

## The influence of language background and exposure on phonetic accommodation

Naomi Enzinna\*

**Abstract.** This study examines whether language background, short-term exposure to monolingual and bilingual speech, and long-term exposure to monolingual and bilingual speech influences speech accommodation. To address this question, I examine whether English monolinguals and Spanish-English bilinguals, either from a predominately monolingual community or a predominately bilingual community vary their speech when interacting with a monolingual English speaker versus a Spanish-English bilingual speaker. Additionally, I examine whether speakers are more likely to converge after being primed with monolingual English or Spanish-English bilingual speech. To test this, participants complete an interactive communication task, where they are presented with a 6x6 board on a computer screen and are asked questions about the words on the board, which contain variables that differ in English and Spanish. Results show that both language background and long-term exposure to monolingual or bilingual speech in a speaker's speech community influence accommodation.

**Keywords.** accommodation; voice onset time; exposure; priming; speech community; bilingualism; Miami, Spanish; English; referential communication task

**1. Introduction.** The demographics of the United States are quickly changing. According to a study by the U.S. Census, “By 2044, more than half of all Americans are projected to belong to a minority group (any group other than non-Hispanic White alone); and by 2060, nearly one in five of the nation's total population is projected to be foreign born” (Colby & Ortman 2015:1). Regarding language, this demographic shift could mean that more Americans will be bilingual or will speak a non-standard variety of English. Further, this change may foreshadow more variation in how people living in the U.S. speak due to increased contact between bilingual and monolingual English speakers. For this reason, it is imperative to study how increased exposure to bilingual speech influences language variation in both bilingual and monolingual speech.

From a sociolinguistic perspective, we know that speakers vary their speech for reasons related to accommodation (Beebe & Giles 1984, Giles et al. 1991) and audience design (Bell 1984, Bell 2001). According to speech accommodation theory, speakers will adjust their speech style in order to increase or decrease perceived social distance from a social group and/or to improve communication efficiency. According to audience design, how a speaker adjusts their speech will depend on their audience. Thus, a speaker will vary their speech based on (1) who the listener is and (2) how the speaker wants to be perceived by the listener.

For example, Hwang et al. (2015) examined whether Korean-English bilinguals would adjust their speech when speaking with a monolingual American-English speaker and a Korean-English bilingual. They found that, when speaking English with a monolingual English speaker, Korean-English bilinguals adjusted their speech in order to sound more English-like. When speaking English with a Korean-English bilingual, however, they did not adjust their speech to sound more bilingual/Korean, nor did they adjust their speech to sound more English-like.

---

\* I am grateful for the help I have received from Draga Zec, Sam Tilsen, Abby Cohn, Marie Huffman, Phillip M. Carter, the members of the Cornell Phonetics Lab, and the Cornell Linguistics Department.  
Author: Naomi Enzinna, Cornell University ([nre23@cornell.edu](mailto:nre23@cornell.edu)).

Therefore, depending on the listener, the speakers either adjusted or did not adjust their speech, depending on how they wanted to be perceived by that listener.

Additionally, Hwang et al. (2015) examined whether speakers would be more likely to produce English or Korean segments directly after being exposed to (i.e., primed by) those segments. Priming studies have shown that priming speakers with various elements of speech (e.g., fine-grained phonetic detail, morphemes, sentence structure, etc.) facilitates use of those elements (Enzina 2017). In Hwang et al. 2015, Korean-English bilingual speakers were more likely to produce an English segment when they were primed with that segment. However, when the Korean-English bilingual speakers were primed with Korean-like segments, they were not more likely to produce those segments than when they were not primed. This research shows that short-term exposure influences speech accommodation, but also that the effect of short-term exposure is determined by how a speaker wants to be perceived by their audience.

If this is the case, we might expect that speakers from different speech communities, where speakers have different language backgrounds, linguistic exposure, and social ties, aim to be perceived by their audience differently. Of concern to the present study are speakers who have had long-term exposure to either English monolingual or bilingual speech in their community. This is important to address because long-term exposure has been shown to influence production. For example, Sancier & Fowler (1997) compared the speech of a Brazilian Portuguese speaker after they spent several months in Brazil versus the United States. They found that the speaker had shorter VOTs in both their English and Brazilian Portuguese after spending time in Brazil. Additionally, we know that language contact can encourage the formation of new language varieties. For example, both monolingual English speakers and Spanish-English bilinguals in Miami, which has a majority Hispanic population, speak a variety of English with Spanish-influenced properties, such rhythm and pitch (Enzina 2015, 2016; Carter & Lynch 2015).

Therefore, if long-term exposure to speech influences production, speech accommodation is likely to be influenced as well. How long-term exposure influences accommodation, however, will be dependent on the speech that a speaker has been exposed to within their speech community. According to Bell (2001), “speakers draw on the range of linguistic resources available in their speech community to respond to different kinds of audiences” (145). In other words, a speaker’s linguistic experiences are an essential part of determining how a speaker adjusts their speech for different audiences. For example, speakers are more likely to converge with a listener who speaks the same dialect as them (Kim et al. 2011). This is because they have experience with that dialect and how they want to be perceived by speakers of that dialect. Assuming this, speakers who have had long-term exposure to monolingual speech in their community should be more likely to adjust their speech in order to accommodate to monolingual speech, and speakers who have had long-term exposure to bilingual speech in their community should be more likely to accommodate to bilingual speech.

With this in mind, in this study, I investigate whether language background, short-term exposure to monolingual/bilingual speech, and long-term exposure to monolingual/bilingual speech influences speech accommodation. To address this question, I examine whether English monolinguals and Spanish-English bilinguals, either from a predominately monolingual community (Ithaca, NY, which was 6.85% Hispanic in 2010)<sup>1</sup> or a predominately bilingual community (Miami, FL, which was 67.7% Hispanic in 2016)<sup>2</sup> vary their speech when interacting with a

---

<sup>1</sup> Census statistic retrieved from <http://censusviewer.com/city/NY/Ithaca>.

<sup>2</sup> Census statistic retrieved from <https://www.census.gov/quickfacts/fact/table/miamidadecountyflorida/POP060210>.

monolingual English speaker versus a Spanish-English bilingual speaker. Additionally, I examine whether speakers are more likely to converge after being primed with monolingual English or Spanish-English bilingual speech.

Specifically, I examine accommodation of Voice Onset Time (VOT) after word-initial voiceless stops, a phonetic feature that differs in English and Spanish: English has long lag VOTs, averaging around 60-120 milliseconds, whereas Spanish has short lag VOTs, averaging around 0 to 30 milliseconds (Yavaş & Byers 2014). Previous research on VOT in bilingual speech shows the possibility of L1 “interference”, where a bilingual’s L1 phonology causes the speaker to be unable to achieve monolingual-like VOTs (Flege 1992, 1995; Sancier & Fowler 1997). However, previous research also shows that L2 dominant bilinguals are able to achieve monolingual-like speech (Flege & MacKay 2004, Guion et al. 2000, Yavaş & Byers 2014). Therefore, the VOTs of the bilingual speakers may fall anywhere between that of a monolingual Spanish speaker and a monolingual English speaker. My expectation is that a bilingual speaker’s long-term exposure to bilingual or monolingual speech will influence their ability to produce monolingual English-like VOTs.

