On comparison groups in heritage phonetics/phonology research: The case of bilingual Spanish vowels

Megan Solon
Nyssa Knarvik
Josh DeClerck
University at Albany

Abstract
Previous research on heritage Spanish vowel production (e.g., Ronquest, 2012, 2013; Willis, 2005) has revealed consistent and systematic differences—including asymmetry in the vowel space, condensing and fronting of back vowels, and reduction and centralization of unstressed vowels—as compared to traditional descriptions of the monolingual Spanish vowel triangle. The present study takes another look at heritage Spanish vowels (both quality and quantity), using a group of “homeland” native Spanish-speaking late Spanish-English bilinguals for comparison purposes. Data for both groups were collected via a dyadic, meaning-focused task. Results revealed no significant differences between heritage and homeland groups in overall vowel quality or vowel quantity. Both speaker groups produced some atonic vowels with distinct quality and shorter durations than their tonic counterparts, but, again, few differences were found between groups. The heritage speaker data alone appear comparable to those of previous studies; however, the comparison to native Spanish-speaking late bilinguals points to important similarities to other native speakers’ vowels not previously signaled in studies that compare production to monolingual norms.

Heritage speakers have been defined in various ways, with broader definitions often emphasizing a cultural connection to a language through familial relevance (e.g., Fishman, 2001; Van Deussen-Scholl, 2003) and narrower definitions stressing heritage speakers’ functional proficiency (at least to some degree) in the heritage and majority languages (e.g., Montrul, 2012; Valdés, 2000). Following Valdés (2000), Henriksen (2015) defines heritage language speakers as “bilingual speakers who speak ethnolinguistically minority languages during infancy and early childhood prior to formal education or significant societal exposure to another (i.e., majority) language” (p. 286). Some heritage speakers may be simultaneous bilinguals (i.e., acquiring the heritage and majority language from birth), whereas others are considered sequential bilinguals (i.e., acquiring the heritage language first and, then, around school age, acquiring the majority language). Nevertheless, by (a narrower) definition, and the one adopted in the present study, heritage speakers are early bilinguals living in a majority language context who often become more dominant in the majority language by adulthood (Montrul, 2012).

Recently, much attention has been paid to understanding, documenting, and explaining the linguistic profiles and patterns of heritage speakers. Studies pursuing knowledge of heritage linguistic systems often compare heritage speakers to other groups of (many times monolin-
gual) native speakers and/or to groups of second language (L2) learners of the heritage language to gather information regarding the structure of, variation within, and limits of the heritage language system as compared to those of other speakers of the same language with different language learning histories. To date, the majority of research on heritage linguistic systems, including that which specifically explores heritage Spanish, has focused on morphosyntactic structures (e.g., Rao & Ronquest, 2015), a bias potentially based on “the general impression... that even basilectal heritage speakers sound native like” (Polinsky & Kagan, 2007, p. 378). Still, as recent and ongoing research continues to reveal, significant differences in heritage speaker phonological systems and phonetic productions can often be observed and may reveal additional important information regarding heritage linguistic systems.

The present study adds to the growing body of literature on heritage phonetic and phonological systems via its examination of the phonetic realization of vowels by heritage speakers of Spanish. Additionally, the present study builds upon previous studies on heritage Spanish vowels by incorporating a comparison group comprised of Spanish-English bilingual “homeland” native speakers of Spanish (Potowski, 2014), also referred to here as “late” bilinguals due to their later acquisition of English as compared to “early” bilinguals, or heritage speakers. As detailed in later sections, this comparison group—as opposed to previous studies’ use of traditional accounts of monolingual Spanish vowel systems for comparative purposes—permits us to narrow in on the potential influence of differences regarding the specific language background and experiences of heritage speakers as well as to reject potential false conclusions regarding the uniqueness of heritage speaker linguistic outcomes that diverge from monolingual norms. In what follows, we briefly review general findings regarding heritage phonetics and phonology (with an emphasis on studies of heritage Spanish), highlight investigations of vowels in other heritage languages, and discuss those studies that have specifically explored heritage Spanish vowels before moving on to a description of our own study.

Background

Heritage Phonetics and Phonology

As mentioned, within research specifically focused on heritage speakers’ linguistic systems, empirical investigations of heritage phonetics and phonology have lagged behind investigations of morphosyntax. Research on heritage Spanish is no exception to this trend (Rao & Ronquest, 2015), although, in recent years, research on heritage Spanish phonetics and phonology has grown noticeably. Given the present study’s focus on segmental aspects of the heritage phonetic/phonological system, our review will center around previous research examining segments in heritage Spanish; nevertheless, it is prudent to recognize that recent empirical work has also contributed to our growing understanding of heritage Spanish suprasegmental phenomena (e.g., Carter & Wolford, 2016; Colantoni, Cuza, & Mazzaro, 2016; Kim, 2015; Rao, 2016; Robles-Puente, 2014) as well as of the phonetic/phonological patterns in other heritage languages (e.g., Asherov, Fishman, & Cohen 2016, on Israeli heritage Russian; Łyskawa, Maddesux, Melara, & Nagy, 2016, on heritage Polish in Canada, to name just a few).

Although some work (e.g., Mazzaro, Cuza, & Colantoni, 2016) has suggested no significant differences between heritage and (other types of) native speakers in the ability to discriminate phonological contrasts in perception, the general impression that heritage speakers’ phonetic production mirrors that of other groups of native speakers (Polinsky & Kagan, 2007, p 378) is
one that has been (at least partially) refuted by recent studies. Henriksen (2015) and Amengual
(2016), for example, both explored rhotic realization in heritage Spanish. In both studies, heri-
tage speakers of Spanish maintained a phonological contrast between tap and trill rhotic pho-
- nemes, but exhibited much individual variation in the phonetic realization of rhotic phonemes
(Henriksen, 2015) and utilized different cues to maintain the phonological contrast depending
on dominance in the heritage versus majority language (Amengual, 2016). Au, Oh, Knightly, Jun,
and Romo (2008) found that heritage speakers’ production of Spanish voiceless and voiced stops
were rated by native speakers as significantly more “accented” than native speaker productions.
Finally, results of Rao’s (2014, 2015) investigations of intervocalic /b/ and /b d g/, respectively,
showed that heritage speakers with more at-home exposure to Spanish produced pure approx-
imants (the realization more common in prescriptive accounts of Spanish pronunciation) more
frequently, and speakers with less at-home exposure were more likely to produce tense approx-
imants and, overall, to exhibit much greater variability in production patterns. Similar findings
showing distinct heritage phonetic patterns have been attested in other heritage languages as
well (e.g., Kang, George, & Soo, 2016, for stop voicing in heritage Tagalog; Łyskawa et al., 2016, for
word-final obstruent devoicing in heritage Polish).

