The role of lexical stress in spoken English word recognition by listeners of English and Taiwan Mandarin

Shu-chen Ou
National Sun Yat-sen University

Two perceptual experiments investigated how the suprasegmental information of monosyllables is perceived and exploited in spoken English word recognition by listeners of English and Taiwan Mandarin (TM). Using an auditory lexical decision task in which correctly stressed English words and mis-stressed nonwords (e.g. *camPAIGN* vs. **CAMpaign*) were presented for lexical decisions, Experiment I demonstrated that TM listeners could perceive the differences between stressed and unstressed syllables with native-like accuracy and rapidity. To examine how the perceived suprasegmental contrast would constrain English lexical access, Experiment II was conducted. It used a cross-modal fragment priming task in which a lexical decision had to be made for a visually presented English word or nonword following an auditory prime, which was a spoken word-initial syllable. The results showed that English and TM listeners recognized the displayed word (e.g. *campus*) faster both after a stress-matching (e.g. *CAM-*) prime and a stress-mismatching (e.g. *cam-*) prime than after a control prime (e.g. *MOUN-*, with mismatching segments). This indicates that suprasegmental information does not inhibit a segmentally matching but suprasegmentally mismatching word candidate for both the two groups, although TM is a language where lexical prosody is expressed syllabically and its listeners tend to interpret lexical stress tonally. Yet, the two groups' responses were slower after the stressed primes than after the unstressed ones, presumably because the former generally had more possible continuations than the latter do. It is therefore concluded that when recognizing spoken English words, both the native and non-native (TM-speaking) listeners can exploit the suprasegmental cues of monosyllables, which, however, are not so effective that they will outweigh the segmental cues.

Keywords: spoken word recognition, English lexical stress, lexical tone language, cross-modal fragment priming experiment

https://doi.org/10.1075/lali.00049.ou
Language and Linguistics 20:4 (2019), **569–600**
ISSN 1606-822X / E-ISSN 2309-5067

1. Introduction

To understand speech, one has to recognize the individual words. In most cases, the listener can identify the words of their native language accurately and rapidly although there are no consistent markers of lexical boundaries in the speech signal itself (Lehiste 1972; Nakatani & Dukes 1977; Woodward & Aslin 1990). The ability to single out a word in a mental lexicon consisting of a myriad of entries has invited substantial research into phonetic cues that may subserve spoken word recognition. One of these cues is lexical stress – a contrastive suprasegmental feature of words in languages like English (e.g. *INsert* vs. *inSERT*, where the uppercase letters indicate the stressed[1] syllables). If such a feature is exploited in on-line speech processing, the English words *MUsic* and *muSEUM*, for example, can be told apart before the fourth segment ([ɪ] in *MUsic* or [i] in *muSEUM*), the point at which the words begin to differ segmentally. The suprasegmental[2] difference in the initial syllable already differentiates the words.

Lexical stress is of course not a linguistic universal. This fact raises the questions whether and how this suprasegmental feature is used by non-native listeners to recognize the words of languages like English. There is much research on perception of lexical stress by listeners whose native language does not possess variable word stress (e.g. French-speaking listeners: Dupoux et al. 1997; Peperkamp 1997; Peperkamp & Dupoux 2002; Dupoux et al. 2008). However, the consequences of their perceptual success or failure for spoken word recognition have not been extensively investigated (e.g. Tremblay 2008). An additional impetus for exploring non-native listeners' use of lexical stress is provided by the typology of lexical prosody: while cross-linguistic word recognition studies have been conducted on closely related languages such as English and Dutch (e.g. Cutler et al. 2007), there is little to no relevant work delving into languages which express lexical distinctions at the suprasegmental level by means other than lexical stress. For this reason, the present study investigates the role of lexical stress in English word recognition by listeners of Taiwan Mandarin (henceforth TM), a lexical tone language typologically distant from English. The investigation would provide insight into how experience with the

1. It has been proposed that there are multiple levels (e.g. primary, secondary, and tertiary) of stress in English (Chomsky & Halle 1968). However, since a distinction between primary and non-primary stress is sufficient for the discussion in this paper, the terms "stressed" and "unstressed" will be used to refer to this distinction for convenience.

2. Although the term "suprasegmental" covers a wide range of prosodic features (e.g. lexical stress, phrasal pitch accent, intonation, etc.), this study focuses only on, and uses the term to refer to, features associated with lexical prosody.

lexical prosody of the first language (L1) may or may not influence the processing of the spoken words of a second language (L2).

The remainder of this paper is organized as follows. §2 first gives a brief sketch of spoken word recognition and one way by which this process can be experimentally investigated. A review of previous studies probing into the use of lexical stress in recognizing spoken English words follows, and two hypotheses concerning the extent to which the suprasegmental information of English monosyllables is exploited by Mandarin listeners are proposed. §3 and 4 report the designs and the results of two perceptual experiments carried out to address the hypotheses. §5 presents a general discussion of the findings, followed by some concluding remarks and further issues.

2. Spoken word recognition: The role of lexical stress

Recognition of spoken words requires continuous processing of the speech signals. In general, it involves two dynamic processes: multiple candidate words are activated (Zwitserlood 1989; Marslen-Wilson 1990; Connine et al. 1994) and then competition among these words ensues (Goldinger et al. 1989; Gaskell & Marslen-Wilson 2002). Further information in the speech input is used at the first opportunity to winnow down the set of candidate words (Marslen-Wilson & Warren 1994; McQueen et al. 1999; Soto-Faroco et al. 2001). A particular word is soon recognized once its uniqueness point (e.g. Marslen-Wilson 1980), the point at which the accrued information supports no other possible candidates, has been reached. Therefore, for a word to outcompete the others, the amount of evidence in favor of it is crucial. What arise from such a view of the word recognition process are empirical questions concerning which aspects of a lexical item constitute the evidence that can help reduce the concurrently activated competitors and also the extent to which they are helpful. Numerous related studies have been conducted to explore the use of lexical stress in languages such as English and Dutch, focusing on cases where segmental structures cannot further reduce word candidates but the suprasegmental differences among the words possibly can.

In English and Dutch, primary stress freely falls on a particular syllable of a word and can sometimes distinguish between words that are segmentally identical (e.g. *INsert*, a noun vs. *inSERT*, a verb). The prominence of the stressed syllable is realized with a combination of acoustic-phonetic features, including higher pitch, longer duration, greater intensity, and sometimes fuller vowel quality (Fry 1955; Gay 1978; Kochanski & Orphanidou 2008). Various methods have been devised to test whether the cues that are purely suprasegmental (i.e. pitch, duration, and intensity) are relevant for spoken word recognition. Among them is the cross-modal priming technique, which is used in many previous studies (e.g. Swinney 1979; Marslen-Wilson & Zwitserlood 1989; Cooper et al. 2002; van Donselaar et al. 2005) and in the present work (Experiment II). In a cross-modal priming experiment, subjects may be asked to carry out certain tasks after being presented with an auditory prime and then a visual stimulus. For example, in Experiment II of this paper, subjects decided whether a visually displayed string of letter (e.g. *campus*) was a real English word or not as soon as possible after hearing the prime (e.g. *CAM-*, the stressed initial syllable of *campus*). If there is effective exploitation of the stress cues in the prime, decision latencies would be expected to be shorter when the auditory and visual stimuli contain matching stress information (e.g. hear *CAM-* and see *campus*) but longer when the stimuli contain mismatching stress information (e.g. hear *CAM-* and see *campaign*, the first syllable of which is unstressed), in comparison to decision latencies in a control condition where the prime mismatches the visual target segmentally (e.g. hear *MOUN-* and see *campaign*).

Cooper et al. (2002) have employed this priming technique to investigate how lexical stress is used by native English listeners in recognizing their own words. In one of their experiments, spoken monosyllables excised from disyllabic words, which were either stressed (e.g. *MU-* of *music*) or unstressed (e.g. *mu-* of *museum*), served as auditory primes to visual target words such as *music* or *museum* or to some nonwords. The results indicated that the English-speaking participants' lexical decisions were significantly faster in a stress-matching condition where the monosyllabic prime matched the first syllable of the visual target both segmentally and suprasegmentally than in a control condition. There was also a stress-mismatching condition where the auditory and visual stimuli matched segmentally but not suprasegmentally. Nevertheless, it was shown that the participants' responses in this condition were still significantly faster than those in the control one. These findings suggest that insofar as monosyllabic primes are concerned, recognition of words such as *music* and *museum* is facilitated as long as their initial syllables are segmentally identical to the prime; suprasegmental mismatch does not inhibit either of the candidates. Cooper et al. (2002) conclude that segmental cues outweigh suprasegmental ones in native English listeners' on-line word processing.

Although the findings reported by Cooper et al. (2002) may appear paradoxical given that English is a lexical stress language, they are in fact explicable and consistent with earlier studies. Cutler (1986) carried out an associate priming experiment with English-speaking participants, who heard one member from a stress minimal pair (e.g. *FORbear*, which formed such a pair with *forBEAR*) and saw a word that might or might not be semantically related. It was found that word pairs like *FORbear* and *forBEAR* primed recognition of each other's semantic associate (i.e. "ancestor" and "tolerate," respectively). Cutler explains that for English listeners, processing the distinction between syllables differing in lexical stress is usually tantamount to processing the difference between reduced and unreduced vowels. While it is possible for English word stress contrast to involve only pitch, duration, and intensity, this rarely happens: the syllable with weak stress tends to be accompanied by vowel quality reduction. For example, in *SUBject* and *subJECT*, the unstressed vowel of the first syllable of *subJECT* (i.e. [ə]) is reduced and differs from the stressed vowel of the first syllable of *SUBject* (i.e. [ʌ]). The reduction is so prevalent in English that its native listeners usually profit from the segmental information in telling one word from another; there is no compelling need for them to further attend to suprasegmental differences. This view is supported by Cutler & Pasveer's (2006) lexical statistics, which show that there are few stress minimal pairs in English. It is then not surprising that English listeners' word identification would be impeded to a greater extent by segmental mispronunciations than by misplacement of lexical stress (e.g. *INsert* produced as *inSERT*, or vice versa) (Bond & Small 1983; Small et al. 1988; Slowiaczek 1990). Also, their acceptability ratings of cross-spliced word tokens significantly declined for cross-splicings involving alternation of vowel quality (e.g. *autumn* with its *au-* replaced with the *a-* of *atomic*, which is a reduced vowel), but not for those involving change in only the level of stress (e.g. *autumn* with the initial vowel *au-* replaced with that of *automation*) (Fear et al. 1995).

