delimitation process, a Praat script automated the extraction of the dependent variables: VOT and COG. Firstly, the script measured the duration of VOT. Secondly, in order to accurately calculate the COG, the script eliminated the lowest 300 Hz of each token with a pass Hann band filter, as voicing substantially lowers the COG of aspiration, and voicing is predominantly manifested in the lowest frequencies (Silbert & de Jong, 2008). While these stops should not have had any voicing, as they are “voiceless” stops, the decision to only use frequencies above 300 Hz eliminated the possibility of a masking effect on the COG by any flanking or aberrant voicing during the articulation of these stops. See Figure 1 for an example of the coding process in Praat with English “twenties” produced by an English speaker. The interval labeled “14” in the first tier below the spectrogram delimits the aspiration between the release of the stop and the beginning of voicing of the following segment. The interval labeled “w14” in the second tier delimits the duration of the word “twenties”. The “14” is the identification number of the specific token of /p/, /t/, or /k/.

In addition to the two dependent variables (VOT and COG), various independent factors were analyzed. Of primary concern in this study was the linguistic background of the speaker. However, other independent variables were also included in order to measure the importance of linguistic background when compared with other factors. These included the stop, the preceding and following phonological contexts, prosodic stress, the task during which the stop was produced, speaking rate, any previous mention of the word with the stop, cognate status of the word, and lexical frequency of the word.

The linguistic background of the speakers was coded as either early bilingual, English speaker (late L2 Spanish), or Spanish speaker (late L2 English). The specific stop was coded as /p/, /t/, or /k/. The preceding phonological context was categorized as high vowel, non-high vowel, obstruent, sonorant, or pause. The following phonological context was classified as high vowel, non-high vowel, or sonorant (/l/, /r/). The task in which the stop was produced was coded as either reading or conversation. The local speaking rate immediately near each token was calculated by dividing the duration of the word in which the stop occurred by the number of segments in that word. The need to control for speaking rate around each token is due to the fact that speaking rate can fluctuate during a conversation or reading task, and would likely have an effect on VOT and COG (cf. Magloire & Green, 1999). It should be noted that the stop itself was excluded from the calculation of the local speaking rate in order to avoid circularity in the prediction of the stop. In order to

control for a possible priming effect, lexical items were coded for any previous mention in the recording session. The lexical frequency of the words was taken from two corpora: the Corpus del Español (Davies, 2002) and the Corpus of Contemporary American English (Davies, 2008). The frequencies of the words produced during the Spanish reading task and conversation were taken from the 20th century of the Corpus del Español, from the written sections and the oral section, respectively. Similarly, the written and oral sections of the Corpus of Contemporary American English were used to obtain the frequencies of the words produced during the English reading task and conversation. Each frequency was then normalized to a frequency per million words. It should be noted that in the regressions reported below, rather than enter the normalized frequency numbers themselves into the models, a logarithm of base 10 of the normalized frequencies was entered in order to assuage the huge disparities between the outlier frequencies of the most frequent words and all other words (cf. Gries, 2013, p. 254). Cognates were identified by whether there were more similarities than differences in phones and graphemes as well as semantic relatedness of the two lexical items. This allowed for the inclusion as cognates of pairs such as *common* – *común* “common” and false cognates such as *college* – *colegio* “high school” (cf. Amengual, 2012; Brown & Harper, 2009; Mora & Nadeu, 2009).

The independent variables mentioned so far were entered into the mixed effects linear regressions reported below as fixed effects. In addition, the individual speaker and the specific word type in which the stops occurred were entered as random effects in the statistical models.

The coding process began with 1651 tokens; however, some tokens were excluded for several reasons. Some tokens exhibited no aspiration between the release of the stop and the start of voicing, as the voicing started immediately when the stop was released ($N = 63$).² The majority of these tokens were produced in the English preposition *to* ($N = 52$). Several tokens were excluded because they were unnaturally elongated ($N = 9$). Finally, one token of English *to* was excluded because the aspiration of the stop constituted the entirety of the word, as the rest of the word was elided. These exclusions left 1578 tokens from 296 word types (150 word types in English, 146 in Spanish) for the analyses reported below.

Results

Voice onset time

When producing English and Spanish, the VOT of /p/, /t/, /k/ of early bilinguals closely mirrors that of the English speakers and Spanish speakers, respectively (see Figure 2).

The left-hand side of the figure shows the VOT of /p/, /t/, /k/ when all three groups of speakers produce English. The median duration of each group is represented by the dark horizontal line in the corresponding box, in the inner-most portion between the notches. The notch on the box that represents the English speakers’ production and the notch on the box of the early bilinguals overlap, while neither overlaps with the notch on the Spanish speakers’ box. This is *prima facie* support that the median VOT of the English speakers and that of the early bilinguals are not significantly different, but that the median VOT of both groups is significantly different from that of Spanish speakers (cf. Gries, 2013, p. 127). Specifically, while producing English, the median VOT of the English speakers is 52 milliseconds ($SD = 23$), while the median of the early bilinguals is 54 milliseconds ($SD = 21$). As the data violate both the assumption of normality and the assumption of variance homogeneity, instead of the more common *t*-test, a *U*-test was calculated (cf. Gries, 2009, p. 209). The *U*-test shows that the difference between the VOT of the two groups is not significant ($W = 33,154$, $p_{\text{two-tailed}} = 0.3$). Conversely, while producing English, the median VOT of the Spanish speakers is 35 milliseconds ($SD = 24$). The result of a second *U*-test, this time between the early