Heritage Vowel Production across Languages

Specifically within work on heritage vowels, research across languages has consistently pro-
vided evidence for heritage speaker patterns in vowel production that diverge from prescriptive
and monolingual norms. For example, Asherov, Fishman, and Cohen’s (2016) study of Israeli
heritage speakers of Russian found that Russian vowel reduction patterns showed evidence of
a combination of Russian and Hebrew characteristics, with heritage speakers exhibiting reduc-
tion of /o/ but without the expected height contrast in nonce word productions. Godson (2004)
similarly found differences in the phonetic realization of Western Armenian vowels by early
and late English-Armenian bilinguals toward English vowels, but only for those Western Ar-
menian vowels that are already similar to US English vowels. Finally, in her dissertation on
the production of Arabic vowels, Saadah (2011) compared the speech of Arabic native speakers,
Arabic heritage speakers, and Arabic L2 learners. She found that heritage speakers patterned
similarly to native speakers with regard to both first and second formant (F1 and F2; indices for
vowel height and frontness/backness respectively) values for /i:/ and /i/ and F1 values for /u:/ but
patterned more similarly to L2 learners in F1 values of /a:/ and /a/. She also found that heritage
speakers showed intermediate values (between native speakers and L2 learners) for F2 of /a:/, /a/,
and /u:/, and F1 of /u/. In terms of duration, the heritage speakers exhibited intermediate values
in the length of all short vowels and /i/ and /u/ but patterned very similarly to native speakers
in the duration of /a/. The present study examines similar questions regarding vowel realiza-
tion for Spanish heritage speakers in the US. In what follows, we offer a brief overview of the
prototypical Spanish vowel system as described in previous literature and a review of existing
research on heritage Spanish vowel production before describing the present study and its con-
tributions to our understanding of heritage (Spanish) phonetics and phonology.

Spanish Vowel System

Spanish possesses a relatively simple and symmetrical five-vowel system comprised of the
vowels /i e a o u/ (Hualde, 2014). Spanish vowels are typically considered to be pure monoph-
thongs and, in comparison to English, shorter in duration (Hualde, 2014). Additionally, although dialectal differences in vocalic realization have been noted (e.g., Chládková, Escudero, & Boersma, 2011; Willis, 2005, 2008), these differences are relatively minor in nature. Spanish vowels are also generally thought to show few differences with regard to stress context, although recent research has suggested evidence of centralization (e.g., Cobb & Simonet, 2015; Menke & Face, 2010) and reduction (Marín Gálvez, 1994-1995) of unstressed vowels in native Spanish.

Many recent studies on Spanish vowel production (and heritage Spanish vowel production, as described in the next section) have relied on traditional descriptions of the vowel space—such as those offered by Quilis and Esgueva (1983), Martínez Celdrán (1995), and Marín Gálvez (1994-1995)—for comparison purposes. Quilis and Esgueva, for example, offer an acoustic analysis of Spanish vowels produced by 22 monolingual native Spanish speakers from Spain and Latin America; their findings have, for decades, served as primary evidence for the claim of the triangular shape and (relative) symmetry of the Spanish vowel system. Figure 1 recreates the vowel chart presented in Quilis and Esgueva (1983) for their male and female informants (represented separately).

Figure 1. Vowel plot showing average F1 and F2 values by vowel and speaker sex from Quilis & Esgueva (1983)

Marín Gálvez (1994-1995) likewise serves as a seminal study of Spanish vowel duration. His results—taken from acoustic measures of the read speech of two native Spanish speakers—provide quantifiable evidence both of the average duration(s) of vowels in Spanish (i.e., 60-70 ms), the intrinsic differences between Spanish vowels in average duration (i.e., with /a/ having the longest average duration and /i/ and /u/ the shortest), and the effect of stress on vowel duration (i.e., with stressed vowels averaging approximately 12 ms longer than vowels in unstressed position). Figure 2 recreates the average vowel durations for each Spanish vowel in stressed and unstressed positions presented in Marín Gálvez.
As described in the next section, these and similar descriptions of (mostly monolingual) Spanish vowel systems have largely shaped the manner in which heritage Spanish vowels have been described and evaluated.

**Heritage Spanish Vowels**

Due, at least in part, to the relative stability of Spanish vowels across varieties, greater empirical attention in work on native Spanish phonetics/phonology and sociolinguistics, and even on the acquisition of Spanish phonetics/phonology by L2 learners, has been dedicated to consonants than to vowels. Nevertheless, several studies have explored heritage Spanish vowel systems in both production (e.g., Alvord & Rogers, 2014; Ronquest, 2012, 2013, 2016; Willis, 2005) and perception (e.g., Boomershine, 2013; Mazzaro et al., 2016), although this review (given the focus of the present study) will concentrate on previous studies of heritage Spanish vowel production.

Two studies (Ronquest, 2012; Willis, 2005) investigated the production of Spanish vowels by heritage speakers and compared this production to the traditional monolingual Spanish vowel triangle as described, for example, in Quilis and Esgueva (1983). Both studies encountered patterns in heritage speaker vowel production that are distinct from traditional monolingual accounts. Willis (2005), for example, found that heritage speakers produced a lower and more fronted /u/, a lower /o/, and a more fronted /a/. Ronquest (2012) observed a general asymmetry in her heritage speakers’ vowel spaces as compared to traditional accounts and, notably, a more backed /e/ and a more fronted /u/.

Both studies (Ronquest, 2012; Willis, 2005) as well as Ronquest (2013) also explored the role of lexical stress in vowel production by Spanish heritage speakers in the US, where the majority language, English, is known to exhibit considerable vowel reduction in unstressed syllables (e.g., Ladefoged, 2006). Whereas Willis (2005) found no significant differences with regard to F1 and F2 for stressed versus unstressed /a/ (the only vowel for which he explored stress effects), the other two studies did indicate evidence of quality and quantity differences between stressed
and unstressed Spanish vowels by heritage speakers. For example, Ronquest (2012) found a general tendency of centralization and reduction (i.e., shorter duration) for unstressed vowels in comparison to their stressed counterparts in her 16 heritage speaker participants. In a follow-up study exploring data from 13 of the original 16 speakers, Ronquest (2013) found similar results, specifically showing that unstressed /e/, /a/, and /o/ were produced higher than their stressed counterparts, and unstressed /i/, /e/, /o/, and /u/ were more centralized. All five vowels were also significantly shorter in duration in atonic position than in tonic position. Though not reviewed in detail here, Ronquest (2016) also found significant effects for speech style (i.e., elicitation task) in vowel quality and quantity for the same heritage speakers, with more expanded vowel spaces and longer durations in more controlled speech.

Taken together, these studies suggest that heritage Spanish vowel productions differ in substantial ways from common symmetrical descriptions of the Spanish vowel system and accounts of the stability of vowel quality and duration regardless of tonicity. Nevertheless, it is important to note that all of these studies (with the exception of Ronquest, 2016, which, given the focus on the effect of speech style, makes no direct comparison to any other speaker group, including from previous studies) compared heritage speaker productions to previous, traditional accounts of monolingual Spanish vowel production; that is, none of these studies included their own comparison group(s), and all comparisons made were to vowels produced by monolingual Spanish speakers. Heritage speakers are, by definition, (early) bilinguals; thus, we argue that a more appropriate comparison group may be other bilinguals with differing language learning histories or experiences, given that, as Pascual y Cabo and Rothman (2012) write, “monolingual and bilingual realities are distinctive” (p. 454; see also Ortega, 2013).