However, the mounting evidence for segments being a more effective cue than lexical stress in spoken word recognition cannot be taken to mean that listeners of whatever lexical stress language would pattern with those of English. A rather different picture emerges in listeners of Dutch (van Donselaar et al. 2005) and Spanish (Soto-Faraco et al. 2001), to whom the suprasegmental dimensions of spoken words are of greater relevance and utility. Van Donselaar et al. conducted one cross-modal fragment priming experiment where Dutch-speaking participants were asked to make lexical decisions for a visually presented Dutch word (e.g. *OCTOPUS*) after hearing a monosyllabic word onset fragment (e.g. *OC-*). As expected, their decisions were significantly more rapid in the stress-matching condition than in the control one. Nevertheless, the response times in the stress-mismatching condition were not significantly different from those in the control condition; that is, mismatch of lexical stress leads to neither facilitation nor inhibition. Taken together, the findings from van Donselaar et al. (2005) and Cooper et al. (2002) suggest that when recognizing the words of their own language, Dutch listeners are more sensitive to the presence of suprasegmental information than English listeners are.

Some studies thus examine whether L2-English listeners such as those whose L1 is Dutch can exploit suprasegmental information more effectively than native ones do. Off-line perceptual tasks have demonstrated that Dutch listeners' word identification is impeded by mis-stressing (van Heuven 1985; van Leyden & van Heuven 1996). This finding is consistent with the less strong tendency for Dutch unstressed syllables to be associated with reduced vowels (Booij 1995). Furthermore, Dutch listeners are actually found to use lexical stress even better than English native ones did in identification tasks where they had to ascribe a heard fragment of an English word to its source word (Cutler et al. 2007). As a result, it would not be unreasonable to expect that they can utilize surpasegmental cues to a greater extent than native English listeners do in recognizing English words. This possibility was investigated in the aforementioned cross-modal fragment priming experiment by Cooper et al. (2002), which presented all their participants with English materials. The results, however, indicated that the findings obtained by van Donselaar et al. (2005) with Dutch materials were not replicated: compared with those in the control condition, Dutch listeners' lexical decisions were significantly faster both in the stress-matching and in the stress-mismatching conditions. There was facilitation as long as the monosyllabic prime and the visual target word contained matching segments, and this is exactly the same pattern of results observed for the English group. Cooper et al. (2002) conclude that for Dutch listeners, lexical stress is not as helpful for constraining lexical access in English as it is in their native language.

Dutch listeners' failure to extend the efficiency with which they use the suprasegmental cues of monosyllables to their L2 (English) is compatible with the explanation offered by Cutler (1986). As has been mentioned, few words can be told from each other on a purely suprasegmental basis because the contrast between stressed and unstressed syllables is highly correlated with difference in vowel quality in English. This then suggests that the correlation in the language per se is responsible for the less effective use of lexical stress and that when presented with monosyllabic primes, listeners are likely to pattern with the English native ones in Cooper et al. (2002), regardless of their linguistic backgrounds. This argument is consistent with the findings from Tremblay's (2008) cross-modal priming experiment with L2-English learners whose native language was Canadian French – a language that lacks variable stress. They heard a single word onset syllable (e.g. *MYS*-) and selected the English word that matched it (e.g. *mystery*, with the mismatching alternative being *mistake*). The results revealed that the ability to constrain English word access with the help of suprasegmental cues is hardly in evidence in this population of English learners.

It is currently not clear whether the argument also holds for listeners of Mandarin, a lexical tone language which uses pitch in a lexically contrastive fashion. Two hypotheses are proposed for these listeners. First, they would simply pattern with the native English and non-native (Dutch) participants of Cooper et al. (2002) in an analogous cross-modal fragment priming experiment. That is, after a monosyllabic prime is presented, there is facilitation of responses as long as it matches the first syllable of the visual target segmentally even if the stimuli mismatched suprasegmentally. This possibility is supported by the finding that for both Mandarin and English listeners, vowel quality is the primary cue in off-line identification of stressed syllables whereas pitch is only a secondary one (Chrabaszcz et al. 2014). In fact, even when listening to their own words, listeners of tone languages such as Cantonese process segmental distinctions more efficiently than tonal information (Cutler & Chen 1997). For example, in two auditory priming experiments by Lee (2007), Mandarin listeners' lexical decisions to spoken monosyllabic Mandarin words are faster if the previously heard monosyllabic prime word is segmentally identical but tonally distinct than if it is tonally identical but segmentally distinct, indicating that matching segmental information seems to outweigh matching suprasegmental information.[3]

Nevertheless, although tonal cues are less efficiently processed than segmental ones, it is possible that Mandarin listeners display an asymmetry in native versus L2 spoken word recognition (just as Dutch-speaking ones do). An alternative hypothesis may be that Mandarin listeners actually outperform English native ones in using the suprasegmental information of monosyllables to recognize spoken English words. If this is the case, it would be predicted that they exhibit no facilitation or even inhibition when their responses in the stress-mismatching condition are compared with those in the control. Support for such a prediction – and also the impetus for exploring Mandarin listeners' exploitation of monosyllabic primes – are provided by the fact that lexical tones are pitch patterns over single syllables. As shown by the oft-cited example in (1), the four lexical tones of Mandarin function in much the same way as distinctive segments do: a change to the tone of a syllable will result in a change in meaning. Experience with such lexical prosody would promote sensitivity to the suprasegmental dimensions of individual syllables. This view is substantiated by Schaefer & Darcy (2014), whose cross-linguistic perceptual experiment reveals that Mandarin listeners are more accurate in categorizing monosyllables with various pitch patterns (which are actually Thai lexical tones) than listeners of a lexical stress language (e.g. English) or a lexical pitch accent language (e.g. Japanese) do. The lexical prosody of Mandarin encourages the language's listeners to take advantage of the suprasegmental information of monosyllables, perhaps even when these syllables are from an L2.

(1) Mandarin lexical tones. The letters H and L represent high and low pitch values, respectively.[4]

HH (Tone 1)	LH (Tone 2)	LL (Tone 3)	HL (Tone 4)
ma 'mother'	ma 'hemp'	ma 'horse'	ma 'scold'

3. We would like to thank two anonymous reviewers for bringing the study by Chrabaszcz et al. (2014) and that by Lee (2007) to our attention.

4. Pitch heights in Mandarin can also be indicated by using numerals (1, 2, 3, 4, and 5), as in Chao (1968), where Tone 1 is 55, Tone 2 is 35, Tone 3 is 214, and Tone 4 is 51.

This alternative hypothesis is reinforced by the well-documented observation that Mandarin listeners are inclined to make a tonal interpretation of lexical stress. The native language has a profound impact on discrimination of both L2 segmental and suprasegmental contrasts (e.g. Best 1994; Flege 1995; So & Best 2010; Lukyanchenko et al. 2011) and Mandarin listeners are no exception. They are reported to be biased toward the cue of pitch in their production and perception of English stressed versus unstressed syllables (Cheng 1968; Juffs 1990; Archibald 1997; Wang 2008a). For example, in Cheng's code-switching study, it is suggested that English unstressed syllables had been treated by his Mandarin-speaking participants as bearing the [−high] tone because they triggered Mandarin Tone 3 Sandhi, whereby a Tone 3 (a [−high] tone) is produced as a Tone 2 (a rising tone) when followed by another Tone 3. The tendency to equate differences in lexical stress to tonal contrasts is corroborated by listening experiments showing that identification of stress patterns is particularly difficult when pitch is no longer a reliable cue. In Ou (2016), English-learning listeners of TM heard one member of an English disyllabic stress minimal pair (e.g. *PERmit* and *perMIT*) and determined whether it was a noun, which had a trochaic stress pattern, or a verb, which had an iambic stress pattern. The listeners accurately identified the lexical category of the spoken word when it was excised from a statement sentence (i.e. *I said* ______.), a context where the stressed syllable was simultaneously signaled by higher pitch, longer duration, and stronger intensity. Nevertheless, regardless of their English proficiency level, they failed to do the same when the word was excised from a yes/no question (i.e. *Did you say* ______*?*), a context where its second syllable invariably had higher pitch than the first one. Analogous results have been obtained from comparable experiments using nonword stress minimal pairs (Ou 2010). In brief, the difference in lexical stress is treated by Mandarin listeners as the contrast between [+high] and [−high] tone, which is crucial for word identity in their native language.

The possibility that non-native listeners can make better use of suprasegemental information than native ones do in recognizing English words would be worth testing. Most previous studies (e.g. Cutler & Otake 2004; Broersma & Cutler 2008, 2011) are focused on cases where non-native listeners have poor discrimination of certain English segmental contrasts. One example is Dutch listeners' confusion of the [ɛ] and [æ] sounds in English. Using a series of cross-modal priming experiments, Broersma and Cutler (2011) found that, not surprisingly, English-speaking subjects' recognition of English words like *deaf* was facilitated by a monosyllabic prime that contained a matching vowel (e.g. *def-*, excised from *definite*) but inhibited by one that contained a mismatching vowel (*daff-*, excised from *daffodil*), when compared with their responses after a control prime. Yet, a different pattern was found for the Dutch-speaking group: both *def-* and *daff-* primed their recognition of *deaf*, producing facilitation when compared with the control prime. This is a sign of ineffective use of the vowel quality difference between [ɛ] and [æ], which presumably stems from the fact that the vowels are not contrastive in Dutch. In Mandarin, however, the distinction between a [–high] tone and a [+high] tone is phonologically relevant. It is of interest to see whether lexical stress, being interpreted tonally, can effectively constrain Mandarin listeners' English word access.

The two competing hypotheses are recapitulated as follows. First, it is hypothesized that Mandarin listeners' exploitation of lexical stress would not be superior to that of English-speaking ones and thus predicted that the two groups would show the same response pattern: compared with those following a control prime, lexical decisions to a displayed English word are facilitated by a prime that is segmentally identical to the word's initial syllable, whether the prime and the syllable suprasegmentally match (e.g. hear *CAM*- and see *campus*) or mismatch (e.g. hear *CAM*- and see *campaign*). However, the tone language listeners may be more effective users of lexical stress. If this is the case, it is predicted that their recognition responses would be inhibited or at least not facilitated by a suprasegmentally mismatching prime, when compared with the responses after a control prime. Conducted to test the predictions was one cross-modal fragment priming experiment where the auditory primes were some English monosyllables. The participants included native English listeners as well as non-native listeners with their L1 being TM.