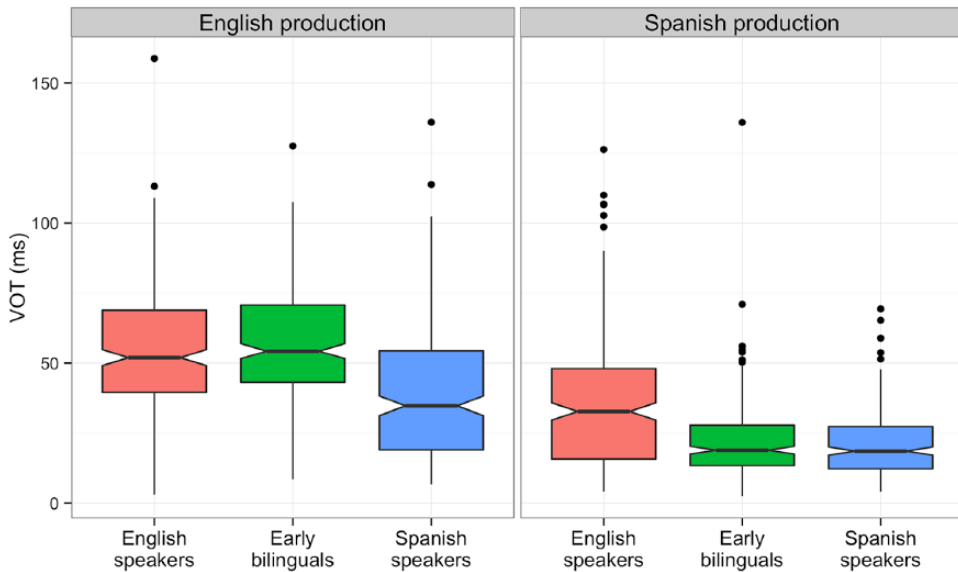


Figure 2. Voice onset time (VOT) of /p/, /t/, /k/ by linguistic background, grouped by language of production.

bilinguals and the Spanish speakers, shows that the difference in VOT of these two groups is significant ($W = 48,766$, $p_{\text{two-tailed}} \leq 0.001$). In summary, these monofactorial analyses suggest that, when producing English, the VOT of early bilingual speakers mirrors that of English speakers, but the VOT of Spanish speakers does not.

In order to measure if linguistic background still exerts a significant conditioning effect when all independent factors are considered, a multifactorial mixed effects model in which VOT was entered as the dependent variable was run. Linguistic background was still selected as significant, as were other factors.³ With the reference level set to English speakers, the model did not select early bilingual speakers as significantly different from English speakers, while the Spanish speakers were significantly different (see Table 1). This is viewed as further evidence that VOT has been closely approximated by the early bilingual speakers.

The early bilingual speakers also mirror the Spanish speakers in Spanish. The right-hand side of Figure 2 shows the VOT of the three groups while producing Spanish. The notches on the boxes that represent the durations of early bilinguals and Spanish speakers overlap but do not overlap with the notch on the box representing the English speakers' VOT. Specifically, while producing Spanish, the median VOT of Spanish speakers is 19 milliseconds ($SD = 11$), as is the median VOT of early bilingual speakers ($SD = 13$). A U -test verifies that the difference between these two groups is not significant ($W = 36,285$, $p_{\text{two-tailed}} = 0.5$). In contrast, the median VOT of the English speakers, when producing Spanish, is 33 milliseconds ($SD = 23$). A second U -test between early bilingual speakers and English speakers shows that the difference is significant ($W = 47,728$, $p_{\text{two-tailed}} \leq 0.001$).

In the mixed effects model with the tokens produced in Spanish, the reference level of the linguistic background variable was Spanish speaker. In this model, the VOT of early bilinguals was not significantly different from that of Spanish speakers, while the VOT of English speakers was significantly different (see Table 2).⁴

Table 1. Minimal adequate model of mixed effects linear regression of voice onset time of /p/, /t/, /k/ produced in English.*Random effects*

Group	Name	Variance	SD
Word	(Intercept)	0.0001	0.01
Speaker	(Intercept)	0.0001	0.009
Residual		0.0003	0.08

Fixed effects

Variable	Estimate	SE	df	t	p
(Intercept)	0.056	0.006	37.146	10.032	0
Stop: /p/	-0.013	0.003	108.684	-3.937	0
Stop: /t/	0.007	0.003	92.333	2.136	0.035
Speaker: early bilingual	0.001	0.006	12.038	0.121	0.905
Speaker: Spanish speaker	-0.019	0.006	12.183	-3.186	0.008
Speaking rate	0.082	0.015	723.091	5.423	0
Lex. freq. (logged)	-0.002	0.001	64.276	-3.142	0.003
Stress: tonic	0.01	0.003	92.051	3.377	0.001

Reference levels of categorical fixed effects: stop = /k/; speaker: English speaker

Non-significant variables: preceding phonological context, following phonological context, task, previous mention, cognate status

Log likelihood: 1948.87; $df = 11$; restricted maximum likelihood (REML): -3897.7

