The influence of pitch contour on Mandarin speakers' perception of English stress

Yaobin Liu
Stony Brook University

The influence of pitch contour on Mandarin speakers' perception of English stress

Abstract

Previous studies on L2 stress perception have mainly focused on words in isolation or in single intonational contexts. This paper reports on a study exploring the influence of different intonation contours, falling (declarative) and rising (yes/no question), on nonnative speakers' stress identification. The study compared the perception of stress position in English words by native speakers of Mandarin, a tone language, and English, a stress language. The results showed that Mandarin speakers exhibited misperception of stress position when high tones did not coincide with the stressed syllable. As a control condition, native English speakers also displayed misperception of stress, but to a lesser extent in the condition of initial stress. Tonal transfer and asymmetrical cue usage are believed to be responsible for the perceptual differences.

The influence of pitch contour on Mandarin speakers' perception of English stress

Yaobin Liu*

1 Introduction

Speakers often misperceive a foreign language because of differences between the native and the foreign language. In addition to misperception at the segmental level, the perceptual asymmetry caused by language background difference also manifests itself in the domain of suprasegmentals, such as tone, stress and intonation. English is a stress language in which every word has a syllable that is most prominent. Mandarin is a tone language in which every syllable has lexically marked pitch. Pitch is used very differently in the two languages: pitch signals lexical contrasts in Mandarin but signals sentence-level intonation in English. This paper is an experimental investigation into the nature of one perceptual consequence that this typological asymmetry may lead to.

Stress in English, from the speaker's perspective, usually corresponds to "the amount of effort expended"; from the listener's point of view, the perceptual prominence of stressed syllables generally means "longer, louder and higher in pitch than unstressed syllables" (Yavaş 2011). Fry's (1955, 1958, 1965) seminal work on the acoustic properties of English lexical stress have included duration, intensity (or amplitude), pitch (F_0) and vowel reduction (formant structure) in the cue inventory. The relative weighting among these cues has been under extensive debate in the literature. Fry (1958) and Lehiste (1970) suggested fundamental frequency as most influential. Fry (1955) found a stronger effect of duration than intensity, while Lieberman (1960) found that "the envelope amplitude seems more important than duration". Sluijter and van Heuven's (1996) experiment based on Dutch, a ''stress-accent'' language similar to English (Beckman 1986), showed that duration is the "most effective correlate of stress". Zhang & Francis (2010) concluded from their experimental results that vowel quality is weighted more than other cues. It is possible that the diversity of findings results from differences in dimensions, parameters and measurements that were adopted in these studies. Despite the divergences on which cue ranks higher, they converge on the point that all of the four acoustic correlates are active in native speakers' stress perception.

While English stress is signaled by multiple cues, pitch and stress may not interact in an unambiguous way. When intonation contour is imposed on stress contour, the stressed syllable may not have the highest pitch in a sentence as in citation forms. Often the sentence pitch contour, or intonation, can be represented by "a sequence of high (H) and low (L) target pitches" (Ladefoged & Johnson 2011), using the ToBI system, and pitch accents (designated with *) map to stressed syllables. In citation forms and declarative sentences (intonation: H* L-L%), the nuclear pitch accent bears a high tone. However, at the end of a yes/no question (YNQ) (intonation: L* H-H%), the nuclear pitch accent bears a low tone with high pitch on the following phrase accent and boundary tone (Ladefoged & Johnson 2011), as exemplified in (1).

(1) a. declarative: falling contour

I like MOvies.
H* L- L%

b. yes/no & echo question: rising contour

Do you like MOvies?
L* H- H%

You like MOvies?
L* H- H%

Now let's turn to Mandarin, a classical tone language. By definition, tone is phonetically realized by pitch variation. Empirical studies on both production and perception also have found fundamental frequency (F0) to be the primary acoustic correlate of Mandarin tones, specifically F0 height and F0 contour, both of which are considered important cues (Howie 1970, 1976, Moore & Jongman 1997, Jongman et al 2006, Duanmu 2007). The significance of pitch in Mandarin would be expected to render Mandarin speakers more sensitive to pitch than to other cues when they are exposed to English lexical stress.

*I am indebted to Ellen Broselow, Michael Becker and Jiwon Yun for their guidance and support. Many thanks to the audience at NACCL 29 and PLC 42 for comments and suggestions.