To our knowledge, only one study of (US) heritage Spanish vowel production exists that incorporates its own comparison group of other Spanish-English bilinguals: Alvord and Rogers (2014) examined the influence of English on Spanish vowel realization by 10 Spanish-English bilinguals and one Spanish monolingual all of Cuban descent and living in Miami, Florida. These authors divided their speakers into three groups on the basis of place of birth and age of immigration to the United States (Group 1: born in Cuba, immigrated to US after age 11; Group 2: born in Cuba, immigrated to US before age 6 or child of at least one parent from Group 1; Group 3: born in US to at least one parent from Group 2). Although some differences between groups were observed, no clear-cut findings for influence of English on Spanish vowel production were observed. For instance, the third group—those born in the United States to at least one parent in the second immigrant group—showed some evidence of fronting of Spanish /o/ (which was not present in the productions of Groups 1 and 2), but Spanish /o/ as realized by this group did not occupy the much more fronted and lower position of English /æ/, as had been claimed in previous research. The relative stability of /u/ across all three groups provided, according to the authors, additional evidence for an overall lack of influence from English. Alvord and Rogers (2014) also found that their data “overwhelmingly indicated that stressed and unstressed vowel spaces are inherently different, with unstressed vowel space almost always being more reduced than stressed vowel space” (p. 165). Nevertheless, the authors emphasize that this centralization “cannot be simply written off as transfer from English” (p. 165) because even the most Spanish-dominant group showed nearly categorical centralization of unstressed vowels.

In the present study, we aim to follow Alvord and Rogers’s (2014) lead and continue to advance research on heritage Spanish vowel productions by presenting an analysis of heritage Spanish vowel realization as compared to monolingual Spanish vowels but rather to vowels produced by a comparable group of late Spanish-English bilinguals who spoke Spanish during
childhood and adolescence and began learning English as adults (see Colantoni et al., 2016, for a similar approach to comparison groups [heritage speakers and long-term immigrants] in the study of Spanish intonational contours). We take a slightly different approach than Alvord and Rogers (or Colantoni et al.): Rather than focusing on one dialectal group, we explore heritage speakers from various dialectal (and national origin) backgrounds together and compare their production to an equally diverse group of late Spanish-English bilingual (homeland native Spanish) speakers. Although we recognize that such diversity in participant backgrounds may bring into question issues of dialectal differences (a topic we return to in the Discussion), we also believe that this diversity is reflective of heritage Spanish university populations and the varied input to which they are exposed (especially, for the present heritage participants, most of whom grew up in New York City and surrounding areas). Thus, we compare a varied group of heritage speakers of Spanish to a varied group of homeland native Spanish speakers (also bilingual) to explore whether the distinct properties of heritage Spanish vowel systems previously discovered are, in fact, particular to heritage speakers or are, instead, characteristics of other bilingual Spanish speakers’ systems as well. Going forward and to maintain in focus the comparison between two bilingual groups, we refer to our heritage speakers as early (Spanish-English) bilinguals and to our homeland speakers as late (Spanish-English) bilinguals.

Present Study

The present study examines Spanish vowel production in two groups of Spanish speakers: early Spanish-English bilinguals (i.e., heritage speakers) and late Spanish-English bilinguals (i.e., homeland Spanish speakers). The study is guided by the following research questions:

1. How do the Spanish vowels of early Spanish-English bilinguals compare to those of late Spanish-English bilinguals with regard to both vowel quality (i.e., F1 and F2 values) and vowel quantity (i.e., duration)?

2. Are there differences in early and late bilingual vowel productions—with regard to both vowel quality and quantity—by stress (i.e., tonic vs. atonic)?

Methodology

Participants

The participants in this study comprised two groups: early Spanish-English bilinguals (i.e., heritage speakers) and late Spanish-English bilinguals (homeland native speakers). The early bilingual group comprised 10 speakers, ranging in age from 18 to 27 years old, with an average age of 20.1 years. All participants in this group were female. Eight of the 10 participants were born and grew up in the United States, all in New York City. The other two participants were born in Ecuador and the Dominican Republic and moved to the United States permanently at 1.5 and 9 years of age, respectively.¹

An anonymous reviewer questioned whether someone who immigrated to the US at the age of 9 could/should qualify as a heritage speaker. Although we recognize that this participant’s language learning background may differ with regard to age of significant exposure to English from that of the other heritage speaker participants in the present study, we believe this participant’s inclusion in the study is warranted for several reasons. First, this participant fits the definition of a heritage speaker adopted in the present study and presented in the introduction. Second, although this participant may have received some formal education in Spanish (i.e., from ages 5-9), the majority of her formal education was/has been in English (i.e., from ages 9+). Finally, this participant considered...
ican, Puerto Rican, Ecuadorian, and Salvadoran (as well as a mixture therein) descent. Each participant was exposed to Spanish before the age of 4 and to a significant amount of English by the age of 9. At the time of the study, nine of these 10 participants were enrolled in a university Spanish for Bilinguals course, and the other participant was enrolled in a third semester Spanish as a foreign language course. Each participant self-identified as a heritage speaker according to the definition (as provided on the background questionnaire): “A heritage speaker of Spanish in the US is often considered to be someone who learned Spanish as their first language [or one of their first languages] in childhood, but who, as an adult, may be dominant in English”.

The comparison group for this study was a group of four late Spanish-English bilinguals who were long-term immigrants to the United States. These participants—two females and two males—ranged in age from 32 to 54 years old. Two participants grew up in Colombia, one in Puerto Rico, and one in Mexico. Their ages of arrival to the United States, which was the point at which each of them was significantly exposed to English, were between 14 and 32 years.

Tasks and Procedures

The data used in this study were collected during the completion of a dyadic, meaning-focused production task. Each early and late bilingual participant was paired with a L2 Spanish learner of approximately intermediate level to complete an interactive map task modeled after a similar task in Solon, Long, and Gurzynski-Weiss (2017). The context of the task was as follows: The participants had been hired by a tour company to lead a tour through a city in a Spanish-speaking country, and the company had given them maps to prepare. The maps, however, were incomplete: One participant received a map that indicated the order of the stops on the tour and the route but lacked the names of any locations on the map, and the other had the names of locations of sites throughout the city but did not indicate the tour route or the order of stops (see example maps in Appendix A). The participant “tour guides” had to “call” each other in order to exchange information to complete each of their maps. After completing the first map task, participants were given another similar map task in which they had opposite roles. To facilitate the elicitation of all Spanish vowels within particular phonetic contexts, the street names on these maps were vocalic minimal pairs (e.g. Calle Tido and Calle Tedo) and city landmarks were strategically assigned invented (often nonce) monikers. During the completion of the task, the speech production of each participant was recorded via individual head-mounted microphones and a USBPre external sound card directly onto a laptop computer. The recordings were sampled at 44.1 kHz.

Following the completion of the production task, all participants also completed a background questionnaire, which aided in the gathering of relevant personal data such as that summarized in the previous section.

