LANGUAGE LEARNING
A Journal of Research in Language Studies

Language Learning ISSN 0023-8333

EMPIRICAL STUDY

English Learners' Use of Segmental and Suprasegmental Cues to Stress in Lexical Access: An Eye-Tracking Study

Katrina Connell,[a] Simone Hüls,[b] Maria Teresa Martínez-García,[c] Zhen Qin,[d] Seulgi Shin,[b] Hanbo Yan,[e] and Annie Tremblay[b]

[a]The Hong Kong Polytechnic University, [b]University of Kansas, [c]Hankuk University of Foreign Studies, [d]Shanghai Jiao Tong University, and [e]Shanghai International Studies University

This study investigated the use of segmental and suprasegmental cues to lexical stress in word recognition by Mandarin-speaking English learners, Korean-speaking English learners, and native English listeners. Unlike English and Mandarin, Korean does not have lexical stress. Participants completed a visual-world eye-tracking experiment that examined whether listeners' word recognition is constrained by suprasegmental cues to stress alone or by a combination of segmental and suprasegmental cues. Results showed that English listeners used both suprasegmental cues alone and segmental and suprasegmental cues together to recognize English words, with the effect of stress being greater for combined cues. Conversely, Mandarin listeners used stress in lexical access only when stress was signaled by suprasegmental cues alone, and Korean listeners did so only when stress was signaled by segmental and suprasegmental cues together. These

This study was a collaborative research endeavor that stemmed from a graduate seminar taught by the last author. The first six authors, ordered alphabetically, made equal contributions to the research. We would like to thank the editor and the anonymous reviewers for their careful reading of our manuscript and for their insightful comments.

OPEN MATERIALS This article has been awarded an Open Materials badge. All materials are publicly accessible via the IRIS Repository at https://www.iris-database.org. Learn more about the Open Practices badges from the Center for Open Science: https://osf.io/tvyxz/wiki.

Correspondence concerning this article should be addressed to Annie Tremblay, University of Kansas, Department of Linguistics, 1541 Lilac Lane, Blake Hall Room 427, Lawrence, KS 66045. E-mail: atrembla@ku.edu

DOI: 10.1111/lang.12288

results highlight the importance of a cue-based approach to the study of stress in word recognition.

Keywords stress; lexical access; eye tracking; English; Mandarin; Korean

Introduction

Spoken word recognition is influenced not only by segmental information (e.g., acoustic information from which consonants and vowels are extracted from the signal) but also by suprasegmental information, such as fundamental frequency (F0), duration, and intensity. For example, suprasegmental cues to lexical stress distinguish words and thus constrain lexical access in a number of languages, including Dutch (e.g., van Donselaar, Koster, & Cutler, 2005), Spanish (e.g., Soto-Faraco, Sebastián-Gallés, & Cutler, 2001), Italian (e.g., Sulpizio & McQueen, 2012; Tagliapietra & Tabossi, 2005), Greek (e.g., Protopapas, Panagaki, Andrikopoulou, Gutierrez Palma, & Arvaniti, 2016), and English (e.g., Cooper, Cutler, & Wales, 2002). These and other findings have been interpreted as suggesting that suprasegmental information is processed in parallel with segmental information and thus immediately constrains lexical activation and competition (e.g., T. Cho, McQueen, & Cox, 2007). Accordingly, computational models of spoken word recognition have been developed to simulate the use of both types of information in word recognition (e.g., Shuai & Malins, 2017).

In order to recognize words accurately and efficiently, second language (L2) learners must therefore be able to extract both segmental and suprasegmental information from the speech signal and use this information in lexical access. The present study sheds further light on this issue by investigating native and nonnative English listeners' use of segmental and suprasegmental cues to lexical stress in online word recognition. Languages differ in whether or not they have lexical stress—that is, whether the position of the most prominent syllable in the word differs across words—and if so, what types of cues signal stress in the language. Hence, learning to use stress in L2 word recognition is not a trivial task. In English, stress has both segmental and suprasegmental correlates: Stressed syllables contain a full vowel, and unstressed syllables tend to contain a reduced vowel (e.g., Gay, 1978; Lindblom, 1963).[1] Everything else being equal, stressed syllables have higher F0 (in pitch-accented words), longer duration, and higher intensity than unstressed syllables (e.g., Beckman, 1986; Lieberman, 1960). English is thus an ideal language for examining whether L2 learners' ability to use stress in lexical access is contingent on the types of cues that signal stress.

Existing research has shown that L2 learners' use of lexical stress in word recognition is strongly influenced by whether or not their first language (L1) has lexical stress. Using sequence recall and lexical decision tasks, Dupoux, Sebastián-Gallés, Navarrete, and Peperkamp (2008) found that native speakers of Metropolitan French, a language that does not have word-level stress, had difficulty encoding Spanish nonwords that differed in stress placement, where stress was realized with suprasegmental cues (e.g., *NUmi* vs. *nuMI*, with the capitalized letters representing the stressed syllable). Furthermore, these French speakers failed to reject Spanish words realized with the incorrect stress placement, for example, **SAlud* ("health"). Importantly, French speakers had difficulty encoding Spanish stress irrespective of their proficiency in Spanish (for similar findings with French Canadian L2 learners of English, see Tremblay, 2008). On the basis of these and other results, Dupoux and colleagues proposed that French listeners are deaf to stress because they have not set the stress parameter to encode stress in their phonological representation of lexical words (e.g., Dupoux, Peperkamp, & Sebastián-Gallés, 2001; Dupoux et al., 2008; Peperkamp & Dupoux, 2002).

What is unclear from these findings, however, is whether L2 learners whose L1 does not have lexical stress would be able to use stress in L2 word recognition if stress were also encoded with segmental cues. One might expect that the presence of vowel reduction as a cue to English stress would enhance L2 learners' use of stress if they encode this segmental information as part of their lexical representation. L2 learners' ability to do so would depend on whether they can encode English reduced vowels as different from English full vowels.

Lin, Wang, Idsardi, and Xu (2014) shed some light on this question. Using sequence recall and lexical decision tasks, they examined whether Mandarin-speaking and Korean-speaking L2 learners of English could encode nonwords that differed in stress (e.g., *MIpa* vs. *miPA*) and reject English words that were incorrectly stressed. Standard Mandarin has been claimed to have stress distinctions in a limited number of words (e.g., Chao, 1968; Duanmu, 2007) whereas Korean has been argued not to have word-level stress (e.g., Jun, 1998, 2000). Importantly, in their lexical decision task, Lin et al. manipulated whether the incorrect stress placement in English words was signaled by suprasegmental cues alone (e.g., **DIsease*) or by both segmental and suprasegmental cues (e.g., **CONfess* realized as /kɔnfɛs/). As in English, words in Standard Mandarin whose first syllable is stressed and whose second syllable is unstressed have a phonologically reduced vowel in the unstressed syllable (e.g., Chao, 1968; Duanmu, 2007). In contrast, Korean does not have a phonological process of vowel reduction (Cho & Park, 2006).

The results of the sequence recall task revealed that Mandarin listeners were more accurate at encoding stress in English nonwords than both English and Korean listeners (English: 40.1%; Mandarin: 46.5%; Korean: 31.5%). The better performance of Mandarin listeners was attributed to their enhanced use of suprasegmental cues to stress due to the presence of lexical tones in their L1 (see also Qin, Chien, & Tremblay, 2016). Conversely, the results of the lexical decision task indicated that Mandarin listeners were less accurate than English listeners but more accurate than Korean listeners at rejecting incorrectly stressed English words (English: 81.6%; Mandarin: 60.1%; Korean: 40.2%). Crucially, only the English listeners showed an effect of segmental information, with changes in vowel quality increasing their ability to reject incorrectly stressed English words (vowel quality change: 85.4%; no vowel quality change: 77.8%).

A priori, these results appear to suggest that segmental cues to stress do not enhance L2 learners' use of stress in word recognition. However, given the preponderance of L1 transfer effects in the use of stress in L2 word recognition (cf. Cooper et al., 2002; Dupoux et al., 2008; Tremblay, 2008) and the lack of effect of segmental cues in the Mandarin listeners' results, there are good reasons to question this conclusion. One possibility is that a lexical decision task is not sufficiently sensitive for capturing L2 learners' use of segmental cues to stress in English.

The findings of Chrabaszcz, Winn, Lin, and Idsardi (2014) suggest that this may indeed be the case. Native English listeners and Mandarin-speaking L2 learners of English (among other L2 learners) completed a perception task in which they were asked to identify the stressed syllable in disyllabic words. The authors manipulated the acoustic cues to stress, independently crossing segmental cues (full vs. reduced) with suprasegmental cues (low vs. high F0, short vs. long duration, low vs. high intensity) for each syllable. This resulted in stimuli where cues sometimes coincided and sometimes conflicted with a particular stress pattern (e.g., participants could hear /mɑbə/ with longer duration on the first syllable but with higher F0 and intensity on the second syllable). The results were clear in showing that the strongest cue to the perception of stress for both English and Mandarin listeners was vowel quality (for similar results using a similar type of experiment with real English words, see Zhang & Francis, 2010).

Hence, it does not appear to be the case that Mandarin-speaking L2 learners of English do not attend to segmental cues to English stress (cf. Lin et al., 2014). However, because the experiments conducted by Chrabaszcz et al. (2014) and Zhang and Francis (2010) used an explicit stress judgment task, it remains to be seen whether Mandarin listeners can use segmental cues to English stress in the course of spoken word recognition. As for Korean-speaking L2 learners of English, it is still unclear whether they can show sensitivity to segmental cues to English stress and use these cues in lexical access if one assumes that a lexical decision task may not be sufficiently sensitive to tap into L2 learners' sensitivity to vowel reduction as a cue to stress (cf. Lin et al., 2014).