In the remainder of this paper, I discuss the following: research questions (Section 2), hypotheses (Section 3), research methods (Section 4), results (Section 5), conclusion and discussion (Section 6).

**2. Research Questions.** In this study, I aim to address the following three questions:

- *Is phonetic accommodation of VOT influenced by a speaker’s language background?* To address this question, I examine whether English monolinguals and Spanish-English bilinguals adjust their VOT when speaking with either a monolingual English speaker or a late Spanish-English bilingual.
- *Is phonetic accommodation of VOT influenced by long-term exposure to monolingual or bilingual speech in their speech community?* I compare accommodation in the speech of speakers from a majority English-monolingual community (Ithaca) and from a majority Spanish-English bilingual community (Miami).
- *Is phonetic accommodation of VOT influenced by short-term exposure to monolingual or bilingual speech?* I examine the influence of short-term exposure to English monolingual and Spanish-English bilingual speech by examining whether speakers adjust their VOT more when primed with a voiceless stop with either a short-lag or long-lag VOT.

**3. Hypotheses.** In response to the research questions presented in Section 2, I hypothesize the following:

- Bilinguals will accommodate to a bilingual speaker more than the monolinguals will, and the monolinguals will accommodate to a monolingual speaker more than the bilinguals will.
- Speakers from the majority bilingual community (Miami) will accommodate to a bilingual speaker more than speakers from the majority monolingual community (Ithaca) will, and speakers from the majority monolingual community will accommodate to a monolingual speaker more than the speakers from the majority bilingual community will.
- Speakers will accommodate more when they are primed, compared to when they are not primed, with a voiceless stop with either short-lag or long-lag VOT.

**4. Methods.** In this section, I discuss the following methodological components of this study: the participants (Section 4.1), the referential communication task (Section 4.2), the pre-recorded monolingual English speaker and Spanish-English bilingual voices used in the task (Section 4.3), the words used as stimuli (Section 4.4), the experimental procedures (Section 4.5), and the data processing procedures (Section 4.6).

4.1. PARTICIPANTS. There were 20 total participants in this study, with five in each of the following participant groups: (1) English monolinguals from Ithaca (M-Ith), (2) Spanish-English bilinguals from Ithaca (B-Ith), (3) English Monolinguals from Miami (M-Mia), and (4) Spanish-English Bilinguals from Miami (B-Mia). All participants from Ithaca had been living in Ithaca for at least a year at the time of the study. Also, all participants from Miami had been living in Miami for at least a year at the time of the study. All English monolinguals are monolingual English speakers who are self-reported monolinguals.<sup>3</sup> All Spanish-English bilinguals speak both Spanish and English fluently. All participants were between the ages of 18 and 35 (in other words, “millennials”) at the time of the study.

4.2. TASK. Participants completed a referential communication task, similar to the one used in Hwang et al. 2015. In the task, participants see a board with 6x6 squares (Figure 1b). In some of the squares there are words. Participants are asked about the words on the board by a pre-recorded monolingual English speaker or Spanish-English bilingual voice (henceforth, Monolingual Speaker and Bilingual Speaker, described further in Section 4.3). Participants are told that the Speakers have incomplete boards (Figure 1a); and in order to complete their boards, the Speakers are going to ask them what words are on their boards by referencing other words on the boards. The reference word is always next to the word being asked about, and is always in a square of the same color.

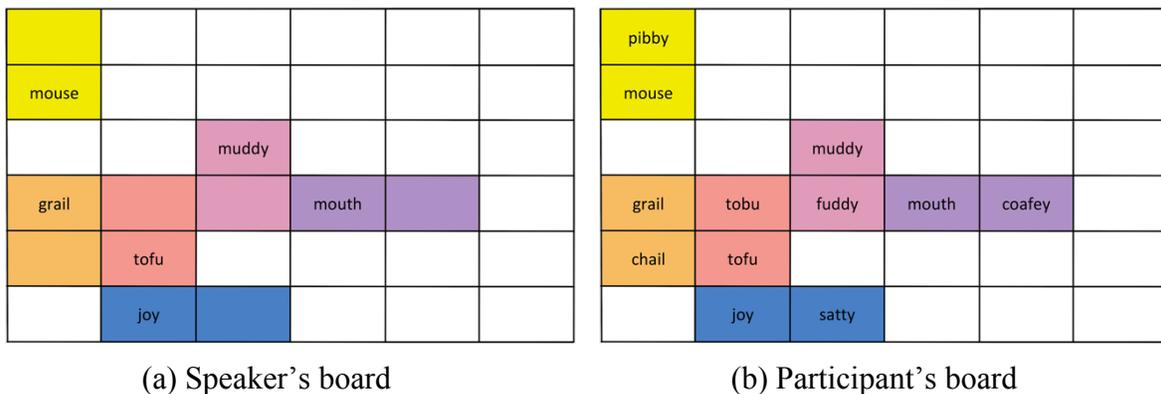


Figure 1. In the task, the participant helps the Speaker fill in the missing words on their board.

Using Figure 1 as an example, the Speaker asks, “What is by the word *mouse*?” Both *mouse* and *pibby* are in yellow squares and next to each other on the participant’s board, so the participant responds, “*Pibby* is by the word *mouse*.” After responding, the participant clicks on

<sup>3</sup> Many of the English monolinguals reported learning a second language at school, but they reported having a very limited vocabulary. Unsurprisingly, many of the Miami “monolinguals” reported using some Spanish in their daily lives, particularly at work. This is to be expected, as there is frequent interaction between monolinguals and bilinguals in Miami and it is a requirement of any job that involves speaking with other people to know some basic vocabulary. Still, I refer to these speakers as monolinguals because they identify as being monolingual in their language background questionnaire.

the square containing the answer (*pibby*). This triggers the pre-recorded voice to ask about another word on the board. Once the participant has been asked about all of the words on the board, a new board begins.

There is a total of 75 boards in the study: 3 for the practice trials, 36 with the Bilingual Speaker, and 36 with the Monolingual Speaker. This task is run in Matlab, and participants listen and respond to the Speakers over a headset with audio and recording capabilities.

4.3. PRE-RECORDED MONOLINGUAL SPEAKER AND BILINGUAL SPEAKER VOICES. Participants interacted with two pre-recorded voices: (1) a Spanish-English bilingual voice (“Bilingual Speaker”) and (2) a monolingual English speaker voice (“Monolingual Speaker”). The Bilingual Speaker is a 40-year-old male from Mexico City, Mexico. He started learning English in elementary school in Mexico, but did not begin speaking English regularly until moving to South Florida at age 30. The Monolingual Speaker is a 29-year-old male who had been living in Ithaca, New York, for five years and has lived the majority of his life in northeastern U.S. The Native and Bilingual Speakers had significantly different VOT durations for /p/, /t/, and /k/.