---

2 As previously mentioned, the diversity of backgrounds both in the heritage and homeland speaker groups was intentional and is, we believe, representative of the diversity of input and exposure that heritage Spanish speakers, especially those enrolled in university Spanish courses (as the present participants were) receive. Although dialectal differences in vowel production are possible, previous research has shown that such differences are generally minor (e.g., Chládková et al., 2011; see also Hualde, 2014).
Analysis

Acoustic

Using the acoustic analysis software Praat (Boersma & Weenink, 2014), each vowel that occurred between two stop phonemes or between a stop and a pause was isolated and coded for lexical item (i.e., the word in which the vowel appeared), vowel (i.e., /i e a o u/), and stress (i.e., tonic vs. atonic). In Spanish, the stop phonemes /b d ɡ/ may be realized as their approximant allophones [β ð ɣ], especially in intervocalic contexts. In some cases, the presence of an approximant allophone made the precise determination of the location of the boundary between a vowel and the preceding or following consonant difficult; as will be detailed later, in all such cases, these tokens were excluded from the vowel duration analysis. Additionally, given that (a) it is known that a vowel’s position in a word (e.g., word-medial as compared to word-final position) can impact vowel realization, especially with regard to duration (e.g., Ronquest, 2016) and (b) the data set was unbalanced with regard to vowel phoneme, position in the word, and stress (e.g., the majority of /u/ tokens and most tonic vowels occurred word-medially), the present analysis excludes vowels in word-final position.

First and second formant (i.e., F1 and F2, respectively) measurements taken from the midpoint of each vowel and vowel duration measurements (in milliseconds) were extracted from each token using a script and the following Praat settings: a window length of 0.025 s for all tokens and five formants estimated under 5500 Hz for females and under 5000 Hz for males. The formants were then normalized using the speaker-intrinsic, vowel-extrinsic Lobanov method in the NORM suite (Thomas & Kendall, 2012), which is a web-based interface to the vowels package for the statistical software R. The Lobanov method in NORM uses each unique speaker’s raw vowel measurements to calculate a normalized formant measurement for each vowel based on the mean and standard deviation for that formant throughout the speaker’s vowel system.

Statistical

To answer our research questions, the relationships between vowel quality and quantity, speaker group, and stress were analyzed statistically using the lmerTest package (Kuznetsova, Brockhoff, & Christensen, 2016) in R (R Core Team, 2016). Two linear mixed effects analyses (one for F1 and one for F2) were performed for each of the five Spanish vowels (/i e a o u/). In these models, speaker and token were included as random effects and speaker group (early or late bilingual), stress (tonic or atonic), and a Speaker Group × Stress interaction were included as fixed effects. Duration of the vowel was also analyzed using linear mixed effects analyses for each vowel. As with the formant analyses, in these models, speaker and token were included as random effects and speaker group, stress, and a Speaker Group × Stress interaction were included as fixed effects.

Results

A total of 1,853 non-word-final vowel tokens were extracted from the data for analysis; 1,386

---

3 Tokens included both the street and landmark names purposefully included in the design of the task as well as any vowels that were produced during task completion within the target contexts described. As a result, the data analyzed come from a mixture of nonce (i.e., invented street and landmark names) and real words.
produced by the early bilingual (heritage Spanish speaker) group and 467 by the late bilingual (homeland Spanish speaker) group. Given that the elicitation task was open-ended and dyadic, the amount of production and the resulting number of vowel tokens per participant varied (early bilinguals: \(M = 123.5, SD = 46.8; range = 53 \text{–} 184\); late bilingual: \(M = 116.8, SD = 60.8; range = 65 \text{–} 189\)). A total of 127 tokens were excluded due to creak, devoicing, and issues with the microphone (78 exclusions from the early bilingual data; 49 from the late bilingual data). Thus, for the formant analysis, 1,726 vowel tokens were examined (\(n = 1,308\) early bilingual; \(n = 418\) late bilingual). As previously mentioned, some additional exclusions were made prior to the duration analyses; these will be discussed in the section on vowel quantity.

### Vowel Quality

Average F1 and F2 values for each vowel and each speaker group were calculated. Figure 3 presents a vowel chart, which visually represents the vowel space for both speaker groups by plotting the average normalized F1 (y axis) and F2 (x axis) values for each vowel and group.

![Vowel plot showing average normalized F1 and F2 values and standard deviations by vowel and speaker group.](Figure 3)

As can be observed in Figure 3, at least descriptively, vowel productions for the early and late bilingual groups differ in a few, relatively minor, ways: Early bilingual /i/ and /e/ are visually more backed (i.e., have lower F2s) than their late bilingual counterpart vowels and are somewhat lower in the vowel space (i.e., have higher F1s). The non-front vowels /a/, /o/, and /u/ also differ between the groups, but not as much: Though not identical, these vowels share roughly the same location within the vowel space, respectively, across the two groups.

Results of the linear mixed models for each measurement (i.e., F1 and F2) and each vowel revealed no statistical differences between speaker groups along the F1 (i.e., height) or F2 (front-back) dimension for any vowel (although the difference observed in the production of /e/ on
the F2 dimension approached significance, \( p = .074 \); see Tables B.1 and B.2 in the Appendix for complete statistical information).

The linear mixed models also examined the role of lexical stress in vowel quality. Figure 4 presents a vowel chart that visually represents the vowel space for both speaker groups together, plotting the average F1 and F2 values for each vowel in each stress context (tonic, atonic).

![Vowel Chart](image)

*Figure 4. Vowel plot showing average normalized F1 and F2 values (both speaker groups combined) by stress.*

As can be observed in Figure 4, overall (i.e., productions for both speaker groups averaged together), nonback vowels (i.e., /i e a/) produced in atonic contexts were visually more centralized than their tonic counterparts, atonic /o/ was visually raised as compared to tonic /o/, and atonic /u/ was (somewhat unexpectedly) backed in comparison to tonic /u/. Statistically, only those differences between stressed and unstressed /i/ and /o/ along the F1 (high-low) dimension \( (p = .031 \text{ and } p = .015, \text{ respectively}) \) and /e/ along the F2 (front-back) dimension \( (p = .030) \) were significant. The difference between stressed and unstressed /e/ on the F1 dimension approached significance \( (p = .054; \text{ see Appendix Tables B.1 and B.2 for complete statistical results}) \).

Finally, we also explored whether the effect of stress differed by speaker group—that is, whether the vowel productions of either the early bilingual or late bilingual group were more affected by lexical stress than those of the other group. Figure 5 presents a vowel chart that depicts the vowel space for both speaker groups, plotting the average F1 and F2 values for each vowel in each stress context (tonic, atonic), this time separated for each vowel by speaker group (early vs. late bilingual).
Although visual differences in the impact of lexical stress on individual vowels by group are apparent in Figure 5, no significant interactions between stress and speaker group were found for any vowel (see Tables B.1 and B.2). Thus, the effects of lexical stress on /i/, /e/, and /o/ productions are similar across the early and late bilingual groups.

In summary, with regard to vowel quality, no statistical differences were observed between early and late bilinguals’ vowel productions with regard to vowel height (F1) or vowel frontness-backness (F2). Both speaker groups exhibited differences in vowel production by lexical stress—/i/ and /o/ in atonic positions were statistically raised and atonic /e/ moved toward the center along the F2 dimension. Nevertheless, there were no differences between speaker groups in these effects of stress.