We might expect Korean listeners to be able to encode English stress in the presence of segmental cues if the changes in vowel quality entailed by such cues result in their assimilating English full and reduced vowels to different Korean vowels. For example, Korean listeners may assimilate the initial, full vowel in *PArrot* to the Korean vowel /ɛ/ or /e/, and they may assimilate the initial, reduced vowel in *paRADE* to the Korean vowel /ʌ/. This would result in their ability to distinguish *PArrot* from *paRADE* early on in the word recognition process without needing to encode suprasegmental cues to stress in lexical access. Alternatively, Korean-speaking L2 learners of English may develop a new phonetic category for representing the English schwa instead of assimilating the schwa to their L1 Korean vowels, in which case they would also be able to use segmental cues to stress in English.

The Present Study

The present study aimed to shed further light on these issues. It employed a visual-world eye-tracking experiment with printed words to investigate the use of segmental and suprasegmental cues to stress in lexical access by native English listeners and Mandarin- and Korean-speaking L2 learners of English (for other visual-world studies on the use of stress in spoken word recognition, see Brown, Salverda, Dilley, & Tanenhaus, 2015; Reinisch, Jesse, & McQueen, 2010; Reinisch & Weber, 2012; Sulpizio & McQueen, 2012). Tremblay (2008) showed that knowledge of stress placement is necessary for L2 learners to be able to use stress in word recognition, though knowledge of stress does not ensure that stress will be used in word recognition. Hence, in addition to our matching the two L2 groups for their proficiency in and experience with English, the data analysis included only those words that our L2 learners were sufficiently familiar with and knew how to stress (see below). Accordingly, any difference found between the L2 groups should not be attributable to L2 learners' different proficiency in or experience with English or to their different levels of knowledge of the words used in the experiment.

In addition to examining L2 learners' processing of English stress, this study provided a refined investigation of native English listeners' use of segmental and suprasegmental cues to stress in lexical access. Early research suggested that English listeners showed reduced sensitivity or no sensitivity to stress in word recognition when stress was not realized with segmental cues (e.g., Cutler, 1986; Fear, Cutler, & Butterfield, 1995). However, in a more recent, crossmodal priming study, Cooper et al. (2002) demonstrated significant facilitative priming when listeners heard a monosyllabic or disyllabic word fragment that matched the target word in its segmental and suprasegmental properties (e.g., stressed [mɪs] matching *MYStery*, unstressed [mɪs] matching *misTAKE*) and no priming when the word fragment matched the target in its segmental properties but mismatched it in its suprasegmental properties (e.g., stressed [mɪs] mismatching *misTAKE*, unstressed [mɪs] mismatching *MYStery*). Although these results were weaker than the corresponding results in Dutch (where van Donselaar et al., 2005, found inhibitory priming with mismatching disyllabic fragments), Spanish (where Soto-Faraco et al., 2001, found inhibitory priming with mismatching monosyllabic and disyllabic fragments), and Italian (where Tagliapietra & Tabossi, 2005, found inhibitory priming with disyllabic fragments that matched only one possible lexical word), they nonetheless suggest that native English listeners' lexical access is constrained, at least in part, by suprasegmental information (see Brown et al., 2015, for evidence that English listeners' initial interpretation of suprasegmental cues to stress in lexical access is also constrained by the prosody of the preceding sentential context).

The experimental design used in this study was inspired by Cooper et al. (2002) and by Tremblay (2008). It includes two types of stress-mismatch (experimental) conditions. In the no-vowel-reduction condition, the initial syllable of the target and competitor words was the same segmentally but differed suprasegmentally—that is, it was stressed in the target word and unstressed in the competitor word (e.g., stressed [kɑɹ] in *CARpet* and unstressed [kɑɹ] in *carTOON*). In the vowel-reduction condition, the initial syllable of the target and competitor words differed both segmentally—that is, it had a full vowel in the target word and a reduced vowel in the competitor word—and also suprasegmentally (e.g., stressed [pæɹ] in *PArrot* vs. unstressed [pəɹ] in *paRADE*).[2] Eye fixations to target over competitor words in these two conditions were compared to those in two stress-match (control) conditions where the initial syllable of the target and competitor words were the same segmentally and suprasegmentally (e.g., no-vowel-reduction condition: stressed [kɑɹ] in *CARpet* and *CARton*; vowel-reduction condition: stressed [pæɹ] in *PArrot* and *PArish*).

If stress constrains lexical access, listeners should show greater activation of the target over the competitor word in the stress-mismatch condition than in the stress-match condition. This should be true at least for English listeners

Table 1 Means and standard deviations for L2 learners' language background and proficiency variables

Variable	L1 Mandarin ($n = 29$)	L1 Korean ($n = 30$)
Age of first exposure to English	10.3 (2.6)	10.8 (2.3)
Years of English instruction	13.6 (3.9)	12.1 (5.0)
Time in English speaking environment (months)	38.2 (33.3)	36.4 (48.1)
English proficiency cloze test ($max = 50$)	33.2 (8.1)	32.6 (9.6)

(e.g., Cooper et al., 2002) and Mandarin listeners (e.g., Lin et al., 2014) but possibly also for Korean listeners if they use vowel reduction as a segmental cue to stress (cf. Lin et al., 2014). Furthermore, if the presence of segmental cues to stress enhances lexical access, stress and vowel reduction should interact, with the effect of stress being larger in the vowel-reduction than in the no-vowel-reduction condition. We expected to find such results for English listeners (Chrabaszcz et al., 2014; Fear et al., 1995; Zhang & Francis, 2010) and possibly for Mandarin listeners (Chrabaszcz et al., 2014; Zhang & Francis, 2010; cf. Lin et al., 2014) and Korean listeners (cf. Lin et al., 2014).

Method

Participants

A total of 33 native English listeners (14 females; $M_{\text{age}} = 25.0$ years, $SD = 4.6$), 29 Mandarin-speaking L2 learners of English (17 females; $M_{\text{age}} = 25.7$ years, $SD = 4.7$), and 30 Korean-speaking L2 learners of English (17 females; $M_{\text{age}} = 29.4$ years, $SD = 5.8$) participated in this study. None of the participants reported speech, hearing, or other language impairments. All Mandarin listeners identified Standard Mandarin to be their native dialect, alone or together with another Chinese dialect (Northern dialect = 2, Xiang dialect = 1, Wu dialect = 2).[3] The majority of Korean listeners (19) considered Seoul Korean to be their native dialect. The native dialects of the remaining Korean listeners, all of whom were familiar with Seoul Korean, were Gyeongsang Korean (5), Chungcheong Korean (4), Jeju Korean (1), and Jeolla Korean (1).[4]

L2 learners' age of first exposure to English, number of years of English instruction, and number of months spent in an English-speaking environment are reported in Table 1, together with their performance on an open-ended cloze test that served as a measure of their English proficiency (Brown, 1980). Wilcoxon rank sum tests did not reveal a significant difference between the two L2 groups on any of these nonnormally distributed measures ($p > .20$).

Table 2 Experimental design of visual-world eye tracking experiment

	Stress mismatch[a]		Stress match[b]	
Condition	Target	Competitor	Target	Competitor
No vowel reduction[c]	CARpet	carTOON	CARpet	CARton
Vowel reduction[d]	PArrot	paRADE	PArrot	PArish

Notes. [a]Suprasegmentally different first syllable; [b]suprasegmentally identical first syllable; [c]segmentally identical first syllable; [d]segmentally different first syllable in stress-mismatch condition.

Materials

A total of 32 disyllabic and trisyllabic critical target and competitor word pairs were selected for this experiment, 16 of which were heard in the no-vowel-reduction condition and 16 of which were heard in the vowel-reduction condition. In the no-vowel-reduction condition, the target word was paired with a competitor word whose first syllable was the same segmentally but differed suprasegmentally from that of the target word (stress-mismatch, experimental condition), or it was paired with a competitor word whose first syllable was the same both segmentally and suprasegmentally (stress-match, control condition). In the vowel reduction condition, the target word was paired with a competitor word whose first syllable differed both segmentally and suprasegmentally from that of the target word (stress-mismatch, experimental condition), or it was paired with a competitor word whose first syllable was the same both segmentally and suprasegmentally (stress-match, control condition). The experimental design is summarized and illustrated with sample test items in Table 2. Each target word was heard only once, with the test items in the stress-mismatch and stress-match conditions being counterbalanced in two lists across participants.

In the critical trials, all target words were stressed on the initial syllable and all competitor words whose first syllable differed suprasegmentally from that of the target were stressed on the second syllable. (The experiment also included filler items where the target word was not stressed on the initial syllable.) In both the no-vowel-reduction and vowel-reduction conditions, the first syllable of the target word always contained a full vowel. The no-vowel-reduction and vowel-reduction conditions differed in whether the first syllable of the stress-mismatch competitor (not heard in the critical stimuli) contained a full or reduced vowel.

The visual display contained the target word, a competitor word, and two distracter words. The words in the display were presented orthographically because not all the words were easily imageable (for a validation of the visual-world paradigm with printed words, see Huettig & McQueen, 2007; McQueen & Viebahn, 2007). All the words had the same number of syllables (two or three). The distracter words did not overlap segmentally with the target or competitor words, but they overlapped segmentally with one another in their initial syllable, with half of them having suprasegmentally identical first syllables and half of them having suprasegmentally different first syllables.

The target, competitor, and distracter words within each condition and between the no-vowel-reduction and vowel-reduction conditions were matched in (log transformed) token frequency (established from the Corpus of Contemporary American English, 2008), number of letters, and number of syllables. Furthermore, the competitor words in the stress-match and stress-mismatch conditions were matched in their percent of orthographic overlap with the target. The complete set of critical target, competitor, and distracter words are provided in Appendix S1 in the Supporting Information online. The mean log-transformed frequency, number of letters, number of syllables, and (for competitor words) the percent of orthographic overlap with the target are provided for each condition in Appendix S2 in the Supporting Information online.

