10th International Conference on Speech Prosody 2020
25-28 May 2020, Tokyo, Japan

Adult Early-Bilingual Speech Rhythm: Evidence from Spanish and English

Alexander C. Aldrich[1]

[1]Concordia College—Moorhead
aaldrich@cord.edu

Abstract

This study examines the effect of language on rhythm production in adult early Spanish/English bilinguals. Nine subjects were recorded speaking sentences in Spanish, a so-called syllable-timed language, and in English, a so-called stress-timed language. From these data, various rhythm metrics, in addition to speech rate, are analyzed. Additionally, a detailed proficiency and bilingual profile of the speakers is given. The descriptive statistics of the latter suggest that the speakers are slightly English dominant. The inferential statistics of the rhythm metrics indicate that the bilinguals display robust evidence of language-specific rhythm production, in the expected direction: In English, speech rhythm is more variable than in Spanish, and this is true across most measures. Implications for theory and the field of bilingualism in general are discussed.

Index Terms: bilingualism, Spanish, English, rhythm, speech rate

1. Introduction

The present study deals with the speech rhythm of adult, early-bilingual Spanish/English speakers. Cross-linguistic rhythm variation has been well documented dating back to the mid-20th century [1]. On the one hand, [2] proposed that languages rhythmically similar to English, or so-called *stress-timed* languages, differ from those similar to Spanish, or so-called *syllable-timed* languages, in terms of *isochrony,* or the phonetic manifestation of prominent rhythm units (stress feet in the former and syllables in the latter). On the other hand, in [3], rhythm is attributable to a language's syllable structure, its propensity toward vowel reduction, and the phonetic realization of stress. Thus, English—a language that employs a large number of consonants in the onset and coda of its syllables, phonological vowel reduction, and large effects of lengthening in lexical stress—differs typologically from Spanish, which does not. Furthermore, studies have shown that a strict interpretation of isochrony in either stress- or syllable-timed languages is not warranted [3], [4], [5]. Later attempts to operationalize speech rhythm using various measures of duration [6], [7], [8] reignited interest in the possibility of observing cross-linguistic rhythm variation via empirical methods, although not all researchers are in agreement with the methodologies employed [9], [10].

In any event, previous research went on to observe speech rhythm in some bilingual populations that speak typologically different languages, although a great deal of which has been carried out on adult late-learners of an L2. Although no relevant theoretical frameworks dealing with the L2 acquisition of prosodic phenomena have been proposed, one can hypothesize outcomes based on the theories that describe L2 segmental development, such as that of [11]'s Second Language Linguistic Perception (L2LP) model. Under this framework, the L2 perception grammar of late learners is, initially, a copy of that of the L1; as the learner is increasingly exposed to the L2, the grammar is updated, and its state is observable via speech production. Indeed, studies have found that late learners initially exhibit rhythm patterns resembling those of their native language, but, with increased linguistic experience, they are capable of making gains toward target-like rhythm production [12], [13], [14].

Along a similar vein, early bilinguals' rhythm production appears to also benefit from increased linguistic experience, at least in childhood. In [15]'s research on the development of rhythm in early Spanish/English bilingual children, it was found that older children show greater evidence of unique rhythm production between their two typologically different languages than do younger early bilinguals. The authors also investigated the rhythm of adult Spanish/English bilinguals and found that they produce both languages with unique rhythm patterns. Similar results were also found in [16].

Some pertinent issues regarding early adult bilingual rhythm production, however, were not addressed by [15] or [16]. For instance, linguistic experience, particularly language dominance, has been shown to be relevant to the production of rhythm in adult early bilinguals [17]. However, the language background of the speakers was not detailed in these studies. Furthermore, prior research on the effectiveness of commonly used rhythm measures have found some inconsistencies between metrics [10], [18], prompting researchers to encourage future work to not "rely on a single metric," [18, p. 1566], but rather use a combination of rhythm measures to avoid over interpreting any results that may come as the consequence of high variation in the data. Both [15] and [16], however, used only one type of rhythm score, and thus any conclusions relevant to the present study cannot be drawn.