Table 2. Minimal adequate model of mixed effects linear regression of voice onset time of /p/, /t/, /k/ produced in Spanish.*Random effects*

Group	Name	Variance	SD
Word	(Intercept)	0.0001	0.0099
Speaker	(Intercept)	0.0001	0.0074

Fixed effects

Variable	Estimate	SE	df	t	p
(Intercept)	0.031	0.005	37.124	6.472	0.000
Stop: /p/	-0.008	0.003	64.012	-2.994	0.004
Stop: /t/	-0.005	0.003	60.821	-1.907	0.061
Speaker: English speaker	0.014	0.005	12.022	2.927	0.013
Speaker: early bilingual	0.002	0.005	12.001	0.383	0.708
Following: non-high vowel	-0.007	0.003	58.979	-2.254	0.028
Following: sonorant	0.004	0.004	64.172	0.952	0.344

Reference levels for categorical fixed effects: stop = /k/; speaker = Spanish speaker; following = high vowel

Non-significant variables: preceding phonological context, task, speaking rate, previous mention, lexical frequency, cognate status

Log likelihood: 2260.48; $df = 10$; restricted maximum likelihood (REML): -4520.95

In summary, both monofactorial and multifactorial analyses suggest that the VOT of early bilingual speakers mirrors that of English speakers when producing English but mirrors that of Spanish speakers when producing Spanish.

Center of gravity

In contrast to the results of VOT, the COG of /p/, /t/, /k/ of early Spanish-English bilingual speakers always mirrors the COG of Spanish speakers, regardless of the language of production (see Figure 3).

The left-hand side of the figure displays the COG when all three groups produce English. As can be seen, the median COG of early bilingual speakers (2541 Hz, $SD = 1427$) and that of Spanish speakers (2482 Hz, $SD = 1509$) are similar, while that of English speakers is substantially higher (3149 Hz, $SD = 1694$). Further, the notches on the boxes representing the COG of early bilingual speakers and that of the Spanish speakers overlap, while these notches do not overlap with the notch on the box representing the English speakers. A U -test between early bilingual speakers and Spanish speakers finds that the difference between their respective COG is not significant ($W = 36,008$, $p_{\text{two-tailed}} = 0.2$). The result of a second U -test between early bilingual speakers and English speakers finds that the difference in COG of these two groups is significant ($W = 41,265$, $p_{\text{two-tailed}} \leq 0.001$).

The multifactorial mixed effects model of COG when all three groups produce English paints a similar picture (see Table 3). With the reference level for linguistic background set to English speaker, the mixed effect model indicates that the COG of both early bilinguals and Spanish speakers is significantly different from that of English speakers.⁵

The right-hand side of Figure 3 shows the COG of the three groups in Spanish. The notches on the box that represent the early bilingual speakers' production and those on the box that represents the Spanish speakers' production almost completely overlap, while the notches of the early bilingual speakers and the English speakers display only a slight amount of overlap. The median COG of early bilingual speakers (1551 Hz, $SD = 922$) and that of Spanish speakers (1574 Hz, $SD = 943$) are similar, while that of English speakers is higher (1719 Hz, $SD = 1278$). A U -test between early bilinguals and Spanish speakers finds that the difference between their respective COG measurements is not significant ($W = 33,871$, $p_{\text{two-tailed}} = 0.5$). However, a second U -test between early bilingual speakers and English speakers finds that the difference in COG of these two groups is significant ($W = 42,941$, $p_{\text{two-tailed}} \leq 0.001$).

The multifactorial mixed effects model of COG in Spanish returns similar results. Among the independent variables selected as significant in the minimal adequate model is linguistic background.⁶ With the reference level set to Spanish speakers, the mixed effect model indicates that the COG of early bilinguals is not significantly different from that of Spanish speakers, but that the COG of English speakers is significantly higher (see Table 4).

To summarize, both monofactorial and multifactorial tests show that the COG of early Spanish-English bilingual speakers mirrors that of Spanish speakers when producing English or Spanish, and that the COG of English speakers is significantly different when producing either language.

Importantly, comparison of the early bilinguals' COG in Spanish with their values in English further shows that the early bilinguals exhibit a significantly higher COG in English (median = 2541 Hz, $SD = 1427$) than in Spanish (median = 1551 Hz, $SD = 922$). The results of a mixed effects model of the tokens of /p/, /t/, /k/ produced by early bilingual speakers in which the original two random effects of speaker and word were entered, along with only one fixed effect variable, that of language of production, show that this difference is significant (language of production reference level = Spanish: coefficient estimate = -706.42 , $SE = 172.26$, $df = 142.56$, $t = -4.1$, $p \leq 0.001$).

