Effects of native dialect on Mandarin listeners' use of prosodic cues to English stress

Zhen Qin, Annie Tremblay

Department of Linguistics, University of Kansas, United States

qinzhenquentin2@ku.edu, atrembla@ku.edu

Abstract

This study investigates the effect of native dialect on the use of prosodic cues to English stress by Standard Mandarin (SM) listeners, Taiwanese Mandarin (TM) listeners, and English listeners. Both SM and TM use fundamental frequency (F0) to realize lexical tones, but only SM uses duration together with F0 to realize lexically contrastive full-full vs. full-reduced stress patterns. Native English listeners and second language learners of English who spoke SM or TM as native language and were at similar proficiencies in English completed a sequence-recall task. English disyllabic non-words that differed in stress placement were resynthesized to contain only F0 cues, only duration cues, or converging F0 and duration cues. The results showed that SM-speaking learners used duration more than TM-speaking learners to recall English non-words. Native dialect is suggested to be considered in second language speech processing models.

Index Terms: lexical stress, prosodic cues, native dialect.

1. Introduction

Languages differ in the extent to which prosodic cues contribute relevant information for lexical access. Second-language (L2) learners weigh prosodic cues to stress as a function of how these cues are used in the native language (L1) (for discussion, see [1]). The present study investigates how the L1 dialect influences Mandarin-speaking L2 learners' use of prosodic cues in the processing of English stress.

In English, although prosodic cues such as fundamental frequency (F0), intensity, and duration signal stressed syllables in accented words [2,3,4], listeners make limited use of these cues, as English stress is also signaled by segmental information (e.g., full vs. reduced vowels; for discussion, see [1]). Although English listeners can use prosodic cues to English stress [5,6], their word recognition is not inhibited by errors in stress placement unless these errors also result in segmental (i.e., vowel quality) changes in the word [7,8,9,10,11].

Prosodic cues have a greater functional load in Mandarin than in English, in part because Mandarin is a tonal language in which lexical identity is signaled primarily by F0 cues (e.g., [mā] 'mother,' [má] 'hemp,' [mǎ] 'horse,' and [mà] 'scold') [12,13]. While all dialects of Mandarin share this tonal system, some dialects differ in whether or not they also have lexical stress. For example, Standard Mandarin (SM) has a stress pattern that contrasts full-full and full-reduced disyllabic words (e.g., respectively, *dongxi* 'west and east' vs. *dongxi* 'stuff') [14,15]. Whereas words with the full-full pattern are stressed on both syllables, those with the full-reduced pattern are stressed only on the first syllable (word-initial/trochaic stress). The reduced vowel in the latter carries a neutral tone rather than a lexical tone and it is shorter than the full vowel [14,15]. These two stress patterns are thus signaled with duration and F0, with these two cues perhaps contributing to lexical access in SM [16,17]. By contrast, Taiwan Mandarin (TM) does not have this stress distinction: Words have the full-full pattern, and the second syllable of disyllabic words is not reduced (e.g., *dongxi* 'west and east' or 'stuff') [13]. Given the absence of the full-reduced stress pattern in TM, F0 may play a more important role than duration for lexical access in TM [18].

Research on Mandarin listeners' use of prosodic cues when processing English words has yielded inconsistent findings. Some speech perception studies found that, like native English speakers, Mandarin-speaking L2 learners of English from mainland China but tested in the US relied more on segmental cues than on prosodic cues to English stress and did not differ from native English speakers in their use of F0 and duration cues to English stress [19]. Other research, however, suggests that Mandarin-speaking L2 learners of English from mainland China and tested in China could use only F0 cues to English tress [20]. Finally, and perhaps more surprisingly, it has been shown that Mandarin-speaking L2 learners of English from Taiwan but tested in the US could use duration alone in the perception of English stress [21].

