J Psycholinguist Res (2010) 39:323–344
DOI 10.1007/s10936-009-9142-2

A Cross-Language Study of Perception of Lexical Stress in English

Vickie Y. Yu · Jean E. Andruski

Abstract This study investigates the question of whether language background affects the perception of lexical stress in English. Thirty native English speakers and 30 native Chinese learners of English participated in a stressed-syllable identification task and a discrimination task involving three types of stimuli (real words/pseudowords/hums). The results show that both language groups were able to identify and discriminate stress patterns. Lexical and segmental information affected the English and Chinese speakers in varying degrees. English and Chinese speakers showed different response patterns to trochaic vs. iambic stress across the three types of stimuli. An acoustic analysis revealed that two language groups used different acoustic cues to process lexical stress. The findings suggest that the different degrees of lexical and segmental effects can be explained by language background, which in turn supports the hypothesis that language background affects the perception of lexical stress in English.

Keywords Lexical stress · Speech perception · Linguistic experience

Introduction

The way in which speech sounds are perceived depends on the nature of listener's language experience. Considerable attention has been devoted to increasing our understanding of the difficulties which second language learners encounter as they try to learn nonnative segmental contrasts. Research in cross-language speech perception has demonstrated that adult native speakers and infants who have gained some familiarity with their native language have difficulty perceptually differentiating phonetic contrasts that are not distinctive in their native language (e.g., Best et al. 1988; Jusczyk 1993; Nusbaum and Goodman 1994; Polka and Werker 1994; Kuhl 2000; Goto 1971; Miyakawa et al. 1981). A typical example is that

---

V. Y. Yu (✉) · J. E. Andruski
Department of Communication Sciences and Disorders, Wayne State University, 207 Rackham Bldg,
60 Farnsworth, Detroit, MI 48202, USA
e-mail: vickie.yu@wayne.edu

Japanese listeners persistently have difficulty discriminating the English /r/-/l/ contrast due to the lack of the /l/ phoneme in their phonetic inventory (Goto 1971; Miyakawa et al. 1981). Similarly, consistent findings have been published in the areas of cross-language and second language acquisition (L2A) studies (e.g., Hume et al. 1999; Best 1994, 1995; Flege 1995; Best et al. 2001.). In general, L2 learners have difficulty discriminating nonnative speech segments and their language background affects how those nonnative speech sounds is processed.

In addition to differences in the repertoire of phonemes, languages differ in their suprasegmental properties. Although some languages do not use suprasegmental properties to distinguish lexical items (e.g., French), the majority appear to use at least one suprasegmental property to this end. The result is that, just as languages divide segmental space differently, they also divide suprasegmental space differently. For example, some languages use tones to distinguish meaning (Mandarin), while others use pitch accent (Japanese, Norwegian), or duration (Finnish), to make lexical distinctions. Studies in this regard have shown that listeners' native language affects the way they perceive nonnative suprasegmental information (Lehiste and Fox 1992; Dupoux et al. 1997; Peperkamp and Dupoux 2002). Lehiste and Fox (1992) manipulated duration and amplitude independently in a sequence of reiterant speech. English and Estonian listeners gave significantly different answers when asked which syllable in a sequence was the most prominent. English listeners were more sensitive to amplitude cues, whereas Estonian listeners were more sensitive to duration cues. The authors suggested that this difference might reflect language background. Estonian is a quantity-sensitive language, therefore the Estonian listeners would presumably rely more on duration cues.

A cross-language study conducted by Dupoux et al. (1997) found that native speakers of French, a language with fixed word-final stress, have difficulties discriminating nonwords that differ only in the position of stress (e.g., [vaˈsuma] vs. [vasuˈma] vs. [vasumaˈ]). In contrast, native speakers of Spanish do not have difficulties with this task, since stress is contrastive in their language. On the basis of this finding, the authors argue that French listeners are "deaf" to stress contrasts because French, unlike Spanish, does not have lexical stress. Subsequently, Peperkamp and Dupoux (2002) proposed a typology of stress-deafness by testing stress perception in adult speakers of several languages: French, Finnish, Hungarian, and Polish. Speakers of some languages showed more robust "stress deafness" effects than did speakers of other languages. They found that French speakers exhibited the strongest effect of stress deafness among these four languages, since French is a non-stress language, whereas Spanish speakers had significantly lower scores than other languages on the stress deafness index, since Spanish is a stress language like English.

Additionally, L2A studies have revealed consistently different patterns for learning English lexical prosody, depending on native language background (Archibald 1992, 1993, 1997; Wayland et al. 2006). In a series of L2A studies, Archibald (1992, 1993, 1997) found that linguistic experience influences not only speech perception but also speech production. He concluded that speakers of stress languages are more likely to show patterned stress behavior than speakers of non-stress languages. In his studies, he proposed that the errors of stress placemen produced by native Polish speakers (Archibald 1992) and native Spanish speakers (Archibald 1993) were due to the transfer of their own native language systems (i.e., first language, L1). In his 1992 study, Archibald investigated the acquisition of English stress patterns by examining adult native Polish speakers' productions and perceptions of English (L2) stress patterns. The production task was to read English words in isolation and in sentences. The perception task involved an identification paradigm where participants listened to English words and identified the stress placement. He observed that Polish speakers transferred their L1 metrical stress patterns (i.e., primary stress always falls on the penult) to the production and perception of two-syllable English words. For example, for English words, such as ‘mainTAIN,’ and ‘apPEAR,’ Polish learners of English tended to produce as ‘MAINtain,’ and ‘APpear,’ respectively. In a later study, Archibald (1993) found that, as was the case with Polish learners of English, Spanish learners of English transferred the stress patterns of their L1 when producing and perceiving English words. Like English, Polish and Spanish are both stress languages. Learners of English from both languages demonstrated stress pattern behavior in which their performance patterns showed consistent influence from their native language.

In contrast, different results were found for learners of English whose native language is not a stress language. Archibald (1997) had Chinese and Japanese learners of English participate in his study with the same task paradigm and stimuli as his previous studies (1992, 1993). Chinese is a tone language and Japanese is a pitch-accent language, therefore both are non-stress languages. Archibald found that both language groups had difficulties placing stress correctly in English words and the errors they made did not have any readily discernable pattern.

This observation was explained later by Wayland et al. (2006). Instead of only looking at L2 learners’ production and perception of two-syllable English words, they also examined the influence of syllabic structure, lexical class and stress pattern of known words on the acquisition of the English stress system. Ten native Thai learners of English were asked to produce and give perceptual judgments on 40 English nonwords of varying syllabic structures in noun and verb sentence frames (i.e., ‘I’d like a ___,’ ‘I’d like to ___’). In the production task, participants were asked to say each nonword in both sentence frames. Their production data were coded for first or second syllable stress by a trained phonetician and a native English speaker. In the perceptual task, the same 40 nonwords were produced with stress on the initial and final syllable in the two carrier frames ‘I’d like a ____’ and ‘I’d like to ____.’ Participants were asked to listen to the prerecorded phrases in pairs that varied only in the stress placement on the nonwords (e.g., ‘I’d like a TOOkips’ vs. ‘I’d like a tooKIPS’). They were instructed to listen to the two sentences and indicate which sentence sounded most like a real English sentence. The results for both the production and perceptual tasks showed that participants’ performance was influenced by their native language. Among three explanatory factors they examined (syllabic structure, lexical class and stress patterns of known words), Thai learners’ patterns of stress assignment on nonwords were significantly influenced by the stress patterns of phonologically similar known words. The authors concluded that speakers of non-stress languages may rely more heavily on word-by-word learning of stress patterns and are less likely to abstract generalities about stress placement using syllabic structure or lexical class than speakers of stress languages, since tone is a lexical property and thus has to be acquired item by item. It might be reasonable to assume that native speakers of other tonal languages such as Mandarin Chinese, Cantonese and Vietnamese would use a similar approach when acquiring the English stress system. Importantly, these authors also pointed out that unlike the production task, in perceptual judgments, Thai participants appeared to prefer final stress over initial stress regardless of syllable structure or lexical class. They suggested that individual variation might provide the best explanation for this, since the data could not be easily explained by any of the participants’ language background.