**Vowel Quantity**

An additional 31 tokens were excluded from the analysis of vowel quantity due to difficulty in precisely determining a boundary between the vowel and an adjacent consonant (e.g., between a vowel and a preceding or following approximant). Thus, the duration analysis included 1,695 vowel tokens ($n = 1,298$ early bilingual; $n = 397$ late bilingual).

Average vowel durations (in milliseconds) for each vowel and speaker group were calculated. Figure 6 presents the overall vowel duration means by vowel separated by speaker group.4

---

4 It should be acknowledged that, although we believe many interspeaker differences in duration are accounted for by including speaker as a random intercept in the linear mixed model analysis, the duration measurements were not normalized by speaking rate due to difficulty in reliably performing this calculation with the dialogic production data.
As can be observed in Figure 6, mean durations for Spanish vowels /i/, /e/, and /o/ produced by the early bilingual group were numerically longer than those produced by the late bilingual group, whereas the mean duration of late bilinguals’ /a/ and /u/ were longer than those of the early bilinguals. Results of the linear mixed models for the duration of each vowel revealed no statistical differences between speaker groups for the duration of any vowel (see Table B.3 in the Appendix for complete statistical information).

We next explored the effect of stress on vowel duration by both groups of Spanish speakers. The overall (with both speaker groups combined) means for vowel durations by stress (tonic vs. atonic) for each vowel are presented in Figure 7.

As can be observed in Figure 7, when considering both speaker groups averaged together, all vowels were numerically longer in duration in tonic position than in atonic position, with the
biggest visual and numerical differences observed for /a/ and /o/. Statistically, only those duration differences between stressed and unstressed /e/ and /o/ were significant ($p < .001$; see Appendix Table B.3 for complete statistical results).

Finally, we also explored whether the effect of stress on vowel duration differed by speaker group—that is, whether the vowel quantity of either the early bilingual or late bilingual group was more affected by lexical stress than that of the other group. Figure 8 presents the mean vowel durations in milliseconds for each vowel by stress position (tonic vs. atonic) and speaker group (early vs. late bilingual).

![Figure 8. Bar chart showing mean vowel durations for each vowel by stress and speaker group.](image)

Although visual differences in the impact of stress by speaker group are apparent in Figure 8, only one interaction between stress and speaker group was significant: The duration difference between tonic and atonic /a/ is significantly greater for the late bilingual group than for the early bilingual group ($p = .007$).

In summary, no overall differences were observed in vowel quantity between the early and late bilingual groups for any vowel. Significant effects of stress were found for two vowels: /e/ and /o/ (with these vowels in tonic position being significantly longer than their atonic counterparts). However, when stress differences were compared between speaker groups, only the differences between tonic and atonic /a/ exhibited significant group differences, with the late bilingual (homeland native) group showing a greater durational difference by stress for /a/ than the early bilingual (heritage) group.
Discussion

The present study explored the production of Spanish vowels by heritage speakers of Spanish (i.e., early Spanish-English bilinguals) in the US and compared their vowel systems with that of a group of “homeland” native Spanish-speaking late Spanish-English bilinguals. Results revealed no significant differences in overall vowel quality or quantity between the two groups. With regard to the potential relationship between lexical stress and vowel production, although unstressed /i/ and /o/ (along the F1 dimension) and /e/ (along the F2 dimension) were statistically distinct in atonic as compared to tonic position overall (i.e., for both speaker groups), no significant differences in vowel production by stress were found between speaker groups. Similarly, although stressed /e/ and /o/ were longer overall than their unstressed counterparts, only differences in the duration of /a/ by stress position were found between speaker groups, with late bilinguals exhibiting greater differences by stress than the early bilingual group.

Compared to previous accounts of heritage Spanish vowels, our results are, in many ways quite similar. For example, considering only our heritage speaker group, the appearance of the vowel space for our 10 heritage speakers is visually similar to that described by and presented in Ronquest (2012). Figures 9 and 10 present the vowel space of the present study’s heritage speakers and the vowel space of Ronquest’s 16 heritage speakers, respectively.5

![Figure 9. Vowel plot showing average normalized F1 and F2 values for the present study’s early bilingual (heritage speaker) group.](image)

5 It should be noted, as may be apparent by the units and scale of the two figures, that Figure 10 from Ronquest (2012) was created using nonnormalized formant values, whereas the vowel space depicted in Figure 9 from the present study is based on normalized formant values.
Both figures picture similarly asymmetrical vowel distributions with condensed back vowel spaces and some fronting of /u/ (although more fronting is apparent in Figure 10 from Ronquest, 2012), especially as compared to the monolingual Spanish vowel space from Quilis and Esgueva (1983) as represented in Figure 2. Nevertheless, when the present study’s heritage speaker vowel space (Figure 9) was statistically compared to that of a group of native Spanish speakers who grew up in Spanish-speaking countries but are now Spanish-English bilinguals (having been introduced to English in late adolescence/adulthood), no significant differences in vowel quality between speaker groups were observed. Thus, it is possible that some of the asymmetry and/or overall differences that have previously been described for heritage Spanish vowel systems (in comparison to the traditional account of the monolingual Spanish vowel space) may not be particular to heritage Spanish speakers but may instead be due (a) to inadequate (as suggested by Ronquest, 2012) or outdated descriptions of the typical Spanish vowel space, which may not be as symmetrical as previously described (as suggested by Willis, 2005, 2008) or (b) to influence from English or from bilingualism in general (as such characteristics were also evident in the vowel spaces of the present study’s native Spanish-speaking late Spanish-English bilinguals). Considered in light of existing models of speech learning, such as Flege’s (1995) Speech Learning Model (SLM), such similarity between bilingual groups (and apparent differences from monolingual groups) could potentially be explained both by the early experience with Spanish of both of our bilingual speaker groups—who established categories for Spanish vowels as first languages—and by the presence of a second language in the phonological space. Flege’s SLM postulates that first language phonetic categories continue to evolve over a speaker’s lifetime and are influence by properties of all phones (from the first or additional languages) that are perceived as realizations of those categories. As such, different groups of Spanish-English bilinguals, despite varying input in, exposure to, and experience in their two languages, might be expected (as was shown in the present study) to demonstrate more similarities to each other in Spanish phonetic production than to monolingual Spanish groups. It is interesting to note that Alvord and Rogers (2014) similarly found relatively few differences between their three groups of immigrants (all of whom, expect one speaker, were bilingual). Thus, despite potential dif-
ferences in input and ongoing use of Spanish by heritage speakers in the US as compared to other native speakers of Spanish, heritage Spanish vocalic systems demonstrate overwhelming similarities to other native Spanish vocalic systems when those other native speakers are also Spanish-English bilinguals.