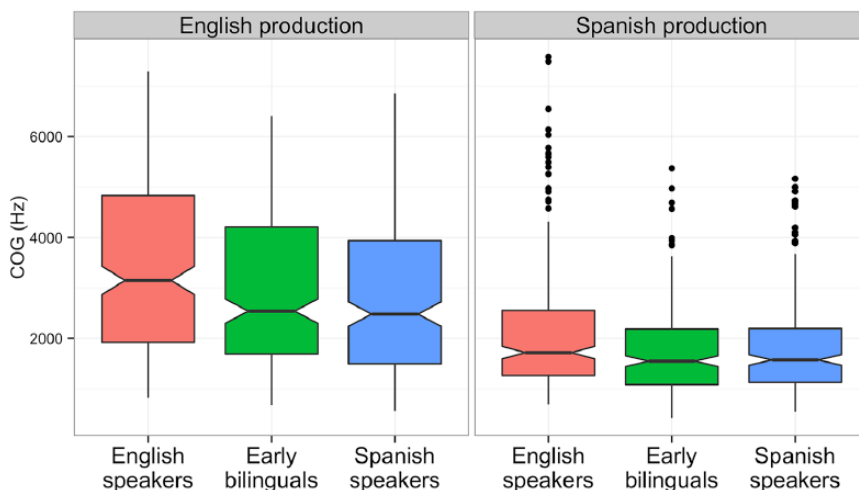


Figure 3. Center of gravity (COG) of /p/, /t/, /k/ by linguistic background, grouped by language of production.

Table 3. Minimal adequate model of mixed effects linear regression of center of gravity of /p/, /t/, /k/ produced in English.

Random effects

Group	Name	Variance	SD
Word	(Intercept)	162,261	402.8
Speaker	(Intercept)	85,575	292.5

Fixed effects

Variable	Estimate	SE	df	t	p
(Intercept)	2905.52	228.963	55.436	12.69	0.000
Stop: /p/	-188.589	153.61	98.545	-1.228	0.222
Stop: /t/	2190.78	143.6	81.295	15.256	0.000
Speaker: early bilingual	-528.04	202.513	12.04	-2.607	0.023
Speaker: Spanish speaker	-867.826	203.211	12.207	-4.271	0.001
Following: non-high vowel	-140.879	190.698	91.142	-0.739	0.462
Following: sonorant	-439.268	214.264	98.793	-2.05	0.043
Cognate: yes	-421.858	130.086	99.03	-3.243	0.002

Reference levels for categorical fixed effects: stop = /k/; speaker = English speaker; following = high vowel; cognate = no

Non-significant variables: preceding phonological context, task, speaking rate, previous mention, lexical frequency

Log likelihood: -6423.47; $df = 11$; restricted maximum likelihood (REML): 12,846.94

These results confirm that the early bilinguals have created an intermediate value in English for COG that is clearly distinguished from the value they use in Spanish.

Table 4. Minimal adequate model of mixed effects linear regression of center of gravity of /p/, /t/, /k/ produced in Spanish.

<i>Random effects</i>					
Group	Name	Variance	SD		
Word	(Intercept)	213,062	462		
Speaker	(Intercept)	25,175	159		
<i>Fixed effects</i>					
Variable	Estimate	SE	df	t	p
(Intercept)	1554.75	133.713	64.811	11.628	0.000
Stop: /p/	-103.861	127.695	125.404	-0.813	0.418
Stop: /t/	1361.22	143.674	117.945	9.474	0.000
Speaker: English speaker	319.914	121.258	12.993	2.638	0.020
Speaker: early bilingual	-18.113	119.201	12.172	-0.152	0.882
Sec. per segment	1364.14	525.726	743.02	2.595	0.010
Cognate: Yes	-309.383	119.523	129.184	-2.588	0.011

Reference levels for categorical fixed effects: stop = /k/; speaker = Spanish speaker; cognate = no
 Non-significant variables: preceding and following phonological contexts, task, previous mention, lexical frequency
 Log likelihood: -6378.53; *df* = 10; restricted maximum likelihood (REML): 12,757.06

Discussion

The first research question (RQ1) concerns whether the primary cue of VOT of /p/, /t/, /k/ of early Spanish-English bilingual speakers is similar to the VOT of native speakers of English and Spanish in their respective languages. The results provide evidence that the early bilinguals adjust VOT to approximate native speech in both languages, offering support for the native-like results of previous studies (Antoniou et al., 2010; Larsen-Freeman & Long, 1991; Magloire & Green, 1999; McCarthy et al., 2014; Thornburgh & Ryalls, 1998).

The second research question (RQ2) explores whether the COG, a secondary cue, of early Spanish-English bilingual speakers is similar to that of native speakers of English and Spanish in their respective languages. In English and Spanish, the early bilingual speakers' COG mirrors that of the Spanish speakers and is significantly different from that of the English speakers. This finding parallels results found in previous studies, including the retention of L1 patterns even if the speakers are L2 dominant (Antoniou et al., 2011) or the establishment of intermediate values (Flege & Eefting, 1987). To offer an explanation for these seemingly contradictory responses to the first two research questions, we turn to discussion of the third research question (RQ3): the possibility of two phonological systems in early Spanish-English bilingual speakers.

