Speech Prosody 2016
31 May - 3 Jun 2106, Boston, USA

Convergence effects in Spanish-English bilingual rhythm

Nicholas Henriksen

University of Michigan
nhenriks@umich.edu

Abstract

This study examines the rhythmic properties of highly proficient Spanish-English bilinguals. Ten Spanish-English bilinguals who live in the United States read aloud sentences in Spanish (their L1) and English (their L2), and ten Spanish and ten English monolingual control speakers read aloud sentences in their respective native languages. We performed an acoustic analysis of vocalic and consonantal durations and applied four timing metrics with the intent of measuring rhythmic variability across groups. We used two Pairwise Variability Index (PVI) metrics (nPVI-V and rPVI-C) and two variation coefficient (Varco) metrics (VarcoV and VarcoC). Results show that the Spanish and English monolingual control groups pattern as expected for syllable- and stress-timed languages, respectively. The bilingual speakers show separate statistical distributions in their L1 and L2 rhythms, implying that they produce separate rhythms in their two languages. Further analysis revealed that the bilingual speakers were consistently English-like in English (their L2) but were not consistently Spanish-like in Spanish (their L1). These data support the notion that extensive exposure to an L2 facilitates development toward native-like command of L2 rhythmic patterns, and that this can further lead to L2-to-L1 convergence effects.

Index Terms: rhythm, bilingualism, English, Spanish, convergence

1. Introduction

Rhythm is a feature of suprasegmental phonology (or prosody) that refers to the relative timing of strong and weak prosodic units in speech [19, 21, 23]. Rhythm had initially been imagined as the product of *isochrony*, the maintenance of equal duration between prosodic units. *Syllable-timed* languages (e.g., Spanish) were understood to maintain syllables of equal duration, while *stress-timed* languages (e.g., English) were believed to maintain equal duration between feet. Experimental evidence, however, failed to support affording isochrony a fundamental role in rhythm [7, 8]. Researchers have since put forth the notion that rhythmic differences between languages are the result of the confluence of a series of phonetic and phonological properties, namely syllable structure complexity, vowel reduction, and lengthening of phrase-final vowels [3, 13].

Timing is a set of acoustic metrics (e.g., PVI's, Varco's, etc.) that facilitates the classification of languages according to rhythm, understood here as a typological linguistic trait. These global acoustic metrics typically quantify a language's consonantal and vocalic variability, and, as argued in Arvaniti [1], timing plays an important role in the understanding of rhythm. This is because languages can be more meaningfully classified according to the categories of rhythm (e.g., stress-timing, syllable-timing, etc.) by means of timing metrics. Timing metrics have proven useful for documenting within-language dialect variation [10, 26] as well as for understanding differences between groups of bilingual and monolingual speakers of the same language [11, 29].

Regarding L2 rhythm, research shows that there are L1-to-L2 effects such that bilingual speakers produce intermediate rhythm values in their L2 [12, 24, 29]. Robles-Puente [24] shows that adult Spanish-English bilinguals who moved to the United States in early childhood display syllable-timed rhythmic patterns in both languages. Carter [9] and Thomas & Carter [27] show that US-based heritage speakers of Spanish display intermediate timing values (i.e., between L1 and L2) in both their English and Spanish. The latter two studies confirm a recurrent finding in bilingualism research, namely that bilingual speakers exhibit bidirectional (or convergence) effects in their phonology: L1 sound patterns affect L2 production, and L2 sound patterns affect L1 production [15, 16, 17]. This is because, as predicted by models of L2 phonological acquisition [4, 16], L1 phonological patterns are gradually restructured with increased L2 experience (i.e., L1 converges toward L2), and vice versa. In this study we consider to what extent bidirectional influences are possible for highly proficient Spanish-speaking L2 learners of English who moved to an L2-dominant community (United States) well beyond adolescence. Since our speakers are bilingual in languages that are rhythmically different, they present an ideal test case for potential L1-L2 interactions with regard to speech rhythm. Given the evidence for bidirectional influence between L1 and L2 in other domains of prosodic structure [20, 25], we hypothesize that bidirectional influence should also be present at the level of speech rhythm for our Spanish-English bilingual speakers. In other words, we predict that there will be L1-to-L2 as well as L2-to-L1 transfer effects.