Wang's (2008) study of Mandarin speakers' use of acoustic cues in the perception of English lexical stress experimentally confirmed the above conjecture. The testing materials in this study were disyllabic nonsense words with two stress patterns, which were perceived in isolation with manipulated stress cues, including F0, duration and intensity. The results showed that "only F0 has a decisive effect on stress judgments by Chinese learners of English", unlike native speakers, serving as evidence for tonal transfer in their L2 speech perception. Similarly, the results of Archibald's (1997) longitudinal study of Chinese speakers' acquisition of English stress showed that Chinese learners did not progress, and actually regressed, on stress perception over a span of five months. He suggested that the learners paid attention only to pitch, which was carried over as part of the lexical entry from L1 to L2. Chao (1980) also noted that speakers of Cantonese, another highly tonal language, "associate high and low tones with stressed and unstressed syllables" (directly from Wang 2008).

In contrast, Zhang & Francis' (2010) comparative study of native English speakers' and Mandarin learners' cue usage in stress perception arrived at the conclusion that both language groups "consistently weighted vowel quality more than other cues." The study elicited stress location judgments on stimuli such as "DEsert (noun) and deSERT (verb)" by asking participants to judge the lexical category of each word. Vowel quality, pitch, duration, and intensity were manipulated on a seven-step continuum, e.g. from [ɛ] to [ɪ] for vowel quality in the first syllable. However, because the noun-verb pairs used in the study contrasted *full* vowels, the primary cue that the subjects relied upon was vowel quality, not necessarily stress. So, whether this could serve as evidence for the claim that vowel quality is the dominating cue to stress should be taken with a pinch of salt. The present study will not attempt to tease apart the degrees of importance of vowel reduction in Mandarin speakers' stress identification, but only focus on empirically verifying the significance of pitch in that regard.

If Mandarin speakers are rigidly tuned to pitch in stress identification, they will possibly misperceive stress when the stressed syllable bears a low tone, as in the rising pitch contour, while native speakers are less likely to do so as they rely on multiple cues. The present study examined if this pattern is borne out empirically.

2 The Present Study

2.1 Hypothesis

Based on the assumptions outlined above, it can be predicted that Mandarin speakers will be more likely than English speakers to be influenced by pitch contour in stress perception; specifically, they would misperceive stress under a rising pitch contour significantly more than English speakers would.

2.2 Research Questions

We can test the proposed hypothesis by answering the questions in (2). The inclusion of (2b) was to set the baseline for (2a), since even if the answers to (2a) suggest that there is a disparity between Mandarin speakers and English speakers regarding the perceptual pattern possibly conditioned by pitch contour, this may not serve as strong evidence for the hypothesis that Mandarin speakers are more influenced by pitch than English speakers are in stress identification. It could be the case that the group of tested subjects overall have extremely low proficiency in English. The expected results of (2b) would ensure that on the falling contour conditions, Mandarin speakers perform nearly as well as English speakers and the effect of pitch contour, if any, can be thus accentuated and substantiated.

(2) a. Do Mandarin speakers misperceive stress in a rising pitch contour significantly more than in a falling pitch contour? Do English speakers display a similar pattern?
b. Do Mandarin speakers perceive stress in a rising pitch contour significantly differently than English speakers? Does the falling pitch contour display a similar pattern?

2.3 Methodology

The falling contour of neutral declaratives and the rising contour of yes/no questions (echo questions in particular)[1] were used, where words with varying stress patterns were placed to see whether the effect, if any, of global prosody on stress perception is consistent. The question of interest can be boiled down to whether the subjects perceive stress as coinciding with high pitch or not.

In order to elicit participants' stress perception in a naturalistic fashion, the current study made use of the fact that stress in English may sometimes indicate the part of speech (POS) of a word. There is a set of lexical items such as *súbject* /*subjéct* and *permit/permit* that can behave as either nouns or verbs, where the major phonological feature that signals their category is stress placement. The typical pattern is that, for disyllabics, the noun forms tend to carry stress on the initial syllable and the verb forms on the final syllable, consistent with the overall statistical patterns of nouns and verbs in general (Kelly 1988, Hammond 1999), etc. This phenomenon is generally familiar to foreign learners, as well as to native speakers. Instead of asking which syllable is most prominent, we asked participants to identify the POS of the target word. In the absence of syntactic/semantic contexts, participants must rely on stress for POS judgments. The method has been used in Smith 2016, in a slightly different form, for a baseline condition where the effect of stress on syntactic category identification was confirmed. However, the use of noun-verb pairs as stimuli in stress research goes back at least to Lieberman 1960.