To examine this, for each voiceless stop (/p, t, k/), a one-way ANOVA was conducted to compare the effects of Speaker Type (Native or Bilingual) on VOT duration (in seconds). An analysis of variance showed that the effect of Speaker Type on VOT for /p/ was significant,  $F(1,31)$ ,  $p = .0002$ . Post hoc comparisons using the Tukey HSD test indicated that the Monolingual Speaker ( $M = 0.059$ ,  $SD = 0.029$ ) had longer VOTs for /p/ than the Bilingual Speaker ( $M = 0.034$ ,  $SD = 0.013$ ). Similarly, an analysis of variance showed that the effect of Speaker Type on VOT for /t/ was significant,  $F(1,31)$ ,  $p < .001$ . Post hoc comparisons using the Tukey HSD test indicated that the Monolingual Speaker ( $M = 0.072$ ,  $SD = 0.022$ ) had longer VOTs for /t/ than the Bilingual Speaker ( $M = 0.027$ ,  $SD = 0.007$ ). Last, an analysis of variance showed that the effect of Speaker Type on VOT for /k/ was significant,  $F(1,31)$ ,  $p = 0.05$ . Post hoc comparisons using the Tukey HSD test indicated that the Monolingual Speaker ( $M = 0.051$ ,  $SD = 0.012$ ) had longer VOTs for /k/ than the Bilingual Speaker ( $M = 0.041$ ,  $SD = 0.018$ ). A comparison of the Speakers’ VOTs is presented in Figure 2.

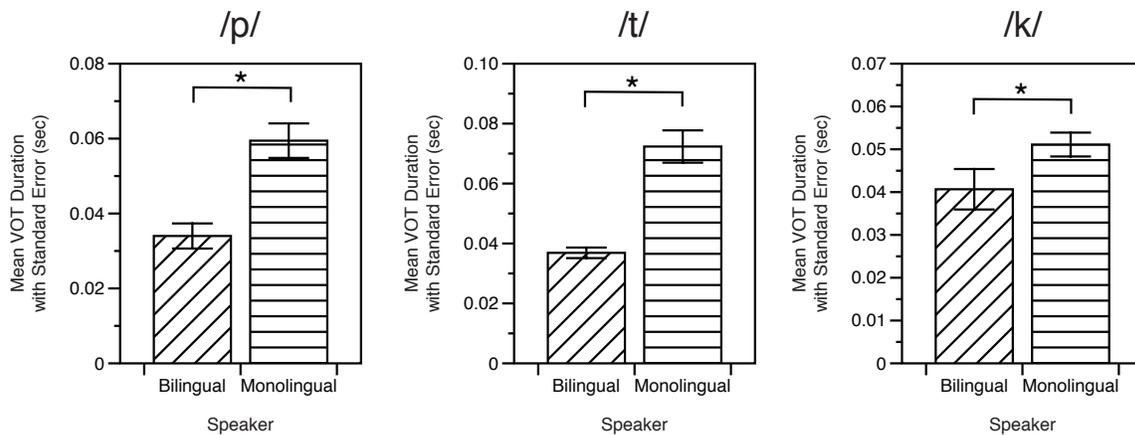


Figure 2. Comparison of the Speakers’ Voice Onset Times (sec) by voiceless stop

4.4. STIMULI WORDS. The words used in this study come in pairs (henceforth, “word pairs”). These word pairs consist of a prime word and a target word. There is a total of 234 word pairs in the study: 18 in the practice boards and 216 that occur once per pre-recorded Speaker. The word pairs were designed to allow for examination of the following target variables: duration of Voice

Onset Time (VOT) after a voiceless stop, velarization of word-final /l/, duration of intervocalic /t/ and /d/ (flapping), vowel quality differences for /ɪ ɛ æ ʌ i e ə o u/, rhythm, and pitch. These variables were selected because they differ in English and Spanish (VOT differences are described in Section 1). In this study, I will be examining VOT only. The remaining variables will be examined in future studies.

There are 108 target words total. All target words contain two of the aforementioned target variables each: one target consonant and one target vowel. 54 of the target words contain a voiceless stop. All of the target words are nonce words or very-low frequency words (in cases where there were no nonce options). The target words are the words missing from the Speakers' boards, which means the participant does not hear the Speakers say the target words. For example, for Figure 1, the Speaker asks, "What is by the word *mouse*?" and the participant responds, "*Pibby* is by the word *mouse*." In this example, the target word is *pibby*, which begins with a voiceless stop, contains vowel /ɪ/, and is a nonce word.

All target words occur once with a "target prime" and once with an "unrelated prime." The target primes contain the same target variables as the target word. The unrelated primes do not contain any of the target variables. All of the priming words are real words. The target primes are low-frequency words that share the same target vowel and consonant as the target word it is paired with, differing from the target word as little as a possible. For example, for the target word *tassy*, the target prime is *taffy*. It shares the vowel /æ/ and the consonant /t/, differing only in place of articulation for the second consonant. The unrelated primes are words that do not contain a target consonant or vowel, and word frequency is not restricted. For example, for the target word *tassy*, the unrelated prime is *roy*. Per board, half of the target words are primed with a target prime, and half are primed with an unrelated prime.

4.5. PROCEDURE. The procedure of the study is as follows: First, participants complete a consent form. Then, participants are seated in front of a laptop computer in a quiet room. Connected to the laptop computer is a headset, which has both audio and recording capabilities, and a mouse. At this time, participants read the instructions and are encouraged to ask questions. After the instructions, participants complete three practice boards. While they complete the practice boards, I listen and correct them if they make mistakes, and I answer any questions they have. Once the practice boards are complete and all their questions are answered, they begin the complete study.

Once the study begins, they hear one of the two pre-recorded voices and complete 36 boards with that voice. Halfway through, participants are allowed to take a break if they wish. There is no restriction on the amount of time they can take for their break, only that they cannot speak to anyone during this period. After the break, they complete the rest of the study (36 more boards) with the second voice. Which voice is heard first is counterbalanced across participants. After the task, they complete a language-background-and-attitudes survey.

4.6. DATA PROCESSING. Matlab was used to record all speech, board information (specifically, the responses the participant said in the order that they said them), and their click times. The click times and board information were saved in tables, which were then used with a Praat (Borsma & Weenink 2018) script to create TextGrids with boundaries after each response. Those TextGrids and their matching sound files, along with a dictionary containing all of the words used in the study and their pronunciation, were used with the Montreal Forced Aligner (MFA) (McAuliffe et al. 2017) to segment the speech. I used the MFA with a pre-trained acoustic model trained on English. It should be noted that this model does not segment aspiration separately from the preceding stop. Thus, in my dictionary, I added an /h/ (HH) after all target voiceless

stops, and the MFA segmented the aspiration as if it were an /h/. After the MFA aligned all of the speech, I checked the alignments for errors. After checking the alignments, the VOT durations were extracted from the TextGrids and analyzed in Matlab. How the data were compared and analyzed is described in Section 5.