2. Methods

2.1. Speakers

We collected speech data from 30 speakers: ten Spanish-English bilinguals (from Michigan, United States); ten monolingual Spanish controls (from Toledo, Spain); and ten monolingual English controls (from Michigan, United States). There were five men and five women per speaker group. We recorded data from the Spanish-English bilinguals and the English monolinguals in the United States and from the Spanish monolinguals in Spain. The age range for all speakers was 25-50 years old, and the mean age was 35.25 years old. The bilingual and monolingual Spanish speakers were native speakers of north-central Peninsular Spanish.

The bilingual speakers were born in Spain and moved to the United States between the ages of 24 and 28 to pursue

 doi: 10.21437/SpeechProsody.2016-148

graduate studies. At the time of recording, all speakers had lived in the United States for a minimum of 10 years, with a mean of 15 years. In order to assess language dominance, our bilingual speakers completed the Bilingual Language Profile (BLP) assessment [5]. The BLP uses self-reported data on language use, identity, and attitudes to achieve a bilingual score for a given speaker. For our version of the BLP, possible scores ranged from +180 (Spanish-dominant) to -180 (English-dominant). The range of BLP scores for our bilingual speakers was +32 to +139, and the mean BLP score was +50.7, indicating weak Spanish dominance.

2.2. Speech materials

All speakers participated in a sentence reading task. The bilinguals read sentences in Spanish and English, and the monolinguals read sentences in their respective native languages. The bilinguals read all sentences in Spanish before reading all sentences in English, and interacted with the researcher (who was proficient in both languages) in Spanish when reading the Spanish sentences and in English when reading the English sentences. We created the sentences to understand the extent to which rhythm metrics are sensitive to variability based on the segmental material of speech data. As demonstrated in [2, 22, 30], the segmental composition of test sentences can have a significant effect on metric scores, even within a given language. We therefore created three conditions per language to test for the effects of syllable structure on rhythm: CV sentences (with mostly "simpler" CV syllables), CVC sentences (with more "complex" CVC syllables), and uncontrolled sentences (selected from novels). For this study we used the same sentences from Arvaniti [2] for both languages. There were five sentences for each condition, and speakers read each sentence twice, yielding 30 sentences per speaker per language. This resulted in a total of 1200 utterances submitted to acoustic analysis. We predicted the CV sentences to show rhythmic patterns more prototypical of syllable-timed languages, and the CVC sentences to show patterns more like stress-timed languages. Since the uncontrolled sentences represent utterances typical of each language, we expected these sentences to show more stress-timed patterns for English, and more syllable-timed patterns for Spanish.

2.3. Acoustic analysis

For each utterance, measurements of consonantal and vocalic intervals were made by simultaneous inspection of waveforms and wide-band spectrograms in Praat [6] following standard segmentation criteria. Following Arvaniti [2], we relied on the phonetic properties of segments, rather than their phonological status. Glides were classified as consonantal segments if there was evidence of frication in the spectrogram, but were otherwise classified as vowels if there was not. We also included all segments that could be measured accurately, such as prepausal segments. We excluded utterance-initial stops that lacked acoustic evidence of an oral closure.

Timing metrics were subsequently calculated for each sentence separately using a script in Praat. For each utterance, we computed the following four timing metrics: rPVI-C, nPVI-V, VarcoV, and VarcoC [14, 18]. rPVI-C is a raw measure of consonantal variability, calculated as the average pairwise absolute difference between consecutive consonant intervals. nPVI-V is a speech rate normalized measure of vocalic variability, calculated as the average pairwise absolute difference between consecutive vowel intervals divided by their mean duration. VarcoV and VarcoC are normalized standard deviations: overall standard deviations (ΔV or ΔC) are divided by the mean of their vowel or consonant intervals, respectively. We chose these four metrics because past research shows that they are useful for documenting timing differences between Spanish and English [e.g., 2].