We propose that the early Spanish-English bilingual speakers distinguish two systems, but base their L2 phonological system of /p/, /t/, /k/ in English primarily on VOT, rather than on COG. This begs the question: Why are early Spanish-English bilingual speakers more able to replicate VOT than COG when constructing their L2 system of /p/, /t/, /k/ pronunciation in English? We return to the aforementioned concepts of selective attention, variability, reliability and predictability, and importance. We posit that VOT is a more important cue (cf. MacWhinney, 2001) when categorizing sounds within the larger system for early bilingual Spanish-English speakers precisely because

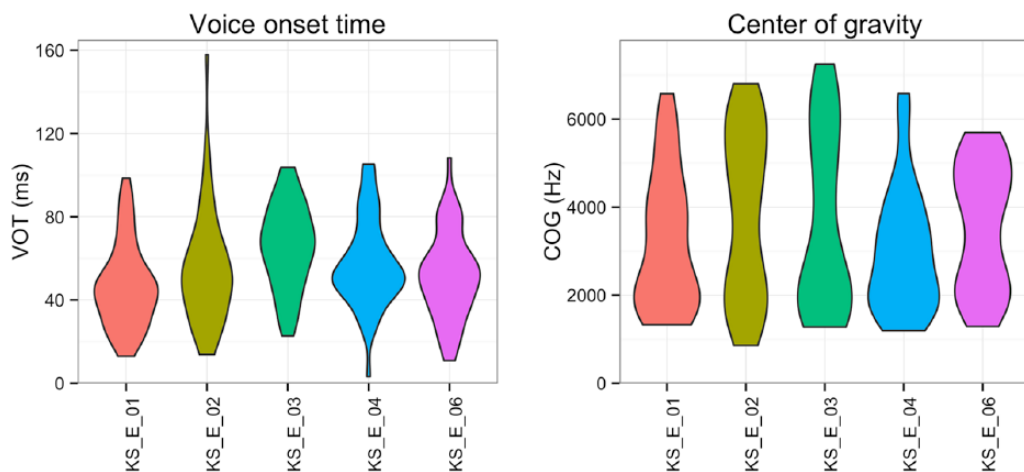


Figure 4. Voice onset time (VOT) and center of gravity (COG) of /p/, /t/, /k/ among English speakers producing English.

it is more prominent for them, drawing selective attention from their initial exposures, and displays less variation across speakers, making it a more reliable and predictable feature. The heightened prominence of VOT is most likely due to the early bilingual speakers' accumulated experience in Spanish to varying VOT realizations across the three stops, as well as in various contexts (initial versus medial position). Since they have previously noticed and deduced VOT patterns in Spanish, their attention to this feature in English from the beginning of their experience was presumably higher and already refined (cf. Flege, 1995; Flege & MacKay, 2004). That is, they would predict that /p/ in initial position would be shorter than /t/ or /k/ based on their Spanish experience, and they would be accustomed to perceiving this distinction and be able to determine if their hypotheses were true (cf. Swain, 2005). In this case, relevant previous experience and the resulting selective attention in the L2 combine to accurately analyze a variable (i.e., it is important) that behaves similarly in the L2 and that is highly predictable, for example, word initial = long lag in English (Ellis, 2013; Pulvermüller, 2013). In the case of VOT, L1 patterns allow for more accurate assessment of L2 patterns and native-like performance results from accumulated exposure for the early bilingual speakers.

In contrast, COG was presumably much harder for the early bilingual speakers to analyze and incorporate into their L2 system, although they have arrived at an intermediate value with higher COG measurements in English than in Spanish. We propose that this intermediate value is due to COG being a secondary cue: it is simultaneously less likely to receive selective attention, more variable and therefore less predictable across speakers, and less important for native-like production (Ellis, 2013; Pulvermüller et al., 2013). Regarding selective attention, COG is not a distinguishing feature of their L1 /p/, /t/, /k/, so attention is less likely to be selectively focused upon it from the outset (Widdison, 1991). In addition, if attention were eventually redirected to COG, it would prove to be a less reliable cue than that of VOT: production of COG varies more widely across the native speakers of English than does VOT. Figure 4 displays a violin plot (a plot that combines a boxplot and a density plot) for each native speaker of English for both VOT and COG when producing English. Figure 5 displays the same for the native speakers of Spanish when producing Spanish.

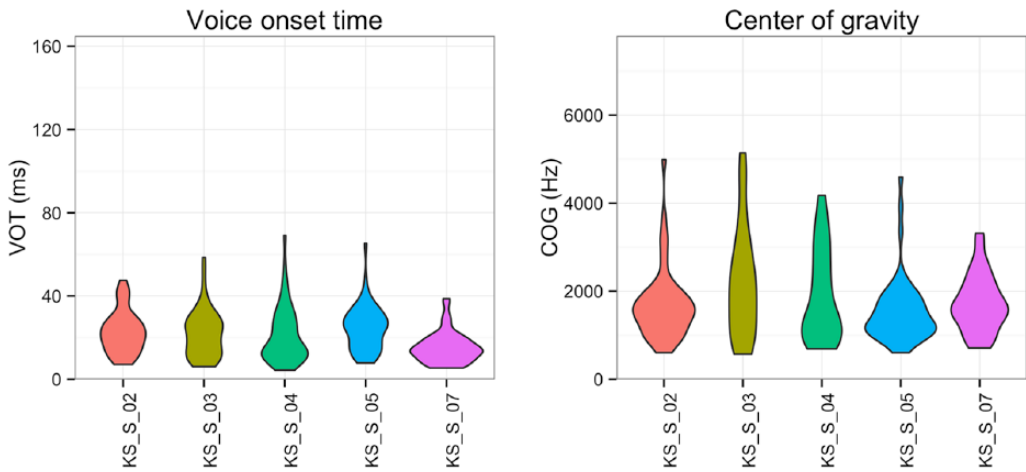


Figure 5. Voice onset time (VOT) and center of gravity (COG) of /p/, /t/, /k/ among Spanish speakers producing Spanish.