**5. Results.** In this section, I compare overall VOT differences between participant groups (Section 5.1), overall VOT differences between participant groups by Speaker (Section 5.2), VOT differences between participant groups at the start of the study versus and at the end of the study (Section 5.3), and VOT change differences between groups (Section 5.4). I also examine the influence of short-term exposure (priming) on VOT (Section 5.5).

**5.1. OVERALL VOT DIFFERENCES BETWEEN GROUPS.** In this section, I examine the overall differences in VOT duration between groups. For each voiceless stop (/p, t, k/), a one-way ANOVA was conducted to compare the effects of participant group on VOT duration (in seconds).

An analysis of variance showed that the effect of participant group on VOT for /p/ was significant,  $F(3,1470)$ ,  $p < .001$ . Post hoc comparisons using the Tukey HSD test indicated that B-Ith ( $M = 0.075$ ,  $SD = 0.030$ ) had longer VOTs for /p/ than M-Ith ( $M = 0.067$ ,  $SD = 0.021$ ), M-Mia ( $M = 0.059$ ,  $SD = 0.022$ ), and B-Mia ( $M = 0.059$ ,  $SD = 0.001$ ). Also, M-Ith had longer VOTs for /p/ than B-Mia and M-Mia, but shorter VOTs than B-Ith. B-Mia and M-Mia did not differ in /p/ VOT.

An analysis of variance showed that the effect of participant group on VOT for /t/ was significant,  $F(3,1468)$ ,  $p < .001$ . Post hoc comparisons using the Tukey HSD test indicated that B-Ith ( $M = 0.075$ ,  $SD = 0.023$ ) and M-Ith ( $M = 0.075$ ,  $SD = 0.020$ ) had longer VOT for /t/ than B-Mia ( $M = 0.068$ ,  $SD = 0.019$ ). M-Mia ( $M = 0.072$ ,  $SD = 0.020$ ) did not differ in /t/ VOT from any group, and M-Ith and B-Ith did not differ from each other.

An analysis of variance showed that the effect of participant group on VOT for /k/ was significant,  $F(3,1464)$ ,  $p < .001$ . Post hoc comparisons using the Tukey HSD test indicated that B-Ith ( $M = 0.087$ ,  $SD = 0.025$ ) had longer VOTs for /k/ than M-Ith ( $M = 0.078$ ,  $SD = 0.022$ ), M-Mia ( $M = 0.072$ ,  $SD = 0.026$ ), and B-Mia ( $M = 0.069$ ,  $SD = 0.023$ ). Also, M-Ith had longer VOTs for /k/ than B-Mia and M-Mia, but shorter VOTs than B-Ith. B-Mia and M-Mia did not differ in /k/ VOT.

According to these results, both monolinguals and bilinguals from the majority monolingual community (M-Ith and B-Ith) had longer VOTs than both monolinguals and bilinguals from the majority bilingual community (M-Mia and B-Mia). In other words, speakers with long-term exposure to monolingual speech in their community had more monolingual-like VOTs, and speakers with long-term exposure to bilingual speech in their community had more bilingual-like VOTs. These results are illustrated in Figure 3.

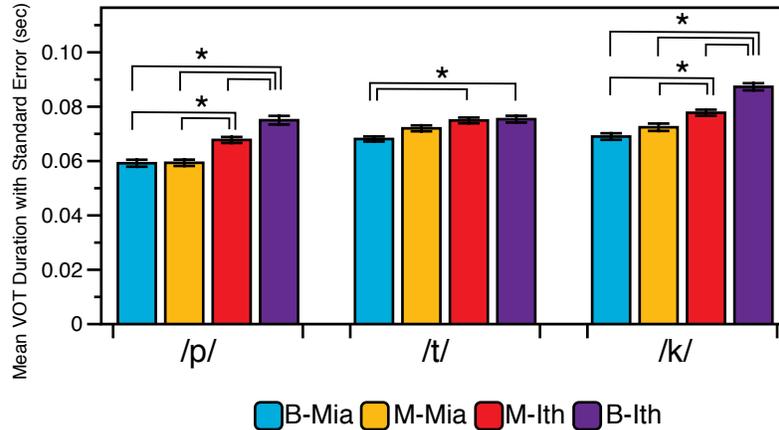


Figure 3. Participants from the monolingual community (M-Ith and B-Ith) had longer VOT durations than participants from the bilingual community (M-Mia and B-Mia).

5.2. OVERALL VOT DIFFERENCES BETWEEN GROUPS BY SPEAKER. In this section, I examine the overall differences in VOT duration between groups by Speaker, by conducting a two-way ANOVA to compare the effects of participant group and Speaker on VOT duration (in seconds).

A two-way analysis of variance yielded a main effect for participant group,  $F(3, 4287)$ ,  $p < 0.001$ . Post hoc comparisons using the Tukey HSD test indicated that B-Ith ( $M = 0.079$ ,  $SD = 0.027$ ) had significantly higher VOTs than all other groups, and M-Ith ( $M = 0.073$ ,  $SD = 0.021$ ) had higher VOTs than M-Mia ( $M = 0.067$ ,  $SD = 0.021$ ) and B-Mia ( $M = 0.065$ ,  $SD = 0.021$ ). M-Mia and B-Mia did not differ. The main effect of Speaker was also significant,  $F(1, 4287)$ ,  $p < 0.001$ ; participants' VOTs were significantly shorter for the Bilingual Speaker ( $M = 0.068$ ,  $SD = 0.024$ ) than the Monolingual Speaker ( $M = 0.074$ ,  $SD = 0.023$ ).

Also, there was a significant interaction between participant group and Speaker,  $F(3, 4287)$ ,  $p = 0.019$ . Post hoc comparisons using the Tukey HSD test indicated that B-Ith, M-Ith, and M-Mia had significantly higher VOTs when speaking with the Monolingual Speaker (B-Ith:  $M = 0.082$ ,  $SD = 0.028$ ; M-Ith:  $M = 0.076$ ,  $SD = 0.021$ ; M-Mia:  $M = 0.071$ ,  $SD = 0.020$ ) than the Bilingual Speaker (B-Ith:  $M = 0.075$ ,  $SD = 0.026$ ; M-Ith:  $M = 0.070$ ,  $SD = 0.021$ ; M-Mia:  $M = 0.063$ ,  $SD = 0.021$ ). B-Mia did not have different VOTs when speaking with the Monolingual Speaker ( $M = 0.066$ ,  $SD = 0.020$ ) versus the Bilingual Speaker ( $M = 0.064$ ,  $SD = 0.023$ ). Also, when speaking with the Monolingual Speaker, all participant groups' VOTs significantly differed from each other. B-Ith had the longest VOTs, followed by M-Ith, followed by M-Mia, and finally followed by B-Mia. When speaking with the Bilingual Speaker, B-Ith again had the longest VOTs, followed by M-Ith, followed by M-Mia and B-Mia. Notably, M-Mia and B-Mia did not significantly differ in overall VOT when speaking with the Bilingual Speaker.