Within the field of suprasegmental perception, researchers have studied how stress cues facilitate speech perception and segmentation and the importance of lexical stress in language processes (e.g., Cutler and Norris 1988; Mattys and Samuel 1997, 2000; Baum 1998; Emmorey 1987; Baum and Pell 1999; Shah and Baum 2006). An important gap remains, however, when it comes to assessing whether listeners can identify stress placement with or without minimal contextual information. It would be pointless to elaborate theories of speech segmentation if there is no evidence to demonstrate that listeners have the perceptual ability to discriminate different stress patterns in words. In the current study, we will examine whether native and non-native English speakers can distinguish stress placement in English words and moreover, address the question of whether language background affects the perception of lexical stress in English. To investigate the role of native language and linguistic experience, native speakers of Mandarin Chinese and American English were chosen as participants. The study uses two experimental tests to address the research question of whether language background affects perception of lexical stress in English. The first experiment investigates the extent to which minimal contextual information (lexical and segmental information) affect lexical stress processing for native speakers of English and Mandarin Chinese. The results of this study are also used to investigate which acoustic cues listeners rely on to perceive lexical stress in English. The second experiment explores whether listeners from these two language backgrounds are able to discriminate lexical stress patterns with or without minimal contextual information.

Experiment 1

In English, words are not evenly distributed across syllable-stress patterns. In a corpus of 20,000 English words used by Cutler and Carter (1987), 90% of content words begin with a stressed syllable. Based on the CELEX database (Baayen, Piepenbrock, & van Rijn 1993), a computational analysis of English word stress by number of syllables reveals that the majority of two-syllable English words have primary stress on the first syllable (hereafter, trochaic stress). The actual proportion of words count is about 74.94%, or 72.97% if frequency of two-syllable words is taken into account. The remaining words have primary stress on the second syllable (hereafter, iambic stress) (Yu 2008). Native English listeners seem sensitive to this uneven distribution of stress patterns. Within the field of suprasegmental perception, researchers have studied how stress cues facilitate speech perception and segmentation and the importance of lexical stress in language processes (e.g., Cutler and Norris 1988; Mattys and Samuel 1997, 2000; Baum 1998; Emmorey 1987; Baum and Pell 1999; Shah and Baum 2006; Cutler and Butterfield 1992; Natkatani and Schaffer 1978; Nooteboom et al. 1978).

Research in child language development also provides evidence that English-learning infants prefer trochaic stress and use it as a cue for word segmentation (Jusczyk and Aslin 1995; Jusczyk et al. 1999; Thiessen and Saffran 2003). Using an artificial language consisting of two-syllable words with trochaic and iambic stress patterns, Thiessen and Saffran found that 9-month-old English infants relied on trochaic patterns as a cue to word segmentation, whereas both 6.5- and 9-month-old infants mis-segmented the words when the stress pattern was iambic.

Thus, based on previous studies, native English listeners use stress for speech segmentation, are sensitive to the uneven distribution of stress patterns of English words, and show particular sensitivity to trochaic stress patterns. If these phenomena are due to language background, can native Chinese listeners be expected to have similar sensitivity to lexical stress in English? Mandarin Chinese differs from English in ways that are particularly interesting for this study. In English, vocal pitch is one of several possible cues to the presence of a stressed syllable. Others include duration, loudness, and vowel quality. In Mandarin Chinese, however, pitch is the primary cue to lexical tone rather than stress. The meaning of a given syllable such as ‘ma’ depends on its tone, or pitch contour. Thus, one important question we will investigate here is whether pitch is of equal importance in perceiving stress for native speakers of tonal versus non-tonal languages. A second question is whether native language affects the degree of reliance on pitch versus loudness and duration as cues to stress. We will also examine whether vowel quality provides a cue for native English and Chinese speakers in perceiving English stress.

In summary, three research questions are addressed in this experiment: (1) Can native English and native Chinese speakers distinguish stress placement in two-syllable English words?, (2) Do lexical and segmental information facilitate or inhibit stress perception in native English and Chinese speakers, and (3) Which acoustic attributes (pitch, duration, intensity, vowel quality) of lexical stress do native English and Chinese speakers rely on to identify stress?

Methods

Participants

Two groups of participants were recruited at Wayne State University: 30 native speakers of English and 30 native speakers of Mandarin Chinese. All participants had normal hearing with no known history of auditory deficiencies. None had linguistic or phonetic training in the past 5 years. Participants from both groups had received at least a Bachelor's degree. Native Chinese participants originated from Mainland China and had Mandarin Chinese as their first language. All Chinese participants began learning English as their second language after 10 years of age and their length of residence in the United States was less than 2 years. All Chinese participants were interviewed about their background in learning English as L2. This information is summarized in Table 1.

Materials

Stimuli consisted of three types of sounds: real words, pseudowords, and hums. All stimuli contained two-syllables. The grammatical categories (nouns, verbs) and stress patterns (trochaic, iambic) of the real word stimuli were counterbalanced. Pseudoword stimuli were also counterbalanced by stress patterns. The real word stimuli were closely matched in terms of word frequency. Word frequency was estimated based on the Francis and Kucera (1982) database. A one-way ANOVA revealed no statistically significant differences in the frequencies of the real word stimuli across stress patterns ($F < 1$). A paired t-test showed no statistically significant mean differences in syllable weight between the first and second syllables of pseudowords ($p = .829$). Recordings of the real words and pseudowords were made in a soundproof booth using a SHURE SM58 microphone and a TASCAM DA-P1 digital audio recorder and a sampling rate of 48 kHz. All the real word and pseudoword stimuli were produced by a native male English speaker. All recorded stimuli were produced at a normal

Table 1 Chinese participants' background of learning English as L2

	Age	AOA	LOR/mo	DEU	Listen/TOEFL[a]
Mean	27.1	12.3	10.4	33.1%	81.3%
StDv	3.9	2.2	11	0.24	5.3

StDv standard deviation, *AOA* age of acquisition of learning L2, *LOR* length of residence in the U.S. (mo = in months), *DEU* daily English usage

[a] Scores of Listening Comprehension of TOEFL by ETS. This average score was converted to percentage. The original full score of Listening Comprehension was 68

pace of speech with falling intonation. Each stimulus was segmented using Cool Edit Pro 2000. The starting and ending points for each segmental stimulus were at zero-crossings in order to prevent clipping artifacts.

Hums were created in Praat by first producing a point process and an intensity object from each sound. (A point process is a set of points along a time axis that show the location of each pitch pulse.) A synthetic hum was then created from the point process, and the intensity object was transformed into an intensity tier. Finally, the hum was multiplied by the intensity tier, producing a synthetic hum with the pitch and amplitude characteristics of the original sound, but with no segmental information.

Prior to the stress identification experiment, three native English speakers who had received linguistic and phonetic training were asked to perceptually verify the correctness of stress placement for all three types of stimuli. In addition to identifying the stress pattern for each stimulus, they were asked to indicate whether any of the pseudowords sounded similar to real words. Stimuli that were identified with 100% accuracy by the three listeners were used for this experiment. Pseudowords that sounded similar to real words were not used. After stimulus verification, it turned out 72 tokens for each type of stimuli (72 real words, 72 pseudowords, 72 hums).

For the stress identification experiment, all stimuli were divided into two subgroups, each of which contained all three types of stimuli. Each stimulus type was presented as a block. The trochaic and iambic trials within each block were counterbalanced. Listeners identified which syllable in each stimulus item was stressed using a computer keyboard. A fixation symbol (a crosshair, '+') was used between blocks and stayed on the computer screen for 2000 ms (i.e., inter-block interval (IBI) = 2000 ms). Within each block, the inter-trial interval (ITI) was 1000 ms. A response period of up to 3000 ms was allowed from the offset of the stimulus. This was a fixed-pace task; that is, the experiment automatically moved to the next trial if the participant did not make a decision within 3000 ms. If the participant did respond, testing moved immediately to the next trial. The order of blocks and trials were randomized for each participant by the computer.