2.4. Statistical analysis

We ran four linear mixed-effects models (LMEMs), each of which corresponded to the one of four dependent variables: nPVI-V, VarcoV, rPVI-C, and VarcoC. Each LMEM included random effects of SPEAKER and SENTENCE within SPEAKER. This approach allowed us to accommodate two levels of correlations among the repeated measurements of the dependent variables: measures within a speaker would have a constant correlation, and measures from the same sentences within the same speaker would have a different correlation. The LMEMs also included fixed effects of GROUP (Spanish monolinguals, Spanish bilinguals, English monolinguals, English bilinguals), SENTENCETYPE (CV, CVC, uncontrolled), GENDER, and all interaction combinations between them.

All models were fitted using the MIXED procedure in SPSS, and degrees of freedom for approximate F-statistics for the fixed effects were computed using a Satterthwaite approximation. Because the degrees of freedom are estimated for the approximate F-statistics, in some cases they may not be whole numbers. The SPSS software only allows users to apply this Satterthwaite approximation, which often results in non-integer degrees of freedom [see 28, pp. 131-132]. Bonferroni adjustments were applied on all pairwise comparisons of means based on the fitted models when fixed factors and/or interaction contrasts were significant.

Our approach to fitting the LMEMs followed the 'top-down' strategy in which we initially included in each model the fixed effects associated with all interactions as well as with the single fixed factors. Whenever a higher-order interaction was not found to be significant, this term was removed and the model was fitted again. We repeated this procedure until a higher-order interaction was found to be significant or until only the single fixed factors remained in the model.

3. Results

In all models there were significant effects of the two-way interaction GROUP*SENTENCETYPE, but not of the two-way interactions GROUP*GENDER or SENTENCETYPE*GENDER or of the three-way interaction GROUP*SENTENCETYPE*GENDER. In this section we therefore plot data and report statistical information for the two-way interaction GROUP*SENTENCETYPE. In Table 1 we summarize the statistical information for the GROUP*SENTENCETYPE interaction for each of the four LMEMs.

DV	df	*F*	*p*
nPVI-V	6, 651.174	2.284	.034
VarcoV	6, 680.908	11.697	$\leq$.001
rPVI-C	6, 644.237	5.739	$\leq$.001
VarcoC	6, 673.576	4.104	$\leq$.001

Table 1. *Statistical information for the GROUP*SENTENCETYPE interaction for the four LMEMs.*

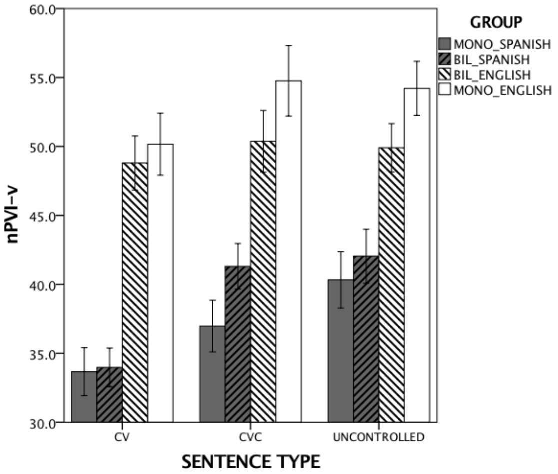

Figure 1: *Mean nPVI-V values.*

GROUP COMPARISON	CV	CVC	Unc.
Mono_Span – Bil_Span	n.s.	*	n.s.
Mono_Span – Bil_Eng	***	***	***
Mono_Span – Mono_Eng	***	***	***
Bil_Span – Bil_Eng	***	***	***
Bil_Span – Mono_Eng	***	***	***
Bil_Eng – Mono_Eng	n.s	n.s.	n.s.