According to these results, all participant groups other than the bilingual speakers from the bilingual community (B-Mia) produced longer VOTs when speaking with the Monolingual Speaker than the Bilingual Speaker. Also, the monolinguals from the bilingual community (M-Mia) were able to achieve bilingual-like VOTs (similar to the B-Mia) when speaking with the Bilingual Speaker. These results are illustrated in Figures 4 and 5.

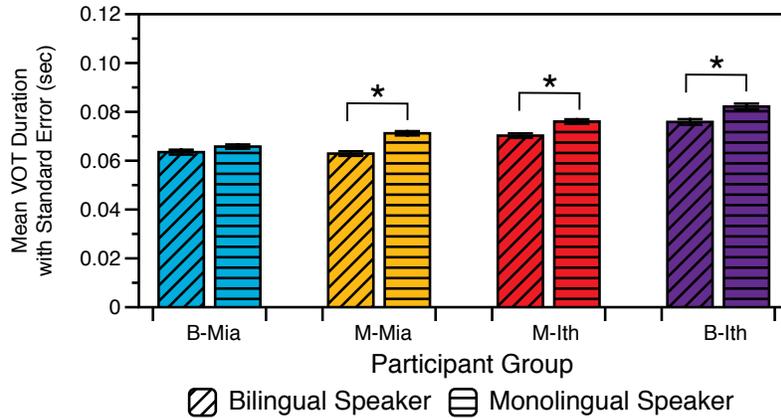


Figure 4. All participant groups other than bilinguals from the bilingual community adjusted their VOTs for either Speaker.

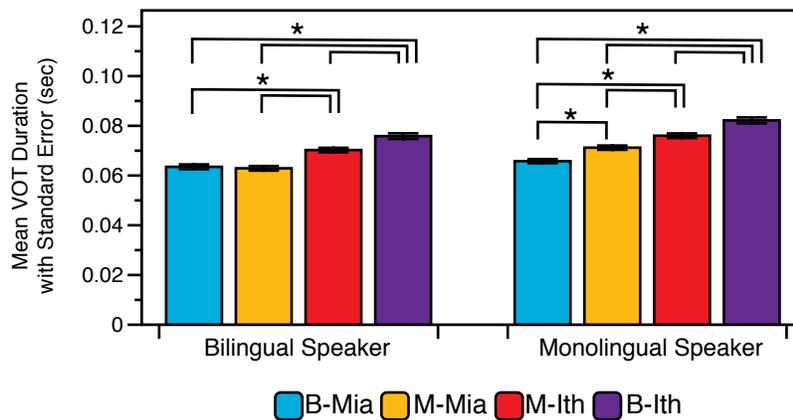


Figure 5. When speaking with the Bilingual Speaker, the monolinguals from the monolingual community achieved bilingual-like VOTs (similar to B-Mia).

5.3. VOT DIFFERENCES BETWEEN GROUPS AT START AND END OF STUDY. To compare whether participant groups' VOTs became more or less similar to one another after speaking with the Monolingual or Bilingual Speaker, I compared participant groups' VOTs for each stop during the first quarter of the study (Block 1) and the final quarter of the study (Block 4) for each Speaker. The change in which groups significantly differ from each other at the start of the study versus the end is used to compare accommodation across groups. For this analysis, a three-way ANOVA was conducted to compare the effects of participant group, voiceless stop, and Speaker on VOT durations (seconds) from Block 1 and then from Block 4.

In Block 1, a three-way analysis of variance yielded a main effect for participant group,  $F(3, 1051)$ ,  $p < 0.001$ . Post hoc comparisons using the Tukey HSD test indicated that VOTs for B-Mia ( $M = 0.067$ ,  $SD = 0.025$ ) and M-Mia ( $M = 0.067$ ,  $SD = 0.019$ ) were significantly lower than VOTs for M-Ith ( $M = 0.075$ ,  $SD = 0.020$ ) and B-Ith ( $M = 0.078$ ,  $SD = 0.028$ ). The main effect of voiceless stop was also significant,  $F(2, 1051)$ ,  $p < 0.001$ ; VOTs for /p/ ( $M = 0.066$ ,  $SD = 0.024$ ) were significantly shorter than VOTs for /t/ ( $M = 0.072$ ,  $SD = 0.020$ ) and /k/ ( $M = 0.077$ ,  $SD = 0.024$ ), and VOTs for /t/ were significantly shorter than VOT for /k/. Last, the main effect of Speaker was significant,  $F(1, 1051)$ ,  $p < 0.001$ ; participants' VOTs were longer when

speaking with the Monolingual Speaker ( $M = 0.074$ ,  $SD = 0.023$ ) than the Bilingual Speaker ( $M = 0.069$ ,  $SD = 0.024$ ).

Further, in Block 1, there was a significant interaction between participant group and Speaker,  $F(3, 1051)$ ,  $p = 0.031$ . Post hoc comparisons using the Tukey HSD test indicated that, when speaking with the Monolingual Speaker, B-Ith ( $M = 0.084$ ,  $SD = 0.028$ ) had longer VOTs than B-Mia ( $M = 0.067$ ,  $SD = 0.021$ ) and M-Mia ( $M = 0.069$ ,  $SD = 0.019$ ); also, M-Ith ( $M = 0.077$ ,  $SD = 0.019$ ) had significantly longer VOTs than B-Mia. When speaking with the Bilingual Speaker, there were no significant differences between B-Ith ( $M = 0.073$ ,  $SD = 0.026$ ), M-Ith ( $M = 0.073$ ,  $SD = 0.020$ ), M-Mia ( $M = 0.065$ ,  $SD = 0.020$ ), and B-Mia ( $M = 0.067$ ,  $SD = 0.028$ ). Lastly, B-Ith had significantly longer VOTs when speaking with the Monolingual Speaker ( $M = 0.084$ ,  $SD = 0.028$ ) than the Bilingual Speaker ( $M = 0.073$ ,  $SD = 0.026$ ). No other interactions were significant.

Therefore, in Block 1, when speaking with the Monolingual Speaker, participants from the monolingual community (B-Ith and M-Ith) had longer VOTs than participants from the bilingual community (B-Mia and M-Mia). When speaking with the Bilingual Speaker, there were no differences between groups.