Paired-samples t tests on the log-transformed frequency, number of letters, and number of syllables of the target, competitor, and distracter words within each condition and between the no-vowel-reduction and vowel-reduction conditions did not reveal significant differences: number of letters of target and stress-mismatch competitor (no vowel reduction), $t(15) = -1.33, p > .10$; number of letters of target and first distracter (no vowel reduction), $t(15) = -1.00$, $p > .10$; number of letters of the stress-mismatch and stress-match competitors (no vowel reduction), $t(15) = 1.16$, $p > .10$; number of letters of the stress-mismatch and stress-match competitors (vowel reduction), $t(15) = 1.03$, $p > .10$; all others, $t(15) < |1.00|$, $p > .10$. Similarly, paired-samples t tests on the percent of orthographic overlap of the stress-mismatch and stress-match competitors with the target did not reveal significant differences between the two types of competitors either in the no-vowel-reduction condition, $t(15) = -1.58$, $p > .10$, or in the vowel-reduction condition, $t(15) = -1.50, p > .10$.

The 32 critical trials were interspersed with 32 filler trials. The target, competitor, and distracter words in the filler trials were similar to those in the critical trials in log-transformed frequency, number of letters, and number of syllables. Of the target words in the filler trials, 23 were stressed on the second syllable, and 19 contained a reduced vowel in the first syllable. Of the competitor words in the filler trials, 20 had the same stress as the target word. All the competitor words in the filler trials partially overlapped with the target word in the segmental content of its first syllable.

The auditory stimuli were recorded by a female native speaker of American English without a strong regional accent. The target words from the critical and filler trials were recorded in the carrier phrase "Click on _____." One recording of the carrier phrase was selected for the experiment. A 200-millisecond silent pause was inserted after the carrier phrase, and the target words were extracted from their original carrier phrase and spliced after this 200-millisecond pause.

The target words in the critical trials were analyzed acoustically. Recall that all the target words in the critical trials were stressed on the initial syllable. Independent-samples *t* tests revealed no significant difference in the mean F0, mean duration, or mean intensity of the first syllable between the target words in the no-vowel-reduction condition ($M_{\mathrm{F0}} = 196$ Hz, $SD = 11$; $M_{\mathrm{duration}} = 190$ milliseconds, $SD = 45$; $M_{\mathrm{intensity}} = 73.1$ dB, $SD = 2.8$) and those in the vowel-reduction condition ($M_{\mathrm{F0}} = 192$ Hz, $SD = 16$; $M_{\mathrm{duration}} = 185$ milliseconds, $SD = 46$; $M_{\mathrm{intensity}} = 73.5$ dB, $SD = 2.7$), all $t(30) < |1.00|, p > .10$. Moreover, an independent-samples *t* test revealed no significant difference in the total duration of the target words in the no-vowel-reduction condition ($M = 666$ milliseconds, $SD = 86$) and those in the vowel-reduction condition ($M = 700$ milliseconds, $SD = 106$), $t(30) < |1.00|, p > .10$. This indicated that the target words in the no-vowel-reduction and vowel-reduction conditions were comparable in the suprasegmental information that they provided to the listeners.

To validate that the full and reduced English vowel in the first syllable of the target and competitor words of the vowel-reduction condition could be assimilated to different Korean vowels, we had the same female native speaker of American English record all competitor words (which were not heard in the experiment). Then, we asked three native Korean listeners at a low proficiency in English who did not participate in this study to transcribe all target and competitor words using Korean orthography. These listeners heard the words in isolation without seeing them orthographically and wrote the closest approximation of what they heard using Korean orthography. For each competitor transcription, we coded whether the first vowel was the same as or different from that in the target transcription (same = 0, different = 1). In the no-vowel-reduction condition, the results showed that a similarly low proportion of the stress-mismatch and stress-match competitors were transcribed using a different vowel from that of the target (stress mismatch = 0.29, stress match = 0.25). Conversely, in the vowel-reduction condition, more stress-mismatch than stress-match competitors were transcribed using a different vowel from that of the target (stress mismatch = 0.75, stress match = 0.29). Paired-samples *t* tests conducted on the item means revealed a significant effect of stress in the vowel-reduction condition, $t(15) = 5.75, p < .001$, but not in the no-vowel-reduction condition, $t(15) = 1.46, p > .10$. This confirmed that if Korean listeners do not have a phonetic category for the English schwa, they should nonetheless be able to assimilate the full and reduced English vowels in the first syllable of the target and competitor words of the vowel-reduction condition to different Korean vowels, at least some of the time.

Procedure

The eye-tracking experiment was compiled with the Experiment Builder (2015) software, and the participants' eye movements were recorded with a head-mounted EyeLink II (2006) eye tracker. The signal from the eye tracker was sampled every 4 milliseconds. An Audio Stream Input/Output–compatible sound card was used on the display computer to ensure that the audio timing would be accurate. The experiment began with a calibration of the eye tracker using the participants' right eye, or left eye if the right eye could not be tracked. This initial calibration was followed by a practice session of four trials and by the main experiment with 64 trials. In each trial, participants saw four printed words in a nondisplayed 2 × 2 grid for 4,000 milliseconds. This preview time ensured that all listeners, including the L2 learners, who use a different writing system in their L1, would have sufficient time to activate the phonological representations of the words on the screen. The words then disappeared and a fixation cross appeared in the middle of the screen for 500 milliseconds. This fixation cross served the purpose of bringing listeners' fixations back to the middle of the screen. As the fixation cross disappeared, the four words reappeared on the screen in the same location, the auditory stimulus was heard synchronously over headphones, and participants' eye movements were measured from the target word onset. Participants were instructed to click on the target word with the mouse as soon as they heard the word. The trial ended with the participants' response, with an intertrial interval of 1,000 milliseconds. The test items were presented in four blocks, each consisting of 16 trials, with each block containing two trials from each condition. The order of the experimental and filler trials within a block and the order of blocks were randomized across participants, with each participant encountering no more than two consecutive experimental trials throughout the experiment. The eye tracker was calibrated before each block and whenever it was necessary during the experiment. The participants completed the experiment in approximately 10 minutes.

After the experiment, L2 learners completed the cloze test that served as a measure of their English proficiency (see Table 1). Next, they completed a questionnaire in which they rated their familiarity with the target and competitor words from the critical trials (32 targets, 64 competitor words) on a scale from 0 (“I have never seen/heard this word”) to 4 (“I have frequently seen/heard this word, I know what it means, and I can provide a definition for it”). In the same questionnaire, L2 learners were also asked to circle the stressed syllable in the target and competitor words.

Data Analysis

Experimental trials that received distracter responses, competitor responses, or no response, or for which eye movements could not reliably be tracked, were excluded from the analyses. This resulted in the exclusion of 0.7% of the trials: 0.0% for English listeners, 0.6% for Mandarin listeners, and 0.1% for Korean listeners. Of the remaining trials, we included in the analysis only those for which L2 learners rated their knowledge of the target and competitor word as at least 3 (“I have frequently seen/heard this word and I know what it means in context, but I could not provide a definition for it”) and correctly identified the stressed syllable in target and competitor words. This resulted in the exclusion of an additional 35.2% of the trials from L2 learners’ data: 19.0% for Mandarin listeners and 16.3% for Korean listeners.

The remaining trials for Mandarin listeners were distributed as follows: 93 in the stress-mismatch, no-vowel-reduction condition; 104 in the stress-mismatch, vowel-reduction condition; 108 in the stress-match, no-vowel-reduction condition; and 92 in the stress-match, vowel-reduction condition. The remaining trials for Korean listeners were distributed as follows: 104 in the stress-mismatch, no-vowel-reduction condition; 132 in the stress-mismatch, vowel-reduction condition; 126 in the stress-match, no-vowel-reduction condition; and 106 in the stress-match, vowel-reduction condition. The Mandarin and Korean listeners did not differ in the distribution of the remaining trials, $\chi^2(3) = 0.486$, $p > .10$. Mandarin listeners’ familiarity ratings for the words included in the analyses were as follows: $M = 3.93$, $SD = 0.26$, for target words and $M = 3.86$, $SD = 0.34$, for competitor words in the no-vowel-reduction condition; $M = 3.90$, $SD = 0.30$, for target words and $M = 3.90$, $SD = 0.30$, for competitor words in the vowel-reduction condition. Korean listeners’ familiarity ratings for the words included in the analyses were as follows: $M = 3.89$, $SD = 0.32$, for target words and $M = 3.90$, $SD = 0.31$, for competitor words in the no-vowel-reduction condition; $M = 3.90$, $SD = 0.30$, for target words and $M = 3.94$, $SD = 0.24$, for competitor words in the vowel-reduction condition. The two

L2 groups did not differ from each other on any of these ratings, as shown by independent-samples t tests, all $t(57) < |1.00|$, $p > .10$. Paired-samples t tests also showed no difference between the L2 learners' ratings in the no-vowel-reduction and vowel-reduction conditions, $t(58) < |1.00|$, $p > .10$, or between the target and competitor words, $t(58) < |1.00|$, $p > .10$.

We analyzed participants' eye movements in the four regions of interest corresponding to the four words on the screen. Proportions of fixations to the target, competitor, and distracter words were extracted in 8-millisecond time windows from the onset to roughly the offset of the target word, with a delay of 200 milliseconds as it takes approximately 200 milliseconds for eye movements to reflect speech processing (Hallett, 1986; Salverda, Kleinschmidt, & Tanenhaus, 2014).[5] Statistical analyses were conducted on the difference between the empirical log-transformed proportions of target and competitor fixations, henceforth referred to as the target-over-competitor fixation advantage. This dependent variable was selected because it takes into consideration both target and competitor word activation (for a similar analysis of visual-world eye-tracking data, see Creel, 2014).