Table 2. *Pairwise comparisons for nPVI-V.*

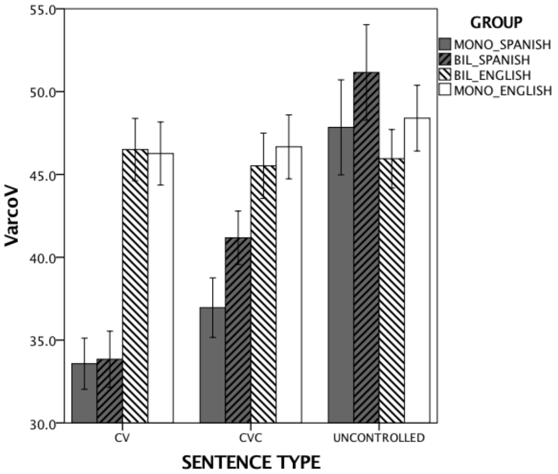

Figure 2: *Mean VarcoV values.*

GROUP COMPARISON	CV	CVC	Unc.
Mono_Span – Bil_Span	n.s.	*	n.s.
Mono_Span – Bil_Eng	***	***	n.s.
Mono_Span – Mono_Eng	***	***	n.s.
Bil_Span – Bil_Eng	***	*	n.s.
Bil_Span – Mono_Eng	***	**	n.s.
Bil_Eng – Mono_Eng	n.s	n.s.	n.s.

Table 3. *Pairwise comparisons for VarcoV.*

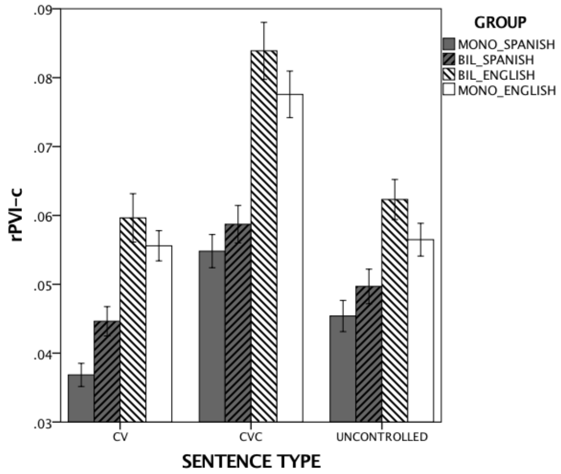

Figure 3: *Mean rPVI-C values.*

GROUP COMPARISON	CV	CVC	Unc.
Mono_Span – Bil_Span	***	n.s.	*
Mono_Span – Bil_Eng	***	***	***
Mono_Span – Mono_Eng	***	***	*
Bil_Span – Bil_Eng	***	***	***
Bil_Span – Mono_Eng	***	***	**
Bil_Eng – Mono_Eng	n.s	n.s.	n.s.

Table 4. *Pairwise comparisons for rPVI-C.*

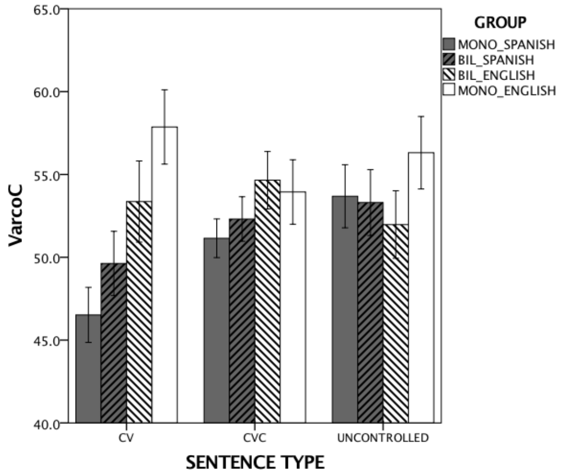

Figure 4: *Mean VarcoC values.*

GROUP COMPARISON	CV	CVC	Unc.
Mono_Span – Bil_Span	*	n.s.	n.s.
Mono_Span – Bil_Eng	*	n.s.	n.s.
Mono_Span – Mono_Eng	***	n.s.	n.s.
Bil_Span – Bil_Eng	*	n.s.	n.s.
Bil_Span – Mono_Eng	***	n.s.	n.s.
Bil_Eng – Mono_Eng	n.s.	n.s.	n.s.