However, before the priming experiment was carried out, it would be necessary to assess the accuracy and speed with which the TM participants could perceive lexical stress contrasts. Although Mandarin listeners can identify the position of a stressed syllable in off-line perceptual tasks (e.g. Altmann & Vogel 2002; Altmann 2006; Wang 2008b; Ou 2016), it has to be proved that they still do so in an on-line task requiring immediate responses, which may be more cognitively demanding. If they cannot, any findings in favor of the first hypothesis could potentially be attributed to the lack of the ability to discern stress differences. A simple auditory lexical decision task (Experiment I) would meet the current need. As used here, the task presented mis-stressed and correctly stressed disyllabic English words truncated from a statement sentence (where the stressed syllables were cued by higher pitch, longer duration, and stronger intensity) for lexical decisions. The TM and English-speaking participants would be compared for response accuracy and rapidity. The cross-modal priming experiment (Experiment II) was then implemented to test the two hypotheses.

3. Experiment I: Auditory lexical decision task

3.1 Materials

The materials of Experiment I were English word-nonword stress minimal pairs, the real words of which were disyllabic content words selected with the following two criteria: they (i) occurred at least five times per million spoken words according to the CELEX English database (Baayen et al. 1993) and (ii) were included in the 7,000-word English vocabulary list for Taiwanese high school students compiled by the College Entrance Examination Center of Taiwan. These criteria were imposed to exclude as many as possible of the words that the non-native participants might not know. Twenty-eight words were selected and their nonword counterparts were derived by shifting the position of primary stress without altering the vowel quality (e.g. *PROblem* [ˈprɑbləm] vs. **proBLEM* [prɑˈbləm]). A full list of the word-nonword stress minimal pairs is in Appendix A.

All the word and nonword items were embedded in the final position of a statement sentence (i.e. *The word is* ______.) to elicit productions under the falling pitch accent context, where stressed syllables were more prominent than unstressed ones in pitch, duration, and intensity. With a printed sheet listing all the items (with phonetic transcriptions for the nonwords), a trained female phonetician and native speaker of English with a North American accent was instructed to read them at least twice in the carrier sentence in a sound-attenuated room. The readings were recorded via a Zoom H4n Handy Recorder, digitized at a sampling rate of 4.4 kHz, and stored on a flash drive. Of all the recorded readings of the same sentence, only one was selected for use and edited to remove its sentential context, leaving only the word or nonword originally in the sentence-final position. The results were 56 single-word stimuli.

E-Prime 2.0 (Psychology Software Tools 2012) was used for stimulus presentation and response recording. The 56 stimuli were presented, one per trial, in two experimental blocks that contained an equal number of trials. Within each block, the order of the trials was randomized. In each trial, a fixation cross (+) was displayed at the center of the screen and remained there for 1500 milliseconds (ms), followed by the presentation of the stimulus. Both response accuracy and response times (RTs) were recorded, and the measurement of RTs began immediately at the offset of the stimulus. Responses longer than 5000 ms were counted as omissions and their RTs were not recorded.

3.2 Acoustic analysis

To see if there were significant acoustic differences of lexical stress in the items recorded by the native English speaker, measurements of pitch, duration, and intensity were carried out by using Praat (Boersma & Weenink 2016) on all the syllables of the 56 stimuli. Shown in Table 1 are the syllables' mean values on these three suprasegmental dimensions, subdivided according to position (i.e. being the first or second syllable of a word or nonword). Since the real words were all disyllabic and each had a mis-stressed counterpart, planned paired *t*-tests were used to compare the stressed and unstressed forms of each syllable (e.g. *cam*- from *camPAIGN* and *CAM*- from **CAMpaign*). The results showed that compared with the corresponding unstressed ones, the stressed syllables had significantly higher pitch ($t(55) = 12.36$, $p < .001$, two-tailed), longer duration ($t(55) = 11.08$, $p < .001$, two-tailed), as well as greater intensity ($t(55) = 10.58$, $p < .001$, two-tailed).

Table 1. Mean values of pitch (in Hz), duration (in ms), and intensity (in dB) for the stressed and unstressed syllables of the 56 stimuli, subdivided according to position. The standard deviations (SDs) are in parentheses

	Stressed			**Unstressed**		
	Pitch	**Duration**	**Intensity**	**Pitch**	**Duration**	**Intensity**
First syllable	216.0 (23.6)	173.4 (63.2)	75.4 (2.2)	180.3 (18.6)	123.1 (50.2)	71.7 (2.9)
Second syllable	188.7 (18.2)	227.6 (56.3)	73.8 (1.7)	149.4 (11.7)	151.9 (51.7)	68.0 (3.1)

3.3 Procedure

All subjects were tested individually in a sound-attenuated chamber with a desktop computer. At the outset of the experiment, they were shown instructions (in English) on the computer monitor, which were accompanied by oral explanations from the experimenters (given in Mandarin or English). They were informed that they would hear a pre-recorded sound file over headphones in each trial and their task was to decide, as soon as possible, whether it was a real English word or not by pressing a YES key or a NO key on a Serial Response Box (model 200a) from Psychology Software Tools. Subjects were made aware that no feedbacks would be provided and that the stimuli to which they should respond NO were nonwords which were derived from real English words by altering the location of primary stress. They were also told that they could take a break and decide when to move on whenever they completed a block. Eight trials were offered for practice. Then, the experimenter remained in the room to give explanations or answer questions until subjects were ready to proceed to the experiment proper, which took about 20 minutes.

3.4 Participants

Twenty-four English native speakers and 23 native TM speakers participated, all with normal hearing and normal (or corrected-to-normal) vision. To ensure that the TM subjects knew the test stimuli, a minimum TOEIC score of 800 and 10 years of English learning experience were required for their participation.

3.5 Results

In general, subjects were able to make correct lexical decisions, or decide if the stimuli were correctly stressed: the mean accuracy rates of the English and TM groups were 87.64% and 89.28%, respectively. To further determine if there was a significant group difference, their responses were analyzed with a linear mixed-effects logistic regression model. This model (and those used in subsequent analyses) was fitted by using the lme4 package (Bates et al. 2015) of R (R Core Team 2015). The dependent variable was the response to each trial, which was either correct or incorrect (coded as 1 and 0, respectively). The predictors were all binary factors each with two levels coded as −0.5 and 0.5; they included *Group, Stress Position*, and *Word Status*. The *Group* predictor indicated whether the response was made by an English speaker (−0,5) or by a TM speaker (0.5). *Stress Position* specified whether the stimuli presented in a trial was stressed on the first syllable (−0.5) or on the second one (0.5). *Word Status* encoded the contrast between real-word (0.5) and nonword (−0.5) stimuli. The fixed effects to be included in the model were selected via a procedure that began with constructing one-variable models that each contained only the main effect of one of the predictors. The model (or one of the models) which revealed a significant main effect was chosen. It was then compared by using a likelihood ratio test with a new, less parsimonious model derived by further including a main effect or interaction to see if this newly introduced fixed effect was necessary. This process was repeated until there was no additional main effect or interaction that could significantly contribute to goodness-of-fit to the data. The selected model consisted of the main effects of *Stress Position* and *Word Status* and their interaction. By-subject and by-item random intercepts were entered as random effects.

The results are as follows. The model revealed that there were two significant main effects: one of *Stress Position* and the other of *Word Status*. The former indicated that as far as nonword stimuli are concerned, the ones with primary stress on the second syllables were responded to significantly more accurately than the ones with primary stress on the initial syllables ($\beta = 1.20$, $SE(\beta) = .19$, $z = 6.35$, $p < .001$) (see Figure 1). The latter indicated that among stimuli with initial primary stress, those with the correct stress pattern (i.e. real words) were responded to with significantly higher accuracy than those with the incorrect stress pattern (i.e. nonwords) ($\beta = 2.02$, $SE(\beta) = .19$, $z = 10.62$, $p < .001$). However, it makes little sense to interpret these two main effects since they were essentially the by-products of the interaction term in the model, which was significant ($\beta = -2.24$, SE(β) = .52, $z = -4.29$, $p < .001$). As is clear from Figure 1, what brought forth the significant interaction was subjects' relatively poor performance on the mis-stressed stimuli with primary stress on the first syllables: they were inclined to respond YES to spoken nonwords such as **CARtoon* and **TECHnique*. This may be attributed to the fact that the majority of English words begin with stressed syllables (Cutler & Carter 1987), leading listeners to assume that a real word typically has initial primary stress. Note that such a bias was found for both the English and TM groups. In addition, the main effect of *Group* and its interaction with the other two predictors had been excluded for little contribution to model fit. There was thus no compelling evidence for suggesting that the two groups were in any way different from each other in terms of response accuracy; the TM participants were able to respond to English lexical stress just as well as the English native ones were.