Procedures

All participants were tested individually using a desktop computer. The test stimuli were presented binaurally through headphones. After each stimulus was presented, participants were asked to identify the stress pattern of the sound they heard; that is, whether the sound had stress on the first syllable or the second syllable. Response choices appeared visually on the screen simultaneously with the auditory presentation of each stimulus. The choice of '1st syllable' was in green at the left side of the screen, and the choice of '2nd syllable' was in red at the right side of the screen. Two matching keyboard response buttons were colored green and red. The visual response choices disappeared from the screen immediately after participants pressed a keyboard response button to make their decision. Experiment instructions were given orally by the experimenter and also displayed on the screen prior to each section of the experiment. A four-minute practice section with feedback was given prior to the test stimuli.

Results

Stimulus Types and Stress Patterns

Tables 2 and 3 summarize the two language groups' performance in terms of accuracy and response time. Accuracy is similar for both groups. Interestingly, the Chinese listeners

Table 2 Accuracy and mean RTs by stimulus type and language group

Language group	Real words	Pseudowords	Hums
English	83.5% (1108.4 ms)	82.6% (906.1 ms)	79.8% (946.6 ms)
Mandarin	82.0% (964.1 ms)	82.6% (911.4 ms)	80.2% (752.8 ms)

Table 3 Accuracy and mean RTs by stress pattern and language group

Language group	Trochaic	Iambic
English	83.0% (967.4 ms)	80.9% (1006.6 ms)
Mandarin	84.8% (903.7 ms)	78.4% (848.6 ms)

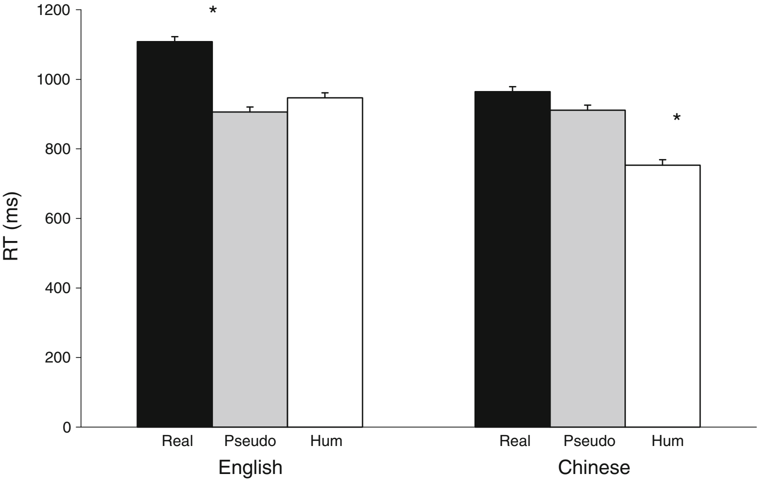

Fig. 1 RTs (ms) for the three types of stimuli by language group

responded faster than English listeners without sacrificing overall accuracy. A three-way repeated measures ANOVA with Stimulus type and Stress pattern as fixed factors, and language group (Group) as a random factor was conducted on accuracy (%) and mean response time (RT). For the accuracy data, no significant main effects or interaction effects were found. However, while there were no differences between the language groups in terms of accuracy, the RT data did show significant differences. No significant main effect was found in RTs between the two groups, but a three-way interaction of Group X Stimulus type X Stress pattern was significant ($F(2, 9958) = 3.947$, $p < .05$).

Figures 1 and 2 compare the English and Chinese groups' RT patterns for the three types of stimuli and two stress patterns, respectively. In terms of stimulus type, English listeners responded fastest to pseudowords and slowest to real words, while Chinese listeners responded fastest to hums and slowest to real words. In terms of stress pattern, English listeners responded fastest to stimuli with trochaic stress, while Chinese listeners responded fastest to stimuli with iambic stress.

To examine the English and Chinese groups' response patterns in greater detail, two separate repeated measures ANOVAs were conducted on accuracy and mean RTs for each language group, with Stimulus type and Stress pattern as fixed factors and Participants as a

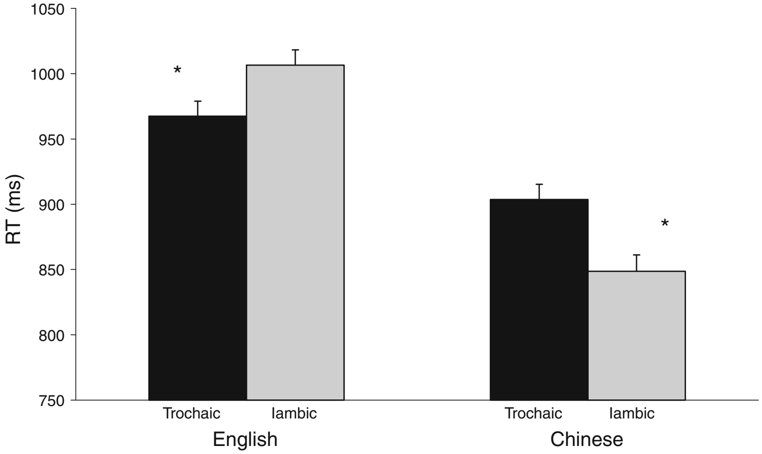

Fig. 2 RTs for trochaic and iambic stress patterns by language group

random factor. Results for the English group will be discussed first. For the accuracy data, the ANOVA revealed no significant main effect for Stress pattern, and Stimulus type only approached significance ($F(2, 58) = 2.496$, $p = .091$). For the RT data, significant mean differences were found for Stimulus type ($F(2, 58) = 16.974$, $p < .001$) and Stress pattern ($F(1, 29) = 4.274$, $p < .05$). As Table 2 shows, the mean RTs to trochaic stress patterns were 39.3 ms shorter than RTs to iambic stress patterns for English listeners. A Tukey's HSD post hoc comparison indicated that RTs to real words differed significantly from pseudowords and hums, where RTs for real words were significantly longer than those for pseudowords and hums.

For the Chinese group, the accuracy data showed a significant main effect for Stress pattern ($F(1, 30) = 13.062$, $p < .001$). Accuracy for trochaic stress patterns was significantly higher than for iambic stress patterns (Table 2). For the RT data, significant main effects were found for Stimulus type ($F(2, 59) = 25.804$, $p < .001$) and Stress pattern ($F(1, 29) = 4.967$, $p < .05$). In addition, the interaction between Stimulus type and Stress pattern was significant ($F(2, 61) = 7.207$, $p < .01$). A Tukey's HSD post hoc comparison showed that RTs to hums differed significantly from real words and pseudowords, with hums having the shortest response latencies. Mean RTs to iambic stress patterns were also significantly shorter than those for trochaic stress patterns. Reaction times for the Chinese listeners are summarized in Fig. 3. As can be seen in the figure, Chinese speakers' responses to different stress patterns were different on real words than on pseudowords and hums. Listeners responded more quickly when real words had a trochaic stress pattern, but more slowly when pseudowords and hums had a trochaic stress pattern.

Results so far show that English and Chinese listeners responded to three types of stimuli and stress patterns differently. The English listeners responded slower but more accurately to real words compared to pseudowords and hums. This may be the result of a lexical effect, where lexical information facilitates their judgment of stress placement but slows down the processing of stress information. In addition, English listeners responded faster and more accurately for pseudowords than for hums, indicating an effect of segmental information. Segmental information seems to facilitate the English listeners' judgment and processing of stress placement.

Chinese listeners performed similarly to English listeners on real words, with longer response times but higher accuracy for real words than for pseudowords and hums. Overall,

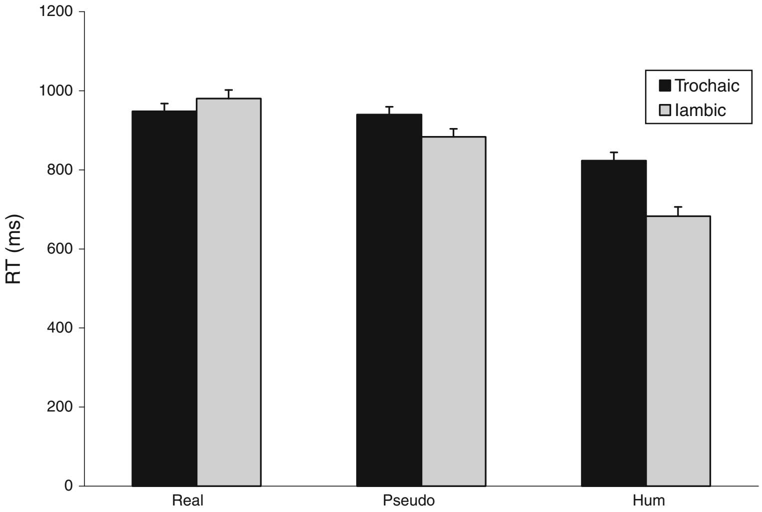

Fig. 3 RTs for Chinese listeners by stress pattern and stimulus type

however, Chinese listeners responded faster and more accurately to hums. The results also indicate that these two language groups have a different stress preference. English listeners are faster and more accurate for stimuli with trochaic stress patterns whereas, Chinese listeners responded faster and more accurately to stimuli with iambic stress patterns.