Based on the above methodology, the prediction is that Mandarin speakers are more likely to misidentify nouns as verbs in the echo question intonation contour than English speakers.

3 Experiment

3.1 Stimuli

Two sets of stimuli were produced by a female and a male native speaker of American English, both of whom were trained linguists, using a Zoom H6 Handy Recorder in a sound-proof recording booth on Stony Brook University campus. The experimental items (listed below) were 12 pairs of noun-verb pair words, all disyllabic except for one,[2] with initial stress in the noun and final stress in the verb, respectively. The speakers were instructed to stress sufficiently on stressed syllables while sounding natural and to read the stimuli at a normal speed. For all the disyllabic forms where pronunciations with reduced vowels in the unstressed syllable are acceptable in English, the speakers were instructed to produce these forms with full vowels rather than reduced vowels, e.g. *impact* ['ɪmpækt]*/impact* [ɪm'pækt].

(3)

1. permit (n)-permit (v)	5. abstract (a)-abstract (v)	9. update (n)-update (v)
2. import (n)-import (v)	6. imprint (n)-imprint (v)	10. ferment (n)-ferment (v)
3. impact (n)-impact (v)	7. intrigue (n)-intrigue (v)	11. indent (n)-indent (v)
4. increase (n)-increase (v)	8. invite (n)-invite (v)	12. overlap (n)-overlap (v)

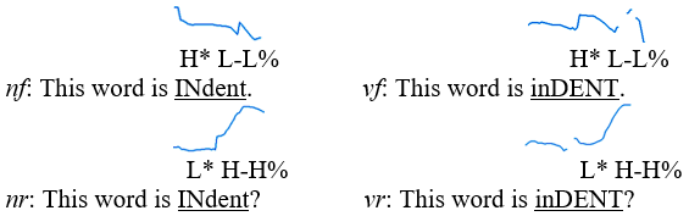
These testing items were placed in sentential contexts, with two kinds of intonation contour, namely falling (declarative) and rising (echo question). Both of them utilize the same syntactic template "This word is ___./?", in which confounding factors from syntax and semantics can be maximally obviated. The two stress patterns, indexed by part of speech (noun ≡ initial stress, verb ≡ final stress), together with the two intonation contexts form a matrix of 4 experimental conditions: noun-in-falling (*nf*), verb-in-falling (*vf*), noun-in-rising (*nr*), verb-in-rising (*vr*). A sample set is given

[1]Different types of echo questions (simple request to repeat, a signal of incredulity, indignation, etc.) may have slightly different intonation contours, and echo questions generally also have slightly different intonation patterns than yes-no questions. What were adopted here were common intonation patterns shared by these sentence types.

[2]I could not find more disyllabic pairs that need not contrast vowel quality, hence the inclusion of "overlap". Since only the initial and final syllables were concerned, the trisyllabicity of "overlap" did not matter.

below; for a complete set of stimuli, refer to Appendix. The acoustic properties of the stimuli produced and used in the experiment showed that they conformed to the general intonation patterns of declarative sentences and echo questions. The stimuli in all four sets were presented with mixed-gender voice to prevent any potentially related bias that might obscure the results.

(4)



It is worth mentioning that the four conditions have distinct pitch accents on the target word and afterwards, as displayed in (4).

The 12 pairs of lexical items were randomly divided into four groups (A, B, C, D), each assigned to one of the 4 conditions above, hence 3 tokens were collected for each condition. These groups were arranged using a Latin square design to counterbalance grouping effects, as shown in Table 1. Accordingly, every subject perceived three trials for each of the four conditions with no repeated items and at least four subjects were required in order for all lexical items on all conditions to be tested upon.

nf	*vf*	*nr*	*vr*
A	B	C	D
B	C	D	A
C	D	A	B
D	A	B	C

Table 1: Latin Square design of stimuli arrangement.

An equal number of filler items such as *pizza* and *achieve*, which do not have ambiguous part of speech nor varying stress position, were included in the experiment. Besides the environments of declarative sentence and echo question, they also appeared in regular yes/no questions, i.e. "Is this word _____?". In total, each subject heard 24 trials and the four sets of trials were randomly but evenly distributed among participants.