These inconsistent findings may stem, at least in part, from the different proficiency of the Mandarin-speaking L2 learners of English that were tested: The L2 learners who lived in the US [19,21] are likely to be more proficient than those who lived in China [20]. Once proficiency is controlled for, it is unclear whether the L1 dialect would influence L2 learners' processing of English stress. Whereas the above studies examined L2 learners who spoke different Mandarin dialects [19,20 vs. 21], they did not investigate whether the particular properties of these dialects—specifically, whether or not the L1 dialect has lexical stress—would affect L2 learners' encoding of English stress. L1 effects on L2 learners' use of prosodic cues to stress are well documented [22, 23, 24]; much less is known about the effect of L1 dialect on non-native listeners' use of such cues in speech processing.

To fill this gap, the present study investigates the use of F0 and duration cues to English stress by English listeners and L2 learners of English who speak SM or TM as L1 dialects and are at a similar proficiency in English. For F0 cues, given that SM and TM have similar tonal systems, we predict that the SM and TM groups will not differ from each other in their use of F0 to encode stress in English words. For duration cues, since SM has a lexical stress contrast signaled in part by duration cues, we predict that SM listeners will perform better than TM listeners in the use of duration cues to English stress. Finally, because SM does not have a reduced-full stress pattern, we predict that SM listeners will differ from TM listeners only on words that are stressed on the initial syllable—not on words that are stressed on the final syllable.

2. Methods

We conducted a sequence recall task adapted from [22,23]. This paradigm requires listeners to encode phonetically variable stimuli that differ in stress placement and hold them in short-term memory for a brief amount of time. Given the

10.21437/SpeechProsody.2014-25

memory load it imposes, this paradigm taps into L2 learners' abstract phonological knowledge and is thus ideally suited for investigating the effect of L1 dialect on the use of prosodic cues in the encoding of L2 stress.

2.1. Participants

Participants included 15 L2 learners of English who spoke SM (mean age: 24.2; SD: 4.0) or TM (mean age: 25.7; SD: 3.4) as L1. Their results were compared to those of 15 native English listeners (mean age: 21.3; SD: 3.9). The two groups of L2 learners did not differ in their age of acquisition of English (p>.1) or in their proficiency in English (p>.1), assessed with a cloze test [25]. No participant had any speech or hearing impairments.

2.2. Stimuli

The stimuli were minimal pairs of English non-words that differed in stress placement (e.g., ['fʌði] vs. [fʌ'ði]). The non-words had a $C_1V_1C_2V_2$ structure. To exclude segmental cues of stress such as vowel reduction, /ɪ/, /ʊ/ and /ʌ/ were used in V_1 position and [i] was used in V_2 position, as they are not reduced to schwas in their respective positions in English. For C_1 and C_2, only fricatives were used to avoid providing segmental cues to stress. In total, four segmental non-words (/sɪvi/, /zʊθi/, /fʌði/, and /hʌfi/) were used. All stimuli were produced by a female native speaker of American English (Midwest). She read each non-word four times in the carrier sentence "Say _ again." Two tokens were selected from the four repetitions of each word type. The stimuli included a total 16 tokens (4 segmental words × 2 stress patterns × 2 tokens). Any given sequence (described in the next section) never contained the same token twice.

As shown in Table 1, the non-words with initial vs. final stress differed significantly in F0 (p<.01) and duration (p<.01) ratios of the first syllable (σ1) to the second syllable (σ2).

Table 1. *Acoustic characteristics of non-words with initial stress and non-words with final stress*.

	F0						Duration					
	Word-initial			Word-final			Word-initial			Word-final		
	σ1	σ2	ratio	σ1	σ2	ratio	σ1	σ2	ratio	σ1	σ2	ratio
Mean	226	162	1.4	174	192	0.9	163	293	2.2	140	328	3.0
SD	22	15	0.1	7	12	0.1	76	26	0.9	73	32	1.4