Acoustic Attributes

A linear mixed model was employed to examine which acoustic attributes (pitch, duration, intensity) were most strongly associated with listeners' responses to stress patterns. For the statistical analysis, the ratio of stressed to unstressed syllable value was taken for each acoustic attribute. Separate linear mixed analyses were performed on each stimulus type for each language group using the accuracy and RT data. In the linear mixed analysis, Stress pattern was a fixed factor, ratio of pitch, ratio of duration and ratio of intensity were covariates, and Participant was a random factor. Only correct responses were included in the analysis of mean RTs.

For the accuracy data, the linear mixed analysis showed only one highly significant relationship to listener response. Specifically, pitch was significantly related to response accuracy for pseudowords with trochaic stress patterns in the Chinese group ($t = 3.491$, $p < .001$).

The RT data were more informative. Results for the linear mixed analysis of mean RTs are summarized in Tables 4 and 5 for the English and Chinese groups, respectively. The results indicate that the English and Chinese groups rely on different acoustic cues to identify stress patterns for the different types of stimuli. In addition, the pattern of reliance on acoustic cues is much simpler for the Chinese listeners: pitch is used to recognize syllable stress in each stimulus type and stress pattern. Duration serves as a secondary cue to iambic stress in pseudowords.

In contrast, for the English group, duration was the most consistent cue to stress. Pitch significantly predicts RTs only to real words with a trochaic stress pattern and pseudowords with an iambic stress pattern. The only stimuli to show no relationship between duration and RTs were real words with trochaic stress. Intensity was the least used cue to stress. Only hum stimuli with iambic stress showed a relationship between intensity and RTs.

Table 4 Results from linear mixed analyses for RTs in the English group

Stimulus type	Model	*B* weight	*t* value	*p* value
Real words				
Trochaic	Pitch	−524.376	−4.085	$p < .001$
Iambic	Duration	−73.052	−2.882	$p < .01$
Pseudo words				
Trochaic	Duration	−6.680	−2.526	$p < .05$
Iambic	Pitch	−461.111	−4.765	$p < .001$
	Duration	−26.403	−2.549	$p < .05$
Hums				
Trochaic	Duration	−185.755	−1.405	$p < .001$
Iambic	Intensity	−1952.900	−2.915	$p < .01$
	Duration	−57.303	−2.196	$p < .05$

Table 5 Results from mixed model analyses for RTs in the mandarin group

Stimulus type	Model	*B* weight	*t* value	*p* value
Real				
Trochaic	Pitch	−440.795	−4.127	$p < .001$
Iambic	Pitch	−625.102	−4.268	$p < .001$
Pseudo				
Trochaic	Pitch	−599.988	−4.789	$p < .001$
Iambic	Pitch	−680.106	−5.816	$p < .001$
	Duration	−23.836	−1.938	$p = .053$
Hum				
Trochaic	Pitch	−1030.600	−5.455	$p < .001$
Iambic	Pitch	−685.427	−2.410	$p < .05$

Vowel Quality Analysis

Vowel quality is often reported to be a cue to stress in word segmentation and recognition in English (e.g., Lehiste 1970; Beckman 1986; Pierrehumbert 1980; Cutler and Clifton 1984; Trommelen and Zonneveld 1999). In English, stressed syllables are usually characterized by full vowels, whereas unstressed syllables often contain reduced vowels which approach schwa in quality. Native English speakers may perceive a syllable that contains a full vowel as the strong/stressed syllable. It is interesting to know whether listeners in this study used vowel quality as a cue to stress in real words and pseudowords. We analyzed the quality of stressed and unstressed vowels in these stimuli by measuring their first and second formant frequencies (F1 and F2) at the durational center of each vowel. The degree of reduction was then measured by finding the distance of each vowel from the centroid of our male speaker's vowel quadrilateral. Linear mixed analyses were performed on real word and pseudoword stimuli separately for each language group using the accuracy and RT data. In the linear mixed analysis, Stress pattern was a fixed factor, distance between the vowel in the first syllable (DV1), second syllable (DV2) and the centroid of the quadrilateral were covariates, and Participant was a random factor. Only correct responses were included in the analysis of mean RTs.

Table 6 Results from vowel analyses for Pseudowords in the English group: Accuracy and RTs

	Model	*B* weight	*t* value	*p* value
Accuracy				
Trochaic	DV1	.0004	2.9	$p < .01$
Iambic	–	–	–	–
RTs				
Trochaic	DV1	18.651	2.76	$p < .01$
Iambic	–	–	–	–

The linear mixed analysis showed only one highly significant relationship to vowel quality in English listeners' responses, specifically, DV1 was significantly related to response accuracy and RTs for pseudowords with trochaic stress patterns, as shown in Table 6. No significant relationship between vowel quality and stress identification was found for Chinese listeners. The results indicate that English listeners used vowel quality as a cue to identify stress for pseudowords with trochaic stress patterns, whereas Chinese listeners did not rely on vowel quality to identify stress placement for real words or pseudowords.

Experiment 2

Results from the stressed-syllable identification experiment indicated that native English and Chinese speakers use different strategies to identify stress patterns for different types of stimuli. Lexical information and segmental cues appeared to affect English and Chinese speakers' accuracy and speed of responses to varying degrees. For the English group, lexical information had a small positive effect on accuracy, where they performed best on real words. However, lexical access appeared to slow their responses, resulting in longer response latencies to real words than to hums and pseudowords. Segmental cues also appeared to affect English speakers' performance, making them slightly faster and more accurate on pseudowords than hums, (although no statistical mean difference was found). For the Chinese group, the effects of lexical information and segmental cues were less clear. Chinese listeners responded significantly more slowly to real words and pseudowords than to hums, suggesting that there may be some general effect from lexical search and/or segmental processing. To further test whether lexical and segmental information influence listeners' processing of stress patterns, Experiment 2 uses a discrimination experiment in which listeners were asked to discriminate the stress patterns of two stimuli presented as a pair.

Methods

Participants

The 30 native speakers of Chinese and of English who completed Experiment 1 also participated in this experiment.

Materials

The discrimination experiment used three types of stimuli: real word, pseudowords and hums, each of which was composed of two-syllables. All stimuli were different from those used in Experiment 1. The grammatical categories (nouns, verbs) and stress patterns (trochaic,

Table 7 Accuracy and RT results by stimulus type and language group

Group	Real words	Pseudowords	Hums
English	70.7% (1348.9 ms)	73.2% (1141.7 ms)	66.0% (1121.1 ms)
Mandarin	67.8% (1279.6 ms)	70.2% (1251.4 ms)	72.0% (1162.4 ms)

iambic) of the real word stimuli were counterbalanced, and real word stimuli were closely matched in terms of word frequency. Frequency estimates were based on the Francis and Kucera (1982) data-base. A one-way ANOVA on word frequency with stress pattern as factor revealed no statistically significant difference among real word stimuli with first and second syllable stress ($F < 1$). A paired t-test on syllable weight showed no significant differences in syllable weight between the first and second syllable for pseudoword stimuli ($p = .882$). Pseudoword stimuli were counterbalanced by stress pattern. Hums were created by the same method as used in Experiment 1. Similar to the procedure for stimulus verification, three native English speakers were asked to perceptually verify the correctness of stress placement for all three types of stimuli. Pseudowords that sounded similar to real words were excluded from the test. All three types of stimuli which were identified over 75% accuracy were used in this experiment.