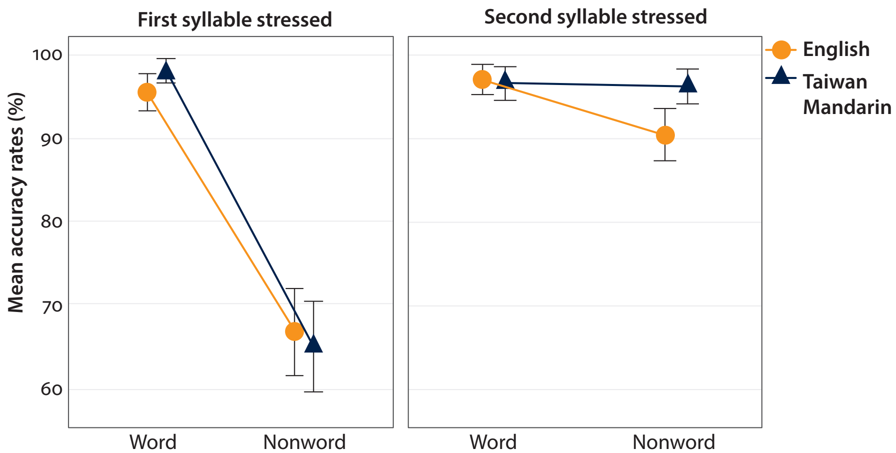

Figure 1. Mean accuracy rates by position of primary stress (i.e. first- or second-syllable stressed) and by word status (i.e. word or nonword) in Experiment I for English and Taiwan Mandarin subjects. The standard error (SE) bars indicate 95% confidence intervals

As the experiment used an on-line task that required immediate lexical decisions, a further analysis was conducted on the RT data of correct responses. RTs that were two-and-a-half SDs above the average RT of all subjects (i.e. longer than 1391 ms) were first discarded as outliers and this removed 11.55% of the data.[5] Figure 2 shows the mean values of the remaining RTs for the two groups. The RTs were natural-log

5. There was no lower limit for excluding potential RT outliers as listeners might be able to decide on the word status of the stimulus before it reached its offset, the point at which RT measurement started.

transformed and used as the dependent variable of a linear mixed-effects regression model, which was selected by following the same procedure as that described above. The model contained the main effect of *Word Status* as the sole fixed effect and by-subject and by-item random intercepts as random effects. The fixed effect was significant: responses to correctly stressed real words were more rapid than those to mis-stressed nonwords ($\beta = -.32$, SE(β) = .02, $t = -18.67$, $p < .001$). The absence of the main effect of *Group* again suggested that there was no evidence for a group difference. In general, the TM group could respond with native-like rapidity.

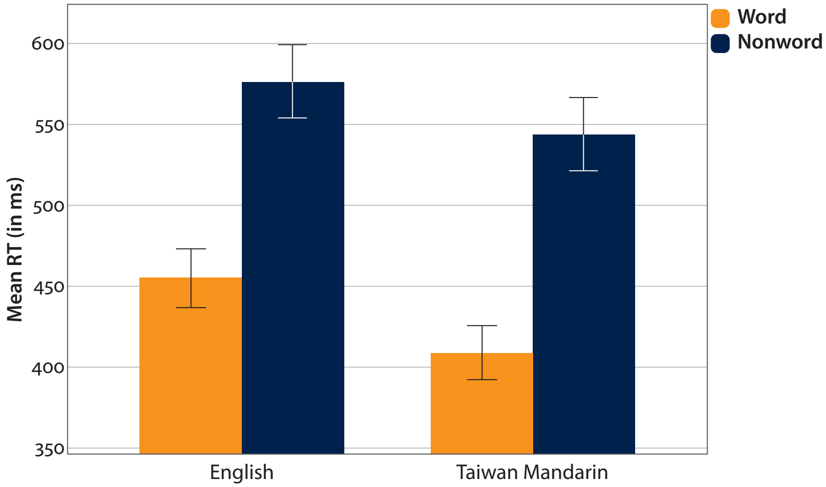

Figure 2. Mean RTs by word status in Experiment I for English and Taiwan Mandarin subjects, with SE bars representing 95% confidence intervals

It is clear from the above findings that the TM listeners could distinguish between stressed and unstressed syllables and thereby determine whether the stimuli had correct word stress patterns as accurately as the native English listeners could. Besides, their lexical decisions were not delayed and even appeared to be slightly faster compared with those of the English group. The TM listeners were not impervious to the suprasegmental information of English words. However, the extent to which such information is exploited to constrain lexical access would be another question, which was addressed in the next step of research. Specifically, it was investigated whether the TM listeners would outperform the English listeners in using the suprasegmental cues of English monosyllables. The experiment conducted is reported below.

4. Experiment II: Cross-modal fragment priming task

4.1 Materials

Twenty-one pairs of disyllabic English words were used in Experiment II. Each of them consisted of two words that had to begin with a segmentally identical first syllable but differ in the position of primary stress (e.g. *CAMpus* and *camPAIGN*). This requirement would exclude word pairs in which differences of vowel quality in the first syllables co-varied with the presence or absence of primary stress (e.g. *SUBject* and *subJECT*). If multiple words could be paired with one another, the two that were the closest in spoken frequency per million words according to the CELEX database formed a pair. Such pairs were the critical items of the experiment. Another disyllabic word was then assigned to each of these pairs as the control item. Its first syllable had the same syllable structure as that of the pair, but comprised distinct segments. For example, *mountain*, with its initial syllable being /maʊn/, was assigned as the control item to the pair *campus* and *campaign*, both of which began with /kæm/. This resulted in 21 sets of words, each consisting of one control and two critical items. A complete list of these sets was in Appendix B.

All the items were inserted in the final position of a non-constraining sentence (e.g. *The word he said was* ______.) and were read by the same English native speaker as Experiment I. The subsequent procedure of speech recording was also the same, except that the recorded sentences were edited to remove the second syllable of the inserted word rather than the sentential context. This left a monosyllabic word onset fragment at the end of each sentence (e.g. *The word he said was CAM-.*), which served as the auditory prime. The upcoming visual target word was either a critical item (which was a real word) from the same set of words as the truncated word used as the prime, or an orthographically possible but non-existent English word retrieved from the ARC nonword database (Rastle et al. 2002). The lengths of nonwords ranged between four and eight letters – the minimal and maximal lengths of the critical items, respectively. In accord with this design, the edited sentences and the visual targets were paired in all possible combinations to construct the trials of the experiment. Depending on the type of the prime and the upcoming target, the trials belonged to either of the following conditions: (i) the stress-matching condition, in which the auditory prime and the first syllable of the visual word matched segmentally and suprasegmentally (e.g. hear *CAM-* and see *campus*), (ii) the stress-mismatching condition, in which the two stimuli matched segmentally but not suprasegmentally (e.g. hear *cam-* and see *campus*), (iii) the control condition, in which the auditory prime was a fragment of the control item (e.g. hear *MOUN-* and see *campaign*), and (iv) the nonword condition, in which the visual target was a nonword (e.g. hear *CAM-* and see *broc*). There were 126 trials (three trial conditions × two possible visual targets × 21) for the first three conditions and another 126 trials for the nonword condition.

As with Experiment I, the cross-modal fragment priming experiment was programmed and run by E-Prime 2.0. The trials divided into two parts that each contained three blocks. In each trial, a fixation cross was displayed for 1,500 ms and the fragmentary prime, embedded in its carrier sentence, was presented. The visual target then appeared immediately at prime offset, the point at which recording of response and RTs began, and remained on the screen for 5,000 ms. Responses longer than 5,000 ms were regarded as omissions and their RTs were not recorded. Within each block, there was an equal number of trials in which the visual target was a real word and trials in which it was a nonword. All the trials were presented in a randomized order.

4.2 Acoustic analysis

An acoustic analysis was conducted on the monosyllabic auditory primes in the stress-matching and stress-mismatching conditions. Table 2 shows the mean values of pitch, intensity, and duration for the stressed and unstressed primes. Pairwise comparisons using planned paired t-tests indicated that the stressed primes (e.g. *CAM-*) generally had significantly higher pitch ($t(20) = 9.92$, $p < .001$, two-tailed), longer duration ($t(20) = 5.12$, $p < .001$, two-tailed), and stronger intensity ($t(20) = 7.84$, $p < .001$, two-tailed) than their unstressed counterparts (e.g. *cam-*).

Table 2. Mean values of pitch (in Hz), duration (in ms), and intensity (in dB) for the stressed and unstressed primes presented in the stress-matching and stress-mismatching conditions, each with its SD in parentheses

Stressed			Unstressed		
Pitch	**Duration**	**Intensity**	**Pitch**	**Duration**	**Intensity**
235.7 (25.3)	159.6 (50.5)	78.7 (1.7)	180.3 (17.2)	124.7 (42.1)	73.8 (2.0)

4.3 Procedure

Participants were randomly allocated to complete one part of the experiment first. They were informed that in each trial, they would hear a pre-recorded sentence with part of its final word truncated and then see a string of letters on a computer screen. All they needed to do was to decide, as soon as possible, whether *the string on the screen* was a real English word or not by pressing the YES key or the NO key on the same serial response box. Eight trials were offered as practices at the outset of each part. Participants proceeded to the other part after completing the one with which they began. The whole experiment lasted approximately 30 minutes.

4.4 Participants

Same as Experiment I.

4.5 Results

All subjects decided on the word status of the visually presented words at a near-perfect accuracy rate (English: 98.90%; TM: 93.67%), except for three TM participants, whose accuracy rates in and only in the control condition were 2.5 SDs below the mean. Their data were discarded. Of relevance to subsequent analyses were the RTs of the correct YES responses of the valid subjects. RTs shorter than 100 ms might be problematic since it is physiologically impossible to make motor responses to perceived stimuli with such rapidity (Luce 1986). Those that were 2.5 SDs above the mean RT (i.e. longer than 1,047 ms) were treated as outliers. Applying these criteria to remove the potentially invalid RTs excluded 2.53% of the correct YES responses.

The RTs of the remaining data were log-transformed and subjected to analyses using linear mixed effects regression models. The predictors of interest included *Group* and *Condition*.[6] Just as in Experiment I, the *Group* predictor was used to examine whether there was a difference between English and TM listeners, coded as −0.5 for the former and 0.5 for the latter. *Condition* was a three-level factor that consisted of the contrasts between the control condition (which was designated as the baseline level and coded as −0.5) on one hand and the stress-matching and mismatching conditions (coded as 0.5) on the other hand. A model selection procedure that was identical to that of Experiment I but began with the main effect of *Condition* was followed here (and in the following analysis). The selected model

6. We are grateful to one anonymous reviewer for pointing out that the design of Experiment II, in which the same visual target word was presented a total of three times, might induce a long-term priming effect: responses to a word would be facilitated on a second encounter. We thus re-fitted the models originally used in the subsequent analyses, adding a fixed effect called *Repetition*, which represented the *n*th time the target word was presented. The effect was significant in the model reported in this paragraph ($\beta = -.05$, SE(β) = .003, $t = -17.90$, $p < .001$) and the one in the next paragraph ($\beta = -.07$, SE(β) = .004, $t = -16.65$, $p < .001$), suggesting that the more times a word had been presented, the faster the responses tended to be. However, even after *Repetition* it was introduced to the models, the significant patterns of the fixed effects of interest to the research questions (e.g. *Condition*) remained unchanged.

had a single fixed effect for *Condition*; the random effects included a by-subject random intercepts and by-item random intercepts for the spoken prime and the visual target. The sole fixed effect was significant: when compared with those in the control condition, responses were significantly faster in the stress matching ($\beta = -.06$, $SE(\beta) = .01$, $t = -7.37$, $p < .001$) and stress-mismatching ($\beta = -.04$, $SE(\beta) = .01$, $t = -4.31$, $p < .001$) condition for the English group (see Figure 3). This pattern would also hold for the TM-speaking participants, since the absence of the main effect of *Group* and its interaction with *Condition* from the model suggested that there was neither a significant group difference (in response speeds of the control condition) nor a need to adjust the effect of *Group* for *Condition*. The findings are therefore consistent with the hypothesis that there would be a parallel between TM and English listeners in the extent to which the stress cues of spoken monosyllables are exploited. That is, recognition of a visual target is facilitated as long as the initial syllable of the target and the preceding prime contain identical segments; a prime with matching segmental content but suprasegmentally mismatching information leads to no inhibition and even produced facilitation.