In sum, although some work has been carried out in regard to adult, early-bilingual rhythm, much remains to be uncovered, particularly the extent to which early bilinguals of rhythmically different languages distinguish between the distinct rhythm patterns in production. In light of this, the present paper is guided by the following research questions: What is the nature of the speech rhythm of adult, early bilinguals of Spanish and English? Specifically, do they exhibit a separation between their Spanish and English rhythm? If so, is this pattern consistent across methods of measurement?

A hypothesis to these questions is that—using the language of the L2LP—adult, early-Spanish/English bilinguals will maintain the perception grammar of the two languages updated to reflect a distinct rhythm production between Spanish and English.

 10.21437/SpeechProsody.2020-108

2. Methods

2.1. Speakers

Nine female speakers took part in the research. All were students at a large research university in southwestern United States. They all reported that a) they were at least 18 years of age, b) they were not fluent in any other language besides English and Spanish, c) Spanish was spoken in their home since they were children, and d) they did not have any vision, hearing, or speech difficulties. The mean age of acquisition of Spanish was 0 (i.e., from birth) and that of English was 3.8 (*SD* = 1.1). In other words, the participants all spoke Spanish from birth and began learning English at around age 4.

Three measures of proficiency were also collected: 1) self-reported speaking proficiency in English and in Spanish, 2) English and Spanish LexTALE, and 3) the Bilingual Language Profile (BLP). The speakers rated their English and Spanish speaking proficiency on a scale of 0–6, with 6 being 'very well'. The LexTALE [19] and LexTALE-Esp [20] are standardized tests of "vocabulary knowledge and proficiency" [19, p. 325] of English and Spanish, respectively, that correlate with one's meta-linguistic knowledge [19]. Lastly, the BLP [21] is a standardized metric of language dominance whose scores range from –218 to +218. The closer the score is to 0 the more balanced the bilingual is, while positive scores suggest English dominance. The results from these measures are presented in Table 1. As observed from the data, the participants appear to be slightly English dominant; however, prior studies reporting these measures for speakers of a similar linguistic background reported comparable scores (see [22] for LexTALE-Esp and [23] for BLP).

Table 1: *Descriptive statistics of proficiency measures.*

Proficiency Measure	**Mean (*SD*)**
Self-reported English proficiency	5.4 (0.7)
Self-reported Spanish proficiency	4.2 (0.8)
LexTALE	81% (14)
LexTALE-Esp	60% (11)
Bilingual Language Profile	50 (28)

2.2. Materials

The data analyzed in this study come from laboratory materials, which are comprised of eight complete sentences spoken in Spanish and eight separate ones in English. An effort was made to control the materials for mean number of syllables between languages (cf., Spanish [*M* = 18.8, *SD* = 1.5] and English [*M* = 17.8, *SD* = 1.3]). Likewise, none of the materials in either language included rhotic (/r, ɾ, ɝ/) or lateral phonemes (/l/), neither did they include glides (i.e., [j, w]) at syllable boundaries. (E.g., Spanish: "Te digo que Pepe nunca pasa tiempo en esta discoteca." *I'm telling you that Pepe never spends time at this club.* / English: "Dad posted on Facebook that he got a new job as a mechanic.")

Each sentence is considered an individual token, and each was repeated three times in a different, random order. Thus, each participant provided 48 tokens (8 sentences × 3 repetitions × 2 languages), for a study total of 432 tokens (48 tokens × 9 participants). For each participant, the most fluent of the three repetitions of each token, defined as that which contained the fewest pauses and the most natural-sounding speech, was selected as the representative sample of that token for that individual and was that which was submitted to analysis. Thus, for each speaker, 16 tokens were submitted to the final data set (8 sentences × 2 languages), for a total of 144 tokens (9 participants × 8 sentences × 2 languages).

2.3. Recordings

The English-language materials were recorded first, followed by the Spanish, using PsychoPy 2 [24] as the method of presentation. Although all instructions were in English (to facilitate another experiment), the participants were told that the materials that they were about to read were in English or Spanish, respectively. Following [25], participants were instructed to repeat each sentence aloud in their normal, conversational voice at a rate that felt natural and comfortable. They were asked to say the entire sentence without pausing while speaking and to repeat the sentence if they made a mistake.