As seen on the left-hand side of both Figures 4 and 5, the violin shapes for VOT of speakers in both languages are relatively uniform, indicating that the variability is smaller and therefore leads to greater predictability and reliability. In contrast, the violin shapes for COG, on the right-hand side of both figures, are more variable for the English speakers (Figure 4) while those for the Spanish speakers are more uniform in shape and size (Figure 5). Assuming that the English speakers in this study are a representative sample of English speakers in general, this variability in COG likely makes it more difficult to analyze and distinguish prototypical exemplars from outliers (Bybee, 2013; Ellis, 2002, 2013; Montrul, 2013). The wide range of “acceptable” values among native speakers of English presumably makes it harder to deduce a pattern for COG production and also suggests that accurate assessment of COG is not necessarily needed for L2 speakers to be considered native-like, as much variability exists even among native speakers. In this case, unpredictability coupled with the lack of importance of the outcome and diminished selective attention leads to imperfect “learning” (cf. Ellis, 2013).

In general, the predictability of a variant makes its analysis easier and repetition in a particular context makes its representation stronger. If that analysis is initially based on L1 criteria and seemingly proves correct during L2 production, that analysis is prioritized and reinforced (cf. Bybee, 2013). This is the case with the early bilingual speakers’ native-like performance of VOT in both languages. However, with COG, multiple factors influence the early bilingual speakers’ incomplete analysis and different realization in English: (1) L1 experience suggests that it would not be relevant, thus it is less likely to be noticed than VOT when first attending to input and constructing the system; (2) due to this, attention is not selectively focused on it when analyzing L2 input; (3) in addition, it is highly variable and is therefore not as reliable a cue as VOT once, and if, attention is later focused on it as the system evolves with accumulated exposure. The high predictability of and accumulated experience with VOT lead to a diminished need to analyze COG as a key feature of these phones. The resulting COG in English represents intermediate values that are different from the early bilinguals’ L1 Spanish COG as well as from the English speakers’ L1 COG. However, as the median COG among the early bilinguals falls within the lowest ranges of the English speakers’ values, it is presumed that these intermediate values are likely interpreted as native-like, and therefore contribute to their L2 phonological system in English.

Conclusion

When learning a L2, speakers naturally draw on their accumulated experiences of learning and speaking their L1 and the system that they have already constructed and use on a daily basis. Features that proved to be variable yet helpful and predictable in the L1 are more likely to be prioritized in the learner's processing of input in the L2, and therefore an attempt is likely made to identify any patterns of variation of those features. If a particular feature played an important role in distinguishing variants in the L1, it is likely to be considered important when processing L2 input, whereas secondary cues that contribute to variation but that do not invite selective attention or are widely variable may go unnoticed or be analyzed imperfectly. In this case, these early bilinguals learned English after their Spanish phonological system was established. As such, they have been less successful in mimicking their L2's COG in the production of /p/, /t/, /k/. However, and importantly, their English phonological system does reflect some recognition of this feature's role in English /p/, /t/, /k/, as they clearly raise COG when changing from Spanish to English, despite the fact that they do not achieve native-like COG.

We do not suggest that, in general, VOT is necessarily more perceptible than COG with /p/, /t/, /k/. Rather, we propose that learners hear, process, and categorize what they have learned to consider important, and the phonological system is consequently constructed based on their deductions that, at least initially, draw heavily from L1 experiences. Learning new features takes time and may be incomplete, particularly if a feature is overshadowed by others judged by speakers to be primarily responsible for the variation. In this case, the early bilingual speakers have privileged VOT over COG, but presumably do not stand out as non-natives for it. They would therefore have less incentive to try to "fix" the COG issue, as they can pass as native with native-like VOT but intermediate COG.

Exposure to variation eventually leads to abstraction of patterns and the differentiation of outliers from prototypical exemplars. We assume that exposure to more speakers in a variety of contexts leads to better access to and analysis of variation. In this study, the early bilingual speakers, who have successfully constructed two systems, have had more consistent and cumulative experience with native accents in both languages than the other two groups of speakers, and perhaps also have a stronger motivation for mastery of each language, as they are long-term members of the communities in which these languages are spoken.

Acknowledgements

We thank Alan Brown, Jenny Dumont, and Mary Kohn for their feedback on an earlier version of this paper as well as the two anonymous reviewers. Also, and especially, we thank the participants for their time and willingness to take part in this project.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

Notes

1. We recognize that most early Spanish-English bilingual speakers in the USA can be considered native speakers of both Spanish and English. Likewise, we recognize that the other two groups of speakers are late sequential Spanish-English bilinguals. However, for ease of expression, in this paper we utilize "early bilingual speakers", "English speakers", and "Spanish speakers" to distinguish the three groups.