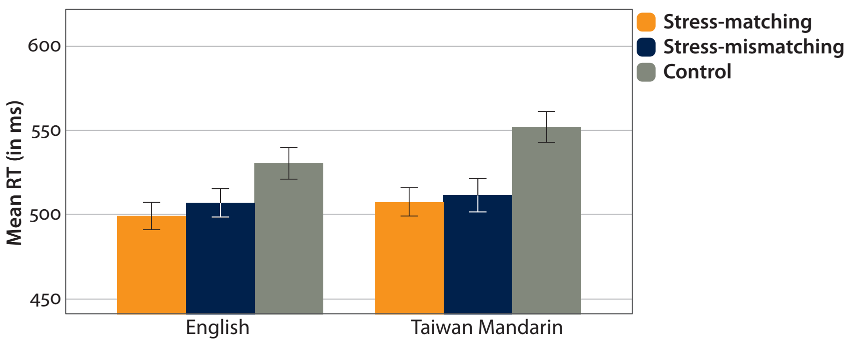

Figure 3. Mean (unlogged) RTs of valid correct YES responses in Experiment II by group and by condition, with SE bars indicating 95% confidence intervals

The above findings, however, cannot be interpreted to mean that suprasegmental information is not useful at all. The contrast between a stressed prime and an unstressed prime is in fact relevant for subjects' English word recognition, as revealed by a further linear mixed-effects regression model fitted to the log-transformed RTs in only the stress-matching and stress-mismatching conditions. The predictors still included *Group*, but this time *Condition* was replaced with two binary factors called *Prime Stress* and *Stress Match*. The former specified whether the monosyllabic prime was stressed (0.5) or not (–0.5), while the latter represented the contrast between stress-matching condition (0.5) and the stress-mismatching (–0.5) one. With the same random-effect structure as the previous one, the selected model contained the main effect of *Prime Stress* as the only fixed effect. Interestingly, the model showed that subjects' responses were significantly slower when the preceding prime was stressed than when it was not ($\beta = .029$, SE(β) = .011, $t = 2.73$, $p < .05$) (see Figure 4). This result was rather unexpected. Nevertheless, as will be seen, it is probably a consequence of an asymmetry between the stressed and unstressed primes in terms of the numbers of English words that can continue the primes. A discussion of the findings from Experiments I and II is given in the next section, with particular focus on those obtained from the second one.

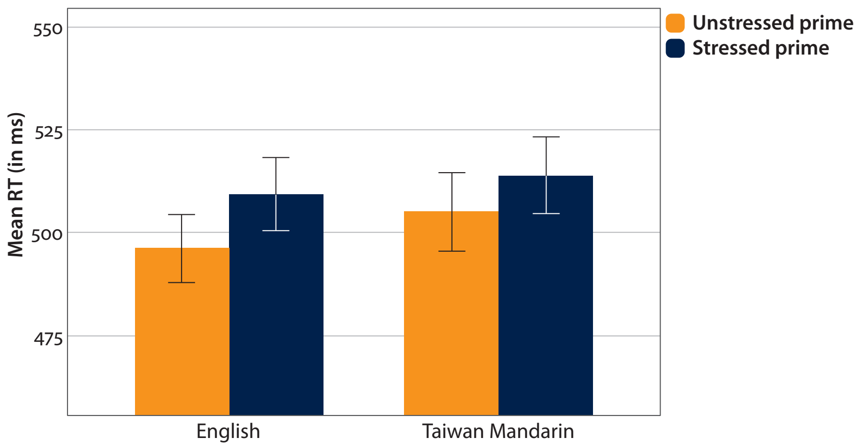

Figure 4. Mean (unlogged) RTs of valid correct YES responses in the stress-matching and mismatching conditions of Experiment, subdivided by group and by prime type condition (i.e. with or without primary stress). The SE bars indicate 95% confidence intervals

5. General discussion and conclusion

A cross-modal fragment priming experiment was conducted to investigate how the suprasegmental information of monosyllables is used in spoken English word recognition by the language's native listeners as well as by non-native listeners whose L1 is TM, a lexical tone language. The finding of Cooper et al.'s (2002) English- and Dutch-speaking subjects is replicated: the monosyllabic primes of both the stress-matching and stress-mismatching conditions produce facilitation in comparison to those of the control condition. The comparable experiment in the current study furnishes evidence that this also holds for listeners of TM, which is typologically distant from English and Dutch and encodes its lexical prosody (i.e. lexical tone) at the syllable level. Such evidence is noteworthy given that listeners of

Mandarin outperform those of English or Japanese in categorizing different pitch patterns realized on monosyllables (Schaefer & Darcy 2014) and tend to equate differences in lexical stress to tonal contrasts (e.g. Cheng 1968; Juffs 1990; Archibald 1997). Nowhere in the results of Experiment II, however, is there any indication that the TM participants make more effective use of the suprasegmental features of English monosyllables than the English or Dutch ones in Cooper et al. (2002). The prediction proposed based on the first hypothesis in § 2 is supported: responses would be facilitated as long as the monosyllabic prime segmentally matches the first syllable of the visual target.

It has to be noted that TM listeners' failure to achieve superior exploitation cannot be attributed to inability to perceive lexical stress contrasts – namely, the "stress deafness" problem, which has been reported for French listeners (e.g. Peperkamp & Dupoux 2002; Dupoux et al. 2008). Experiment I demonstrates that when asked to make on-line lexical decisions involving correctly stressed real words and mis-stressed nonwords, TM listeners can respond as accurately and rapidly as the English ones. The two groups even show the same response bias: both of them are inclined to false alarm to items that are actually nonwords derived by shifting primary stress from the second syllables to the first syllables. There is thus little reason to suspect that TM listeners are less capable of telling a stressed syllable from an unstressed one during on-line word processing (although they may rely heavily on the cue of pitch).

Although their ability to discern differences in lexical stress is intact, listeners of TM, English, and Dutch all find the segmental content of an English monosyllable more useful than its suprasegmental features in spoken word recognition. Such a cross-linguistic parallel would be expected if it is assumed that the strong correlation between vowel quality and lexical stress in English is responsible. As Cutler (1986) explains, English unstressed syllables frequently contain reduced vowels and attention to vowel quality alone suffices to distinguish between stressed and unstressed syllables in most cases. The vocabulary of the language simply does not offer much opportunity to reduce lexical candidates merely on the basis of suprasegmental information. The view is bolstered by statistics garnered by Cutler & Pasveer (2006) and empirically supported by cross-linguistic comparison with Dutch listeners, for whom stress-mismatching monosyllabic primes do not produce facilitation (van Donselaar et al. 2005). Consistent with this view, the findings of Cooper et al. (2002) and Experiment II then converge to suggest that as far as the suprasegmental information of monosyllable is concerned, neither a less strong correlation of vowel reduction and unstressed syllables nor experience with lexical tone would lead to better use of English lexical stress. The segmental cues of single syllables are likely to outweigh the suprasegmental ones in recognition of spoken English words, regardless of the listener's L1 backgrounds.

For TM listeners, however, there is another factor that may have compromised the extent to which the distinction between stressed and unstressed syllables can be useful. As have been mentioned, Mandarin listeners tend to interpret the stressed versus unstressed difference as a tonal contrast – one that is expressed via pitch. However, recall that there are multiple acoustic correlates of lexical stress: in addition to higher pitch, stressed syllables are also cued by longer duration and greater intensity. If pitch is already sufficient for Mandarin listeners' word stress distinction, then the cues of duration and intensity are not quite relevant and perhaps it is their redundancy that removes the benefits that might otherwise be available when there was pitch information alone. That is, while stressed and unstressed syllables are generally treated as belonging to two tonal categories (e.g. [+high] and [−high] tones), their extra durational and amplitudinal differences may have prevented them from being the ideal exemplars of these categories. One consequence of this for spoken word recognition is that the stressed and unstressed primes in Experiment II cannot be exploited so effectively as to inhibit a segmentally matching but suprasegmentally mismatching candidate. Such a view assumes that non-native listeners can somehow perceive acoustic-phonetic features of L2 speech that are not important for phonemic or prosodic distinctions in native languages (e.g. duration for TM listeners). This assumption is not unfounded, however. Previous segmental research has revealed that L2-English listeners (e.g. Japanese listeners) are sensitive to differences in the goodness of fit of two English sounds (e.g. [r] and [l]) that are generally considered to be assimilated into a single native category (e.g. Takagi 1995; Iverson et al. 2003) and that the effect of such sensitivity manifests itself in their English word recognition (e.g. Cutler et al. 2006). As for perception of prosody, Ou (2016) similarly reports that TM-speaking learners of English attend to the fine-grained acoustic differences between tokens of even the same word stress pattern. It would therefore be of interest to examine the priming effects of English monosyllabic primes that have been manipulated to control for duration and intensity.

It may need to be pointed out that the argument just presented is not at odds with the vowel quality account discussed previously. The strong tendency for unstressed syllables to have reduced vowels and TM listeners' sensitivity to the redundant suprasegmental cues (i.e. differences in duration and intensity between stressed and unstressed syllables) may have both come into play, making lexical stress less useful than segments for TM listeners.