In Block 4, a three-way analysis of variance yielded a main effect for participant group,  $F(3, 1048)$ ,  $p < 0.001$ . Post hoc comparisons using the Tukey HSD test indicated that VOTs for B-Mia ( $M = 0.061$ ,  $SD = 0.019$ ) and M-Mia ( $M = 0.067$ ,  $SD = 0.020$ ) were significantly lower than VOTs for M-Ith ( $M = 0.071$ ,  $SD = 0.021$ ) and B-Ith ( $M = 0.081$ ,  $SD = 0.026$ ); and B-Mia had significantly shorter VOTs than M-Mia. The main effect of voiceless stop was also significant,  $F(2, 1048)$ ,  $p < 0.001$ ; VOTs for /p/ ( $M = 0.063$ ,  $SD = 0.023$ ) were significantly shorter than VOTs for /t/ ( $M = 0.071$ ,  $SD = 0.020$ ) and /k/ ( $M = 0.076$ ,  $SD = 0.023$ ), and VOTs for /t/ were significantly shorter than VOT for /k/. Last, the main effect of Speaker was significant,  $F(1, 1048)$ ,  $p < 0.001$ ; participants' VOTs were longer when speaking with the Monolingual Speaker ( $M = 0.072$ ,  $SD = 0.022$ ) than the Bilingual Speaker ( $M = 0.068$ ,  $SD = 0.023$ ).

Further, in Block 4, there was a significant interaction between participant group and Speaker,  $F(3, 1048)$ ,  $p = 0.05$ . When speaking with the Monolingual Speaker, B-Ith ( $M = 0.083$ ,  $SD = 0.026$ ) had longer VOTs than B-Mia ( $M = 0.061$ ,  $SD = 0.018$ ), M-Mia ( $M = 0.072$ ,  $SD = 0.019$ ), and M-Ith ( $M = 0.073$ ,  $SD = 0.020$ ). Also, B-Mia had significantly shorter VOTs than M-Mia, M-Ith, and B-Ith. When speaking with the Bilingual Speaker, B-Ith ( $M = 0.079$ ,  $SD = 0.026$ ) had significantly longer VOTs than M-Ith ( $M = 0.068$ ,  $SD = 0.022$ ), M-Mia ( $M = 0.062$ ,  $SD = 0.019$ ), and B-Mia ( $M = 0.062$ ,  $SD = 0.020$ ). Lastly, M-Mia had significantly longer VOTs when speaking with the Monolingual Speaker than the Bilingual Speaker. No other interactions were significant.

These results show that, by Block 4, the bilinguals from the monolingual community (B-Ith) have begun to diverge when speaking with the Bilingual Speaker, increasing their VOTs to sound more English-like. Further, bilinguals from the bilingual community (B-Mia) diverge from the Monolingual Speaker, decreasing their VOT to sound more bilingual-like. Lastly, the monolinguals from the bilingual community (M-Mia) converge slightly with the Monolingual Speaker. These results are illustrated in Figures 6 and 7.

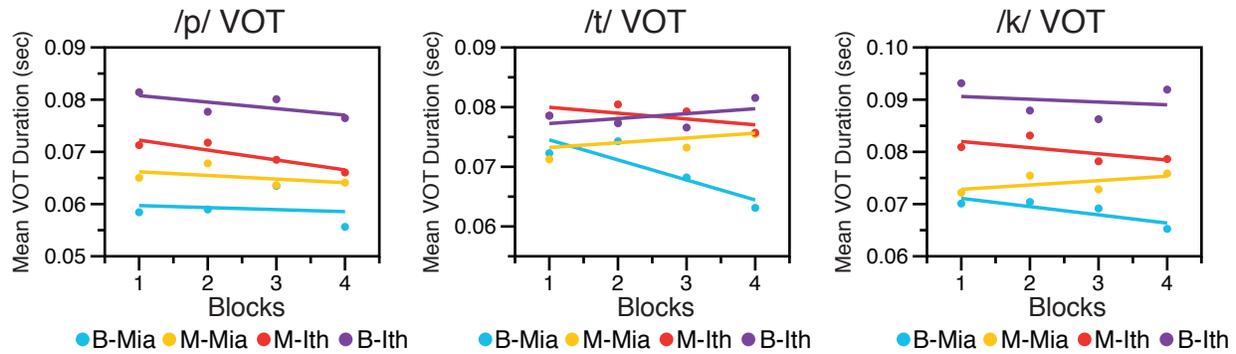


Figure 6. Mean VOT by block and participant group with Monolingual Speaker. Results show B-Mia diverging from, and M-Mia converging with, the Monolingual Speaker.

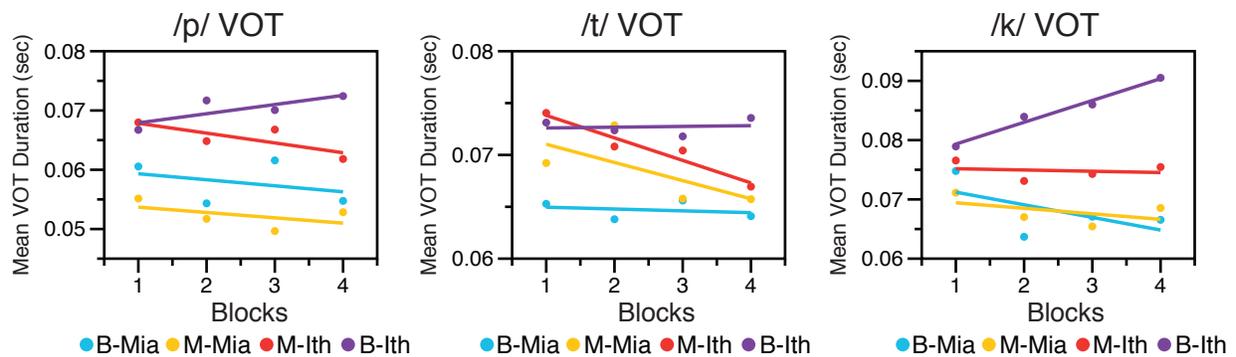


Figure 7. Mean VOT by block and participant group with Bilingual Speaker. Results show B-Ith diverging from the Bilingual Speaker.

5.4. VOT CHANGE DIFFERENCES BETWEEN GROUPS. In this section, I examine whether participant groups accommodated differently by comparing whether the amount and direction of VOT change was different between groups. To compare this, I first calculated the change in VOT for each participant per Speaker (Change VOT = mean(VOT duration in Block 1) – mean(VOT duration in Block 4)). Then, for each Speaker, a two-way ANOVA was conducted to compare the effects of participant group and voiceless stop on VOT change (in seconds).

With the Bilingual Speaker, there was a significant main effect of participant group,  $F(3, 48)$ ,  $p = 0.001$ . Post hoc comparisons using the Tukey HSD test indicated that VOT change for B-Ith ( $M = 0.0061$ ,  $SD = 0.0106$ ) was significantly higher than VOT change for B-Mia ( $M = -0.0053$ ,  $SD = 0.0100$ ). This means that, while speaking with the Bilingual Speaker, the B-Ith increased their VOT while the B-Mia decreased their VOT. The main effect of voiceless stop was not significant,  $F(2, 48)$ ,  $p = 0.4522$ ; and the interaction between voiceless stop and participant group was not significant,  $F(6, 48)$ ,  $p = 0.3602$ .