All recordings were carried out in a sound-attenuated booth using a head-mounted, dynamic microphone and a Fostex DC-R302 digital recorder. Signals were digitized at 44.1 kHz, 16-bit quantization.

2.4. Segmentation of acoustic data

A systematic approach, using a combination of spectral information and waveform shape, was employed to segment the acoustic data by hand in Praat [26]. Contiguous vowels were considered a single vocalic interval, and, likewise, contiguous consonants were considered a single consonantal interval. The initial boundary of all utterances was considered the onset of modal voicing of the first vowel of the utterance and was marked at the upward zero crossing. Due to the difficulty of reliably segmenting phrase-final vowels, which display a decrease of formant intensity, possibly due to their propensity to lengthen in phrase-final position (Klatt, 1976), the final boundary of all utterances was determined to be the offset of the F2 of the second-to-last vowel of the utterance. In other words, all utterance durations began and ended with a vowel. Any intra-sentential pauses produced by the speakers were excluded from the analysis, as were sequences affected by pre-pausal lengthening. For further detail, see the methods described in [27, pp. 131–135].

2.5. Rhythm metrics

Several different rhythm metrics were derived from these data using a Praat script. Some—including ΔV, ΔC, %V [6], VarcoV, and VarcoC [7]—measure rhythm globally, while others—nPVI-V and rPVI-C [8]—measure it locally. Similarly, as speech rate has been shown to affect timing [7], some metrics that normalize for changes in speech rate (VarcoV, VarcoC, nPVI-V) were also included. Lastly, speech rate is measured independently as an additional assessment of the language dominance of the speakers. The names of these measures and how they are calculated are summarized in

Table 2. Similar to previous studies [6], [17], each is multiplied by a metric-specific variable for the sake of comparison. Note that, due to English's tendency for ambisyllabic words and ambiguous syllable boundaries [28], vocalic and consonantal intervals are used to calculate speech rate instead of syllables.

Table 2: *Rhythm metrics used in this study, defined.*

Metric	Calculation
ΔV	*SD* of vocalic intervals, times 1000.
ΔC	*SD* of consonantal intervals, times 1000.
%V	Vocalic intervals summed and divided by utterance duration, times 100.
VarcoV	*SD* of vocalic intervals, divided by the mean of vocalic intervals, times 100.
VarcoC	*SD* of consonantal intervals, divided by the mean of consonantal intervals, times 100.
nPVI-V	Absolute difference in duration between sequential vowel intervals, divided by the mean of the two intervals, summed across all samples, divided by the number of samples, times 100.
rPVI-C	Absolute difference in duration between sequential consonantal intervals, summed across all samples, divided by the number of samples, times 1000.
Speech Rate	Sum of number of vocalic and consonantal intervals divided by utterance duration.

With regard to these metrics, based on the nature of the timing patterns of English and Spanish, one can postulate the direction of outcomes in terms of the hypotheses of this study. English and Spanish differ in their phonetic realization of syllable structure, phonological vowel reduction, and stress, as would be expected in a comparison of stress-timed and syllabled-timed languages [3], respectively. Specifically, English allows up to four consonants in the coda [29, p. 34], whereas only two are permitted in the same position in Spanish [30, p. 101]. Furthermore, English exhibits phonological vowel reduction in unstressed syllables [31, p. 94], whereas Spanish does not [32, pp. 126–127]. Lastly, a substantial effect of stress-induced vowel lengthening is observed in English [33], while this is not the case for Spanish [29, p. 68]. Taken together, these differences lead one to expect that, on the one hand, the values for ΔV, ΔC, VarcoV, VarcoC, nPVI-V, and rPVI-C will be greater in English than in Spanish. On the other hand, one would expect %V to be less in English than in Spanish, and speech rate to not be different between the two languages.

2.6. Analysis

In all, after eliminating two tokens due to mispronunciations, a total of 142 tokens were included in the final analysis ((9 participants × 8 items × 2 languages) – 2 misses). The data were averaged over items for a by-subjects analysis as a function of Language (Spanish, English) and were subsequently submitted to repeated-measures ANOVAs using the *ezANOVA* function from the *ez* package [34] in R [35] via the RStudio [36] graphical user interface. In these analyses, the dependent variable was always the rhythm metric in question, and the within-subjects independent variable was Language (Spanish, English). Generalized Eta-Squared measures of effect size [37] are also reported.