The present study’s findings regarding the effects of lexical stress also, in many ways, corroborate previous findings for heritage Spanish. For example, Alvord and Rogers (2014) found clear evidence of lexical stress effects with regard to vowel quality—with the unstressed vowel space being significantly more centralized than the tonic vowel space. Nevertheless, as with overall quality differences, Alvord and Rogers, again, found no differences in centralization by participant group, suggesting (similar to the present study) that such centralization is not specific to just bilingual speakers born and educated in the majority language environment (i.e., their Group 3 speakers or our early bilingual group). Ronquest (2013) also observed significant lexical stress differences for her heritage speakers as well as variation in the magnitude of lexical stress effects by vowel. Nevertheless, with regard to vowel quantity, whereas Ronquest compares her heritage speakers’ vowel durations with those of Marín Gálvez’s (1994-1995) monolingual peninsular Spanish speakers and concludes that heritage Spanish vowels are longer overall and show a greater degree of reduction in unstressed contexts than monolingual native speakers’ vowels, the present study’s comparison to a homeland native group of speakers who are also bilingual in English found few significant group differences. In fact, only in the duration of /a/ did the early and late bilingual groups differ significantly, and, surprisingly, it was the late bilingual (homeland) group that exhibited greater differences between stressed and unstressed /a/ duration. Such findings recall those of Cobb and Simonet (2015), whose comparison of Spanish vowel quality by lexical stress as produced by English-speaking second language learners and native Spanish speakers revealed lexical stress effects by all participant groups (native and learner) and, additionally, that differences in the quality of Spanish /e/ in stressed versus unstressed positions were greater for their native control group (a group similar to the present study’s homeland, late Spanish-English bilingual group) than for their group of advanced second language learners of Spanish. Thus, overall, the relative lack of differences in vowel production between various bilingual groups again points to the importance of considering a potential general effect of bilingualism or knowledge of English, as opposed to language learning characteristics specific to heritage speakers, in explaining observed heritage phonetic/phonological patterns.

It should be acknowledged that the present study’s comparison group does not necessarily represent a completely appropriate “baseline” for our heritage speaker group. That is, as summarized by Henriksen (2015), in order to gather information about potential changes to the Spanish linguistic system due to changing experiences and exposure to the heritage and majority languages, comparison with first generation immigrants of the same regional variety may be most fruitful, especially considering that previous generations typically represent the input in the heritage language to which heritage speakers are exposed. Thus, additional future research that compares vowel systems of US Spanish speakers by generation, as did Alvord and Rogers (2014), would provide important insight into the presence and trajectory of some of the differences noted in the present study across generations. Nevertheless, we argue that our participants are representative of the growing population of US university Spanish heritage learners who come from a variety of national backgrounds, with varied linguistic profiles, and who often receive ongoing Spanish input from a variety of sources including homeland native speakers, other heritage speakers, and even L2 learners. Thus, the present comparison group comprised of homeland native speakers of various countries of origin is, we argue, representative of (at
least one of) the speech community(ies) to which these speakers belong (i.e., a diverse university student body and faculty) and provides an additional important point of comparison for these heritage speakers.

As alluded to in the Methodology section, it must be acknowledged that grouping heritage as well as homeland native speakers from various backgrounds may mask dialectal as well as individual differences in vowel production stemming from, for example, specific language learning practices or histories. Although a detailed individual analysis is beyond the scope of this study, a cursory look at individual vowel spaces of the present study’s heritage speakers (see Figure 11) shows some variation in vowel quality between speakers, but with no immediately apparent background characteristics that easily explain such variation. The most distinct vowel spaces have been indicated with lines connecting enlarged mean vowels in Figure 11.

As observed in Figure 11, early bilingual participant 2 (EB2) exhibits a backed /i/ and lowered /a/ in comparison to other early bilingual vowel averages. This speaker was born in Queens, New York to two native Spanish-speaking parents from the Dominican Republic, had learned English and Spanish from birth, and had lived in the Dominican Republic for 8 months the previous year. Early bilingual participant 6 (EB6) exhibited an extremely condensed back vowel space, with her /o/ and /u/ both occupying the space of most participants’ /o/. She, too, was born in New York (Manhattan) to two native Spanish-speaking parents from the Dominican Republic and had learned both English and Spanish from birth. Finally, early bilingual participant 8 (EB8), who exhibited a more fronted /u/ than most other early bilingual participants, was also born in New York (Bronx). She indicated that her parents’ native language was English, but she learned English and Spanish from birth and visits family in Puerto Rico every couple of years. Thus, no apparent anomaly in language learning background nor any overarching commonalities between these three speakers explain the distinctive nature of these speakers’ vowel spaces.
It is also interesting to note that those participants born outside of the US (EB3 and EB9) did not differ noticeably in vowel quality from the overall early bilingual group nor can differences by gender/sex explain observed variation as all heritage speakers in the present study were female.

The late bilingual (homeland native) speaker group also showed some individual variation in vowel quality (see Figure 12), most notably with late bilingual speaker 2 (LB2) exhibiting more fronted high vowels (/i/ and /u/) than the other late bilingual participants. This speaker was born in Colombia and moved to the US at the age of 32 (seven years prior to the time of data collection for this study). It should be noted that the differences observed in this speaker’s vowels cannot necessarily be attributed to dialectal differences, as late bilingual participant 4 (LB4) was also from (a similar region of) Colombia and did not exhibit the same fronting of high vowels. No clear differences by speaker sex were observed either (LB1 and LB4 = female; LB 2 and LB 3 = male).

Figure 12. Vowel plot showing individual average normalized F1 and F2 values for the present study’s late bilinguals (homeland native speakers).

Thus, although some individual variation is apparent, the exploration of heritage and homeland native speakers from multiple national backgrounds does not appear to have masked significant dialectal differences nor are differences on the basis of differing experiences on the part of the early bilingual (heritage) speaker group immediately apparent. Instead, by exploring this varied group, this study provides a glimpse into the varied vowel systems of the US’s diverse heritage Spanish-speaking population as well as a picture of their overall commonalities.

Finally, it is worth highlighting the fact that the present study used a different elicitation task than any previous study on heritage Spanish vowels. We believe that the interactive map task used in the present study has many strengths including that it elicited a high number of vowel tokens within particular contexts from meaning-focused speech production. Additionally, although it incorporated written cues (i.e., street names and landmarks), it largely avoided read speech, which has been shown to be less representative of heritage speakers’ bilingual status.
than speech from less-controlled tasks (Colantoni et al., 2016). However, it should be acknowledged that the map task was a dyadic endeavor: All of the present participants interacted with a L2 learner of Spanish during task completion. Although the learner status and approximate proficiency level of the participants’ interlocutors were held constant across participants and speaker groups, it is possible, for example, that our late bilingual group may have reacted differently in speech production toward L2 learners during interaction than our early bilingual group. This possibility is, of course, true of any type of interlocutor (with, for example, a native interlocutor model presenting somewhat different potential limitations), and we believe that the benefits of communicative and meaning-focused speech outweigh the limitations of potential interlocutor effects. Nevertheless, future studies would do well to triangulate data elicitation methods as done, for example, by Ronquest (2012).