The experimental paradigm was similar to that used in Experiment 1 with the exception of number of stimuli and testing trials: Listeners heard pairs of stimuli presented over headphones and were asked to indicate whether their stress patterns were the same or different. A total of 360 stimuli were used in this experiment: 120 real words, 120 pseudowords, and 120 hums. Within each type of stimuli, each individual stimulus was paired with another that had either the same stress pattern or the opposite stress pattern in order to create same and different trials. This resulted in a total of 60 trials for each type of stimuli, including 30 trial pairs that had the same stress pattern and 30 that had different stress patterns. Trials were evenly distributed into two subtests consisting of three types of stimuli, each of which was presented as a block.

Procedures

The procedures in Experiment 2 were the same as those reported in Experiment 1 except that participants were asked to decide whether the two sounds they heard had the same stress pattern or different stress patterns. The response choices were displayed visually on the screen, with the choice of ‘Same’ in green at the left side of the screen, and the choice of ‘Different’ in the red at the right side of the screen. A five-minute practice section with feedback was given prior to the experiment.

Results

Table 7 summarizes the English and Chinese participants’ accuracy and mean RTs for the three types of stimuli. Overall, the English and Chinese groups had similar accuracy in discriminating stress patterns. The English listeners were most accurate in discriminating stress patterns for pseudowords and least accurate for hums. The Chinese listeners were most accurate for hums and least accurate for real words.

A two-way repeated measures ANOVA, with Group and Stimulus type as fixed factors and Participant as a random factor revealed no significant main effects for either the accuracy or the RT data. A significant interaction between Group and Stimulus type was found for

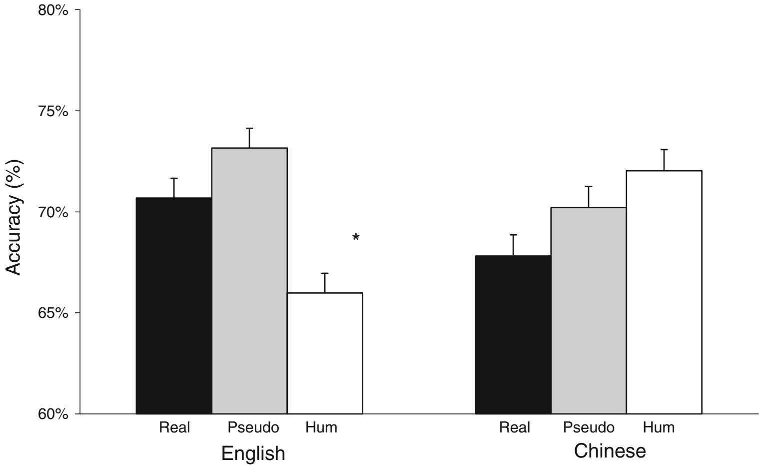

Fig. 4 Accuracy by stimulus type and language group

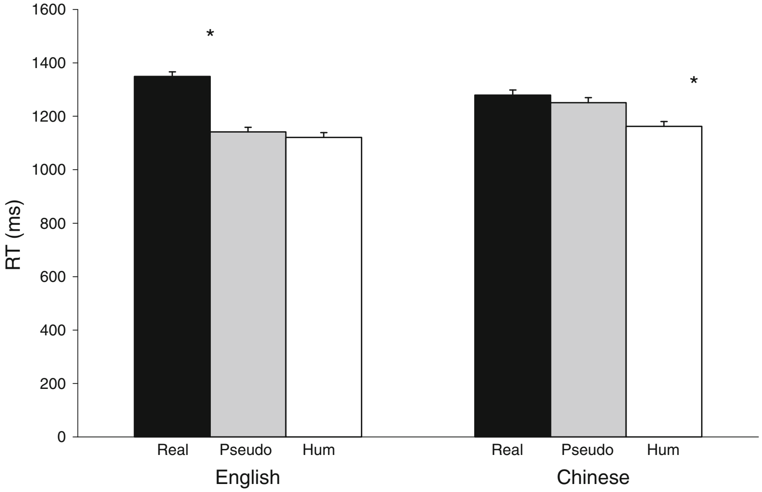

Fig. 5 RTs by stimulus type and language group

both the accuracy data ($F(2, 12321) = 13.052$, $p < .001$) and the RT data ($F(2, 8620) = 12.517$, $p < .001$). Figures 4 and 5 show the accuracy and mean RT data for the different stimulus types for English and Chinese speakers.

Separate one-way repeated measures ANOVAs were performed on accuracy and mean RTs for the English and Chinese groups, with Stimulus type as a fixed factor, and Participants as a random factor. For the English group, the ANOVA revealed a significant main effect for Stimulus type in both the accuracy data ($F(2, 13) = 4.444$, $p < .05$) and the RT data ($F(2, 74) = 11.627$, $p < .001$). A Tukey's HSD post hoc comparison showed that accuracy for hums was significantly lower than for real words and pseudowords, and that RTs to real words were significantly longer than to pseudowords and hums. For the Chinese group, the ANOVA showed a main effect of Stimulus type ($F(2, 62) = 4.352$, $p < .05$) only for the

RT data. A Tukey's HSD post hoc comparison showed that RTs to hums were significantly shorter than to real words and pseudowords.

In this experiment, the results indicate that lexical and segmental information influenced English listeners and Chinese listeners differently. Lexical information facilitated English listeners' judgment but slowed down their processing of stress placement. In contrast, Chinese listeners did not benefit from lexical or segmental information in the discrimination task, with lower accuracy and longer response time in real words and pseudowords as compared to hums.

General Discussion

Lexical and Segmental Effects

The results of the stressed-syllable identification task in Experiment 1 showed that English and Chinese speakers were able to identify placement of stress in all three stimulus types with nearly equal accuracy. The results also revealed some different relationships between accuracy and RTs for the English versus the Chinese speakers. Results for the English speakers will be considered first.

English speakers were slower and more accurate on real words as compared to pseudowords and hums, although the difference in accuracy was not statistically significant. The slightly higher accuracy and significantly slower RTs may be the result of a lexical effect, where lexical information modestly facilitates judgments of stress placement, increasing accuracy, but slows down the processing of stress information, resulting in longer mean RTs.

The results suggest a similar facilitation effect for English speakers for segmental information. Segmental information assisted English speakers in making faster decisions about stress placement. Knowing what sounds are present may slightly facilitate stress pattern identification for English speakers. For pseudowords, which contain segmental information but do not have lexical status, the English group responded faster in comparison with real words and hums.

Like the English speakers, the Chinese speakers' RTs in the stress identification experiment were slowest in response to real words. However, whereas the fastest RTs for the English speakers were in response to pseudowords, the fastest RTs for the Chinese speakers were in response to hums. The Chinese speakers' significantly faster RTs to hums than to real words and pseudowords may result from the absence of lexical and segmental information in the hums. Given the Chinese speakers' consistent reliance on pitch as a cue to English stress, this also suggests that the absence of lexical and segmental information allows these stimuli to be processed as a tonal contour that is similar to Chinese tones.

The results of the discrimination task showed different response patterns for the English and Chinese speakers. The response patterns in the discrimination task also support the findings of the stressed-syllable identification task. English speakers had greater difficulty correctly discriminating hum pairs than real word and pseudoword pairs. This may be due to the lack of lexical and segmental information in hums and again suggests that English speakers benefit from lexical and segmental information when discriminating stress patterns. In addition, English lexical stress is realized on vowels, and full vowels most often occur in stressed syllables. English speakers might use vowel quality as a supplemental cue to discriminate stress patterns, resulting in higher accuracy for real and pseudoword pairs compared to hum pairs. Similar to the results for the stress identification task, lexical and segmental information improved accuracy for the English speakers but lexical information slowed English speakers' RTs to real word pairs.