3. Results

Box-and-whisker plots, shown in Figure 1, graphically depict the median and first and third quartiles (within the box) of the data, data extending 1.5 times the IQR (whiskers), and data (points) lying outside this range.

A series of ANOVAs determined that the value for the speakers' English was significantly greater than that of Spanish for ΔV, ΔC, VarcoV, VarcoC, nPVI-V, and rPVI-C (see Figures 1A–1F, respectively). These inferential statistics are presented with more detail in Table 3.

The findings presented in Table 3 suggest that the early Spanish/English bilingual speakers in this study differentiate their English speech rhythm from that of their Spanish, and that this difference is robust. Furthermore, the findings are in the expected direction: greater durational variability is exhibited among vowel/consonant sequences in English than in Spanish. Importantly, the effect size, as reported by the Generalized Eta-Squared measures, is large [37] for all analyses. In other words, these speakers demonstrated greater global durational variability (ΔC) in English than in Spanish, even after normalizing for speech rate (VarcoC), and the case was the same for vocalic intervals (ΔV and VarcoV, respectively). Similarly, in English, they exhibited greater local durational variability from one vocalic interval (nPVI-V) to the next, as was also true for consonantal intervals (rPVI-C).

Table 3. *Results of repeated-measures ANOVAs with dependent variable (Metric) and independent variable Language (Spanish, English), including Generalized Eta-Squared measures of effect size.*

Metric	*df*	*F* ratio	*p* value	η^2_G
ΔV	(1, 8)	74.66	< .001	0.49
ΔC	(1, 8)	121.50	< .001	0.59
VarcoV	(1, 8)	77.72	< .001	0.66
VarcoC	(1, 8)	102.58	< .001	0.84
nPVI-V	(1, 8)	167.74	< .001	0.71
rPVI-C	(1, 8)	111.40	< .001	0.62

In addition to these rhythm measures, speech rate was also significantly different between Spanish and English, with a slightly higher rate in Spanish than in English (Spanish: M = 10.73, SD = 1.55; English: M = 10.08, SD = 2.18 | $F(1, 8) = 10.06$, $p < .02$, $\eta^2_G = 0.06$). Notice, however, that this effect is small. In terms of %V, the ANOVA did not report a significant difference between the Spanish and English values (Spanish: M = 49.66, SD = 3.45; English: M = 49.44, SD = 5.25 | $F < 1$, $\eta^2_G = 0.00$). Both of these results will be considered further in the discussion.

4. Discussion and Conclusions

In this study, the Spanish and English rhythm production of nine female early Spanish/English bilinguals were analyzed. The speakers' proficiency in both languages was described in detail: Generally, speakers rated themselves as seemingly more proficient in English than in Spanish. Similarly, they obtained apparently higher scores in English on a vocabulary standardized test than in Spanish (LexTALE), and, finally, they appeared to lean slightly more dominant in English (BLP).