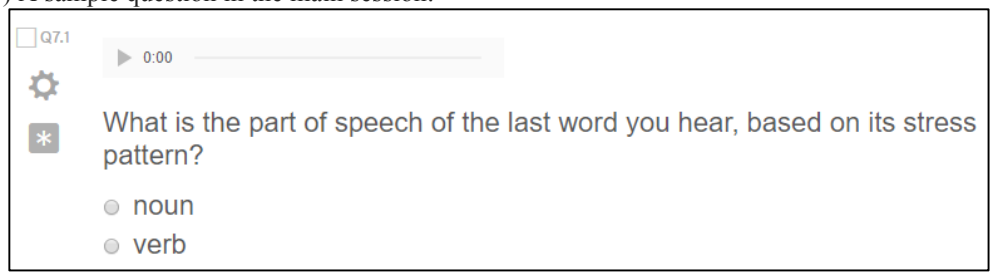
3.2 Procedure

A brief survey of basic information and language background was conducted before the testing sessions, which included two pre-tests and a main session. The first pre-test was to make sure the participant could identify basic categories such as noun, verb and adjective. The second pre-test was to inform/remind the participant of a set of items in the English lexicon that can be both noun and verb and that stress is the key distinguishing feature in the absence of context. They were first trained by listening to the sound files of some of these words accompanied by their POS displayed on the screen and afterwards were explicitly told the pattern: the noun form has initial stress and the verb form has final stress. Then they were tested on five other items to verify that they had learned/reactivated this pattern.

In the main session, participants listened to the stimuli and identified the POS of the last word of each sentence by making a forced choice between "noun" and "verb" based on its stress pattern they heard.[3] The experiment was conducted through the online research platform Qualtrics. It was self-paced and took 17 minutes on average to complete.

[3]It is also reasonable to include a third choice "not sure" to get finer-grained results. But with forced binary choice, we should be able to sufficiently observe which stress location is being perceived more favorable under a certain condition.

(5) A sample question in the main session:



3.3 Participants

73 people took part in the experiment. After filtering out data from participants whose duration of taking the experiment was unexpectedly long (threshold=78 minutes), or whose language background failed to fully match the requirement, or whose responses on the two pre-tests and the sample and practice tests failed to pass the threshold accuracy rate (only one mistake allowed), 53 participants' data was submitted for analysis. The target group consisted of 38 native Mandarin speakers (19 male, 19 female, average age=22.6 years) who learned English as a second language (including early bilinguals) and had resided in the US for at least 1 year and on average 4.77 years. The control group consisted of 15 native speakers of American English (8 male, 7 female, average age=20.7 years), mostly monolingual. All the participants were college students or recent graduates and the majority of them were affiliated with the Stony Brook University community. None reported any vision, hearing, communication or learning problems.[4]

3.4 Results

In total, 636 responses were collected for the experimental items, 159 for each condition. These responses were in the form of either "noun" or "verb", indicating different stress locations perceived by the subjects. The response of "noun" indexes stress on the initial syllable and that of "verb" indexes stress on the final syllable. The responses were converted to accuracy in terms of the actual stress position, correct response coded as 1 and incorrect 0.