However, caution has to be exercised in interpreting the cross-linguistic parallel among TM, English, and Dutch listeners. It would be hasty to claim that segments always outweigh suprasegmentals for all these listeners while it is possible that the participants' ability to exploit lexical stress in Experiment II may have been underestimated due to the use of monosyllabic primes. The information of a single syllable might be too impoverished to permit efficient constraint on lexical access. In fact, when the spoken primes were disyllabic (e.g. *admi-*, excised from *admiration*), Cooper et al. (2002) found that the stress-mismatching condition, compared with the control one, lead to neither facilitation nor inhibition of responses to words such as *admiral* for English listeners (but not for Dutch-speaking ones). One question that arises is then whether TM listeners will show improved ability to use English lexical stress if they are presented with disyllabic primes. Nevertheless, attempts to empirically address this question would be challenged by some practical limitations. For example, there are few English words that have segmentally identical first two syllables but differ only in the stress pattern of these syllables. Even if they can be found, they are not necessarily known to non-native speakers who learn English as a foreign language. In fact, half of the word pairs from which Cooper et al.'s disyllabic primes are derived contain at least one member that does not occur in the 7,000-word list that was used in Experiment II for material selection. In this case, any effects elicited by the auditory primes are potentially spurious.

As far as monosyllables are concerned, it can be concluded that segmental cues are more useful than suprasegmental ones for accessing English words and such a conclusion holds for both English and TM listeners. Yet, a crucial qualification has to be made here: lexical stress is still relevant for spoken word recognition. The second analysis reported in § 4.5 bears on this point; it reveals that there was a significant main effect of *Prime Stress*. As will be explained below, the finding is particularly noteworthy since it is likely to be a reflection of native and non-native listeners' attention to suprasegmental information and also of the structure of the English lexicon.

The main effect of *Prime Stress* indicates that responses were generally slower after the stressed primes than after the unstressed ones. An explanation is obtained when the two types of prime fragments are compared for the numbers and frequencies of the words that can continue them. As has been described, the process of spoken word recognition involves concurrent activation of multiple words supported by the speech input; a word onset fragment would activate a cohort of words that can possibly be its continuations. In Experiment II, fragments like this were the primes, and there is a conspicuous asymmetry between the stressed (e.g. *CAM-*) and unstressed (e.g. *cam-*) primes in terms of the number and spoken frequency of their possible continuations. Specifically, the possible continuations for the former type of primes (e.g. *camp, campus, camping, campfire*, etc.) outnumber and are usually more frequent than those of the latter type (e.g. *campaign*). The asymmetry is confirmed by statistics from the CELEX English database (Baayen et al. 1993). Weighted by spoken frequency per million words, the numbers of all content words that can continue the stressed fragments (mean log value: 6.56 per million words) are generally higher than the numbers of all content words that can continue the corresponding unstressed fragments (mean log value: 5.47 per million words).[7] This can be seen in Figure 5. A stressed syllable thus has the potential of activating a larger cohort of competitors than an unstressed one, and the consequence of this is well-documented in previous research (e.g. Norris et al. 1995; Vroomen & de Gelder 1995; Vitevitch 2007): the more words are activated, the more competition there is. The TM and English listeners tended to respond slower after perceiving the stressed primes presumably because they entertained more competing lexical hypotheses. Such an explanation is also consistent with the fact that most English words begin with stressed syllables (Cutler & Carter 1987) and the response bias that the two participant groups showed in Experiment I.

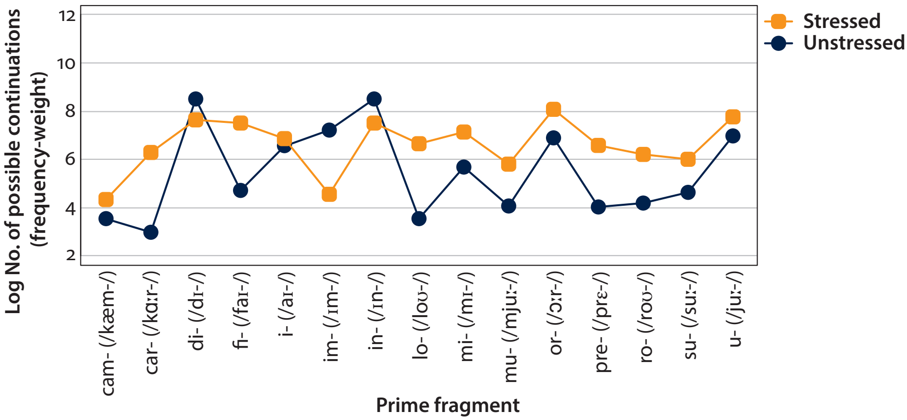

Figure 5. Log numbers of possible continuations (weighted by spoken frequency per million words) for each different pair of stressed and unstressed fragmentary primes in Experiment II, calculated based on data from the CELEX English database

Further evidence for this possible continuation account is furnished by a closer inspection of Figure 5, which reveals that three types of prime fragments are notable exceptions to the overall pattern: /dɪ-/, /ɪm-/, and /ɪn-/. For them, the number of frequency-weighted possible continuations of the unstressed fragments is higher than that of their stressed counterpart. Their anomaly can probably be attributed

7. The count for the unstressed fragments may be an overestimation in the sense that it encompasses all the words without initial primary stress while ignoring the fact that there can be degrees of non-primary stress (e.g. secondary and tertiary stress). However, since only information about the location of primary stress in a word is available from the CELEX database, word-initial syllables with non-primary stress are treated as if they all had the same level of stress. Yet, it can be expected that the difference between the "stressed" and "unstressed" fragments in terms of numbers of possible continuations presented here would be more drastic if finer distinctions are to be drawn among the latter.

to the fact that /dɪ-/, /ɪm-/, and /ɪn-/ happen to be English prefixes and thus can occur more frequently as word-initial syllables than the other non-prefix fragments. However, in words containing these prefixes, especially in those that are relatively long, primary stress tends to fall on syllables other than the initial (e.g. *deforesTA-tion, indispenSAble*, etc.). Therefore, many of the morphologically complex words beginning with /dɪ-/, /ɪm-/, and /ɪn-/ are likely not initially stressed, driving up the numbers of possible continuations for unstressed /dɪ-/, /ɪm-/, and /ɪn-/.

One way to assess the current explanation for the significant main effect of *Prime Stress* is to conduct separate re-analyses for the prefix and non-prefix prime fragments. There were nine pairs of prefix fragments: six beginning with /ɪn-/, two with /dɪ-/, and one with /ɪm-/. The TM and English listeners' log-transformed RTs to visual targets following this subset of items were subjected to a linear mixed-effects model identical to the one used for analyzing their RT data in only the stress-matching and stress-mismatching conditions. The model therefore contained a single fixed-effect for *Prime Stress*, which, however, was not significant in this case ($\beta = .01$, SE(β) = .01, $t = .82$, $p > .10$). While one may expect a significant main effect of *Prime Stress* indicating that responses were slower after the unstressed fragments than after the stressed ones, this finding should not be surprising because for most of the prefix pairs (e.g. the six /ɪn-/ pairs), the number of possible continuations of the unstressed fragment is only slightly higher than that of the stressed one, as can be seen in Figure 5. Yet, a markedly different result was obtained when a model was fitted to log-transformed RTs after the remaining non-prefix prime fragments (e.g. /kæm-/). The model was the same as the previous one except it contained the main effects of *Prime Stress* and *Stress Match* as fixed effects. As expected, the difference between the stressed and unstressed primes was statistically robust: responses were made significantly quicker after the unstressed ones than after the stressed ones ($\beta = .03$, SE(β) = .01, $t = 2.40$, $p < .05$). These two follow-up analyses thus lend well-founded support for the possible continuation account, suggesting that both TM and English listeners are sensitive to the suprasegmental information of the spoken primes.

Although the general tendency for the stressed primes to induce slower lexical decisions can be explained and successfully replicated, it may be an unexpected discovery in view of a study by Tremblay (2008), who has reported the opposite. In her cross-modal word identification experiment, participants (who were native speakers of English or Canadian French) heard a monosyllabic prime (e.g. *MYS-*) and saw two English words (e.g. *mystery* and *mistake*). Their task was to select the word that matched the prime both segmentally and suprasegmentally (in this case, *mystery*). In contrast to what was found in Experiment II, the results showed that responses were generally more rapid after stressed primes (e.g. *MYS-*) than after unstressed ones (e.g. *mys-*). Tremblay explains that recognition of initially stressed words can be facilitated by stressed syllables because such metrically strong syllables, as previous segmentation studies (e.g. Cutler & Norris 1988; Cutler & Butterfield 1992) have suggested, guide English speech segmentation and are points at which lexical search is initiated. While the conflicting findings appear puzzling, it has to be noted that there are appreciable differences between the cross-modal experiment of Tremblay and that of the current study. What Tremblay's participants completed was a selection task that presented two visual targets and the information of the preceding prime was crucial for making correct responses. In contrast, the participants in the present Experiment II (and also in Cooper et al. (2002) or in van Donselaar et al. (2005)) decided on the word status of a single visual target and the prime was irrelevant to this decision. It thus cannot be conclusively determined whether or not the discrepancy in findings is merely an artifact of the difference in visual stimuli. However, if the findings are to be reconciled, then one explanation may be that Tremblay's subjects had to pay focal attention to the detailed phonetic differences between stressed and unstressed primes in order to respond correctly. In this way, the acoustically more prominent stressed primes would lead to faster recognition simply because they were more audible.

Based on the above discussion, it is concluded that the findings support the prediction of the first hypothesis – that Mandarin listeners do not show better exploitation of the suprasegmental information of English words than English-speaking ones do. This cannot be attributed to a lack of ability to discern differences in lexical stress, as Experiment I revealed that TM listeners are as rapid and accurate in telling correctly stressed English words from mis-stressed nonwords. Yet, Experiment II demonstrates that while they still use suprasegmental information of the monosyllabic primes, the information is not effective enough to rule out stress-mismatching but segmentally matching lexical candidates during spoken word recognition. Exactly the same pattern is found for the language's native listeners and another group of non-native listeners with a typologically similar L1 (i.e. Dutch). Therefore, as far as the use of single syllables in constraining lexical access is concerned, attention to segmental cues seems to yield a greater payoff than attention to suprasegmental cues and this may be an experience universal to all listeners to English. Briefly discussed below are two further issues related to the cross-linguistic use of lexical stress in spoken word recognition.