Listeners' target-over-competitor fixation advantage was modeled using growth curve analysis (Mirman, 2014; Mirman, Dixon, & Magnuson, 2008). Growth curve analysis enables researchers to model the curvilinear relationship between proportions of fixations and time. Growth curve analysis is less subjective than a time-window analysis in that it models the shape of participants' overall fixation curve rather than their average fixations at arbitrary points in time. However, the interpretation of growth curve analysis results can be difficult in the presence of baseline preferences in the data, that is, where proportions of fixations are higher in one condition than in another due to factors other than the speech signal. In the present study, to conclude that the stress pattern of the competitor word matching or mismatching that of the target had an effect on participants' fixations, in addition to showing a significant effect of stress, the growth curve analysis should reveal *interactions* between stress and at least one of the time polynomials. In the absence of a baseline preference for either stress-mismatch or stress-match items, such interactions would indicate that the different shape of the participants' fixation curve for the stress-mismatch and stress-match conditions can be attributed to their intake of the speech signal.

The growth curve analyses were run on the difference between participants' transformed proportions of target and competitor fixations using the lme4 package for fitting linear mixed-effects models in R (Bates, Maechler, Bolker, & Walker, 2015). For the sake of clarity, we first present the analysis of the results of the individual groups. These analyses included stress (mismatch and match, coded as 0 and 1), vowel reduction (no vowel reduction and vowel reduction, coded as –0.5 and 0.5), and time (linear, quadratic, cubic), and their interaction as fixed effects. The vowel-reduction variable was contrast coded so that the effect of stress could be interpreted on both levels of the vowel-reduction variable. The stress variable was not contrast coded because the effect of vowel reduction can be seen only on the stress-mismatch level of the stress variable, where the competitor word that differs in its stress from the target word contains or does not contain a reduced vowel in its initial syllable. In the stress-match condition, the competitor word has the same stress as the target word, so there should not be an effect of vowel. Hence, the effect of vowel reduction should be interpreted as a simple effect on items in the stress-mismatch condition. A backward-fitting function from the package LMERConvenienceFunctions (Tremblay & Ransijn, 2015) was used to identify the model that accounted for significantly more of the variance than all simpler models, as determined by log-likelihood ratio tests. Only the results of the model with the best fit are presented (for a discussion of this approach, see Mirman, 2014), with *p* values being calculated using the lmerTest package in R (Kuznetsova, Brockhoff, & Christensen, 2016). Analyses yielding significant interactions between stress and vowel reduction were followed up by subsequent growth curve analyses conducted separately on the no-vowel-reduction and vowel-reduction conditions, with the alpha level adjusted to .025. All analyses included participant as random intercept and the time polynomials as random slopes, thus modeling a different curve for each participant. To determine whether the L2 groups differed from native listeners and from each other, larger analyses were conducted that tested three-way interactions for the effects of stress, L1, and time separately for the no-vowel-reduction and vowel-reduction conditions. These analyses are summarized here and explained further in the Supporting Information online (Appendix S4).

Listeners' use of stress in word recognition should result in a higher target-over-competitor fixation advantage in the stress-mismatch than in the stress-match condition, with this effect increasing from the onset to the offset of the target word due to listeners' intake of the speech signal. This increase in the effect of stress over time may affect listeners' target-over-competitor fixation advantage in two ways: Listeners may show a steeper fixation curve (listeners' target-over-competitor fixation advantage may increase more rapidly over time) in the stress-mismatch than in the stress-match condition, and listeners may show a less convex (i.e., less U-shaped) fixation curve (listeners' target-over-competitor fixation advantage may dip less over time) in the stress-mismatch

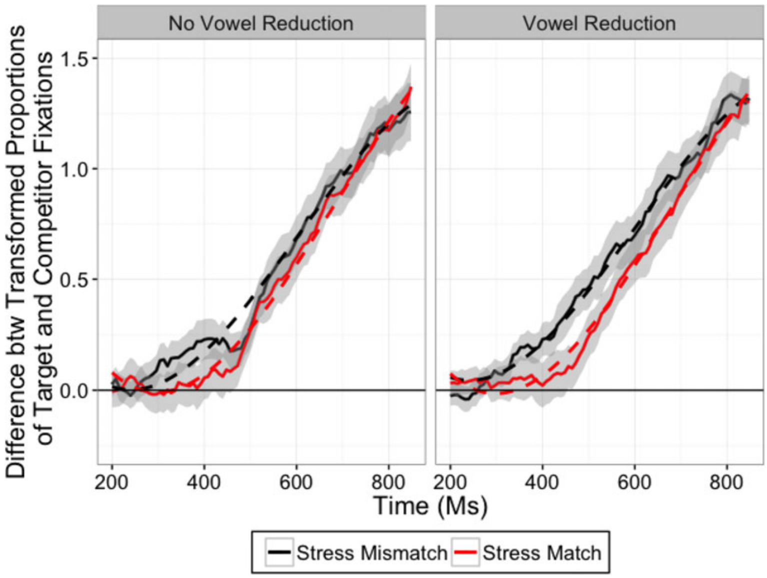

Figure 1 English listeners' target-over-competitor fixation advantage in the stress-mismatch and stress-match conditions for items in the no-vowel-reduction and vowel-reduction conditions. The solid lines represent listeners' data; the dashed lines represent the predicted data based on the growth curve analysis of English listeners' data (Table 3). The shaded area represents one standard error above and below the mean. [Color figure can be viewed at wileyonlinelibrary.com]

than in the stress-match condition. Hence, if stress constrains lexical access, listeners should show a significant interaction between stress and the linear and/or the quadratic time polynomial in addition to showing a significant effect of stress. Furthermore, if stress interacts with vowel reduction, listeners should show significant three-way interactions between stress, vowel reduction, and the linear and/or the quadratic time polynomial in addition to showing a larger effect of stress in the vowel-reduction than in the no-vowel-reduction condition. These three-way interactions should stem from the aforementioned two-way interactions between stress and time being stronger in the vowel-reduction condition than in the no-vowel-reduction condition.

Results

English Listeners

Figure 1 shows English listeners' target-over-competitor fixation advantage in the stress-mismatch and stress-match conditions for items in the

Table 3 Results of growth curve analysis on the difference between English listeners' transformed proportions of target and competitor fixations in all conditions ($\alpha = .05$)

	Estimate	*SE*	*t*	*p*
Intercept	0.556	0.045	12.29	< .001
Time				
Linear	3.880	0.297	13.06	< .001
Quadratic	0.050	0.160	3.12	.004
Cubic	–0.357	0.097	–3.69	< .001
Stress	–0.088	0.008	–10.64	< .001
Vowel reduction	0.043	0.012	3.64	< .001
Time × Stress				
Quadratic	0.535	0.076	7.16	< .001
Stress × Vowel reduction	–0.048	0.017	–2.89	.004

no-vowel-reduction and vowel-reduction conditions (see Appendix S3 in the Supporting Information online for English listeners' separate proportions of target, competitor, and distracter fixations). Recall that all critical target words began with a stressed syllable and thus contained a full vowel; for stress-mismatch items, the first syllable of the target and competitor words in the no-vowel-reduction condition differed only suprasegmentally whereas the first syllable of the target and competitor words in the vowel-reduction condition differed both segmentally and suprasegmentally.

Table 3 presents the results of the growth curve analysis with the best fit on the difference between English listeners' transformed proportions of target and competitor fixations in all conditions. As summarized in this table, the negative estimate for the effect of stress means that English listeners had a lower target-over-competitor fixation advantage in the stress-match than in the stress-mismatch condition. The positive estimate for the effect of vowel reduction indicates that English listeners' target-over-competitor fixation advantage for the stress-mismatch items was higher in the vowel-reduction than in the no-vowel-reduction condition. The positive estimate for the interaction between the quadratic time polynomial and stress indicates that the English listeners' fixation curve had more of a convex (i.e., U) shape in the stress-match than in the stress-mismatch condition. Importantly, the negative estimate for the interaction between stress and vowel reduction indicates that English listeners' (negative) effect of stress was larger (more negative) in the vowel-reduction condition than in the no-vowel-reduction condition.

Table 4 Results of growth curve analyses on the difference between English listeners' transformed proportions of target and competitor fixations separately for the no-vowel-reduction and vowel-reduction conditions ($\alpha = .025$)

	Estimate	*SE*	*t*	*p*
No vowel reduction				
Intercept	0.534	0.054	9.92	< .001
Time				
Linear	3.856	0.343	11.24	< .001
Quadratic	0.612	0.214	2.89	.007
Cubic	–0.422	0.140	–3.02	.005
Stress	–0.064	0.010	–6.19	< .001
Stress × Time				
Quadratic	0.344	0.094	3.68	< .001
Vowel reduction				
Intercept	0.545	0.044	12.38	< .001
Time				
Linear	3.520	0.254	13.86	< .001
Quadratic	0.331	0.196	1.69	.100
Stress	–0.112	0.011	–10.65	< .001
Stress × Time				
Quadratic	0.727	0.095	7.65	< .001

The two-way interaction between stress and vowel reduction warranted additional growth curve analyses testing for the effects of time and stress on the difference between English listeners' transformed proportions of target and competitor fixations separately for the no-vowel-reduction and vowel-reduction conditions. Table 4 presents the results of the follow-up growth curve analyses with the best fit. Importantly, these analyses revealed significant effects of stress and significant interactions between the quadratic time polynomial and stress for both the no-vowel-reduction and vowel-reduction conditions, with English listeners showing a lower target-over-competitor fixation advantage and a more convex fixation curve in the stress-match than in the stress-mismatch condition. The two-way interaction between stress and vowel reduction can therefore be attributed to the stronger effect of stress in the vowel-reduction condition, $t = 7.65$, than in the no-vowel-reduction condition, $t = 3.68$.