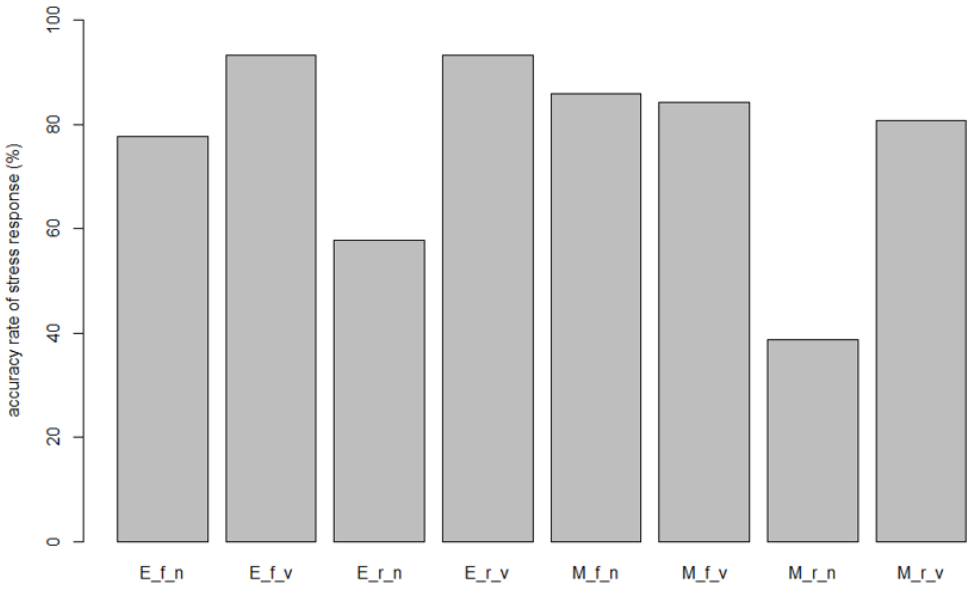


Figure 1: Accuracy rate of stress responses on all conditions in both language groups.

[4]One subject from the control group answered yes to this question, but the accuracy rate of this subject's responses was 85%, significantly higher than chance, hence no removal from the analysis.

For both English speakers and Mandarin speakers, the accuracy rate was well above or close to 80% for all four conditions (*nf*, *vf*, *nr*, *vr*) except for the condition of nouns in rising contours, whose accuracy rate was below 60%. This result suggests: i) pitch contour had an effect on stress perception of both language groups, and ii) there was an interaction between pitch contour and stress position in that the effect of pitch contour was only present in words with initial stress. Specifically, for Mandarin speakers, the difference in accuracy rate of responses on nouns between the falling contour condition (86%) and the rising contour condition (39%) was 47%, while the difference in accuracy rate on verbs was only 4% (falling contour 84% vs. rising contour 81%). In contrast, for English speakers, the difference in accuracy rate of responses on nouns between the falling contour condition (78%) and the rising contour condition (58%) was 20%. On the other hand, there was no difference in accuracy rate on verbs (falling contour 93% vs. rising contour 93%), similar to the Mandarin speakers' pattern. In summary, the major difference between English speakers and Mandarin speakers lay in the degree of influence of pitch contour on the perception of initially stressed words and the degree was reflected by difference in accuracy rate between the falling contour and the rising contour. All these effects are visible in the bar plots in Figure 1.

To test the statistical strength of the above effects, a mixed effects logistic regression model was fitted to the dataset using the *glmer* function from the *lme4* package (Bates et al 2017) in R. The dependent variable was the accuracy of participants' stress responses. The model included a full crossing of *L1* by *pitch contour* by *stress position* as fixed effects predictors and a random effects structure with a random slope for *L1* and a random intercept by *Subject*, and a random intercept by *Item*. The predictors were all binary, contrasting *English* vs. *Mandarin*, *falling* vs. *rising* contour, and *initial* (indexed as "n") vs. *final* stress (indexed as "v") respectively. They were coded with a mean of zero and a range of one, specifically *L1* ranging from -0.28 to +0.72 and the other two predictors both from -0.50 to +0.50. Table 2 shows the model.

This model shows the main effects of pitch contour and stress position, confirming that for both Mandarin and English groups, the rising pitch contour was more likely to cause misperception of stress than the falling contour, and the initial stress was more subject to misperception than the final stress. The interaction between the two variables was also significant, indicating that the effect of pitch contour was significantly stronger for words with initial stress than those with final stress. The effect of native language failed to reach significance either alone or in any interaction. This is not surprising given the great similarity between the two language groups in the accuracy of responses of verbs, which occupy half of the sample space.

Random effects:

Groups	Name	Variance	Std.Dev.
Subject:L1	(Intercept)	0.88	0.94
Item	(Intercept)	0.13	0.36
L1	(Intercept)	0.00	0.00

Fixed effects:

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	1.50	0.21	7.08	< 0.001
L1	0.58	0.41	1.43	0.15
PitchContour	1.23	0.24	5.10	< 0.001
StressPosition	-1.35	0.24	-5.60	< 0.001
L1:PitchContour	-0.97	0.57	-1.72	0.09
L1:StressPosition	-0.95	0.57	-1.68	0.09
PitchContour:StressPosition	2.14	0.48	4.41	< 0.001
L1:PitchContour:StressPosition	-1.45	1.13	-1.28	0.20

Table 2: The random effects structure and the fixed effects coefficients of the mixed model of the whole dataset. Number of observations: 636, groups: Subject, 53; Item, 12. Positive estimate reflects higher probability of correct perception of stress.