However, when analyzing VOT change by voiceless stop, there was an effect of participant group on VOT change. Specifically, for /p/, /t/, and /k/, a one-way ANOVA was conducted to compare the effects of participant group on VOT change (in seconds) when speaking with a Bilingual Speaker. An analysis of variance showed that the effect of participant group on VOT change for /k/ was significant,  $F(3, 16)$ ,  $p = 0.042$ . Post hoc comparisons using the Tukey HSD test indicated that /k/ VOT change for B-Ith ( $M = 0.0121$ ,  $SD = 0.0113$ ) was significantly higher than /k/ VOT change for B-Mia ( $M = -0.0083$ ,  $SD = 0.0160$ ). Next, an analysis of variance showed that the effect of participant group on VOT change for /p/ was near significant,  $F(3, 16)$ ,

$p = 0.089$ . There was not a significant effect of participant group on VOT change for /t/,  $F(3,16)$ ,  $p = 0.302$ . These results are illustrated in Figure 8.

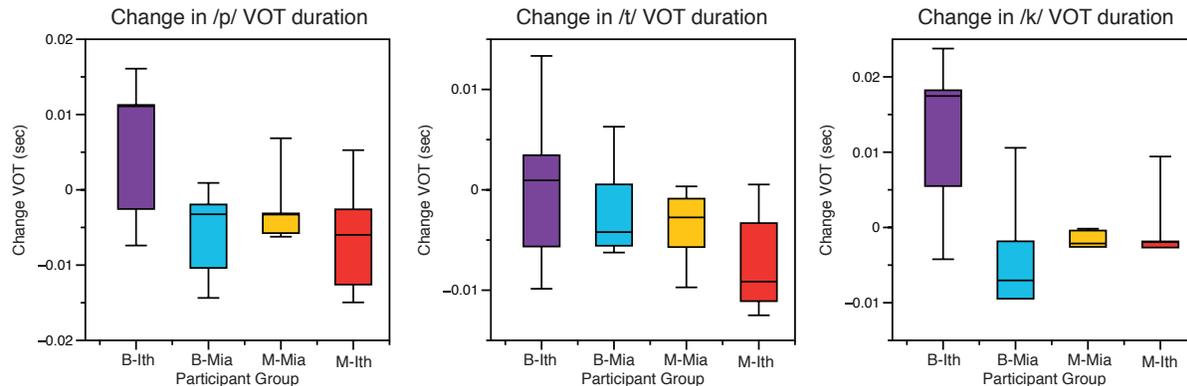


Figure 8. VOT change by participant group with the Bilingual Speaker. B-Ith increased their VOTs and B-Mia decreased their VOTs while speaking with the Bilingual Speaker.

With the Monolingual Speaker, a two-way ANOVA was conducted to compare the effects of participant group and voiceless stop on VOT change (in seconds). There were no significant main effects for participant group,  $F(3,48)$ ,  $p = 0.429$ ; or voiceless stop,  $F(2,48)$ ,  $p = 0.849$ ; nor was there an interaction between participant group and voiceless stop,  $F(6,48)$ ,  $p = 0.948$ . In other words, VOT change did not differ by participant group or voiceless stop when speaking with the Monolingual Speaker.

These results indicate that the direction of VOT change was different for the two bilingual participant groups when they were speaking with the Bilingual Speaker; the bilinguals from the monolingual community (B-Ith) increased their VOTs, diverging from the Bilingual Speaker, while the bilinguals from the bilingual community (B-Mia) decreased their VOTs, converging with the Bilingual Speaker.

5.5. PRIMING INFLUENCE ON VOT. In this section, I examine whether short-term exposure to monolingual English or bilingual speech influenced VOTs. To test this, I conducted a three-way ANOVA to examine the influence of participant group, Speaker, and priming on VOT duration (in seconds). Of interest is the interaction between priming and Speaker.

A three-way analysis of variance yielded a main effect for participant group,  $F(3, 4279)$ ,  $p < 0.001$ . Post hoc comparisons using the Tukey HSD test indicated that B-Ith ( $M = 0.079$ ,  $SD = 0.027$ ) had significantly higher VOTs than all other groups, and M-Ith ( $M = 0.073$ ,  $SD = 0.021$ ) had higher VOTs than M-Mia ( $M = 0.067$ ,  $SD = 0.021$ ) and B-Mia ( $M = 0.065$ ,  $SD = 0.021$ ). M-Mia and B-Mia did not differ. The main effect of priming was also significant,  $F(1, 4279)$ ,  $p = 0.012$ ; primed VOTs ( $M = 0.070$ ,  $SD = 0.024$ ) were significantly shorter than unprimed VOTs ( $M = 0.072$ ,  $SD = 0.023$ ). The main effect of Speaker was also significant,  $F(1, 4279)$ ,  $p < 0.001$ ; VOTs were significantly shorter for the Bilingual Speaker ( $M = 0.068$ ,  $SD = 0.024$ ) than the Monolingual Speaker ( $M = 0.074$ ,  $SD = 0.023$ ).

There was a near significant interaction between participant group and priming,  $F(3, 4279)$ ,  $p = 0.07$ . Notably, B-Ith had shorter VOTs when primed ( $M = 0.077$ ,  $SD = 0.027$ ) than unprimed ( $M = 0.081$ ,  $SD = 0.027$ ). There was also a significant interaction between participant group and Speaker,  $F(3, 4279)$ ,  $p = 0.019$ , as detailed in Section 5.2. No other interactions were significant. Importantly, there was no interaction between Speaker and priming,  $F(1, 4279)$ ,  $p =$

0.395; nor was there an interaction between participant group, Speaker, and priming,  $F(3, 4279)$ ,  $p = 0.764$ . Thus, priming did not influence accommodation.<sup>4</sup>

**6. Conclusions and discussion.** In this study, I examined whether phonetic accommodation is influenced by a speaker's language background, long-term exposure to monolingual or bilingual speech in their speech community, and short-term exposure to monolingual or bilingual speech through priming. The results of this study showed that both language background and long-term exposure to speech in a speaker's speech community influences accommodation.