2. We recognize that tokens with no aspiration can be considered outliers. However, in order to analyze VOT as well as COG, we include only tokens that had measurable duration between the release of the stop and the start of voicing, as COG cannot be measured if there is no aspiration present.
3. Due to space constraints, the full results of this model will not be discussed, but appear in Table 1. The other variables selected as significant were the specific stop, speaking rate, and lexical frequency. Specifically, English /p/ has a significantly shorter VOT than /k/, while /t/ does not have a significantly shorter VOT than that of /k/. As expected, as speaking rate slows, the VOT of /p/, /t/, /k/ increases (cf. Magloire & Green, 1999). Finally, VOT of /p/, /t/, /k/ is shorter as the frequency of the word in which these sounds occur increases.
4. The other independent variables selected as significant were the specific stop and the following phonological context. The stop /p/ is significantly shorter than /k/, whereas /t/ is not significantly shorter than /k/. The VOT of /p/, /t/, /k/ when followed by a non-high vowel is significantly shorter than when followed by a high vowel. However, the VOT of /p/, /t/, /k/ when followed by a consonant, specifically the liquids /l/ and /r/, is not significantly different.
5. Also selected as significant were the specific stop, the following phonological context, and the cognate status of the word with /p/, /t/, /k/. Specifically, /t/ has a significantly higher COG than the reference level /k/, while /p/ is not significantly different from /k/. The COG of the stops when followed by a sonorant, either /l/ or /r/, is significantly lower than the COG of the stops when followed by a high vowel. However, there was no difference between the COG of stops when followed by a high vowel or a non-high vowel. In comparison to non-cognate words, cognate words have a significantly lower COG when producing English.
6. Also selected as significant were the specific stop, the speaking rate, and the cognate status of the word with /p/, /t/, /k/. Specifically, /t/ has a significantly higher COG than the reference level /k/, while /p/ is not significantly different from /k/. As expected, as speaking rate slows, COG increases. In comparison to non-cognate words, cognate words have a significantly lower COG when produced in Spanish.

References

- Abramson, A. S., & Lisker, L. (1973). Voice-timing perception in Spanish word-initial stops. *Journal of Phonetics*, 1, 1–8.
- Amengual, M. (2012). Interlingual influence in bilingual speech: Cognate status effect in a continuum of bilingualism. *Bilingualism: Language and Cognition*, 15, 517–530.
- Antoniou, M., Best, C. T., Tyler, M. D., & Kroos, C. (2010). Language context elicits native-like stop voicing in early bilinguals' productions in both L1 and L2. *Journal of Phonetics*, 38, 640–653.
- Antoniou, M., Best, C. T., Tyler, M. D., & Kroos, C. (2011). Inter-language interference in VOT production by L2-dominant bilinguals: Asymmetries in phonetic code-switching. *Journal of Phonetics*, 39, 558–570.
- Audacity Team. (2013). Audacity: A free multi-track audio editor and recorder (Version 2.0). Retrieved from <http://web.audacityteam.org> (accessed 15 November 2013).
- Balukas, C., & Koops, C. (2015). Spanish-English bilingual voice onset time in spontaneous code-switching. *International Journal of Bilingualism*, 19(4), 423–443.
- Best, C. T. (1995). A direct realist view of cross-language speech perception. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 171–203). Baltimore, MD: York Press.
- Best, C. T., & Tyler, M. (2007). Nonnative and second-language speech perception: Commonalities and complementarities. In M. J. Munro, & O. Bohn (Eds.), *Second language speech learning: The role of language experience in speech perception and production* (pp. 13–24). Amsterdam, The Netherlands: John Benjamins.
- Bhatia, T. K., & Ritchie, W. C. (Eds.). (2012). *The handbook of bilingualism and multilingualism* (2nd ed.). Hoboken, NJ: Wiley-Blackwell.
- Blitt, M. A., Casas, M., & Copple, M. T. (2015). *Exploraciones: Curso intermedio*. Boston, MA: Cengage Learning.