Table 5. *Pairwise comparisons for VarcoC.*

In Figures 1 & 2 and Tables 2 & 3 we report the data for our two vowel metrics, nPVI-V and VarcoV, respectively. In Figures 3 & 4 and Tables 4 & 5 we report the data for our two consonantal metrics, rPVI-C and VarcoC, respectively.

Regarding the Spanish monolingual vs. English monolingual comparison (Mono_Span – Mono_Eng), the English monolinguals displayed higher mean variability scores than the Spanish monolinguals for all sentence types on all metrics. Nine of the twelve pairwise comparisons returned significant differences between the two speaker groups. However, three comparisons did not return significant differences: uncontrolled sentences on VarcoV, as well as CVC and uncontrolled sentences on VarcoC.

Regarding the Spanish monolingual vs. Spanish bilingual comparison (Mono_Span – Bil_Span), the Spanish bilinguals patterned Spanish-like on seven of the twelve comparisons (i.e., there were not significant differences between the Spanish monolinguals and the Spanish bilinguals). However, the two Spanish-speaking groups displayed significant differences for CVC sentences on nPVI-V, CVC sentences on VarcoV, CV and uncontrolled sentences on rPVI-C, and CV sentences on VarcoC. For these five comparisons, the bilinguals' mean values in Spanish were always in the direction of those of the monolingual English speakers, which is suggestive of L2-to-L1 transfer effects.

Regarding the Spanish bilingual vs. English bilingual comparison (Bil_Span – Bil_Eng, i.e., comparing the bilinguals in their two languages), the bilingual speakers patterned like the monolingual control groups in both languages. Specifically, for the nine comparisons in which the monolingual controls showed significant differences between their group means, the bilinguals also showed significant differences between the means of their two languages. For the three comparisons in which the monolingual controls did not show significant differences between their group means, the bilinguals did not show significant differences between the means of their two languages.

Regarding the English bilingual vs. English monolingual comparison (Bil_Eng – Mono_Eng), the bilinguals showed English-like values for all sentence type comparisons on all metrics (i.e., there were never significant differences between the English bilingual means and the English monolingual means). This indicates that although there were inferred L2-to-L1 transfer effects (see above), there was no statistical evidence for L1-to-L2 transfer effects for our bilingual speakers.

4. Discussion

Contra our hypothesis, we did not find full support for bidirectional rhythmic influence between L1 and L2 in the rhythms of our bilingual speakers. On the one hand, the bilinguals did not consistently pattern Spanish-like in Spanish (their L1). On the other hand, the bilinguals consistently patterned English-like in English (their L2). This finding of L2-to-L1 but not L1-to-L2 influence suggests that the L1 rhythmic properties of our bilingual speakers have become gradually (but not totally) restructured with increased L2 experience. This result is in line with Coetzee et al. [11], who studied the rhythmic properties of Afrikaans-Spanish bilinguals who had been living in an L2-dominant community (i.e., Argentina) for more than two-thirds of their lives. Similar to [11], we attribute the lack of L1-to-L2 influence to the fact that our bilinguals have been living in an L2-dominant environment for at least the past decade. This shows that the dominance relationship between a bilingual speaker's languages can affect the extent of influence between the two languages.

The fact that the bilinguals display English-like timing values in English suggests that these speakers exhibit native-like control of processes such as final lengthening and stress-sensitive vowel reduction in their L2. With regard to their Spanish data, the bilinguals displayed intermediate values (i.e., between the Spanish and English monolingual means) in their L1 on five comparisons: CVC sentences on nPVI-V; CVC sentences on VarcoV; CV and uncontrolled sentences on rPVI-C; and CV sentences on VarcoC. One possible explanation for the findings on nPVI-V and VarcoV is that the bilinguals may be applying English-like processes of final lengthening and stress-sensitive vowel reduction when speaking their L1. However, this hypothesis can only be tested with a more focused analysis of the local durational properties of our speakers' rhythmic patterns [e.g., 12].