This study investigates the use of F0 and duration cues to the perception of English stress. The stimuli were resynthesized such that stress would be conveyed by both F0 and duration ("F0+duration cues" condition), by F0 alone ("F0 cues" condition), or by duration alone ("duration cues" condition). The non-words were first normalized for intensity at 70 dB. The F0 and duration cues were then manipulated in PSOLA [26] such that the non-words would have the average F0 and/or duration of the natural stimuli reported in Table 1. Thus, in the conditions where F0 conveyed stress placement, the F0 of σ1 and σ2 would be that corresponding to the averages in Table 1, and in the conditions where duration conveyed stress placement, the duration of σ1 and σ2 would be that corresponding to the averages in Table 1. In the condition where F0 was the only cue to stress placement, the duration of each syllable corresponded to the average duration across stress patterns but not across syllables (i.e., 156 ms for σ1 and 310 ms for σ2). This was done in order to preserve word-final lengthening in the words. In the condition where duration was the only cue to stress placement, the F0 of each syllable corresponded to the average F0 across both syllables and stress patterns (i.e., 189 Hz).

2.3. Procedures

The sequence-recall task had two main phases: a familiarization (or association) phase and a testing phase.

In the familiarization phase, participants were trained to associate 1 and 2 (on a keyboard) with, respectively, words with initial stress and words with final stress. This was done using real words (i.e., *trusty* vs. *trustee*). Participants received feedback on whether or not their responses were correct. After 18 trials, an association test was conducted, in which participants were required to correctly identify the stress pattern of the tokens they heard (as 1, word-initial stress, or 2, word-final stress). Only the participants who received an accuracy score equal to or higher than 95% were invited to complete the testing phase.

In the testing phase, participants were asked to recall sequences of four tokens by pressing 1 and 2 in the correct order. Each test trial began with the auditory presentation of a sequence that included two stress-initial tokens and two stress-final tokens with the same segments (e.g., [fʌ'ði] ['fʌði] ['fʌði] [fʌ'ði]). Each sequence consisted of four different tokens from the same cue condition (F0+duration cues, F0 cues, or duration cues). To shorten the duration of the experiment, three of the six possibilities of number ordering (i.e. [1122], [2211], [1212], [1221], [2121], [2112]) were used for two segmental non-words, and the other three possibilities were used for the other two segmental non-words. The order of sequences and of tokens within each sequence was randomized across participants. The experiment included 36 experimental sequences (4 segmental words × 3 conditions × 3 orderings).

The non-words in the sequences were separated by an interstimulus interval of 50 ms (see [22,23]). The last interstimulus interval in the sequence was followed by the prompt "OK" to prevent participants from using echoic memory to recall the sequences. The inter-trial interval was 1,500 ms. Participants completed a practice session of 12 real-word sequences (e.g., *trustee trusty trustee trusty*) to ensure that they would understand the paradigm before the actual experiment.

2.4. Data analysis

Logit mixed-effects models were conducted on the participants' sequence-encoding and word-encoding accuracies. The sequence-encoding accuracy was computed to examine the overall effect of prosodic cue on the processing of English stress; the word-encoding accuracy was computed to compare the processing of English stress for words with initial stress vs. words with final stress. The models were fitted in R, using the lmer( ) function from the lme4 package for mixed-effects models (for discussion, see [27]).

Model 1 analyzed the sequence-encoding accuracy of the three groups in the first three conditions, with cue condition (F0+duration cues vs. F0 cues or duration cues), L1 (English vs. SM or TM), and the interaction between the two as fixed variables. In this model, the F0+duration cues condition and the English group were used as baselines. Model 2 analyzed only the L2 learners' corresponding sequence-encoding

accuracy, with L1 dialect (SM vs. TM) instead of L1 as fixed variable. The baseline for L1 dialect was SM.

Models 3-4 analyzed the participants' word-encoding accuracy in the first three conditions but separately for words with initial vs. final stress, with cue condition (F0+duration cues vs. F0 cues or duration cues), L1 (English vs. SM or TM), and the interaction between the two as fixed variables. The baselines were the same as in Model 1. Models 5-6 analyzed only the L2 learners' word-encoding accuracy in the first three conditions separately for words with initial vs. final stress, but with L1 dialect (SM vs. TM) instead of L1 as fixed variable. The baselines were the same as in Model 2. In all the models, participant and item were crossed random variables.