These results indicate that stress constrained English listeners' lexical access more when the stress difference between the first syllable of target and competitor words was realized with both segmental and suprasegmental cues than when it was realized only with suprasegmental cues. Crucially, even in the

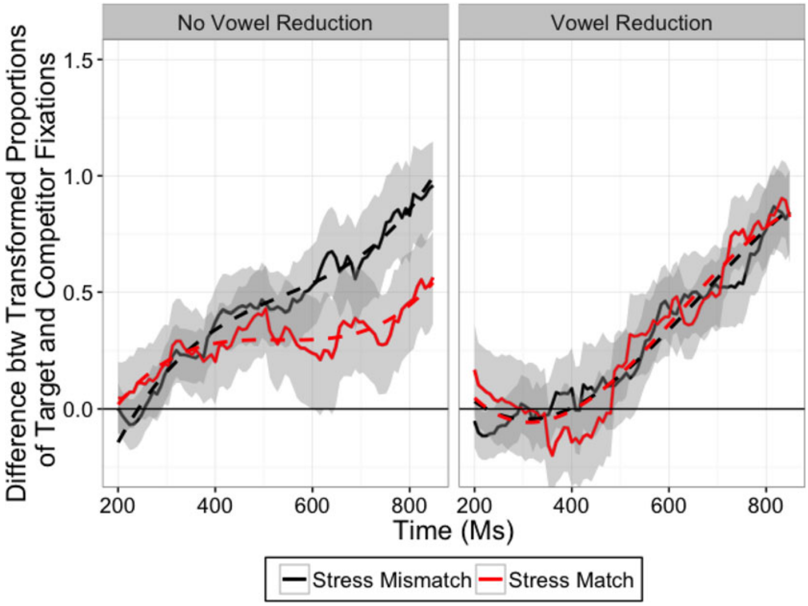

Figure 2 Mandarin listeners' target-over-competitor fixation advantage in the stress-mismatch and stress-match conditions for items in the no-vowel-reduction and vowel-reduction conditions. The solid lines represent listeners' data; the dashed lines represent the predicted data based on the growth curve analysis of Mandarin listeners' data (Table 5). The shaded area represents one standard error above and below the mean. [Color figure can be viewed at wileyonlinelibrary.com]

absence of segmental cues in the competitor word, English listeners could use stress in lexical access.

Mandarin Listeners

Figure 2 shows Mandarin listeners' target-over-competitor fixation advantage in the stress-mismatch and stress-match conditions for items in the no-vowel-reduction and vowel-reduction conditions (see Appendix S3 for Mandarin listeners' separate proportions of target, competitor, and distracter fixations). Again, for stress-mismatch items, only in the vowel reduction did the target and competitor words differ both segmentally and suprasegmentally.

Table 5 presents the results of the growth curve analysis with the best fit on the difference between Mandarin listeners' transformed proportions of target and competitor fixations in all conditions. As shown in this table, the negative estimate for the effect of stress means that Mandarin listeners had a lower

Table 5 Results of growth curve analysis on the difference between Mandarin listeners' transformed proportions of target and competitor fixations in all conditions ($\alpha = .05$)

	Estimate	*SE*	*t*	*p*
Intercept	0.373	0.084	4.42	< .001
Time				
Linear	2.692	0.420	6.41	< .001
Quadratic	0.338	0.234	1.45	> .100
Cubic	0.023	0.188	< |1.00|	> .100
Stress	–0.112	0.016	–6.85	< .001
Vowel reduction	–0.162	0.023	–6.98	< .001
Time × Stress				
Linear	–1.025	0.148	–6.92	< .001
Time × Vowel reduction				
Linear	0.069	0.209	< |1.00|	> .100
Quadratic	0.797	0.147	5.44	< .001
Cubic	–0.658	0.146	–4.49	< .001
Stress × Vowel reduction	0.177	0.032	5.38	< .001
Time × Stress × Vowel reduction				
Linear	1.562	0.296	5.27	< .001

target-over-competitor fixation advantage in the stress-match than in the stress-mismatch condition. The negative estimate for the effect of vowel reduction indicates that Mandarin listeners' target-over-competitor fixation advantage for stress-mismatch items was lower in the vowel reduction than in the no-vowel-reduction condition. The negative estimate for the interaction between the linear time polynomial and stress indicates that the Mandarin listeners' fixation curve was less ascending in the stress-match than in the stress-mismatch condition. The positive and negative estimates for the interaction between vowel reduction and, respectively, the quadratic and cubic time polynomials mean that the Mandarin listeners' stress-mismatch fixation curve had a more convex (i.e., U) and more S shape in the vowel-reduction than in the no-vowel-reduction condition. Crucially, the positive estimate for the interaction between stress and vowel reduction indicates that Mandarin listeners' (negative) effect of stress in the stress-match condition decreased from the no-vowel-reduction to the vowel-reduction condition. Last but not least, the positive estimate for the three-way interaction between the linear time polynomial, stress, and vowel reduction suggests that the no-vowel-reduction and vowel-reduction conditions differed in the strength of the interaction between stress and the linear time polynomial.

Table 6 Results of growth curve analyses for the difference between Mandarin listeners' transformed proportions of target and competitor fixations separately for the no-vowel-reduction and vowel-reduction conditions ($\alpha = .025$)

	Estimate	*SE*	*t*	*p*
No vowel reduction				
Intercept	0.418	0.118	3.56	.001
Time				
Linear	3.122	0.431	7.24	< .001
Stress	–0.186	0.019	–10.02	< .001
Stress × Time				
Linear	–1.902	0.168	–11.34	< .001
Vowel reduction				
Intercept	0.329	0.063	5.23	< .001
Time				
Linear	2.076	0.497	4.18	< .001

Table 6 presents the results of the follow-up growth curve analyses with the best fit separately for the no-vowel-reduction and vowel-reduction conditions. As can be seen in this table, the growth curve analyses yielded a significant effect of stress only in the no-vowel-reduction condition, with the target-over-competitor fixation advantage being lower in the stress-match than in the stress-mismatch condition. The two-way interaction between stress and vowel reduction reported in Table 5 can thus be attributed to the effect of stress being found only in the no-vowel-reduction condition. The growth curve analyses also revealed a significant interaction between the linear time polynomial and stress only in the no-vowel-reduction condition, with Mandarin listeners' fixation curve being less ascending for stress-match than for stress-mismatch items. The occurrence of this two-way interaction in the no-vowel-reduction condition but not in the vowel-reduction condition appears to be driving the three-way interaction between the linear time polynomial, stress, and vowel reduction reported in Table 5.

These results suggest that Mandarin listeners used stress in lexical access only when the stress difference between the first syllable of the target and competitor words was realized with suprasegmental cues alone.

Korean Listeners

Figure 3 shows Korean listeners' target-over-competitor fixation advantage in the stress-mismatch and stress-match conditions for items in the

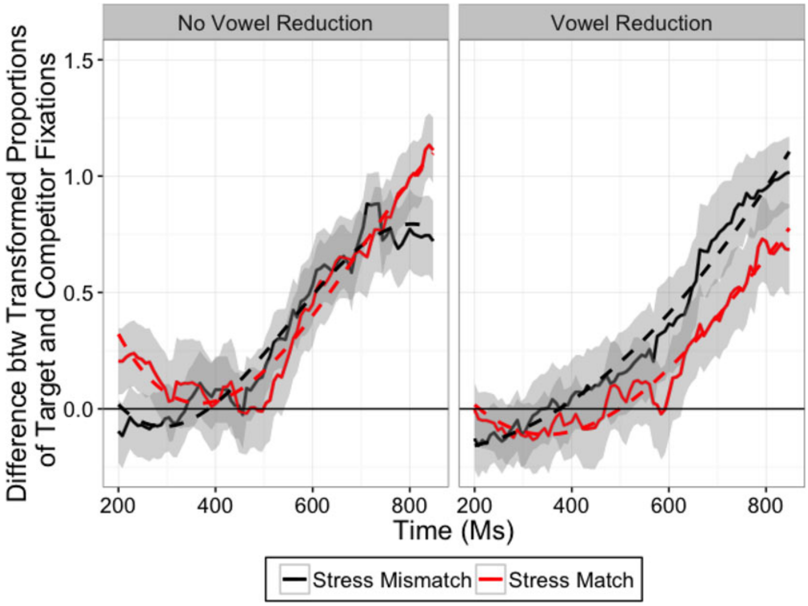

Figure 3 Korean listeners' target-over-competitor fixation advantage in the stress-mismatch and stress-match conditions for items in the no-vowel-reduction and vowel-reduction conditions. The solid lines represent listeners' data. The dashed lines represent the predicted data based on the growth curve analysis of Korean listeners' data (Table 7). The shaded area represents one standard error above and below the mean. [Color figure can be viewed at wileyonlinelibrary.com]

no-vowel-reduction and vowel-reduction conditions (see Appendix S3 for Korean listeners' separate proportions of target, competitor, and distracter fixations).