Long-term exposure to monolingual or bilingual speech had a major impact on the speakers' VOTs and the way they adjusted, or did not adjust, their speech for either the Monolingual or Bilingual Speaker. Speakers from the bilingual community (B-Mia and M-Mia) had overall lower, more bilingual-like VOTs than the speakers from the monolingual community (M-Ith and B-Ith). The bilinguals from the bilingual community (B-Mia) did not accommodate to the Monolingual Speaker; instead, they were shown to diverge slightly from the Monolingual Speaker and slightly converge with the Bilingual Speaker. However, overall, B-Mia's VOTs did not significantly differ for either Speaker; instead, their VOTs remained shorter than other participant groups in the majority of cases. The only participant group that had similar VOTs was the monolingual group from the bilingual community (M-Mia); when speaking with the Bilingual Speaker, M-Mia had VOTs similar to the bilinguals in the same community (B-Mia). Also, M-Mia had the second lowest VOTs overall. Conversely, speakers from the monolingual community (M-Ith and B-Ith) had longer VOTs than the speakers from the bilingual community (M-Mia and B-Mia). Interestingly, the bilingual speakers from the monolingual community (B-Ith) had the longest VOTs of all the participant groups. Also, B-Ith diverged from the Bilingual Speaker, increasing their VOT to sound more English-like.

Language background also influenced how speakers adjusted their speech. Specifically, the bilingual groups (B-Ith and B-Mia) seemed to be influenced by long-term exposure, more so than the monolingual groups. The bilinguals both diverged from the Speaker they were not the majority in their community: B-Ith diverged from the Bilingual Speaker and had the highest overall VOTs in the study, while the B-Mia diverged somewhat from the Monolingual Speaker and had the lowest VOTs in the study. Conversely, the monolinguals showed some convergence with both Speakers (unlike the bilinguals).

Short-term exposure (priming) to bilingual and monolingual speech was not shown to have an immediate effect on accommodation in this study. One possible reason for this is that priming may have occurred for unprimed words on a board through extension. For example, a primed word pair for /p/ may have unintentionally primed an unprimed word pair for another voiceless stop (/t/ or /k/) on the same board. Additionally, the effects of priming may have lasted longer than expected, where a primed word on one board unintentionally primed an unprimed word on another board. Regardless, we do see a change in VOT durations over the course of the study, so interaction with the Speakers is influencing VOTs, at least for some of the participant groups.

These results show that language background and long-term exposure to bilingual or monolingual speech influences speech accommodation. As previous research has shown, a speaker will accommodate differently depending on their listener and how they want to be perceived by their listener. Also, speakers will draw on their linguistic experiences while they are

---

<sup>4</sup> It should be noted that I examined the influence of priming on overall VOTs, VOT change, and VOT differences at the start versus the end of the study. In none of these cases did priming influence accommodation.

speaking, which influences production and, as this study shows, accommodation. The results of this study show that speakers of different language backgrounds and from different speech communities will accommodate differently, largely influenced by their linguistic experience.

## References

- Beebe, Leslie M., & Howard Giles. 1984. Speech-accommodation theories: A discussion in terms of second-language acquisition. *International Journal of the Sociology of Language* 1984(46). 5-32. <https://doi.org/10.1515/ijsl.1984.46.5>.
- Bell, Allan, 1984. Language style as audience design. *Language in Society* 13(2). 145-204. <https://doi.org/10.1017/S004740450001037X>.
- Bell, Allan. 2001. Back in style: Reworking audience design. In Penelope Eckert & John R. Rickford (eds.), *Style and Sociolinguistic Variation* 139-169. Cambridge University Press.
- Boersma, Paul & David Weenink. 2018. Praat: doing phonetics by computer [Computer software]. Version 6.0.37. <http://www.praat.org/> (3 February, 2016.)
- Carter, Phillip M., & Andrew Lynch. 2015. Multilingual Miami: Current trends in sociolinguistic research. *Language and Linguistics Compass* 9(9). 369-385. <http://dx.doi.org/10.1111/lnc3.12157>.
- Colby, Sandra L., & Jennifer M. Ortman. 2017. Projections of the size and composition of the US population: 2014 to 2060: Population estimates and projections. [http://wedocs.unep.org/bitstream/handle/20.500.11822/20152/colby\\_population.pdf](http://wedocs.unep.org/bitstream/handle/20.500.11822/20152/colby_population.pdf) (23 February, 2018.)
- Enzina, Naomi R. 2016. Spanish-influenced rhythm in Miami English. *Proceedings of the Linguistic Society of America* 1. 34-1. <http://dx.doi.org/10.3765/plsa.v1i0.3733>.
- Enzina, Naomi. 2015. Spanish-influenced prosody in Miami English. *Cornell Working Papers in Phonetics and Phonology* 2015. 78-120. <https://pdfs.semanticscholar.org/915c/0be3e615df3cb2534b54ec1ba5af71b72192.pdf> (23 February, 2018.)
- Enzina, Naomi. 2017. How speakers select synthetic and analytic forms of English comparatives: An experimental study. *Proceedings of the Linguistic Society of America* 2. 46-1. <http://dx.doi.org/10.3765/plsa.v2i0.4101>.
- Flege, James E. 1992. Speech learning in a second language. *Phonological development: Models, research, implications*. 565-604. Timonium, MD: York Press.
- Flege, James E. 1995. Second language speech learning: Theory, findings, and problems. *Speech perception and linguistic experience*. 233-277. Timonium, MD: York Press.
- Flege, James E., & Ian R. MacKay. 2004. Perceiving vowels in a second language. *Studies in Second Language Acquisition* 26(1). 1-34. <https://doi.org/10.1017/S0272263104026117>.
- Giles, Howard, Justine Coupland, & Nikolas Coupland (eds.). 1991. *Contexts of accommodation: Developments in applied sociolinguistics*. Cambridge University Press.
- Guion, Susan G., James E. Flege, & Jonathan D. Loftin. 2000. The effect of L1 use on pronunciation in Quichua–Spanish bilinguals. *Journal of Phonetics* 28(1). 27-42. <https://doi.org/10.1006/jpho.2000.0104>.
- Hwang, Jiwon, Susan E. Brennan, & Marie K. Huffman. 2015. Phonetic adaptation in non-native spoken dialogue: Effects of priming and audience design. *Journal of Memory and Language* 81. 72-90. <https://doi.org/10.1016/j.jml.2015.01.001>.

- Kim, Midam, William S. Horton, & Ann R. Bradlow. 2011. Phonetic convergence in spontaneous conversations as a function of interlocutor language distance. *Laboratory Phonology* 2(1). 125-156. <https://doi.org/10.1515/labphon.2011.004>.
- McAuliffe, Michael, Michaela Socolof, Sarah Mihuc, Michael Wagner, & Morgan Sonderegger. 2017. Montreal Forced Aligner [Computer program]. Version 0.9.0. <http://montrealcorpusools.github.io/Montreal-Forced-Aligner/> (17 January, 2017.)
- Sancier, Michele L., & Carol A. Fowler. 1997. Gestural drift in a bilingual speaker of Brazilian Portuguese and English. *Journal of Phonetics* 25(4). 421-436. <https://doi.org/10.1006/jpho.1997.0051>.
- Yavaş, Mehmet & Emily Byers. 2014. Production of voiceless stops in early sequential Spanish-English bilinguals. *Unusual Productions in Phonology: Universals and Language-Specific Considerations*. 242-263.