If our predictions are correct, we should find significant interactions between the effect of duration cues (as compared to the F0+duration baseline) and the effects of L1 and L1 dialect in the sequence-encoding results, with TM listeners showing a larger difference between the duration cues condition and the baseline condition than the English and SM listeners. In the word-encoding results, we should find similar results only for non-words with initial stress.

3. Results

3.1. Sequence-encoding accuracy

The accuracy of the sequence encoding in the three cue conditions is shown for the three groups in Figure 1.

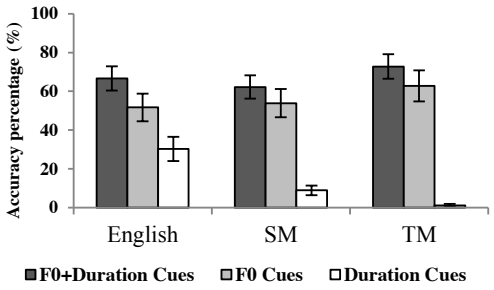

Figure 1. *All participants' mean sequence-encoding accuracy (and standard errors).*

A first logit mixed-effects model was conducted on the participants' sequence-encoding accuracy. The results, reported in Table 3, showed significant effects of cue for F0 and for duration, as well as significant cue (duration) × L1 interactions for both the SM and TM groups. This indicates that both the SM and TM groups showed a larger difference between the F0+duration cues condition and the duration cues condition than the English group did.

Table 2. *Model 1: Logit mixed-effects model on all participants' sequence-encoding accuracy (df=1612).*

Variable	Est.	SE	*z*	*p*
Cue (F0)	–0.7	0.24	–7.1	<.01
Cue (Duration)	–1.8	0.25	–3.2	<.001
L1 (SM)	0.2	0.85	–0.5	>.1
L1 (TM)	0.5	0.47	1.0	>.1
Cue (F0) × L1 (SM)	0.4	0.33	1.1	>.1
Cue (Duration) × L1 (SM)	–1.5	0.42	–3.6	<.001
Cue (F0) × L1 (TM)	0.2	0.35	0.5	>.1
Cue (Duration) × L1 (TM)	–4.8	0.86	–5.5	<.001

A second logit mixed-effect model was conducted on the L2 learners' sequence-encoding accuracy. The results, shown in Table 4, revealed a significant effect of cue for duration, as well as a significant cue (duration) × L1 dialect interaction. This indicates that TM listeners showed a larger difference between the F0+duration cues condition and the duration cues condition than the SM listeners did.

Table 3. *Model 2: Logit mixed-effects model on the L2 learners' sequence-encoding accuracy (df=1615).*

Variable	Est.	SE	*z*	*p*
Cue (F0)	0.4	0.23	–1.7	>.1
Cue (Duration)	–3.3	0.34	–9.6	<.001
L1 dialect	0.7	0.48	1.5	>.1
Cue (F0) × L1 dialect	–0.2	0.34	–0.6	>.1
Cue (Duration) × L1 dialect	–3.3	0.90	–3.7	<.001

To summarize, as we predicted, the SM and TM groups do not differ from each other in the use of F0 cues to encode sequences of English non-words. However, TM listeners differ from both English listeners and SM listeners in the use of duration to encode English non-word sequences.

3.2. Word-encoding accuracy

As mentioned above, word-encoding accuracy was also analyzed to compare listeners' performance on words stressed on the initial syllable and words stressed on the final syllable. Figure 2 displays the accuracy of the individual word encoding for the three cue conditions, with the top panel showing the accuracy for non-words with initial stress and the bottom panel showing the accuracy for non-words with final stress.