Table 7 presents the results of the growth curve analysis with the best fit on the difference between Korean listeners' transformed proportions of target and competitor fixations in all conditions. The negative estimate for the effect of stress indicates that Korean listeners' target-over-competitor fixation advantage was lower for stress-match than for stress-mismatch items. The negative and positive estimates for the interaction between stress and, respectively, the linear and quadratic time polynomials indicate that the Korean listeners' fixation curve was less ascending and more convex (i.e., U) in the stress-match than in the stress-mismatch condition. The positive estimates for the interaction between vowel reduction and both the linear and cubic time polynomials mean that the Korean listeners' stress-mismatch fixation curve was more ascending

Table 7 Results of growth curve analysis on the difference between Korean listeners' transformed proportions of target and competitor fixations in all conditions ($\alpha = .05$)

	Estimate	*SE*	*t*	*p*
Intercept	0.309	0.059	5.28	< .001
Time				
Linear	3.028	0.351	8.62	< .001
Quadratic	0.625	0.342	1.82	.078
Cubic	–0.232	0.195	–1.19	> .100
Stress	–0.045	0.016	–2.76	.006
Vowel reduction	0.031	0.023	1.38	> .100
Time × Stress				
Linear	–0.608	0.146	–4.16	< .001
Quadratic	0.752	0.146	5.16	< .001
Time × Vowel reduction				
Linear	0.679	0.204	3.32	< .001
Cubic	0.430	0.146	2.95	.003
Stress × Vowel reduction	–0.267	0.033	–8.13	< .001
Time × Stress × Vowel reduction				
Linear	–1.241	0.297	–4.18	< .001

and less S shaped in the vowel reduction than in the no-vowel-reduction condition. Importantly, the negative estimate for the interaction between stress and vowel reduction means that Korean listeners' effect of stress increased from the no-vowel-reduction to the vowel-reduction condition. Additionally, the three-way interaction between the linear time polynomial, stress, and vowel reduction suggests that the no-vowel-reduction and vowel-reduction conditions differed in the strength of the interaction between stress and the linear time polynomial.

Follow-up growth curve analyses assessed the effects of time and stress separately in the no-vowel-reduction and vowel-reduction conditions. The growth curve analyses with the best fit are presented in Table 8. As shown in this table, there was a significant effect of stress only in the vowel-reduction condition, with the target-over-competitor fixation advantage being lower in the stress-match than in the stress-mismatch condition. The growth curve analyses also revealed significant interactions between the quadratic time polynomial and stress in both the no-vowel-reduction and vowel-reduction conditions, but this effect appears to be due to an early baseline preference for stress-match items in the no-vowel-reduction condition, with the fixation curve dipping more in the stress-match than in the stress-mismatch condition due to this baseline

Table 8 Results of growth curve analyses on the difference between Korean listeners' transformed proportions of target and competitor fixations separately for the no-vowel-reduction and vowel-reduction conditions ($\alpha = .025$)

	Estimate	*SE*	*t*	*p*
No vowel reduction				
Intercept	0.327	0.072	4.49	< .001
Time				
Linear	2.122	0.350	6.06	< .001
Quadratic	0.550	0.310	1.77	.085
Stress	0.043	0.020	2.19	.028
Stress × Time				
Quadratic	1.030	0.176	5.84	< .001
Vowel reduction				
Intercept	0.324	0.101	3.22	.003
Time				
Linear	3.298	0.523	6.31	< .001
Quadratic	0.758	0.491	1.54	> .100
Stress	–0.267	0.019	–14.06	< .001
Stress × Time				
Linear	–1.037	0.172	–6.04	< .001
Quadratic	0.769	0.171	4.49	< .001

preference. Importantly, only in the vowel-reduction condition was Korean listeners' fixation curve less ascending for stress-match items than for stress-mismatch items.

These results suggest that Korean listeners used stress in lexical access only when the stress difference between the first syllable of target and competitor words was realized with both segmental and suprasegmental cues.

Between-Group Comparisons

To determine whether the L2 groups differed from native listeners and from each other, we also conducted larger analyses that tested three-way interactions for the effects of stress, L1, and time separately for the no-vowel-reduction and vowel-reduction conditions. These analyses are presented in detail in Appendix S4 in the Supporting Information online.

For the no-vowel-reduction condition, the results of these between-group analyses revealed that the L2 groups differed from native listeners and from each other. First, compared to English listeners, Mandarin listeners showed a larger effect of stress whereas Korean listeners showed a smaller effect of stress, and Korean listeners also showed a weaker effect of stress compared to Mandarin listeners. Second, Mandarin listeners differed from both English and Korean listeners by showing a less ascending fixation curve in the stress-match than in the stress-mismatch condition (English and Korean listeners did not show such a difference). Third, Mandarin listeners differed from both English and Korean listeners in not showing a curvature difference in their fixations to stress-match versus stress-mismatch items (English and Korean listeners showed such a difference), and Korean listeners also differed from English listeners in showing a greater curvature difference, with their fixation curve being more convex in the stress-match than in the stress-mismatch condition compared to that of English listeners.

For the vowel-reduction condition, the results of the between-group analyses also revealed that the L2 groups differed from native listeners and from each other. First, compared to English listeners, Mandarin listeners showed a smaller effect of stress whereas Korean listeners showed a larger effect of stress, and compared to Mandarin listeners, Korean listeners also showed a larger effect of stress. Second, Korean listeners differed from English listeners in showing a less ascending fixation curve in the stress-match than in the stress-mismatch condition (English listeners did not show such an effect).

These results demonstrate that the three groups differed in their use of segmental and suprasegmental cues to stress in the recognition of English words (see Appendix S4 for more details).

Discussion

This study investigated whether Mandarin- and Korean-speaking L2 learners of English would differ from each other and from native English listeners in the use of segmental and suprasegmental cues to stress in lexical access. The L2 learners were matched in their proficiency in and experience with English, and they were familiar with the meaning and stress of the experimental words. Participants completed a visual-world eye-tracking experiment with printed words. In the experimental stress-mismatch conditions, the first syllable of the target and competitor words differed in either suprasegmental cues to English stress (no-vowel-reduction condition) or in both segmental and suprasegmental cues to English stress (vowel-reduction condition). In the control conditions, the first syllable of the target and competitor words was identical in both segmental and suprasegmental cues to English stress.

The results of the eye-tracking experiment showed that English listeners used both segmental and suprasegmental cues to stress in word recognition, with the presence of both cues in the competing word resulting in a larger effect of stress than the presence of only suprasegmental cues, but with suprasegmental cues nonetheless constraining lexical access. These results corroborate the findings of Cooper et al. (2002), who demonstrated that English listeners' online word recognition can benefit from suprasegmental cues to stress alone. The present study is the first to replicate these findings with an online lexical task using real English words. The English listeners' results are also in line with those of studies showing that segmental cues to stress play an important role in the recognition of English words, with stress constraining lexical competition more in the presence than in the absence of vowel-reduction cues (Fear et al., 1995; Lin et al., 2014).

The results of the eye-tracking experiment further indicated that, like English listeners, Mandarin listeners were able to use stress in word recognition when the stress difference between the first syllable of the target and competitor words was realized by suprasegmental cues alone. These results were expected. Standard Mandarin has been claimed to have stress distinctions in a limited number of words (e.g., Chao, 1968; Duanmu, 2007). These stress distinctions in the L1 may have enabled Mandarin-speaking L2 learners of English to use stress when accessing English words (e.g., Lin et al., 2014). It is also possible that the existence of lexical tones in Mandarin Chinese further enhanced Mandarin listeners' use of suprasegmental cues to English stress in word recognition (for discussion, see Lin et al., 2014; Qin et al., 2016).

More surprising was the Mandarin listeners' inability to use stress in word recognition when the first syllable of the target and competitor words differed in both segmental and suprasegmental cues (unlike English listeners). Vowel reduction is also a cue to stress in Standard Mandarin: Words whose first syllable is stressed and whose second syllable is unstressed have a phonologically reduced vowel in the unstressed syllable (e.g., Chao, 1968; Duanmu, 2007). Mandarin listeners have also been shown to tune in to vowel quality in their (explicit) perception of English stress (Chrabaszcz et al., 2014; Zhang & Francis, 2010). We had therefore predicted that Mandarin listeners would be able to use stress in the vowel-reduction condition. Even if Mandarin listeners ignored segmental cues to stress, we would have expected them to use the cooccurring suprasegmental cues to stress that distinguished the first syllable of the target and competitor words in the vowel-reduction condition. An important characteristic of Standard Mandarin is that reduced vowels are not permitted in word-initial position (e.g., Chao, 1998; Duanmu, 2007). From this, one might hypothesize that Mandarin-speaking L2 learners of English would fail to represent the vowel in the first syllable of the competitor words as reduced.

However, this would still not explain why Mandarin listeners were unable to use suprasegmental cues to stress in the vowel-reduction condition.

Because the data analysis included only words that Mandarin listeners were familiar with and knew how to stress, word familiarity and knowledge of stress do not explain the present results. Because the first syllable of the target word in our critical stimuli always contained a full vowel, and given that the first syllable of the target words in the no-vowel-reduction and vowel-reduction conditions did not differ in F0, duration, or intensity, Mandarin listeners' nonuse of stress in the vowel-reduction condition can also not be attributed to acoustic characteristics of the first syllable of the target words that they heard. The target words in the vowel-reduction condition ended 34 milliseconds later than those in the no-vowel-reduction condition, but this numerical difference was not significant (see Note 5); even if it had been, it is unclear why this difference would have resulted in the nonuse of stress in the vowel reduction condition.