Regarding methodological implications, we note that the syllabic composition of our stimuli clearly affect the results of timing metrics [2, 22, 30]. In our study there was a consistent effect of the GROUP*SENTENCETYPE interaction in the LMEMs, with generally robust data dispersions between "simpler" CV sentences and more "complex" CVC sentences. However, even when comparing the two monolingual control groups, three of the twelve pairwise comparisons did not return significant results (one for VarcoV and two for VarcoC). Similar to our findings, Tan & Low [26] showed that Varco metrics returned fewer significant differences between varieties of English than vocalic PVI indices. These results thus underscore the importance of considering timing metrics as neither interchangeable for each other nor perhaps even correlated with one another [cf. 10, p. 10]. What is of greater interest for our investigation, perhaps, is that when comparing the bilinguals in their two languages, the results of their between-language comparisons mirrored those of the monolingual control groups. This indicates that our bilingual speakers maintain distinct rhythms to the extent that they contrast in monolingual grammars. Because of this, it is critical that future research on bilingual rhythm take into account the syllabic structure of the stimulus material, and also provide data from comparable monolingual groups of each language.

5. Conclusion

In conclusion, our Spanish-English bilingual speakers' L1 (Spanish) showed partial evidence of rhythmic influence from their L2 (English), but their L2 did not show influence from their L1. Future studies should consider the extent to which the global timing differences that we have documented here can be corroborated by the local durational aspects of bilingual speech [e.g., 12]. It will also be important to compare our findings to those of advanced L2 learners from a non-academic context. Altogether, one important contribution of our study is that we have demonstrated the usefulness of timing metrics for understanding important between-group differences when comparing bilingual and monolingual speakers of a language, as well as for understanding within-group differences when comparing bilingual speakers in their two languages.

6. References

[1] A. Arvaniti, “Rhythm, timing, and the timing of rhythm,” *Journal of Phonetics* 66, pp. 46-63, 2009.

[2] A. Arvaniti, “The usefulness of metrics in the quantification of speech rhythm,” *Journal of Phonetics* 40, pp. 351-373, 2012.

[3] W. Barry and B. Andreeva, “Cross-language similarities and differences in spontaneous speech patterns,” *Journal of the International Phonetic Association* 31(1), pp. 51-66, 2001.

[4] C.T. Best and M.D. Tyler, “Nonnative and second-language speech perception: Commonalities and complementarities,” in Bohn, O.-S., Munro, M.J. (eds.), *Language Experience in Second Language Speech Learning*. Amsterdam: Benjamins, pp.13-34, 2007.

[5] 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. https://sites.la.utexas.edu/bilingual/, 2012.

[6] P. Boersma and D. Weenink, PRAAT: Doing phonetics by computer. <http://www.praat.org/>, 2015.

[7] D. Bolinger, “Pitch accent and sentence rhythm,” in I. Abe & T. Kanekiyo (eds.), *Forms of English: Accent, morpheme, order*, pp. 139-180. Cambridge, MA: Harvard University Press, 1965.

[8] A.M. Borzone de Manrique and A. Singorini, “Segmental duration and rhythm in Spanish,” *Journal of Phonetics* 11, pp. 117-128, 1983.

[9] P. Carter, “Quantifying rhythmic differences between Spanish, English, and Hispanic English,” in Gess, R., Rubin, E. (eds.), *Theoretical and Experimental Approaches to Romance Linguistics.* Amsterdam: John Benjamins, pp. 63-75, 2005.

[10] C. Clopper and R. Smiljanic, “Regional variation in temporal organization in American English,” *Journal of Phonetics* 49, pp. 1-15, 2015.