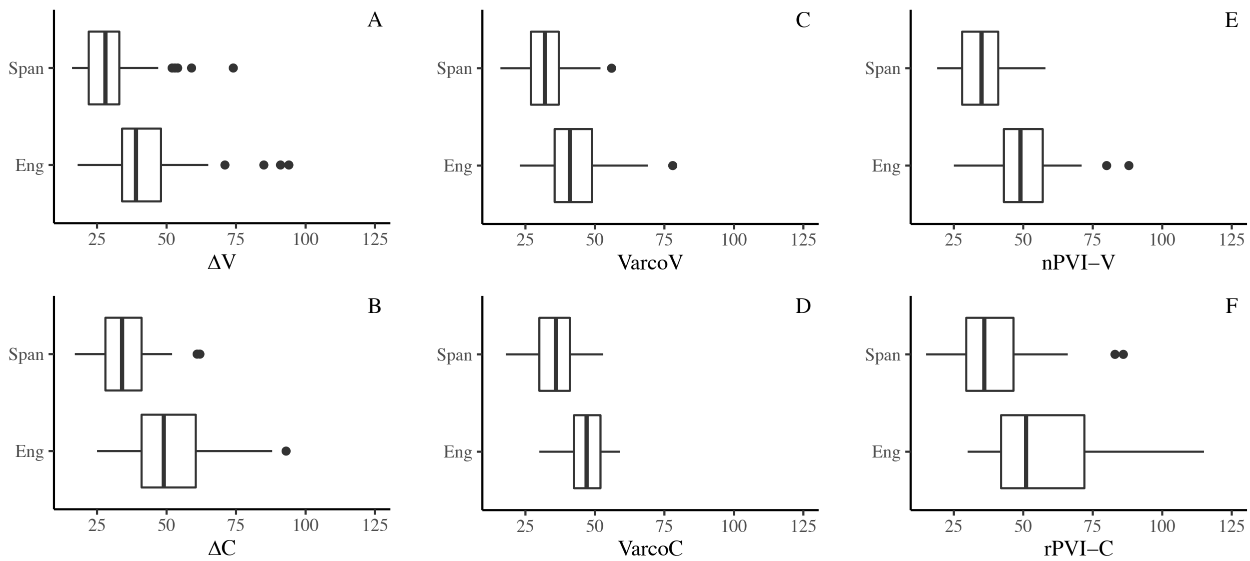

Figure 1: *Effect of language (Spanish, English) on measures of speech rhythm in bilingual speakers' production.*

The results of the inferential statistics suggest that these speakers exhibit language-specific properties in their Spanish and English rhythm production, as the results for ΔV, ΔC, VarcoV, VarcoC, nPVI-V, and rPVI-C all indicated greater variability in English than in Spanish. Applying the language of the L2LP, although designed with segmental acquisition in late-bilinguals in mind, it can be said that these speakers maintain their English and Spanish prosodic grammars updated to reflect the rhythm properties of each—stress-timed (English) on the one hand, and, on the other, syllable-timed (Spanish). Although some evidence from prior research signaled that this might be the case, the present research contributes to the published literature by showing that this effect is large and present across many measures of rhythm both globally and locally.

There was, however, one measure of rhythm in which Spanish and English did not differ: %V, or the proportion of vowels in the utterance. Prior research suggests that Spanish would have a higher %V value than English [8] due to the fact that English has phonological vowel reduction, which leads to a reduction in vowel quantity in unstressed syllables, and has a more complex syllable structure, thus conditioning greater consonant clusters, whereas Spanish does not. As an anonymous reviewer points out, however, most studies in which a rhythmic difference between Spanish and English were found as measured by %V analyzed British English (e.g., [25]), which, at a mean value of 38, appears to be lower than dialects of English of the Western Hemisphere, possibly due to a greater propensity for vowel reduction in the former than in the latter. For instance, %V in Canadian English is reported as a mean of 47 [38], and [13], whose speakers are from the same region of the United States as those of the present study, reports an approximately similar mean %V, both of which align with the reported mean of 49 here. Importantly, the mean %V value of 48 reported elsewhere [25] for native speakers of Spanish is comparable to the mean of 49 reported here for the early Spanish/English bilinguals, as well. In sum, the %V values reported in this study reflect those reported in previous literature, and, in light of the null effect for this sample of early Spanish/English bilinguals, this metric may not be effective at distinguishing rhythm categories in this population.

Lastly, in regard to speech rate, a small effect was found in which the rate was higher in Spanish than in English. As pointed out by an anonymous reviewer, this finding is intuitive when taking into account the rhythm properties of each language. Considering that Spanish has a simpler syllable complexity than English, including simpler consonant clusters, more syllables—or, as rate was measured here, vowel or consonantal sequences—can be produced than in English in the same amount of time, which would lead to a greater rate of speech. Future research on early bilingual speech rhythm should seek to include a sample of monolingual speakers of each language for comparison.

In conclusion, this study presents comprehensive rhythm production data from nine early Spanish/English bilinguals. A detailed description of the speakers' linguistic experience is provided. The results of the inferential statistics indicate a significant effect of language on the speech rhythm in these talkers: For most measures of rhythm, including those that gauge rhythm globally and locally, and those that account for differences in speech rate, English was more variable than Spanish, as would be expected based on the traditional accounts of the phonetic manifestations of rhythm in these languages. These findings have theoretical implications for bilingualism at large: Early Spanish/English bilinguals differentiate their two languages in terms of rhythm, suggesting a possibly unique abstract organization for each at the prosodic level.