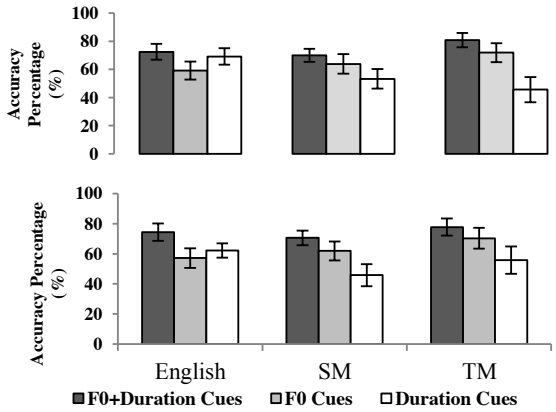

Figure 2. *All participants' word-encoding mean accuracy (and standard errors) for non-words with initial stress (top panel) and non-words with final stress (bottom panel).*

Third and fourth logit mixed-effects models were conducted on the participants' word-encoding accuracy separately for non-words with initial stress and non-words with final stress. The results are reported in Table 4. For non-words with initial stress, the model revealed significant effects of cue for F0, as well as significant cue (duration) × L1 interactions for both the SM and TM groups. This indicates that for non-words with initial stress, both the SM and TM groups showed a larger difference between the F0+duration cues condition and the duration cues condition than the English group did. For non-words with final stress, the model revealed significant effects of cue for F0 and duration, as well

as significant cue (duration) × L1 interactions for both the SM and TM groups. This again indicates that for non-words with final stress, both the SM and TM groups showed a larger difference between the F0+duration cues condition and the duration cues condition than the English group did.

Table 4. *Models 3-4: Logit mixed-effects model on all participants' word-encoding accuracy (df=1612).*

Word-initial stress				
Variable	Est.	SE	*z*	*p*
Cue (F0)	–0.7	0.17	–4.0	<.001
Cue (Duration)	–0.2	0.17	–1.0	>.1
L1 (SM)	–0.2	0.26	–0.7	>.1
L1 (TM)	0.4	0.27	1.6	>.1
Cue (F0) x L1 (SM)	0.4	0.23	1.6	>.1
Cue (Duration) x L1 (SM)	–0.6	0.23	–2.5	<.05
Cue (F0) x L1 (TM)	0.2	0.25	0.6	>.1
Cue (Duration) x L1 (TM)	–1.5	0.24	–6.2	<.001
Word-final stress				
Variable	Est.	SE	*z*	*p*
Cue (F0)	–0.9	0.17	–5.2	<.001
Cue (Duration)	–0.7	0.17	–3.8	<.001
L1 (SM)	–0.1	0.50	–0.2	>.1
L1 (TM)	–0.6	0.50	1.2	>.1
Cue (F0) x L1 (SM)	0.5	0.24	1.8	>.05
Cue (Duration) x L1 (SM)	–0.6	0.24	–2.4	<.05
Cue (F0) x L1 (TM)	0.4	0.26	1.4	>.1
Cue (Duration) x L1 (TM)	–0.8	0.26	–3.0	<.01

Fifth and sixth logit mixed-effects model were conducted on the L2 learners' word-encoding accuracy separately for non-words with initial stress and non-words with final stress. The results are reported in Table 5. For non-words with initial stress, the model revealed a significant effect of cue for duration, a significant effect of L1 dialect, and a significant cue (duration) × L1 dialect. This indicates that TM listeners showed a larger difference between the F0+duration cues condition and the duration cues condition than the SM listeners did for non-words with initial stress. For non-words with final stress, the model revealed significant effects of cue for duration and F0, but no interaction between either cue and L1 dialect. This means that TM listeners did not differ from SM listeners in their use of duration cues in non-words with final stress.