In contrast, Chinese speakers showed higher accuracy and faster RTs for hum pairs compared to real word and pseudoword pairs in the discrimination experiment. Lexical and segmental information did not improve accuracy for Chinese speakers. This might be attributed to the difficulty of processing stimuli which are not part of the native lexicon and which contain non-native sounds and sound sequences. Given the extra processing burden that this may impose, Chinese speakers would not be expected to benefit from lexical or segmental information to the same degree as English speakers. As in the stress identification task, their better performance discriminating the hum stimuli may be attributed to the fact that hums contain only prosodic information, making them similar to pitch contours in Chinese tone sequences. Unlike the higher accuracy for real words as compared with pseudowords and hums in the stress identification task, Chinese speakers performed worst on real words in the discrimination task, with lower accuracy and longer response times. This may be due to the increased difficulty of the discrimination task. In this task, listeners need to process lexical information for two words at the same time as they make judgments regarding stress location.

Stress Preference

Overall, English and Chinese speakers were able to identify both trochaic and iambic stress patterns with equal accuracy. In terms of reaction time, however, English speakers responded faster to trochaic stress patterns, whereas Chinese speakers responded faster to iambic stress patterns. These results indicate that native English speakers recognize trochaic stress patterns faster than iambic stress patterns even when lexical or segmental information is not present. The finding of a “trochaic preference” for English speakers is also supported by published literature on speech segmentation, which has shown that native English listeners prefer to interpret stress patterns as strong-weak (e.g., Cutler and Butterfield 1992; Taft 1984; van Heuven 1985). Results from studies in child language development are also in accordance with this finding. English learning infants demonstrate a preference for trochaic stress patterns in English words (e.g., Jusczyk et al. 1999), and have been shown to rely on trochaic stress patterns for word segmentation (Thiessen and Saffran 2003). The finding of a trochaic preference is also compatible with the statistical distribution of stress patterns in English, where the trochaic stress pattern is the dominant pattern in English two-syllable words (Yu 2008).

In contrast with the English speakers, Chinese speakers identified iambic stress patterns significantly more quickly than trochaic patterns. Although trochaic stress patterns required more time to identify, they were identified significantly more accurately than iambic patterns. The significantly faster RTs to iambic stimuli suggest that Chinese speakers are more sensitive to iambic stress patterns. This finding is consistent with the results of a second language acquisition study by Wayland et al. (2006). They found that Thai speakers, who speak a tone language like Mandarin Chinese, showed a perceptual preference for final stress over initial stress in two-syllable pseudowords. However, the authors did not provide further explanation for these results. Although the study presented here provides no direct evidence of an iambic preference for Chinese speakers, in conjunction with the results of Wayland et al., the present results suggest that the bias of Chinese speakers towards iambic stress and English speakers toward trochaic stress could result from their different language background. Additional support for this conclusion is provided by Chao (1979) Final Stress Theory. This theory states that, in Mandarin Chinese, the final syllable usually receives stress regardless of whether it occurs in a word or a phrase. Empirical support for this theory is provided by several phonetic studies in Mandarin Chinese (e.g., Yan and Lin 1988; Lin et al. 1984), which show that native Chinese listeners favor final syllable stress over initial syllable stress. The convergence of this evidence with the current results suggests that Chinese speakers' perception of lexical stress in English is affected by their native language. The Chinese speakers' more accurate identification of trochaic stress may be a product of learning English as a second language. As with English speakers, greater exposure to trochaic stress patterns may lead to more accurate recognition.

Relationship Between Stress Preference and Stimulus Type

In the present study, English speakers showed significantly faster reaction times to trochaic stress patterns across all three types of stimuli. They also identified trochaic stress patterns more accurately across all three types of stimuli, but this difference did not reach significance. Like English speakers, Chinese speakers responded faster to trochaic stress patterns than iambic stress patterns in real words. In addition, their responses to trochaic stress patterns were significantly more accurate than responses to iambic patterns. Unlike English speakers, they exhibited the reverse RT pattern for pseudowords and hums, with shorter RTs to iambic stimuli than to trochaic stimuli (effect size: 51.3 ms for pseudowords and 140.7 ms for hums). This differing relationship between stress pattern and stimulus type across English and Chinese speakers again suggests an effect of language background. Chinese speakers in this study were all second language learners of English. Although they did not have native proficiency in English, they appear to have been sufficiently familiar with the language to show a similar pattern to native English speakers in responding to real words. However, the Chinese speakers' performance on unfamiliar stimuli (the pseudowords) was similar to their performance on hums, with faster response times to iambic stress patterns. When Chinese speakers heard words that were completely unfamiliar, they may have processed them using their default stress mechanism, yielding faster responses to iambic stress patterns in pseudowords and hums.

Relationship Between Acoustic Attributes, Vowel Quality and Stress Preference

The results of this study also provide insight into English and Chinese listeners' use of acoustic cues related to stress (pitch, duration, intensity, vowel quality). The current data show that English and Chinese speakers rely on different acoustic cues to identify stress patterns in two-syllable stimuli. English speakers showed a complex pattern in which the cue(s) used were dependent on both the stress pattern and the type of stimuli. In contrast, Chinese speakers relied on pitch for all three types of stimuli, adding duration as a secondary cue only for pseudowords with iambic stress. In terms of vowel cues for real word and pseudoword stimuli, the current data showed that English speakers used vowel quality only for pseudowords with trochaic stress. This suggests that English speakers make use of vowel quality cues to process stress when lexical information is not available. Unlike English speakers, Chinese speakers do not use vowel quality to identify stress for real words or pseudowords.

Once again, these results suggest that the difference in acoustic cue use is related to language background. English is a stress language, where lexical stress is realized by the acoustic attributes of pitch, duration and intensity. Stressed syllables tend to receive a combination of higher pitch, longer duration and greater intensity. In addition, stressed syllables are associated with full vowels, while unstressed syllables are often associated with reduced vowels. The results of this study indicate that native English speakers use all of these acoustic cues to process stress patterns. For stimuli with a trochaic pattern, a single cue significantly predicted reaction times for stress identification (pitch in real words, and duration in pseudowords and hums). For stimuli with an iambic pattern, duration alone significantly predicted

RTs in real words, while pseudowords and hums showed the addition of pitch and intensity, respectively. Vowel quality was used as a cue to trochaic stress in pseudowords, indicating that when lexical information is not available, English speakers use vowel quality to assist in processing stress.

In contrast with English, Mandarin Chinese is a tonal language, where the primary acoustic cue to all four tones is pitch contour. Chinese speakers in this study used pitch to identify trochaic and iambic stress patterns in two-syllable English real words, pseudowords and hums. In all but one condition (pseudowords with iambic stress), pitch alone predicted Chinese speakers' RTs for stress identification. For pseudowords with iambic stress, Chinese speakers added duration as a secondary cue. As expected, Chinese speakers did not use vowel quality to process stress for real words and pseudowords, as vowel quality is not an identifying feature for Mandarin Chinese tones. These findings indicate that English and Chinese speakers use different acoustic attributes to identify stress patterns. The Chinese listeners' consistent reliance only on pitch, in particular, provides strong evidence to support the hypothesis that language background affects the way that listeners perceive and process lexical stress.

Conclusion

The results of the two experiments presented here support the hypothesis that language background affects the perception of lexical stress in English. In general, both native English and native Chinese speakers were able to identify English stress with nearly equal accuracy. However, they showed different preferences for English stress patterns. Trochaic stress was preferred by native English speakers, whereas iambic stress was preferred by native Chinese speakers. These different stress preferences were based on language background. Different effects of lexical and segmental information were also found for stress identification for the English and Chinese speakers. Lexical information in the real word stimuli improved English speakers' accuracy but interfered with their response times, most likely due to automatic lexical processing. Segmental information in the pseudoword stimuli improved both accuracy and response times for English speakers.

Like English speakers, Chinese speakers showed a lexical effect for stress identification in real words. However, Chinese speakers did not benefit from segmental information in stress identification. This different effect of lexical and segmental information between English and Chinese speakers was also a result of language background. Finally, English speakers and Chinese speakers used different acoustic cues to identify stress. English speakers showed a complex pattern where the cues used were dependent on the stress pattern and type of stimuli. In contrast, the Chinese speakers relied primarily on pitch for all types of stimuli and stress patterns.