5. Acknowledgments

This study was derived from the author's dissertation work. The research was supported by a Graduate and Professional Student Council (GPSC) Research and Project (ReaP) Grant (RSRCH-216FY'18) from the University of Arizona. The author would like to thank those who participated in this study, Miquel Simonet for his dedicated mentorship on this project, and the four anonymous reviewers for their feedback.

6. References

[1] A. Lloyd James, *Speech signals in telephony*. London, UK: Sir Isaac Pitman & Sons, LTD, 1940.

[2] K. L. Pike, *The Intonation of American English*. Ann Arbor: University of Michigan Press, 1947.

[3] R. M. Dauer, “Stress-timing and syllable-timing reanalyzed,” *J. Phon.*, vol. 11, pp. 51–62, 1983.

[4] A. C. Aldrich and M. Simonet, “Duration of syllable nuclei in Spanish,” *Stud. Hisp. Lusoph. Linguist.*, vol. 12, no. 2, pp. 247–280, 2019.

[5] M. H. Cuenca, “Análisis instrumental de la duración de las vocales en español,” *Philol. Hispalensis*, vol. 11, pp. 295–307, 1996.

[6] F. Ramus, M. Nespor, and J. Mehler, “Correlates of linguistic rhythm in the speech signal,” *Cognition*, vol. 73, no. 3, pp. 265–292, 1999.

[7] V. Dellwo, “Rhythm and Speech Rate: A Variation Coefficient for deltaC,” in *Language and language-processing: Proceedings of the 38th Linguistics Colloquium*, P. Karnowski and I. Szigeti, Eds. Frankfurt, Germany: Peter Lang., 2006, pp. 231–241.

[8] E. Grabe and E. L. Low, “Durational Variability in Speech and the Rhythm Class Hypothesis,” *Pap. Lab. Phonol. 7*, pp. 515–546, 2002.

[9] R. M. Dauer, “Phonetic and phonological components of language rhythm,” *Proc. XIth Int. Congr. Phonetic Sci.*, vol. 5, pp. 447–450, 1987.

[10] A. Arvaniti, “Rhythm, timing and the timing of rhythm,” *Phonetica*, vol. 66, no. 1–2, pp. 46–63, 2009.

[11] J.-W. van Leussen and P. Escudero, “Learning to perceive and recognize a second language: the L2LP model revised,” *Front. Psychol.*, vol. 6, no. August, p. 1000, 2015.

[12] P. Trofimovich and W. Baker, “Learning Second Language Suprasegmentals: Effect of L2 Experience on Prosody and Fluency Characteristics of L2 Speech,” *Stud. Second Lang. Acquis.*, vol. 28, no. 1, pp. 1–30, 2006.

[13] A. Li and B. Post, “L2 Acquisition of Prosodic Properties of Speech Rhythm,” *Stud. Second Lang. Acquis.*, vol. 36, no. 02, pp. 223–255, 2014.

[14] L. van Maastricht, E. Krahmer, M. Swerts, and P. Prieto, “Learning direction matters: A study on L2 rhythm acquisition by Dutch leaners of Spanish and Spanish learners of Dutch,” *Stud. Second Lang. Acquis.*, vol. 41, no. 1, pp. 87–121, 2019.

[15] F. Bunta and D. Ingram, “The Acquisition of Speech Rhythm by Bilingual Spanish- and English-Speaking 4- and 5-Year-Old Children,” *J. Speech, Lang. Hear. Res.*, vol. 50, no. 4, pp. 999–1014, 2007.

[16] P. M. Carter and T. Wolford, “Cross-generational prosodic convergence in South Texas Spanish,” *Spanish Context*, vol. 13, no. 1, pp. 29–52, 2016.

[17] N. Henriksen and S. Fafulas, “Prosodic timing and language contact: Spanish and Yagua in Amazonian Peru,” *Stud. Hisp. Lusoph. Linguist.*, vol. 10, no. 2, pp. 225–257, 2017.