Table 5. *Models 5-6: Logit mixed-effects model on L2 learners' word-encoding accuracy (df=1615).*

Word-initial stress				
Variable	Est.	SE	*z*	*p*
Cue (F0)	–0.3	0.16	–1.8	>.05
Cue (Duration)	–0.7	0.16	–4.7	<.001
L1 dialect	0.6	0.23	2.7	<.01
Cue (F0) x L1 dialect	–0.2	0.24	–0.9	>.1
Cue (Duration) x L1 dialect	–0.9	0.23	–3.9	<.001
Word-final stress				
Variable	Est.	SE	*z*	*p*
Cue (F0)	–0.4	0.17	–2.6	<.01
Cue (Duration)	–1.2	0.17	–7.2	<.001
L1 dialect	–0.7	0.58	1.3	>.1
Cue (F0) x L1 dialect	–0.1	0.26	–0.3	>.1
Cue (Duration) x L1 dialect	0.2	0.26	–0.8	>.1

In summary, the SM and TM groups pattern like each other in the use of F0 to encode stress in individual English non-words, irrespective of whether stress is word-initial or word-final. However, as predicted, SM and TM listeners differ from each other in the use of duration cues to word-initial stress, a stress pattern that exists in SM, but not in the use of duration cues to word-final stress, a stress pattern that does not exist in SM.

4. Discussion and Conclusion

Our results showed that the SM- and TM-speaking L2 learners of English did not differ from each other in the use of F0 cues to process English stress. We attribute these results to the presence of lexical tones in Mandarin, which are signaled primarily by F0 cues.

However, TM listeners differed from both English and SM listeners in the use of duration cues to English stress, with TM listeners being relatively less accurate than the other two groups in the duration cues condition as compared to the F0+duration cues condition. These results were found in the participants' sequence-encoding accuracy and their word-encoding accuracy for non-words with initial stress. These findings are exactly as predicted: In SM but not in TM, duration is used to encode a stress distinction (i.e., full-full vs. full-reduced). This in turns allow SM listeners to use duration for encoding English stress. Importantly, SM does not have a reduced-full stress pattern, and it is precisely in this word-final stress condition that SM and TM patterned similarly.

The results also showed that SM listeners differed from English listeners in the use of duration cues to English stress. One possibility is that the greater occurrence of vowel reduction in English than in SM may lead English listeners to be more sensitive to duration cues than SM listeners. Alternatively, SM may have relatively few words that contrast in stress placement, potentially leading SM listeners to rely less on duration cues as compared to English listeners [13].

Our results are consistent not only with the proposal that non-native listeners weigh prosodic cues to stress as a function of the role of these cues in the L1 [1], but also with more general cue-weighting accounts of L2 speech perception [28,29]. Learning to process English stress depends in large part on the prosodic cues used to realize stress and on whether these cues are similarly used to access words in the L1. The present study also suggests that not only the L1, but also the L1 dialect, plays a crucial role in determining whether non-native listeners can use specific prosodic cues to encode English stress. L2 speech processing models should thus consider the influence of L1 dialect, which may impact how prosodic cues are weighted in L2 processing.

5. Acknowledgements

We would like to thank Yu-fu Chien for help with the data collection in Taiwan; Dr. Allard Jongman, Dr. Joan Sereno, Dr. Jie Zhang, as well as the members of LING 850, for their valuable comments on this research; and all participants.

6. References

[1] Cutler, A. Native listening: Language experience and the recognition of spoken words, MIT Press, 2012.

[2] Beckman, M. E. Stress and Non-Stress Accent, Foris, 1986.

[3] Fry, D. "Duration and intensity as physical correlates of linguistic stress", Journal of the Acoustic Society of America, 27:765-768, 1955.

[4] Lieberman, P. "Some acoustic correlates of word stress in American English", Journal of the Acoustic Society of America 32:451-454, 1961.

[5] Cooper, N., Cutler, A. and Wales, R. "Constraints of lexical stress on lexical access in English: Evidence from native and non-native listeners", Language and Speech, 45:207-228, 2002.

[6] Cutler, A., Wales, R., Cooper, N. and Janssen, J. "Dutch listeners' use of suprasegmental cues to English stress", in J. Trouvain and W. J. Barry [Eds], Proceedings of the 16th International Congress for Phonetic Sciences, 1913-1916, Pirrot, 2007.

[7] Cutler, A. "Forbear is a homophone: Lexical prosody does not constrain lexical access", Language and Speech, 29:201-220, 1986.