Conclusion

The present study analyzed the Spanish vowel realizations of 10 US heritage speakers of Spanish and compared their Spanish vocalic systems with those of four homeland native Spanish-speaking late Spanish-English bilinguals. Results revealed no significant differences in vowel quality or vowel quantity by speaker group. Nevertheless, both groups exhibited vowel production differences by lexical stress with regard both to vowel quality and vowel quantity. The comparison with a group of Spanish speakers whose main differences from the heritage group were those of formal education in Spanish and age at which English was acquired revealed some important differences in findings as compared to previous research on heritage Spanish vowels and bring to the forefront the important issue of comparison groups for research on heritage linguistic systems. Although comparisons to monolingual Spanish do provide information regarding differences between the linguistic systems of such speakers (and, in the case of heritage Spanish vowels, such comparisons have provided important initial evidence that heritage phonetic and phonological systems are not identical to those of monolingual speakers), such comparisons are also limited in their abilities to narrow in on the potential sources of such difference and may lend themselves to, arguably unfair, interpretations of faulty or incomplete acquisition, as opposed to simply patterns related to the distinct realities of mono- versus bi-lingualism (e.g., Pascual y Cabo & Rothman, 2012). The present study’s comparison group—that of native Spanish-speaking late Spanish-English bilinguals—permitted the comparison of two groups whose main difference was not mono- as compared to bi-lingualism, but rather early as compared to late bilingualism with English. Results of this comparison indicated many similarities and few significant differences between early and late Spanish-English bilinguals’ Spanish vowel productions, similar to findings in Alvord and Rogers’s (2014) cross-generational production study, Mazzaro et al.’s (2016) study on the discrimination of Spanish consonants and vowels by heritage speakers as compared to long-term immigrants (similar to the present study’s homeland native group), and Colantoni et al.’s (2016) findings for the realization of Spanish pitch accents (but only in less controlled speech). These commonalities, despite differences in language learning histories, suggest that the characteristics previously ascribed to heritage speakers may, in fact, not be specific to heritage speakers but rather may be characteristics of bilingual systems or, at least, of Spanish-English bilinguals with early exposure to Spanish. Future research, then, should further explore the nature and extent of these similarities and differences as they relate to other types of bilinguals (e.g., L2 learners). Such research will advance our knowledge of the intricacies of bilingual linguistic systems as they relate to particular linguistic targets or foci as
well as to unique and varied linguistic and language learning backgrounds.
Works Cited


Appendix A

Sample maps from elicitation task

Part A

Part B
Appendix B

Table B. 1

Summary of Linear Mixed-Effects Statistics for Models Predicting Normalized F1 Values

<table>
<thead>
<tr>
<th>Vowel</th>
<th>n</th>
<th>Fixed Effects</th>
<th>$\hat{b}$</th>
<th>SE</th>
<th>t</th>
<th>p</th>
<th>Random effects</th>
<th>Variance</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>/y</td>
<td>453</td>
<td>Intercept</td>
<td>1.35</td>
<td>0.11</td>
<td>12.66</td>
<td>&lt;.001</td>
<td>Participant</td>
<td>0.03</td>
<td>0.18</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Speaker group (late bilingual)</td>
<td>0.15</td>
<td>0.17</td>
<td>0.84</td>
<td>.405</td>
<td>Token</td>
<td>0.04</td>
<td>0.19</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Stress (tonic)</td>
<td>0.10</td>
<td>0.10</td>
<td>1.08</td>
<td>.281</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Speaker group × Stress (late, tonic)</td>
<td>0.03</td>
<td>0.16</td>
<td>0.18</td>
<td>.855</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>/e</td>
<td>234</td>
<td>Intercept</td>
<td>0.50</td>
<td>0.11</td>
<td>4.42</td>
<td>&lt;.001</td>
<td>Participant</td>
<td>0.01</td>
<td>0.11</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Speaker group (late bilingual)</td>
<td>0.23</td>
<td>0.12</td>
<td>1.83</td>
<td>.074</td>
<td>Token</td>
<td>0.02</td>
<td>0.14</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Stress (tonic)</td>
<td>0.27</td>
<td>0.12</td>
<td>2.29</td>
<td>.030</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Speaker group × Stress (late, tonic)</td>
<td>-0.03</td>
<td>0.14</td>
<td>-0.18</td>
<td>.859</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>/a</td>
<td>375</td>
<td>Intercept</td>
<td>-0.11</td>
<td>0.05</td>
<td>-2.22</td>
<td>.034</td>
<td>Participant</td>
<td>0.01</td>
<td>0.11</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Speaker group (late bilingual)</td>
<td>0.12</td>
<td>0.08</td>
<td>1.43</td>
<td>.166</td>
<td>Token</td>
<td>0.01</td>
<td>0.11</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Stress (tonic)</td>
<td>0.06</td>
<td>0.05</td>
<td>1.96</td>
<td>.284</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Speaker group × Stress (late, tonic)</td>
<td>-0.00</td>
<td>0.08</td>
<td>-0.06</td>
<td>.955</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>/o</td>
<td>279</td>
<td>Intercept</td>
<td>-0.96</td>
<td>0.06</td>
<td>-17.56</td>
<td>&lt;.001</td>
<td>Participant</td>
<td>0.00</td>
<td>0.06</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Speaker group (late bilingual)</td>
<td>0.15</td>
<td>0.10</td>
<td>1.44</td>
<td>.157</td>
<td>Token</td>
<td>0.03</td>
<td>0.17</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Stress (tonic)</td>
<td>-0.08</td>
<td>0.06</td>
<td>-1.41</td>
<td>.163</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Speaker group × Stress (late, tonic)</td>
<td>-0.07</td>
<td>0.11</td>
<td>-0.64</td>
<td>.524</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>/u</td>
<td>385</td>
<td>Intercept</td>
<td>-1.28</td>
<td>0.10</td>
<td>-13.41</td>
<td>&lt;.001</td>
<td>Participant</td>
<td>0.02</td>
<td>0.15</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Speaker group (late bilingual)</td>
<td>0.04</td>
<td>0.14</td>
<td>0.32</td>
<td>.754</td>
<td>Token</td>
<td>0.07</td>
<td>0.26</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Stress (tonic)</td>
<td>0.17</td>
<td>0.10</td>
<td>1.64</td>
<td>.106</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Speaker group × Stress (late, tonic)</td>
<td>-0.06</td>
<td>0.13</td>
<td>-0.47</td>
<td>.640</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Note.* Reference categories indicated in parentheses.
Table B. 2

Summary of Linear Mixed-Effects Statistics for Models Predicting Normalized F2 Values