One methodological consideration may potentially explain Mandarin listeners' results in the vowel-reduction condition. In the eye-tracking experiment, words were presented orthographically in the display because not all the words in the experiment were easily imageable. Mandarin listeners may not have been able to use stress in the vowel-reduction condition if they indeed represented the vowel in the first syllable of the competitor word as phonologically reduced and if the orthographic representation of the word on the screen conflicted with their phonological representation of the word. Of the test items in the vowel reduction condition, 12 contained the orthographic vowel *a* or *o*. These orthographic vowels have very different phonological realizations depending on whether or not they are stressed. It is possible that the orthographic presentation of the words containing these vowels conflicted with listeners' phonological representations, resulting in their inability to use stress in the vowel-reduction condition. This difficulty did not arise in the no-vowel-reduction condition because that condition contained fewer such vowels—six in total—and the quality of the vowels did not change much as a function of stress, resulting in no conflict between orthographic and phonological representations. Mandarin listeners may have been more vulnerable than Korean listeners to such orthographic effects because their L1 writing system is morphosyllabic and does not allow letter-phoneme mappings in the way that the Korean and English writing systems do. Although the Korean writing system is more transparent than that of English, it is not completely transparent due to a variety of phonological processes that take place in Korean. For example, the syllable-final consonants /t^h/, /t/, and /s/ are written as such but realized as [t]. Similarly, the syllable-final consonant /h/ in a word such as /nahta/ ("give birth/bring about") is written as such but phonetically realized as [nat^{h}a]. Hence, Korean listeners may have less difficulty in learning the many-to-one and one-to-many mappings between orthography and sounds in English compared to Mandarin listeners, who do not have a phonetic writing system. Further evidence that, compared to Korean-speaking L2 learners of English, Mandarin-speaking L2 learners rely more on orthographic information and less on prelexical phonological information in visual word recognition can be found in Wang, Koda, and Perfetti (2003).

Importantly, the results of the eye-tracking experiment showed that, like English listeners but unlike Mandarin listeners, Korean-speaking L2 learners of English could use stress in online word recognition when the stress difference between the first syllable of target and competitor words was signaled by both segmental and suprasegmental cues. We expected Korean listeners to be able to encode English stress in the vowel-reduction condition if they had assimilated English full and reduced vowels to different Korean vowels or if they had developed a phonetic category for the English schwa. The present results indicate that Korean listeners can use this segmental information to encode stress and in turn use stress in lexical access. These results differ from those reported by Lin et al. (2014). This discrepancy is likely due to the different methodologies employed in the two studies. On the one hand, eye tracking may provide a more sensitive measure of lexical activation than lexical decisions because eye movements automatically reflect listeners' intake of the speech signal over time without requiring an explicit decision from listeners. On the other hand, the data included in the analysis were from words that the L2 learners were familiar with and knew how to stress, a selection criterion of significant importance when establishing whether L2 learners can use stress in lexical access (see also Tremblay, 2008).

By contrast, when the first syllable of the target and competitor words differed only in suprasegmental cues, Korean listeners could not use stress in lexical access, unlike English and Mandarin listeners. These results are in line with those of previous L2 studies showing that Korean-speaking L2 learners of English have difficulty using suprasegmental cues to stress in the L2 because Korean does not have lexical stress (Lin et al., 2014). This suggests that segmental cues to stress may in fact provide the only means by which listeners from languages without lexical stress can use stress in L2 lexical access. Further research should establish whether this prediction is indeed correct.

Theoretical Implications

An important implication of the current findings is that models of L2 word recognition should consider not only whether the L1 has lexical stress but also *how*, that is, with what types of cues, stress is realized in the L2 (cf. Dupoux et al., 2008). If specific cues to stress in the L2 can be mapped onto different lexical representations in the L1, then it is likely that listeners will be able to use these cues to recognize words in the L2. To illustrate, Qin et al. (2016) demonstrated that when F0 and duration cues to stress conflicted in a sequence recall task with English nonwords, Mandarin-speaking L2 learners of English more often relied on F0 than on duration whereas native English listeners relied on each cue about half of the time. The Mandarin listeners' results were attributed to the importance of F0 cues for identifying lexical tones in the L1. In other words, L2 learners can transfer a phonetic cue from one domain (lexical tones) to another domain (lexical stress) if that cue allows them to encode words differently.

The present results provide additional evidence for this claim: Korean does not have lexical stress or a phonological process of vowel reduction, yet Korean listeners can lexicalize vowel-quality information in their representations (e.g., via phonetic assimilation) in a way that enables them to uniquely encode words that differ in stress placement. Assuming that the encoding of stress requires at least some reliance on suprasegmental cues, these results do not imply that Korean listeners encode lexical stress per se, but they do suggest that Korean listeners can make use of segmental cues to distinguish words that differ in their stress placement. These results highlight the importance of a cue-based approach to phonological encoding for understanding how nonnative listeners recognize and encode words in the L2. Further research should seek to provide a precise quantification of how L2 learners of English weigh different cues to English stress using implicit tasks that shed direct light on lexical activation in online word recognition.

The results of the present study, particularly those of the native English and Mandarin listeners, also have important implications for computational models of spoken word recognition (e.g., McClelland & Elman, 1986; Norris, 1994; Norris & McQueen, 2008) because they suggest that suprasegmental information affects the process of lexical activation and competition (see also Reinisch et al., 2010). To simulate the effect of suprasegmental information on word recognition, computational models of spoken word recognition would have to simultaneously take as input both segmental and suprasegmental cues to lexical identity and modulate lexical activation accordingly. Although existing models of word recognition can readily account for the use of segmental cues to stress in word recognition (see Norris, 1994, for an explicit attempt at modulating the use of metrical cues in speech segmentation), their architectures make it difficult to model the use of suprasegmental information because these models do not simultaneously take as input a variety of fine-grained acoustic cues (but see Shuai & Malins, 2017, for an attempt with lexical tones in Mandarin Chinese). Further research is necessary to develop computational models whose architecture would make it possible to model the use of both segmental and suprasegmental information in word recognition.

Finally, it is necessary to acknowledge the potential limitations of growth curve analysis for the analysis of visual-world eye-tracking data. Although growth curve analysis is a powerful tool for analyzing visual-world eye-tracking data, the interpretation of growth curve analysis results can be difficult in the presence of baseline preferences in the data because the same interactions between condition and time can be due to baseline preferences rather than to the speech signal (see Mirman et al., 2008, for discussion). Our Korean listeners' data in the no-vowel-reduction condition presented such a case, with their fixation line dipping more in the stress-match than in the stress-mismatch condition because of an early baseline preference for stress-match items. Whereas Korean listeners' stress-by-time interaction in the no-vowel-reduction condition was relatively straightforward to interpret, baseline preferences and effects of the speech signal can, in some cases, be difficult to tease apart (see Mirman et al., 2008, for an example). This suggests that growth curve analysis is not a one-size-fits-all method that can be used in all visual-world eye tracking studies, and its results should be interpreted conservatively when it is used.

Conclusion

This study investigated the use of segmental and suprasegmental cues to lexical stress in word recognition by native English listeners and by Mandarin-speaking and Korean-speaking L2 learners of English who were matched in their English proficiency and experience. The results of a visual-world eye-tracking experiment showed that native English listeners used both suprasegmental cues alone and segmental and suprasegmental cues together to recognize English words, with the effect of stress being greater when the two types of cues were combined. In contrast, Mandarin listeners used stress in lexical access only when stress was signaled by suprasegmental cues alone whereas Korean listeners used stress in lexical access only when stress was signaled by segmental and suprasegmental cues together. Mandarin listeners' ability to use suprasegmental cues to stress was attributed to the importance of these cues in their L1, a language with lexical tones. Korean listeners' ability to use lexical stress in the presence of segmental cues was attributed to their assimilation of reduced and nonreduced English vowels to different phonetic categories, allowing them to use stress in word recognition even if their L1 does not encode lexical stress.

These findings highlight the importance of a cue-based approach to the study of the effect of lexical stress on word recognition.

Final revised version accepted 21 December 2017

Notes

1 Following Bolinger (1986), we assume that the full vowels of American English are [i] (e.g., *keyed*), [ɪ] (e.g., *kid*), [e] (e.g., *cade*), [ɛ] (e.g., *Ked*), [æ] (e.g., *cad*), [ʌ] (e.g., *cud*), [ɑ] (e.g., *cod*), [ɔ] (e.g., *cawed*), [o] (e.g., *code*), [ʊ] (e.g., *could*), and [u] (e.g., *cooed*), and the reduced vowels of American English are [ɨ] (e.g., *Willie*), [ə] (e.g., *Willa*), and [ɵ] (e.g., *willow*) (p. 38). Full vowels are longer and more peripheral in their articulation and acoustic space, and reduced vowels are shorter and more centralized in their articulation and acoustic space.

2 We did not include a condition where stress would be signaled by segmental cues but not by suprasegmental cues because this does not occur in naturally produced speech.

3 We analyzed the results with and without the Mandarin listeners who identified another Chinese dialect as their native language. The pattern of results remained the same in both analyses. We therefore did not exclude Mandarin listeners from the data analysis for dialectal reasons.

4 We also analyzed the results with and without the Korean listeners who did not speak Seoul Korean as their native dialect. The pattern of results was also the same in both analyses. We therefore did not exclude Korean listeners from the data analysis for dialectal reasons.

5 The average target word duration was 666 milliseconds in the no-vowel-reduction condition and 700 milliseconds in the vowel-reduction condition. The data analysis required the time bins to be identical between the no-vowel-reduction and vowel-reduction conditions. Thus, eye fixations were extracted from the target word onset to 650 milliseconds (both with a delay of 200 milliseconds).

References

Bates, D., Maechler, B., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*, 1–48. https://doi.org/10.18637/jss.v067.i01

Beckman, M. (1986). *Stress and non-stress accents*. Dordrecht, Netherlands: Foris. https://doi.org/10.1515/9783110874020

Bolinger, D. (1986). *Intonation and its parts: Melody in spoken English*. Stanford, CA: Stanford University Press.