[11] A.W. Coetzee, L. García-Amaya, N. Henriksen, and D. Wissing, “Bilingual speech rhythm: Spanish-Afrikaans in Patagonia,” in The Scottish Consortium for ICPhS 2015 (ed.), Proceedings of the 18th International Congress of Phonetic Sciences. Glasgow, UK: The University of Glasgow, 2015.

[12] A.W. Coetzee and D. Wissing, “Global and local durational properties in three varieties of South African English,” *Linguistic Review* 24, pp. 263-289, 2007.

[13] R.M. Dauer, “Stress-timing and syllable-timing reanalyzed,” *Journal of Phonetics* 11, pp. 51-62, 1983.

[14] V. Dellwo, “Rhythm and speech rate: A variation coefficient for deltaC,” in Karnowski, P., Szigeti, I. (eds.), *Language and Language-Processing: Proceedings of the 38th Linguistics Colloquium.* Frankfurt am Main, Germany: Peter Lang, pp. 231-241, 2006.

[15] J.E. Flege, “The production of “new” and “similar” phones in a foreign language: Evidence for the effect of equivalence classification,” *Journal of Phonetics* 15, pp. 47-65, 1987.

[16] J.E. Flege, “Second language speech learning,” in Strange, W., Jenkins, J. (eds.), *Speech Perception and Linguistic Experience*. Timonium, MD: York Press, pp 233-273, 1995.

[17] J.E. Flege, “Interactions between the native and second-language phonetic systems,” in P. Burmeister, Piske, T., Rohde, A. (eds.), *An Integrated View of Language Development: Papers in honor of Henning Wode*. Trier, Germany: Wissenschaftlicher Verlag, pp. 217-244, 2002.

[18] E. Grabe and E. L. Low, “Durational variability in speech and the rhythm class hypothesis,” in Gussenhoven, C., Warner, N. (eds.), *Laboratory Phonology 7*. Berlin: Mouton de Gruyter, pp. 515-546, 2002.

[19] I. Lehiste, “Isochrony reconsidered,” *Journal of Phonetics* 5, pp. 253-263, 1977.

[20] I. Mennen, “Bi-directional interference in the intonation of Dutch speakers of Greek,” *Journal of Phonetics* 32, pp. 543-563, 2004.

[21] K. Pike, *The intonation of American English*. Ann Arbor: University of Michigan Press, 1945.

[22] P. Prieto, M. Vanrell, L. Astruc, E. Payne, and B. Post, “Phonotactic and phrasal properties of speech rhythm. Evidence from Catalan, English, and Spanish,” *Speech Communication* 54, pp. 681-702, 2012.

[23] F. Ramus, M. Nespor, and J. Mehler, “Correlates of linguistic rhythm in the speech signal,” *Cognition* 73, pp. 265-292, 1999.

[24] S. Robles-Puente, *Prosody in contact: Spanish in Los Angeles.* PhD Dissertation, University of Southern California, 2014.

[25] M. Simonet, “Intonational convergence in language contact: Utterance-final contours in Catalan-Spanish bilinguals,” *Journal of the International Phonetic Association* 41, pp. 157-184, 2011.

[26] R. Tan and E. Low, “Rhythmic patterning in Malaysian and Singapore English,” *Language and Speech* 57(4), pp. 196-214, 2014.

[27] E.R. Thomas and P.M. Carter, “Prosodic rhythm and African American English,” *English World-Wide* 27, pp. 331-355, 2006.

[28] B.T. West, K. B. Welsh, and A. T. Galecki, *Linear mixed models: A practical guide using statistical software* (2nd ed.). Chapman Hall / CRC Press: Boca Raton, FL, 2014.

[29] L. White and S. Mattys, “Calibrating rhythm: First language and second language studies,” *Journal of Phonetics* 35, pp. 501-522, 2007.

[30] L. Wiget, L. White, B. Schuppler, I. Grenon, O. Rauch, and S.L. Mattys, “How stable are acoustic metrics of contrastive speech rhythm?,” *Journal of the Acoustical Society of America* 127, pp. 1559-1569, 2010.