<table>
<thead>
<tr>
<th>Vowel</th>
<th>n</th>
<th>Fixed Effects</th>
<th>$\beta$</th>
<th>SE</th>
<th>t</th>
<th>p</th>
<th>Random effects</th>
<th>Variance</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>/i/</td>
<td>453</td>
<td>Intercept</td>
<td>1.35</td>
<td>0.11</td>
<td>12.66</td>
<td>&lt;.001</td>
<td>Participant</td>
<td>0.03</td>
<td>0.18</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Speaker group (late bilingual)</td>
<td>0.15</td>
<td>0.17</td>
<td>0.84</td>
<td>0.405</td>
<td>Token</td>
<td>0.04</td>
<td>0.19</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Stress (tonic)</td>
<td>0.10</td>
<td>0.10</td>
<td>1.08</td>
<td>0.281</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Speaker group $\times$ Stress (late, tonic)</td>
<td>0.03</td>
<td>0.16</td>
<td>0.18</td>
<td>0.855</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>/e/</td>
<td>234</td>
<td>Intercept</td>
<td>0.50</td>
<td>0.11</td>
<td>4.42</td>
<td>&lt;.001</td>
<td>Participant</td>
<td>0.01</td>
<td>0.11</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Speaker group (late bilingual)</td>
<td>0.23</td>
<td>0.12</td>
<td>1.83</td>
<td>0.074</td>
<td>Token</td>
<td>0.02</td>
<td>0.14</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Stress (tonic)</td>
<td>0.27</td>
<td>0.12</td>
<td>2.29</td>
<td>0.030</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Speaker group $\times$ Stress (late, tonic)</td>
<td>-0.03</td>
<td>0.14</td>
<td>-0.18</td>
<td>0.859</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>/a/</td>
<td>375</td>
<td>Intercept</td>
<td>-0.11</td>
<td>0.05</td>
<td>-2.22</td>
<td>0.034</td>
<td>Participant</td>
<td>0.01</td>
<td>0.11</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Speaker group (late bilingual)</td>
<td>0.12</td>
<td>0.08</td>
<td>1.43</td>
<td>0.166</td>
<td>Token</td>
<td>0.01</td>
<td>0.11</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Stress (tonic)</td>
<td>0.06</td>
<td>0.05</td>
<td>1.96</td>
<td>0.284</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Speaker group $\times$ Stress (late, tonic)</td>
<td>-0.00</td>
<td>0.08</td>
<td>-0.06</td>
<td>0.955</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>/o/</td>
<td>279</td>
<td>Intercept</td>
<td>-0.96</td>
<td>0.06</td>
<td>-17.56</td>
<td>&lt;.001</td>
<td>Participant</td>
<td>0.00</td>
<td>0.06</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Speaker group (late bilingual)</td>
<td>0.15</td>
<td>0.10</td>
<td>1.44</td>
<td>0.157</td>
<td>Token</td>
<td>0.03</td>
<td>0.17</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Stress (tonic)</td>
<td>-0.08</td>
<td>0.06</td>
<td>-1.41</td>
<td>0.163</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Speaker group $\times$ Stress (late, tonic)</td>
<td>-0.07</td>
<td>0.11</td>
<td>-0.64</td>
<td>0.524</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>/u/</td>
<td>385</td>
<td>Intercept</td>
<td>-1.28</td>
<td>0.10</td>
<td>-13.41</td>
<td>&lt;.001</td>
<td>Participant</td>
<td>0.02</td>
<td>0.15</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Speaker group (late bilingual)</td>
<td>0.04</td>
<td>0.14</td>
<td>0.32</td>
<td>0.754</td>
<td>Token</td>
<td>0.07</td>
<td>0.26</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Stress (tonic)</td>
<td>0.17</td>
<td>0.10</td>
<td>1.64</td>
<td>0.106</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Speaker group $\times$ Stress (late, tonic)</td>
<td>-0.06</td>
<td>0.13</td>
<td>-0.47</td>
<td>0.640</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Note.* Reference categories indicated in parentheses.
### Summary of Linear Mixed-Effects Statistics for Models Predicting Duration

<table>
<thead>
<tr>
<th>Vowel</th>
<th>n</th>
<th>Fixed Effects</th>
<th>β</th>
<th>SE</th>
<th>t</th>
<th>p</th>
<th>Random effects</th>
<th>Variance</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>/i/</td>
<td>451</td>
<td>Intercept</td>
<td>94.64</td>
<td>9.50</td>
<td>9.96</td>
<td>&lt;.001</td>
<td>Participant</td>
<td>174.70</td>
<td>13.22</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Speaker group (late bilingual)</td>
<td>-25.86</td>
<td>14.73</td>
<td>-1.76</td>
<td>.083</td>
<td>Token</td>
<td>416.90</td>
<td>20.42</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Stress (tonic)</td>
<td>15.98</td>
<td>9.00</td>
<td>1.78</td>
<td>.078</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Speaker group × Stress (late, tonic)</td>
<td>0.75</td>
<td>14.44</td>
<td>0.05</td>
<td>.959</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>/e/</td>
<td>229</td>
<td>Intercept</td>
<td>90.31</td>
<td>12.49</td>
<td>7.23</td>
<td>&lt;.001</td>
<td>Participant</td>
<td>201.90</td>
<td>14.21</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Speaker group (late bilingual)</td>
<td>-5.01</td>
<td>13.81</td>
<td>-0.36</td>
<td>.719</td>
<td>Token</td>
<td>267.70</td>
<td>16.36</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Stress (tonic)</td>
<td>41.79</td>
<td>12.77</td>
<td>3.27</td>
<td>.002</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Speaker group × Stress (late, tonic)</td>
<td>-21.34</td>
<td>14.81</td>
<td>-1.44</td>
<td>.153</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>/a/</td>
<td>363</td>
<td>Intercept</td>
<td>115.08</td>
<td>9.57</td>
<td>12.02</td>
<td>&lt;.001</td>
<td>Participant</td>
<td>267.80</td>
<td>16.36</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Speaker group (late bilingual)</td>
<td>-23.83</td>
<td>12.60</td>
<td>-1.89</td>
<td>.071</td>
<td>Token</td>
<td>938.50</td>
<td>30.63</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Stress (tonic)</td>
<td>8.26</td>
<td>11.66</td>
<td>0.71</td>
<td>.481</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Speaker group × Stress (late, tonic)</td>
<td>39.35</td>
<td>14.22</td>
<td>2.77</td>
<td>.007</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>/o/</td>
<td>274</td>
<td>Intercept</td>
<td>90.90</td>
<td>8.22</td>
<td>11.07</td>
<td>&lt;.001</td>
<td>Participant</td>
<td>220.20</td>
<td>14.84</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Speaker group (late bilingual)</td>
<td>-23.82</td>
<td>15.02</td>
<td>-1.59</td>
<td>.120</td>
<td>Token</td>
<td>336.60</td>
<td>18.35</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Stress (tonic)</td>
<td>32.70</td>
<td>7.97</td>
<td>4.11</td>
<td>&lt;.001</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Speaker group × Stress (late, tonic)</td>
<td>10.62</td>
<td>13.80</td>
<td>0.77</td>
<td>.445</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>/u/</td>
<td>378</td>
<td>Intercept</td>
<td>95.01</td>
<td>12.10</td>
<td>7.85</td>
<td>&lt;.001</td>
<td>Participant</td>
<td>413.00</td>
<td>20.32</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Speaker group (late bilingual)</td>
<td>-20.34</td>
<td>18.46</td>
<td>-1.10</td>
<td>.276</td>
<td>Token</td>
<td>980.70</td>
<td>31.32</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Stress (tonic)</td>
<td>8.34</td>
<td>13.06</td>
<td>0.64</td>
<td>.526</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Speaker group × Stress (late, tonic)</td>
<td>16.87</td>
<td>17.47</td>
<td>0.97</td>
<td>.335</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Note.* Reference categories indicated in parentheses.