Although the correlation between unstressed syllables and reduced vowels in English is argued to be an explanation for the cross-linguistic parallel observed in Cooper et al. (2002) and in this study, it cannot be precluded that the TM participants' failure to effectively exploit stress cues stems from experience with their own native language. As has been mentioned, segmental information seems to constrain Mandarin lexical access more effectively than suprasegmental (tonal) information (Lee 2007), and such a tendency may simply extend to English spoken word recognition. One way to assess this possibility is to investigate Mandarin-speaking learners of Dutch in an analogous priming experiment using Dutch material. If the native language experience is responsible, the response patterns of Experiment II are expected to be replicated. But if the correlation between vowel quality and lexical stress in the L2 is responsible, then the learners should pattern with the native Dutch participants in van Donselaar et al. (2005), showing no facilitation and perhaps also no inhibition when their responses in the stress-mismatching and control conditions are compared.

The other issue has to do with exploring the possibility that the stressed versus unstressed syllable distinction has been interpreted by non-native listeners as a contrast in suprasegmental dimensions other than pitch and thereby used to facilitate English word recognition. This cannot be addressed with currently available data as the stressed primes in Experiment II were simultaneously cued by higher pitch, longer duration, and stronger intensity. However, it is possible to conceive an experimentation in which cues to stress difference no longer include pitch height. An example is (2), where the first syllable of *CAMpus* or *camPAIGN* is placed at the end of a question sentence, which has a rising pitch accent context, and used as the auditory prime:

(2) Is the word *CAM-* ? (*CAMpus* or *camPAIGN* with the second syllable
cam- ? truncated)

In this context, duration is one cue for determining whether the trailing syllable is stressed or not. Since Japanese, for example, has a length contrast for vowels and intervocalic obstruents, it is of interest to see how listeners of this language exploit the purely durational difference between the *CAM-* and *cam-* fragments in (3) in English word recognition. Perhaps it can be expected that these listeners would outperform those from languages that do not draw durational contrasts, such as Mandarin. Support for the expectation is the finding that Japanese listeners are better than Mandarin ones at using the duration of a vowel as a cue to the voicing of syllable-final consonants in English words (e.g. *pot* vs. *pod*) (Crowther & Mann 1992). Questions like this can be pursued to gain a deeper understanding of the exact way by which lexical stress subserves non-native listeners' recognition of spoken English words.

Acknowledgements

I would like to thank two anonymous reviewers for their comments and constructive suggestions on a previous draft of this paper. I am also grateful to Zhe-chen Adam Guo for his editorial assistance. I am solely responsible for any errors which may remain. Finally, my thanks go to the Ministry of Science and Technology (MOST), Taiwan. The research reported in this paper was supported by a MOST grant (grant no.: 102-2410-H-110-014-MY2).

Appendix A

List of the word-nonword stress minimal pairs of Experiment I. Syllables with primary stress are indicated by uppercase letters and nonwords by asterisks (*).

beCOME-*BEcome	camPAIGN-*CAMpaign	carTOON-*CARtoon
COUNtry-*counTRY	creATE-*CREate	DIfferent-*diFFERENT
exPECT-*EXpect	HAppen-*haPPEN	hoTEL-*HOtel
iDEA-*Idea	iDEAL-*Ideal	inCLUDE-*INclude
LIttle-*liTTLE	mainTAIN-*MAINtain	misTAKE-*MIStake
MOney-*moNEY	MOther-*moTHER	muSEUM-*MUseum
NUMber-*numBER	PARty-*parTY	PEOple-*peoPLE
PROblem-*proBLEM	QUEstion-*queSTION	SOcial-*soCIAL
techNIQUE-*TECHnique	uNIQUE-*Unique	WAter-*waTER
WOman-*woMAN		

Appendix B

List of the 21 pairs of experimental items of Experiment II, with the control items in parentheses.

campus – campaign (mountain)	carton – cartoon (succeed)
distance – distinct (measure)	district – disturb (pocket)
final – finance (passion)	idle – ideal (artist)
impact – imply (advice)	index – induce (option)
infant – infect (absorb)	injure – inject (expert)
insight – install (access)	instance – instead (admit)
interest – intent (observe)	local – locate (yellow)
missing – mistake (bedroom)	music – museum (region)
order – ordeal (empty)	pressure – prestige (climate)
robot – robust (survey)	super – superb (pursue)
unit – unite (belief)	

References

Altmann, Heidi. 2006. *The perception and production of second language stress: A cross-linguistic experimental study*. Newark: University of Delaware. (Doctoral dissertation.)

Altmann, Heidi & Vogel, Irene. 2002. L2 acquisition of stress: The role of L1. (Paper presented at the DGfS Annual Meeting, Mannheim, 27 February-1 March 2002.)

Archibald, John. 1997. The acquisition of English stress by speakers of nonaccentual languages: Lexical storage versus computation of stress. *Linguistics* 35(1). 167–181. https://doi.org/10.1515/ling.1997.35.1.167

Baayen, R. Harald & Piepenbrock, Richard & van Rijn, Hedderik. 1993. *The CELEX database (CD-ROM)*. Philadelphia: Linguistic Data Consortium.

Bates, Douglas & Mächler, Martin & Bolker, Ben & Walker, Steve. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67(1). 1–48. https://doi.org/10.18637/jss.v067.i01

Best, Catherine T. 1994. The emergence of native-language phonological influence in infants: A perceptual assimilation model. In Goodman, Judith C. & Nusbaum, Howard C. (eds.), *The transition from speech sounds to spoken words: The development of speech perception*, 167–224. Cambridge: The MIT Press.

Boersma, Paul & Weenink, David. 2016. Praat: Doing phonetics by computer (Version 6.0.14). (http://www.praat.org/) (Accessed 2016-02-11).

Bond, Zinny S. & Small, Larry H. 1983. Voicing, vowel, and stress mispronunciations in continuous speech. *Perception & Psychophysics* 34(5). 470–474. https://doi.org/10.3758/BF03203063

Booij, Geert E. 1995. *The phonology of Dutch*. Oxford: Clarendon Press.

Broersma, Mirjam & Cutler, Anne. 2008. Phantom word activation in L2. *System* 36(1). 22–34. https://doi.org/10.1016/j.system.2007.11.003

Broersma, Mirjam & Cutler, Anne. 2011. Competition dynamics of second-language listening. *The Quarterly Journal of Experimental Psychology* 64(1). 74–95. https://doi.org/10.1080/17470218.2010.499174

Chao, Yuen-ren. 1968. *A grammar of spoken Chinese*. Berkeley: University of California Press.

Cheng, Chin-chuan. 1968. English stresses and Chinese tones in Chinese sentences. *Phonetica* 18(2). 77–88. https://doi.org/10.1159/000258601

Chomsky, Noam & Halle, Morris. 1968. *The sound pattern of English*. New York: Harper & Row.

Chrabaszcz, Anna & Winn, Matthew B. & Lin, Candise Y. & Idsardi, William J. 2014. Acoustic cues to perception of word stress by English, Mandarin, and Russian speakers. *Journal of Speech, Language, and Hearing Research* 57(4). 1468–1479. https://doi.org/10.1044/2014_JSLHR-L-13-0279

Connine, Cynthia M. & Blasko, Dawn G. & Wang, Jian. 1994. Vertical similarity in spoken word recognition: Multiple lexical activation, individual differences, and the role of sentence context. *Perception and Psychophysics* 56(6). 624–636. https://doi.org/10.3758/BF03208356

Cooper, Nicole & Cutler, Anne & Wales, Roger. 2002. Constraints of lexical stress on lexical access in English: Evidence from native and non-native listeners. *Language and Speech* 45(3). 207–228. https://doi.org/10.1177/00238309020450030101

Crowther, Court S. & Mann, Virginia. 1992. Native language factors affecting use of vocalic cues to final consonant voicing in English. *The Journal of the Acoustical Society of America* 92(2). 711–722. https://doi.org/10.1121/1.403996

Cutler, Anne. 1986. Forbear is a homophone: Lexical prosody does not constrain lexical access. *Language and Speech* 29(3). 201–220. https://doi.org/10.1177/002383098602900302

Cutler, Anne & Butterfield, Sally. 1992. Rhythmic cues to speech segmentation: Evidence from juncture misperception. *Journal of Memory and Language* 31(2). 218–236. https://doi.org/10.1016/0749-596X(92)90012-M

Cutler, Anne & Carter, David M. 1987. The predominance of strong initial syllables in the English vocabulary. *Computer Speech & Language* 2(3). 133–142. https://doi.org/10.1016/0885-2308(87)90004-0

Cutler, Anne & Chen, Hsuan-chih. 1997. Lexical tone in Cantonese spoken-word processing. *Perception and Psychophysics* 59(2). 165–179. https://doi.org/10.3758/BF03211886

Cutler, Anne & Norris, Dennis. 1988. The role of strong syllables in segmentation for lexical access. *Journal of Experimental Psychology: Human Perception and Performance* 14(1). 113–121.

Cutler, Anne & Otake, Takashi. 2004. Pseudo-homophony in non-native listening. *The Journal of the Acoustical Society of America* 115(5). 2392. https://doi.org/10.1121/1.4780547

Cutler, Anne & Pasveer, Dennis. 2006. Explaining cross-linguistic differences in effects of lexical stress on spoken-word recognition. In Hoffman, Rüdiger & Mixdorff, Hansjörg (eds.), *Proceedings of the Third International Conference on Speech Prosody*, 250–254. Dresden: TUD press.

Cutler, Anne & Wales, Roger & Cooper, Nicole & Janssen, Joris. 2007. Dutch listeners' use of suprasegmental cues to English stress. In Barry, William J. & Trouvain, Jürgen (eds.), *Proceedings of the 16th International Congress of Phonetic Sciences* (ICPhS XVI), 1913–1916. Saarbrücken: Universität des Saarlandes.