A cross-language comparison of the real word and hum results for both of the experiments presented in this paper invites some additional speculation about differences between English and Chinese listeners' processing of English lexical stress. Figure 6 summarizes RT results for the real word and hum data. RTs to real words are consistently slower than hums for speakers of both languages, but on both real word tasks, Chinese listeners respond faster than English listeners. This is somewhat surprising, given that they are not native speakers. In spite of their faster RTs, Chinese listeners are only about 3% less accurate than English listeners on both tasks. For stressed syllable discrimination, Chinese listeners' average RTs to real words are 69 ms faster than English listeners, and for stressed syllable identification, Chinese listeners' average RTs are 144 ms faster. As discussed above, the fact that RTs for real words

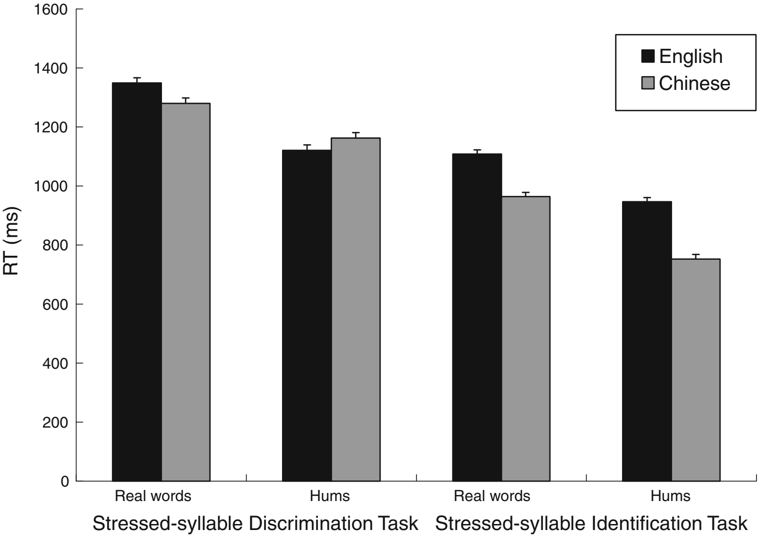

Fig. 6 RTs for real words and hums for English and Chinese listeners in the stressed-syllable discrimination and identification tasks

in general are longer than for hums suggests that lexical access takes place automatically for the real words and slows both listener groups. The fact that Chinese listeners are faster than English listeners in judging real word stress suggests that they are able to make their stress judgments earlier in the lexical access process than English listeners. One possibility here is that English listeners obtain information about real words' stress from the lexicon and must therefore wait until a lexical entry is retrieved, while Chinese listeners determine stress before a lexical entry is retrieved.

If we turn to differences between the discrimination and identification task results, we see that in general, discrimination requires more time than identification. This is presumably because two stimuli must be analyzed and compared for the discrimination task, while only one is analyzed for the identification task. A more detailed cross-language comparison shows that for real words, Chinese listeners require about 316 ms longer to discriminate than to identify stressed syllables, while English listeners require about 241 ms longer. For the hum stimuli, the average increase in RTs from identification to discrimination is 410 ms for the Chinese listeners, but only 175 ms for the English listeners. Across languages then, the shift from identifying to discriminating stress costs Chinese listeners more than English listeners in processing time, particularly for the hum stimuli. What might explain this difference?

Neuroimaging studies have shown (e.g., Gandour et al. 2000, 2003, 2004; Klein et al. 2001; Wang et al. 2001, 2004) that speakers of tone languages process pitch on the left side of the brain as a meaningful element of language. For English speakers the data are less clear, but it appears that when multiple acoustic cues are processed to determine lexical or emphatic stress placement, this processing largely takes place in the right hemisphere, suggesting that it is being processed as separate, prosodic information (Baum 1998). This suggests that pitch information regarding tonal contrasts becomes available to Chinese listeners with segmental information, prior to activation of lexical entries, while it may not be available this early for English speakers. The acoustic results presented here indicate that Chinese listeners also use pitch to process English stress. Since the same acoustic cue is used to evaluate both tone and stress, it is possible that Chinese learners of English process pitch in English stress in the same way as lexical tone, and that this information is already available when lexical entries are being activated.

For hums, Chinese listeners may again process pitch information in the same manner, making pitch contour 'identities' available at the same speed that segmental information becomes available for English listeners. If, however, English pitch contours are classified in a similar way to Chinese tonal contours, the actual decision regarding which syllable is stressed may be more complex. A comparison of two-syllables in a single word could, for example, require a 'high-level' pitch to be compared with a 'high-falling' pitch in order to decide which syllable is stressed. In a discrimination task, the complexity of this decision could be further compounded by the amount of detail that is available regarding pitch. For example a hum with a 'low rising + high rising' pitch pattern could potentially be compared with a hum that has a 'high-level + high-falling' pitch pattern. In such a case, determining whether the hums have the same stress pattern would be complex enough to require extra time.

English listeners, on the other hand, do not refer to a single acoustic feature, or even a fixed set of acoustic features, to decide which syllable is stressed. This may lengthen the time it takes for English listeners to identify the stressed syllable in real words and hums. Although Mattys (2000) showed that for four-syllable words, English listeners could make fine distinctions between primary and secondary stress without accessing the lexicon, Mattys' listeners made these distinctions by comparing syllables with primary or secondary stress to "a single, durable standard – namely, a reduced syllable" (2000, p. 262). Since the two-syllable real words used in this study did not contain fully reduced vowels, it may have been more efficient to determine stress by accessing the lexicon. This could either be a task artifact or the usual strategy for determining real word stress in two-syllable words. For hums, in contrast, listeners in this study could only judge stress from the acoustic signal, making decision times longer for English listeners than Chinese listeners. Once the decision has been made, however, a word or hum can presumably be classified as having first or second syllable stress, making discrimination decisions faster and easier for English listeners than for Chinese listeners.

These speculations suggest several directions for further research. Noun/verb pairs that are distinguished by stress (e.g., INsert/inSERT) constitute a group of words for which it might be necessary to know which syllable is stressed in order to access the correct lexical entry. This link between stress and grammatical category also persists in words that are not otherwise identical. When nouns and verbs are in a grammatical context, therefore, grammatical category may be a better predictor of stress than acoustic cues. Future studies should examine how grammatical category interacts with stress perception, as well as how language background affects perception of stress in nouns and verbs.

The question of whether Chinese listeners process English stress in the same way as Mandarin Chinese tone or in a way that is similar to English listeners should also be examined using neuroimaging techniques. Neuroimaging studies that address this question would also address the more general question of whether distinct neural networks exist for primary and secondary language acquisition and processing in bilingual individuals.

Behavioral studies that require Chinese listeners to decide which syllable is stressed in hums that carry Mandarin Chinese tone pitch contours, and that require them to decide which Chinese tones should be assigned to hums that carry English stress patterns, could also reveal whether Chinese listeners equate pitch cues to English stress to pitch cues for lexical tone.

Studies of this nature could provide practical information for teaching stress production and perception to Chinese learners of English, and possibly also for teaching tone production and perception to English learners of Chinese. Information of this nature may also have applications in clinical settings for improving the efficacy of language rehabilitation in bilingual speakers.

References

Archibald, J. (1992). Transfer of L1 parameter settings: some empirical evidence from Polish metrics. *Canadian Journal of Linguistics, 37*, 301–339.

Archibald, J. (1993). The learnability of English metrical parameters by adult Spanish speakers. *International Review of Applied Linguistics and Language Teaching, 31/32*, 129–142.

Archibald, J. (1997). The acquisition of English stress by speakers of nonaccentual languages: Lexical storage versus computation of stress. *Linguistics, 35*, 167–181.

Baayen, R. H., Piepenbrock, R., & van Rijn, H. (1993). *The CELEX lexical database [CD-ROM]. Philadelphia: University of Pennsylvania, Linguistic Data Consortium.*

Baum, S. R., & Pell, M. D. (1999). The neural bases of prosody: Insights from lesion studies and neuroimaging. *Aphasiology, 13*, 581–608.

Baum, S. R. (1998). The role of fundamental frequency and duration in the perception of linguistics stress by individuals with brain damage. *Journal of Speech Language Hearing Research, 41*, 31–40.

Beckman, M. E. (1986). *Stress and non-stress accent. Netherlands Phonetic Archives 7*. Dordrecht: Foris.