Brown, J. D. (1980). Relative merits of four methods for scoring cloze tests. *The Modern Language Journal*, *64*, 311–317. https://doi.org/10.1111/j.1540-4781.1980.tb05198.x

Brown, M., Salverda, A. P., Dilley, L. C., & Tanenhaus, M. K. (2015). Metrical expectations from preceding prosody influence perception of lexical stress. *Journal of Experimental Psychology: Human Perception and Performance*, *41*, 306–323. https://doi.org/10.1037/a0038689

Chao, Y.-R. (1968). *A grammar of spoken Chinese*. Oakland: University of California Press. https://doi.org/10.1017/S0041977X00129933

Cho, J., & Park, H.-K. (2006). A comparative analysis of Korean-English phonological structures and processes for pronunciation pedagogy in interpretation training. *Meta: Translators' Journal*, *51*, 229–246. https://doi.org/10.7202/013253ar

Cho, T., McQueen, J. M., & Cox, E. A. (2007). Prosodically driven phonetic detail in speech processing: The case of domain-initial strengthening in English. *Journal of Phonetics*, *35*, 210–243. https://doi.org/10.1016/j.wocn.2006.03.003

Chrabaszcz, A., Winn, M., Lin, C. Y., & Idsardi, W. J. (2014). Acoustic cues to perception of word stress by English, Mandarin, and Russian speakers. *Journal of Speech, Language, and Hearing Research*, *57*, 1468–1479. https://doi.org/10.1044/2014_JSLHR-L-13-0279

Cooper, N., Cutler, A., & Wales, R. (2002). Constraints of lexical stress on lexical access in English: Evidence from native and non-native listeners. *Language and Speech*, *45*, 207–228. https://doi.org/10.1177/00238309020450030101

Corpus of contemporary American English. (2008). Retrieved March 5, 2015, from http://corpus.byu.edu/coca/.

Creel, S. C. (2014). Tipping the scales: Auditory cue weighting changes over development. *Journal of Experimental Psychology: Human Perception and Performance*, *40*, 1146–1160. https://doi.org/10.1037/a0036057

Cutler, A. (1986). Forbear is a homophone: Lexical prosody does not constrain lexical access. *Language and Speech*, *29*, 201–220. https://doi.org/10.1177/002383098602900302

Duanmu, S. (2007). *The phonology of standard Chinese*. Oxford, UK: Oxford University Press.

Dupoux, E., Peperkamp, S., & Sebastián-Gallés, N. (2001). A robust method to study stress "deafness." *The Journal of the Acoustical Society of America*, *110*, 1606–1618. https://doi.org/10.1121/1.1380437

Dupoux, E., Sebastián-Gallés, N., Navarrete, E., & Peperkamp, S. (2008). Persistent stress "deafness": The case of French learners of Spanish. *Cognition*, *106*, 682–706. https://doi.org/10.1016/j.cognition.2007.04.001

Experiment Builder [Computer software]. (2015). Mississauga, Canada: SR Research.

Eyelink II [Apparatus and software]. (2006). Mississauga, Canada: SR Research.

Fear, B. D., Cutler, A., & Butterfield, S. (1995). The strong/weak syllable distinction in English. *Journal of the Acoustical Society of America*, *97*, 1893–1904. https://doi.org/10.1121/1.412063

Gay, T. (1978). Physiological and acoustic correlates of perceived stress. *Language and Speech*, *21*, 347–353. https://doi.org/10.1177/002383097802100409

Hallett, P. E. (1986). Eye movements. In K. R. Boff, L. Kaufman, & J. P. Thomas (Eds.), *Handbook of perception and human performance* (pp. 10.11–10.112). New York: Wiley.

Huettig, F., & McQueen, J. M. (2007). The tug of war between phonological, semantic and shape information in language-mediated visual search. *Journal of Memory and Language*, *57*, 460–482. https://doi.org/10.1016/j.jml.2007.02.001

Jun, S.-A. (1998). The accentual phrase in the Korean prosodic hierarchy. *Phonology*, *15*, 189–226. https://doi.org/10.1017/s0952675798003571

Jun, S.-A. (2000). K-ToBI (Korean ToBI) labeling conventions. *UCLA Working Papers in Phonetics*, *99*, 149–173.

Kuznetsova, A., Brockhoff, B., & Christensen, H. (2016). Tests in linear mixed effects models. Version 2.0.32. Retrieved April 15, 2017, from https://cran.r-project.org/web/packages/lmerTest/index.html

Lieberman, P. (1960). Acoustic correlates of word stress in American English. *Journal of the Acoustical Society of America*, *32*, 451–454. https://doi.org/10.1121/1.1908095

Lin, C. Y., Wang, M. I. N., Idsardi, W. J., & Xu, Y. I. (2014). Stress processing in Mandarin and Korean second language learners of English. *Bilingualism: Language and Cognition*, *17*, 316–346. https://doi.org/10.1017/s1366728913000333

Lindblom, B. (1963). Spectrographic study of vowel reduction. *Journal of the Acoustical Society of America*, *35*, 1173–1181. https://doi.org/10.1121/1.2142410

McClelland, J. L., & Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology*, *18*, 1–86. https://doi.org/10.1016/0010-0285(86)90015-0

McQueen, J. M., & Viebahn, M. C. (2007). Tracking recognition of spoken words by tracking looks to printed words. *Quarterly Journal of Experimental Psychology*, *60*, 661–671. https://doi.org/10.1080/17470210601183890

Mirman, D. (2014). *Growth curve analysis and visualization using R*. Boca Raton, FL: Taylor & Francis.

Mirman, D., Dixon, J. A., & Magnuson, J. S. (2008). Statistical and computational models of the visual world paradigm: Growth curves and individual differences. *Journal of Memory and Language*, *59*, 475–494. https://doi.org/10.1016/j.jml.2007.11.006

Norris, D. (1994). Shortlist: A connectionist model of continuous speech recognition. *Cognition*, *52*, 189–234. https://doi.org/10.1016/0010-0277(94)90043-4

Norris, D., & McQueen, J. M. (2008). Shortlist B: A Bayesian model of continuous speech recognition. *Psychological Review*, *115*, 357–395. https://doi.org/10.1037/0033-295X.115.2.357

Peperkamp, S., & Dupoux, E. (2002). A typological study of stress deafness. In C. Gussenhoven (Ed.), *Laboratory phonology 7* (pp. 203–240). Berlin, Germany: de Gruyter. https://doi.org/10.1515/9783110197105.203

Protopapas, A., Panagaki, E., Andrikopoulou, A., Gutierrez Palma, N., & Arvaniti, A. (2016). Priming stress patterns in word recognition. *Journal of Experimental*

Psycholology: Human Perception and Performance, *42*, 1739–1760. https://doi.org/10.1037/xhp0000259

Qin, Z., Chien, Y.-F., & Tremblay, A. [Annie]. (2016). Processing of word-level stress by Mandarin-speaking second language learners of English. *Applied Psycholinguistics*, *38*, 541–570. https://doi.org/10.1017/s0142716416000321

Reinisch, E., Jesse, A., & McQueen, J. M. (2010). Early use of phonetic information in spoken word recognition: Lexical stress drives eye movements immediately. *Quarterly Journal of Experimental Psychology*, *63*, 772–783. https://doi.org/10.1080/17470210903104412

Reinisch, E., & Weber, A. (2012). Adapting to suprasegmental lexical stress errors in foreign-accented speech. *Journal of the Acoustical Society of America*, *132*, 1165–1176. https://doi.org/10.1121/1.4730884

Salverda, A. P., Kleinschmidt, D., & Tanenhaus, M. K. (2014). Immediate effects of anticipatory coarticulation in spoken-word recognition. *Journal of Memory and Language*, *71*, 145–163. https://doi.org/10.1016/j.jml.2013.11.002

Shuai, L., & Malins, J. G. (2017). Encoding lexical tones in jTRACE: A simulation of monosyllabic spoken word recognition in Mandarin Chinese. *Behavior Research Methods*, *49*, 230–241. https://doi.org/10.3758/s13428-015-0690-0

Soto-Faraco, S., Sebastián-Gallés, N., & Cutler, A. (2001). Segmental and suprasegmental mismatch in lexical access. *Journal of Memory and Language*, *45*, 412–432. https://doi.org/10.1006/jmla.2000.2783

Sulpizio, S., & McQueen, J. M. (2012). Italians use abstract knowledge about lexical stress during spoken-word recognition. *Journal of Memory and Language*, *66*, 177–193. https://doi.org/10.1016/j.jml.2011.08.001

Tagliapietra, L., & Tabossi, P. (2005). Lexical stress effects in Italian spoken word recognition. In B. G. Bara, L. W. Barsalou, & M. Bucciarelli (Eds.), *Proceedings of the XXVII Annual Conference of the Cognitive Science Society* (pp. 2140–2144). Mahwah, NJ: Erlbaum.

Tremblay, A. [Annie]. (2008). Is second language lexical access prosodically constrained? Processing of word stress by French Canadian second language learners of English. *Applied Psycholinguistics*, *29*, 553–584. https://doi.org/10.1017/s0142716408080247

Tremblay, A. [Antoine], & Ransijn, J. (2015). Model selection and post-hoc analysis for (G)LMER models. Retrieved June 1, 2017, from https://cran.r-project.org/web/packages/LMERConvenienceFunctions

van Donselaar, W., Koster, M., & Cutler, A. (2005). Exploring the role of lexical stress in lexical recognition. *Quarterly Journal of Experimental Psychology*, *58*, 251–273. https://doi.org/10.1080/02724980343000927

Wang, M., Koda, K., & Perfetti, C. A. (2003). Alphabetic and nonalphabetic L1 effects in English word identification: A comparison of Korean and Chinese English L2 learners. *Cognition*, *87*, 129–149. https://doi.org/10.1016/s0010-0277(02)00232-9

Zhang, Y., & Francis, A. (2010). The weighting of vowel quality in native and non-native listeners' perception of English lexical stress. *Journal of Phonetics*, *38*, 260–271. https://doi.org/10.1016/j.wocn.2009.11.002

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's website:

Appendix S1. Target, Competitor, and Distracter Words in Critical Test Items.
Appendix S2. Properties of Target, Competitor, and Distracter Words in Critical Test Items.
Appendix S3. Listeners' Proportions of Target, Competitor, and Distracter Fixations.
Appendix S4. Results of Between-Group Comparisons.