[18] L. Wiget, L. White, B. Schuppler, I. Grenon, O. Rauch, and S. L. Mattys, “How stable are acoustic metrics of contrastive speech rhythm?,” *J. Acoust. Soc. Am.*, vol. 127, no. 3, pp. 1559–1569, 2010.

[19] K. Lemhöfer and M. Broersma, “Introducing LexTALE: A quick and valid Lexical Test for Advanced Learners of English,” *Behav. Res. Methods*, vol. 44, no. 2, pp. 325–343, 2012.

[20] C. Izura, F. Cuetos, and M. Brysbaert, “Lextale-Esp: A Test to Rapidly and Efficiently Assess the Spanish Vocabulary Size,” *Psicológica Int. J. Methodol. Exp. Psychol.*, vol. 35, no. 1, pp. 49–66, 2013.

[21] D. Birdsong, L. M. Gertken, and M. Amengual, “Bilingual Language Profile: An Easy-to-Use Instrument to Assess Bilingualism,” *COERLL, University of Texas at Austin*, 2012. [Online]. Available: https://sites.la.utexas.edu/bilingual/. [Accessed: 01-May-2018].

[22] L. J. Hoversten, T. Brothers, T. Y. Swaab, and M. J. Traxler, “Early processing of orthographic language membership information in bilingual visual word recognition: Evidence from ERPs,” *Neuropsychologia*, vol. 103, no. June, pp. 183–190, 2017.

[23] J. V. Casillas, “The Longitudinal Development of Fine Phonetic Detail in Late Learners of Spanish,” (Doctoral dissertation.) University of Arizona, Tucson, AZ, 2016.

[24] J. Peirce *et al.*, “PsychoPy2: Experiments in behavior made easy,” *Behav. Res. Methods*, vol. 51, pp. 195–203, 2019.

[25] L. White and S. L. Mattys, “Calibrating rhythm: First language and second language studies,” *J. Phon.*, vol. 35, no. 4, pp. 501–522, 2007.

[26] P. Boersma and D. Weenink, “Praat: Doing phonetics by computer [Computer Program].” 2018.

[27] A. C. Aldrich, “Gestural Development in Adult Second Language Acquisition: Targets and Timing,” (Doctoral dissertation.) University of Arizona, Tucson, AZ, 2019.

[28] A. Cutler, J. Mehler, D. Norris, and J. Segui, “The syllable’s differing role in the segmentation of French and English,” *J. Mem. Lang.*, vol. 25, no. 4, pp. 385–400, 1986.

[29] S. M. Whitley, *Spanish/English Contrasts: A Course in Spanish Linguistics*. Washington, D.C.: Georgetown University Press, 2002.

[30] J. I. Hualde, A. Olarrea, A. M. Escobar, and C. Travis, *Introducción a la lingüística hispánica*, 2nd Ed. Cambridge, UK;New York; Cambridge University Press, 2010.

[31] P. Ladefoged, *A Course in Phonetics*, 5th Ed. Boston, MA: Michael Rosenberg, 2006.

[32] J. I. Hualde, *The sounds of Spanish*. New York, NY: Cambridge University Press, 2005.

[33] K. de Jong, “Stress, lexical focus, and segmental focus in English: Patterns of variation in vowel duration,” *J. Phon.*, vol. 32, no. 4, pp. 493–516, 2004.

[34] M. A. Lawrence, “ez: Easy Analysis and Visualization of Factorial Experiments.” R package version 4.4-0, 2016.

[35] R Core Team, “R: A Language and Environment for Statistical Computing.” R Foundation for Statistical Computing, Vienna, Austria, 2019.

[36] RStudio Team, “RStudio: Integrated Development for R.” RStudio, Inc., Boston, MA, 2019.

[37] R. Bakeman, “Recommended Effect Size Statistic,” *Behav. Res. Methods*, vol. 37, no. 3, pp. 379–384, 2005.

[38] I. Grenon and L. White, “Acquiring rhythm: A comparison of L1 and L2 speakers of Canadian English and Japanese,” in *BUCLD 32: Proceedings of the 32nd annual Boston University Conference on Language Development*, H. Chan, H. Jacob, and E. Kapia, Eds. Somerville, MA: Cascadilla Press, 2008, pp. 155–166.