Best, C. T., McRoberts, G., & Sithole, N. (1988). Examination of perceptual reorganization for non-native speech contrasts: Zulu click discrimination by English speaker adults and infants. *Journal of Experimental Psychology, 14*, 345–360.

Best, C. T. (1994). The emergence of native-language phonological influence in infants: perceptual assimilation model. In J. Goodman & H. Nusbaum (Eds.), *The development of speech perception: The transition from speech sounds to spoken words*. Cambridge, MA: MIT Press.

Best, C. T. (1995). A direct realist view of cross-language speech perception. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 171–204). Timonium, MD: York Press.

Best, C. T., McRoberts, G. W., & Goodell, E. (2001). Discrimination of non-native consonant contrast varying in perceptual assimilation to the listener's native phonological system. *Journal of the Acoustical Society of America, 109*, 775–794.

Chao, Y. R. (1979). *A Grammar of spoken Chinese*. CA: University of California Press.

Cutler, A., & Clifton, C. E. (1984). The use of prosodic information in word recognition. In H. Bouma & D. G. Bouwhuis (Eds.), *Attention and performance X: Control of language processes*. Mahwah, NJ: Lawrence Erlbaum Associates.

Cutler, A., & Carter, D. M. (1987). The predominance of strong initial syllables in the English vocabulary. *Computer Speech and Language, 2*, 133–142.

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

Cutler, A., & Butterfield, S. (1992). Rhythmic cues to speech segmentation: evidence from juncture misperception. *Journal of Memory & Language, 31*, 218–236.

Dupoux, E., Pallier, C., Sebastian, N., & Mehler, J. (1997). A distressing "deafness" in French? *Journal of Memory and Language, 36*, 406–421.

Emmorey, K. (1987). The neurological substrates for prosodic aspects of speech. *Brain and Language, 30*, 305–320.

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

Francis, N., & Kucera, H. (1982). *Frequency analysis of English usage: Lexicon and grammar*. New York: Houghton Mifflin.

Gandour, J., Tong, Y., Talavage, T., Wong, D., Dzemidzic, M., Xu, Y., Li, X., & Lowe, M. (2004). Hemispheric roles in the perception of speech prosody. *NeuroImage, 23*, 344–357.

Gandour, J., Xu, Y., Wong, D., Dzemidzic, M., Lowe, M., Li, X., & Tong, Y. (2003). Neural correlates of segmental and tonal information in speech perception. *Human Brain Mapping, 20*, 185–200.

Gandour, J., Wong, D., Hsieh, L., Weinzapfel, B., Van Lancker, D., & Hutchins, G. D. (2000). A crosslinguistics PET study of tone perception. *Journal of Cognitive Neuroscience, 12*, 207–222.

Goto, H. (1971). Auditory perception by normal Japanese adults of sounds 'l' and 'r'. *Neuropsychologia, 9*, 317–323.

Hume, E., Johnson, K., Seo, M., Tserdanelis, G., & Winters, S. (1999). A cross linguistic study of stop place perception. In *Proceedings of the XIVth international congress of phonetic sciences* (pp. 2069–2072). Berkeley: Congress Organizers.

Jusczyk, P. W. (1993). From general to language-specific capacities: the WRAPSA model of how speech perception develops. *Journal of Phonetics, 21*, 3–28.

Jusczyk, P. W., & Aslin, R. N. (1995). Infants' detection of the sound patterns of words in fluent speech. *Cognitive Psychology, 29*, 1–23.

Jusczyk, P. W., Houston, D., & Newsome, M. (1999). The beginnings of word segmentation in English learning infants. *Cognitive Psychology, 39*, 159–207.

Klein, D., Zatorre, R. J., Milner, B., & Zhao, V. (2001). A crosslinguistic PET study of tone perception in Mandarin Chinese and English speakers. *NeuroImaging, 13*, 636–653.

Kuhl, P. K. (2000). Language, mind, and brain: Experience alters perception. In M. Gazzaniga (Ed.), *The new cognition neuroscience*. Cambridge, MA: MIT Press.

Lehiste, I. (1970). *Suprasegmentals*. Cambridge, MA: MIT Press.

Lehiste, I., & Fox, R. A. (1992). Perception of prominence by Estonia and English listeners. *Language and Speech, 35*, 419–434.

Lin, M., Yan, Z., & Sun, G. (1984). A preliminary experimental study on normal stress of disyllabic words in Chinese. *Dialects, 1984*, 57–73.

Mattys, S. L. (2000). The perception of primary and secondary stress in English. *Perception & Psychophysics, 62*, 253–265.

Mattys, S. L., & Samuel, A. G. (1997). How lexical stress affects speech recognition and interactivity: Evidence from the migration paradigm. *Journal of Memory and Language, 36*, 87–116.

Mattys, S. L., & Samuel, A. G. (2000). Implications of stress-pattern differences in spoken-word recognition. *Journal of Memory and Language, 42*, 571–596.

Miyakawa, K., Strange, W., Vwebrugge, R., Liberman, A. M., Jenkins, J. J., & Fujimora, O. (1981). An effect of linguistic experience: The discrimination of [r] and [l] by native speakers of Japanese and English. *Perception & Psychophysics, 18*, 331–340.

Natkatani, L. H., & Schaffer, J. A. (1978). Hearing 'words' without words: Prosodic cues for word perception. *Journal of the Acoustical Society of America, 63*, 234–245.

Nooteboom, S. G., Brokx, J. P. L., & de Rooij, J. J. (1978). Contribution of prosody to speech perception. In W. J. M. Levelt & G. B. F. d'Arcais (Eds.), *Studies in the perception of language* (pp. 75–107). New York: Wiley.

Nusbaum, H., & Goodman, J. (1994). Learning to hear speech as spoken language. In H. Nusbaum & J. Goodman (Eds.), *The development of speech perception: The transition from speech sounds to spoken words* (pp. 299–338). Cambridge, MA: MIT Press.

Peperkamp, S., & Dupoux, E. (2002). A typological study of stress deafness. In C. Gussenhoven & N. Warner (Eds.), *Laboratory phonology VII* (pp. 203–240). Berlin: Mouton.

Pierrehumbert, J. (1980). The phonology and phonetics of English intonation. Doctoral dissertation, Cambridge: MIT.

Polka, L., & Werker, J. (1994). Developmental changes in perception of non-native vowel contrasts. *Journal of Experimental Psychology: Human Perception and Performance, 20*, 421–435.

Shah, A. P., & Baum, S. R. (2006). Perception of lexical stress by brain-damaged individuals: effects on lexical-semantic activation. *Applied Psycholinguistics, 27*, 143–156.

Taft, L. A. (1984). *Prosodic constraints and lexical parsing strategies*. Doctoral Dissertation, Indiana University, Bloomington.

Thiessen, E. D., & Saffran, J. R. (2003). When cues collide: Use of stress and statistical cues to word boundaries by 7- to 9-month-old infants. *Developmental Psychology, 39*, 706–716.

Trommelen, M., & Zonneveld, W. (1999). Word stress in West Germanic languages. In H. van der Hulst (Ed.), *Word prosodic systems in the languages of Europe* (pp. 478–515). Berlin, Germany: Mouton de Gruyter.

van Heuven, V. J. (1985). Perception of stress pattern and recognition: Recognition of Dutch words with incorrect stress position. *Journal of the Acoustical Society of America, 78*, S21.

Wang, Y., Behne, D. M., Jongman, A., & Sereno, J. A. (2004). The role of linguistic experience in the hemispheric processing of lexical tone. *Applied Psycholinguistics, 25*, 449–466.

Wang, Y., Jongman, A., & Sereno, J. A. (2001). Dichotic perception of Mandarin tones by Chinese and American listeners. *Brain and Language, 78*, 332–348.

Wayland, R., Guion, S. G., & Landfair, B. L. (2006). Native Thai speakers' acquisition of English word stress patterns. *Journal of Psycholinguist Research, 35*, 285–304.

Yan, J., & Lin, M. (1988). Acoustic characteristics of the stress in Bejining trisyllabies. *Fangyan, 3*, 227–237.

Yu, Y. -J. (2008). A cross-language study in perception of lexical stress in English. Doctoral Dissertation, Wayne State University, Detroit.