Cutler, Anne & Weber, Andrea & Otake, Takashi. 2006. Asymmetric mapping from phonetic to lexical representations in second-language listening. *Journal of Phonetics* 34(2). 269–284. https://doi.org/10.1016/j.wocn.2005.06.002

Dupoux, Emmanuel & Pallier, Christophe & Sebastián-Gallés, Núria & Mehler, Jacques. 1997. A destressing "deafness" in French? *Journal of Memory and Language* 36(3). 406–421. https://doi.org/10.1006/jmla.1996.2500

Dupoux, Emmanuel & Sebastián-Gallés, Núria & Navarrete, Eduardo & Peperkamp, Sharon. 2008. Persistent stress "deafness": The case of French learners of Spanish. *Cognition* 106(2). 682–706. https://doi.org/10.1016/j.cognition.2007.04.001

Fear, Beverley D. & Cutler, Anne & Butterfield, Sally. 1995. The strong/weak syllable distinction in English. *The Journal of the Acoustical Society of America* 97(3). 1893–1904. https://doi.org/10.1121/1.412063

Flege, James E. 1995. Second language speech learning: Theory, findings, and problems. In Strange, Winifred (ed.), *Speech perception and linguistic experience: Issues in cross-language research*, 233–277. Timonium: York Press.

Fry, Dennis B. 1955. Duration and intensity as physical correlates of linguistic stress. *The Journal of the Acoustical Society of America* 27(4). 765–768. https://doi.org/10.1121/1.1908022

Gaskell, M. Gareth & Marslen-Wilson, William D. 2002. Representation and competition in the perception of spoken words. *Cognitive Psychology* 45(2). 220–266. https://doi.org/10.1016/S0010-0285(02)00003-8

Gay, Thomas. 1978. Physiological and acoustic correlates of perceived stress. *Language and Speech* 21(4). 347–353. https://doi.org/10.1177/002383097802100409

Goldinger, Stephen D. & Luce, Paul A. & Pisoni, David B. 1989. Priming lexical neighbors of spoken words: Effects of competition and inhibition. *Journal of Memory and Language* 28(5). 501–518. https://doi.org/10.1016/0749-596X(89)90009-0

Iverson, Paul & Kuhl, Patricia K. & Akahane-Yamada, Reiko & Diesch, Eugen & Tohkura, Yoh'ich & Kettermann, Andreas & Siebert, Claudia. 2003. A perceptual interference account of acquisition difficulties for non-native phonemes. *Cognition* 87(1). B47–B57. https://doi.org/10.1016/S0010-0277(02)00198-1

Juffs, Alan. 1990. Tone, syllable structure and interlanguage phonology: Chinese learners' stress errors. *International Review of Applied Linguistics in Language Teaching* 28(2). 99–117. https://doi.org/10.1515/iral.1990.28.2.99

Kochanski, Greg & Orphanidou, Christina. 2008. What marks the beat of speech? *The Journal of the Acoustical Society of America* 123(5). 2780–2791. https://doi.org/10.1121/1.2890742

Lee, Chao-yang. 2007. Does horse activate mother? Processing lexical tone in form priming. *Language and Speech* 50(1). 101–123. https://doi.org/10.1177/00238309070500010501

Lehiste, Ilse. 1972. The timing of utterances and linguistic boundaries. *The Journal of the Acoustical Society of America* 51(6B). 2018–2024. https://doi.org/10.1121/1.1913062

Luce, R. Duncan. 1986. *Response times: Their role in inferring elementary mental organization*. New York: Oxford University Press.

Lukyanchenko, Anna & Idsardi, William J. & Jiang, Nan. 2011. Opening your ears: The role of L1 in processing of nonnative prosodic contrasts. In Granena, Gisela & Koeth, Joel & Lee-Ellis, Sunyoung & Lukyanchenko, Anna & Prieto Botana, Goretti & Rhoades, Elizabeth (eds.), *Selected proceedings of the Second Language Research Forum 2010*, 50–62. Somerville: Cascadilla Press.

Marslen-Wilson, William D. 1980. Speech understanding as a psychological process. In Simon, Jean-Claude (ed.), *Spoken Language Generation and Understanding*, 39–67. Dordrecht: D. Reidel. https://doi.org/10.1007/978-94-009-9091-3_2

Marslen-Wilson, William D. 1990. Activation, competition, and frequency in lexical access. In Altmann, Gerry T. M. (ed.), *Cognitive models of speech processing: Psycholinguistic and computational perspectives*, 148–172. Cambridge: The MIT Press.

Marslen-Wilson, William D. & Warren, Paul. 1994. Levels of perceptual representation and process in lexical access: Words, phonemes, and features. *Psychological Review* 101(4). 653–675. https://doi.org/10.1037/0033-295X.101.4.653

Marslen-Wilson, William D. & Zwitserlood, Pienie. 1989. Accessing spoken words: The importance of word onsets. *Journal of Experimental Psychology: Human Perception and Performance* 15(3). 576–585.

McQueen, James M. & Norris, Dennis & Cutler, Anne. 1999. Lexical influence in phonetic decision making: Evidence from subcategorical mismatches. *Journal of Experimental Psychology: Human Perception and Performance* 25(5). 1363–1389.

Nakatani, Lloyd H. & Dukes, Kathleen D. 1977. Locus of segmental cues for word juncture. *The Journal of the Acoustical Society of America* 62(3). 714–719. https://doi.org/10.1121/1.381583

Norris, Dennis & McQueen, James M. & Cutler, Anne. 1995. Competition and segmentation in spoken-word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 21(5). 1209–1228.

Ou, Shu-chen. 2010. Taiwanese EFL learners' perception of English word stress. *Concentric: Studies in Linguistics* 36(1). 1–23.

Ou, Shu-chen. 2016. Perception of English lexical stress with a marked pitch accent by native speakers of Mandarin. *Taiwan Journal of Linguistics* 14(2). 1–31.

Peperkamp, Sharon. 1997. Prosodic words (HIL Dissertations 34). The Hague: Holland Academic Graphics.

Peperkamp, Sharon & Dupoux, Emmanuel. 2002. A typological study of stress "deafness". *Laboratory Phonology* 7. 203–240.

Psychology Software Tools. 2012. E-Prime 2.0. (Computer Software) Pittsburgh: Author.

R Core Team. 2015. *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing. (https://www.R-project.org/).

Rastle, Kathleen & Harrington, Jonathan & Coltheart, Max. 2002. 358, 534 nonwords: The ARC nonword database. *The Quarterly Journal of Experimental Psychology: Section A* 55(4). 1339–1362. https://doi.org/10.1080/02724980244000099

Schaefer, Vance & Darcy, Isabelle. 2014. Lexical function of pitch in the first language shapes cross-linguistic perception of Thai tones. *Laboratory Phonology* 5(4). 489–522. https://doi.org/10.1515/lp-2014-0016

Slowiaczek, Louisa M. 1990. Effects of lexical stress in auditory word recognition. *Language and Speech* 33(1). 47–68. https://doi.org/10.1177/002383099003300104

Small, Larry H. & Simon, Stephen D. & Goldberg, Jill S. 1988. Lexical stress and lexical access: Homographs and non-homographs. *Perception and Psychophysics* 44(3). 272–280. https://doi.org/10.3758/BF03206295

So, Connie K. & Best, Catherine T. 2010. Cross-language perception of non-native tonal contrasts: Effects of native phonological and phonetic influences. *Language and Speech* 53(2). 273–293. https://doi.org/10.1177/0023830909357156

Soto-Faraco, Salvador & Sebastián-Gallés, Núria & Cutler, Anne. 2001. Segmental and suprasegmental mismatch in lexical access. *Journal of Memory and Language* 45(3). 412–432. https://doi.org/10.1006/jmla.2000.2783

Swinney, David A. 1979. Lexical access during sentence comprehension: (Re)consideration of context effects. *Journal of Verbal Learning and Verbal Behavior* 18(6). 645–659. https://doi.org/10.1016/S0022-5371(79)90355-4

Takagi, Naoyuki. 1995. Signal detection modeling of Japanese listeners' /r/–/l/ labeling behavior in a one-interval identification task. *The Journal of the Acoustical Society of America* 97(1). 563–574. https://doi.org/10.1121/1.413059

Tremblay, Annie. 2008. Is second language lexical access prosodically constrained? Processing of word stress by French Canadian second language learners of English. *Applied Psycholinguistics* 29(4). 553–584. https://doi.org/10.1017/S0142716408080247

van Donselaar, Wilma & Koster, Mariëtte & Cutler, Anne. 2005. Exploring the role of lexical stress in lexical recognition. *The Quarterly Journal of Experimental Psychology Section A* 58(2). 251–273. https://doi.org/10.1080/02724980343000927

Van Heuven, Vincent J. 1985. Perception of stress pattern and word recognition: Recognition of Dutch words with incorrect stress position. *Journal of the Acoustical Society of America* 78. S21. https://doi.org/10.1121/1.2022696

Van Leyden, Klaske & van Heuven, Vincent J. 1996. Lexical stress and spoken word recognition: Dutch vs. English. *Linguistics in the Netherlands* 13(1). 159–170.

Vitevitch, Michael S. 2007. The spread of the phonological neighborhood influences spoken word recognition. *Memory & Cognition* 35(1). 166–175. https://doi.org/10.3758/BF03195952

Vroomen, Jean & de Gelder, Béatrice. 1995. Metrical segmentation and lexical inhibition in spoken word recognition. *Journal of Experimental Psychology: Human Perception and Performance* 21(1). 98–108.

Wang, Qian. 2008a. *Perception of English stress by Mandarin Chinese learners of English: An acoustic study*. Victoria: University of Victoria. (Doctoral dissertation.)

Wang, Qian. 2008b. L2 stress perception: The reliance on different acoustic cues. In Barbosa, Plinio A. & Madureira, Sandra & Reis, César (eds.), *Proceedings of the 4th International Conference on Speech Prosody*, 635–638.

Woodward, Julide Z. & Aslin, Richard N. 1990. Segmentation cues in maternal speech to infants. (Paper presented at the 7th biennial meeting of the International Conference on Infant Studies, Montreal, April 1990.)

Zwitserlood, Pienie. 1989. The locus of the effects of sentential-semantic context in spoken-word processing. *Cognition* 32(1). 25–64. https://doi.org/10.1016/0010-0277(89)90013-9

Author's address

Shu-chen Ou
Department of Foreign Languages and Literature
National Sun Yat-sen University
No. 70, Lien-Hai Road
Kaohsiung City
Taiwan (R. O. C.)
sherryou@mail.nsysu.edu.tw

Publication history

Date received: 5 April 2017
Date accepted: 15 July 2017