[8] Braun, B., Lemhöfer, K. and Mani, N. "Perceiving unstressed vowels in foreign-accented English", Journal of the Acoustical Society of America, 129:376-387, 2011.

[9] Cutler, A. and Clifton, C. E. "The use of prosodic information in word recognition", in H. Bouma and D. G. Bouwhuis [Eds], Attention and performance X, 183-196, Erlbaum, 1984.

[10] Fear, B. D., Cutler, A. and Butterfield, S. "The strong/weak syllable distinction in English", Journal of the Acoustical Society of America, 97:1893-1904, 1995.

[11] Small, L. H., Simon, S. D. and Goldberg, J. S. "Lexical stress and lexical access: Homographs versus nonhomographs", Perception and Psychophysics, 44:272-280, 1988.

[12] Chao, Y.-R. A grammar of spoken Chinese. University of California Press, 1968.

[13] Duanmu, S. The phonology of standard Chinese. Oxford University Press, 2007.

[14] Chen, Y. and Xu, Y. "Production of weak elements in speech—evidence from F0 patterns of neutral tone in Standard Chinese", Phonetica, 63:47-75, 2006.

[15] Lin, M. and Yan, J. "Beijinghua qingsheng de shengxue xingzhi" [Acoustic characteristics of neutral tone in Beijing Mandarin], Fangyan, 3:166-178, 1980.

[16] Shen, X. S. "Relative duration as a perceptual cue to stress in Mandarin", Language and Speech, 36:415-433, 1993.

[17] Ye, Y. and Connine, C. M. "Processing spoken Chinese: The role of tone information", Language and Cognitive Processes Special Issue: Processing East Asian Languages, 14:609-630, 1999.

[18] Hallé, P. A., Chang, Y.-C. and Best, C. T. "Identification and discrimination of Mandarin Chinese tones by Mandarin Chinese vs. French listeners", Journal of Phonetics, 32: 395-421, 2004.

[19] Zhang, Y. and Francis, A.L. "The weighting of vowel quality in native and non-native listeners' perception of English lexical stress", Journal of Phonetics, 38, 260-271, 2010.

[20] Wang, Q. "L2 stress perception: The reliance on different acoustic cues", in P. Barbosa, S. Madureira, and C. Reis [Eds], Proceedings of Speech Prosody 2008, 135-138, Campinas, Brazil, 2008.

[21] Lai, Y. Acoustical Realization and Perception of English Lexical Stress by Mandarin Learners, Unpublished PhD Dissertation. University of Kansas, 2008.

[22] Dupoux, E., Peperkamp, S. and Sebastián-Gallés, N. "A robust method to study stress 'deafness'", Journal of the Acoustical Society of America, 110:1606-1618, 2001.

[23] Dupoux, E., Sebastián-Gallés, N., Navarrete, E. and Peperkamp, S. "Persistent stress 'deafness': The case of French learners of Spanish", Cognition, 106: 682-706, 2008.

[24] Tremblay, A. "Is second language lexical access prosodically constrained? Processing of word stress by French Canadian second language learners of English", Applied Psycholinguistics, 29:553-584.

[25] Brown, J. D. "Relative merits of four methods for scoring cloze tests", Modern Language Journal, 64:311-317, 1980.

[26] Boersma, P. and Weenink, D. Praat: Doing phonetics by computer [computer program], version 5.3, retrieved from http://www.praat.org, December 2012.

[27] Baayen, R. H. Analyzing linguistic data: A practical introduction to statistics. Cambridge University Press, 2008.

[28] Francis, A.L. and Nusbaum, H.C. "Selective attention and the acquisition of new phonetic categories", Journal of Experimental Psychology: Human Perception and Performance, 28(2): 349-366, 2002.

[29] Holt, L.L. and Lotto, A.J. "Cue weighting in auditory categorization: Implications for first and second language acquisition", Journal of the Acoustical Society of America, 119: 3059-3071, 2006.