- Boersma, P., & Weenink, D. (2014). Praat: Doing phonetics by computer (Version 5.4.01). Retrieved from www.praat.org
- Brown, E. L., & Harper, D. (2009). Phonological evidence of interlingual exemplar connections. *Studies in Hispanic and Lusophone Linguistics*, 2, 257–274.
- Bybee, J. (2013). Usage-based theory and exemplar representations of constructions. In T. Hoffman, & G. Trousdale (Eds.), *The Oxford handbook on construction grammar* (pp. 49–69). Oxford, UK: Oxford University Press.
- Davies, M. (2002). Corpus del Español: 100 million words, 1200s-1900s. Retrieved from www.corpusdelespanol.org
- Davies, M. (2008). The Corpus of contemporary American English: 450 million words, 1990-present. Retrieved from corpus.byu.edu/coca
- Docherty, G. J. (1992). *The timing of voicing in English obstruents*. Berlin, Germany: Foris.
- Ellis, N. C. (2002). Frequency effects in language processing: A review with implications for theories of implicit and explicit language acquisition. *Studies in Second Language Acquisition*, 24, 143–188.
- Ellis, N. C. (2013). Construction grammar and second language acquisition. In T. Hoffman, & G. Trousdale (Eds.), *The Oxford handbook on construction grammar* (pp. 365–378). Oxford, UK: Oxford University Press.
- Fabiano-Smith, L., & Bunta, F. (2012). Voice onset time of voiceless bilabial and velar stops in 3-year-old bilingual children and their age-matched monolingual peers. *Clinical Linguistics & Phonetics*, 26(2), 148–163.
- Flege, J. E. (1995). Second language speech learning: Theory, findings, and problems. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 233–277). Timonium, MD: York Press.
- Flege, J. E., & Eefting, W. (1987). Production and perception of English stops by native Spanish speakers. *Journal of Phonetics*, 15, 67–83.
- Flege, J. E., & MacKay, I. (2004). Perceiving vowels in a second language. *Studies in Second Language Acquisition*, 26, 1–34.
- Garcia-Sierra, A., Diehl, R. L., & Champlin, C. (2009). Testing the double phonemic boundary in bilinguals. *Speech Communication*, 51(4), 369–378.
- Goldberg, A. (2013). Constructionist approaches. In T. Hoffman, & G. Trousdale (Eds.), *The Oxford handbook on construction grammar* (pp. 15–31). Oxford, UK: Oxford University Press.
- Gries, S. T. (2009). *Quantitative corpus linguistics with R: A practical introduction*. New York, NY: Routledge.
- Gries, S. T. (2013). *Statistics for linguistics with R: A practical introduction* (2nd ed.). Berlin, Germany: De Gruyter Mouton.
- Labov, W. (1984). Field methods of the project on linguistic change and variation. In J. Baugh, & J. Sherzer (Eds.), *Language in use: Readings in sociolinguistics* (pp. 28–53). Englewood Cliffs, NJ: Prentice Hall.
- Larsen-Freeman, D., & Long, M. H. (1991). *An introduction to second language acquisition research*. New York, NY: Longman.
- Lisker, L., & Abramson, A. S. (1964). A cross-language study of voicing in initial stops: Acoustical measurements. *Word*, 20, 384–422.
- Lopez, V. G. (2012). Spanish and English word-initial voiceless stop production in code-switched vs. monolingual structures. *Second Language Research*, 28(2), 243–263.
- MacWhinney, B. (2001). The Competition Model: The input, the context, and the brain. In P. R. Robinson (Ed.), *Cognition and second language instruction* (pp. 69–90). Cambridge, UK: Cambridge University Press.
- Magloire, J., & Green, K. P. (1999). A cross-language comparison of speaking rate effects on the production of voice onset time in English and Spanish. *Phonetica*, 56, 158–185.
- Major, R. C. (1992). Losing English as a first language. *The Modern Language Journal*, 76, 190–208.
- McCarthy, K., Mahon, M., Rosen, S., & Evans, B. (2014). Speech perception and production by sequential bilingual children: a longitudinal study of voice onset time acquisition. *Child Development*, 85, 1965–1980.

- Montrul, S. A. (2013). *El bilingüismo en el mundo hispanohablante*. Malden, MA: Wiley-Blackwell.
- Mora, J. C., & Nadeu, M. (2009). Experience effects on the categorization of a native vowel contrast in highly proficient Catalan-Spanish bilinguals. *Journal of the Acoustical Society of America*, 125, 2775.
- Olson, D. J. (2013). Bilingual language switching and selection at the phonetic level: Asymmetrical transfer in VOT production. *Journal of Phonetics*, 41(6), 407–420.
- Piccinini, P., & Arvaniti, A. (2015). Voice onset time in Spanish-English spontaneous code-switching. *Journal of Phonetics*, 52(September), 121–137.
- Pulvermüller, F., Cappelle, B., & Shtyrov, Y. (2013). Brain basis of meaning, words, constructions, and grammar. In T. Hoffman, & G. Trousdale (Eds.), *The Oxford handbook on construction grammar* (pp. 397–416). Oxford, UK: Oxford University Press.
- Silbert, N., & de Jong, K. (2008). Focus, prosodic context, and phonological feature specification: Patterns of variation in fricative production. *Journal of the Acoustical Society of America*, 123, 2769–2779.
- Swain, M. (2005). The output hypothesis: theory and research. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning* (pp. 471–484). Mahwah, NJ: Lawrence Erlbaum.
- Thornburgh, D. F., & Ryalls, J. H. (1998). Voice onset time in Spanish-English bilinguals: Early versus late learners of English. *Journal of Communication Disorders*, 31(3), 215–229.
- Widdison, K. A. (1991). *The phonetic basis for s-aspiration in Spanish*. PhD thesis. University of California, Berkeley.
- Wikipedia. (n.d.). Television. Retrieved from <http://en.wikipedia.org/wiki/Television>
- Yavas, M. (2002). Voice onset time patterns in bilingual phonological development. In F. Windsor, M. L. Kelly, & N. Hewlett (Eds.), *Investigations in clinical phonetics and linguistics* (pp. 341–350). Mahwah, NJ: Lawrence Erlbaum.
- Yavas, M., & Byers, E. (2015). Production of voiceless stops in early sequential Spanish-English bilinguals. In M. Yavas (Ed.), *Unusual productions in phonology* (pp. 242–260). New York: Psychology Press.
- Zampini, M. L., & Green, K. P. (2001). The voicing contrast in English and Spanish: The relationship between perception and production. In J. Nicol (Ed.), *One mind, two languages: Bilingual language processing* (pp. 23–48). Cambridge, UK: Blackwell.

Author biographies

Earl K. Brown received a PhD in Hispanic Linguistics in 2008 at the University of New Mexico and currently works as an Associate Professor of Spanish (Linguistics) at Kansas State University. His research centers on language variation in Spanish, especially its sound system. He frequently uses the programming language R to search for strings in text files, visualize data, and perform statistical analyses.

Mary T. Copple received a PhD in Hispanic Linguistics in 2009 at the University of New Mexico and currently works as an Associate Professor of Spanish (Linguistics) at Kansas State University. Her research centers on language variation and change in Spanish and L2 acquisition of Spanish.