To remove the "noise" of verbs, another mixed model was fitted to the noun subset only, with a full crossing of *L1* by *pitch contour* and everything else remaining the same as in the first model. This model, as reported in Table 3, shows a significant interaction between L1 and pitch contour, indicating that the effect of pitch contour was stronger among Mandarin speakers than English speakers, when the stress was initial.

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	0.75	0.17	4.41	< 0.001
L1	0.12	0.34	0.37	0.71
PitchContour	2.03	0.30	6.66	< 0.001
L1:PitchContour	-1.43	0.60	-2.40	< 0.05

Table 3: The coefficients of the mixed model of the noun data only. Positive estimate reflects higher probability of correct perception of stress.

In contrast, the counterpart model fitted to the verb subset, as reported in Table 4, shows no significant effects of either the two predictors or their interaction, suggesting that the effect of pitch contour on the perception of final stress was simply invisible and that the two language groups behaved indistinguishably in this regard.

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	3.16	0.62	5.07	< 0.001
L1	1.38	0.93	1.48	0.14
PitchContour	0.23	0.42	0.56	0.58
L1:PitchContour	-0.32	1.04	-0.31	0.76

Table 4: The coefficients of the mixed model of the verb data only. Positive estimate reflects higher probability of correct perception of stress.

4 Discussion

The current study is an experimental investigation of the role sentence prosody plays in the auditory perception of English word stress by native speakers of Mandarin and of English. The goal of this experiment was to investigate the effects of pitch contour on Mandarin and English speakers' perception of stress position, specifically, whether the rising contour would lead to misperception of stress for Mandarin speakers when the actual stress position is on the initial syllable. The results suggest this is indeed the case. First, in the falling declarative pitch contour, Mandarin speakers matched with native speakers of English on stress identification (accuracy rate 85% vs. 86%), setting up the baseline that the Mandarin group were capable of perceiving stress in a nativelike fashion under default intonational circumstances. Second, there was a main effect of pitch contour, suggesting that both Mandarin speakers and English speakers were influenced by pitch contour in stress perception, with decreased accuracy rates (60% vs. 76%) in rising pitch conditions in general. To zoom in, the significance of the effect was found only in nouns, with accuracy rates decreasing by 47% for Mandarin speakers and 20% for English speakers from the falling contour to the rising contour, but not in verbs (4% vs. 0%). Third, although the English group were influenced by pitch contour, which was unexpected, there was a significant interaction between native language and pitch contour for words with initial stress, suggesting that the effect of pitch contour was explicitly stronger for Mandarin speakers than English speakers at least when they were perceiving initial stress.

The perceptual results of the two language groups on each condition were summarized in Table 5, as an approximation to the observed effects in the above models. Both Mandarin speakers and English speakers show some probability of misperceiving stress of initially stressed nouns in rising pitch contours, with the former being more severe than the latter (accuracy rate 39% vs. 58%). The two groups were separated further in terms of the impact of pitch contour on their perception of word-initial stress, which can be measured by the decrease in accuracy from the falling contour to

the rising contour (47% vs. 20%). The difference between them can be possibly attributed to the disproportionate use of pitch by Mandarin speakers in stress identification, while native speakers of English encompass all four cues. The impact of pitch change, such as rising, is thus greater on Mandarin speakers than on native speakers. This is corroborated by the fact that among the four conditions *nf*, *vf*, *nr*, *vr*, only the *nr* condition is problematic. Under this condition, stressed syllables take on a low pitch while unstressed syllables take on a high pitch, as opposed to other conditions where stressed syllables and unstressed syllables are generally aligned with a high pitch and a low pitch respectively. For native speakers, the effect of this "misalignment" in *nr* can be offset possibly by increased duration, or increased intensity, or simply the pitch change itself, or all of these, but can exert itself in Mandarin leaners' interlanguage system and polarize their stress perception, given their susceptibility to tone in native grammar. The current experiment shows only Mandarin speakers' sensitivity to pitch height, since all stressed syllables in the testing words are restricted to the edges, either initial or final. A future study may explore whether they are also sensitive to pitch contour in L2 stress perception by using multisyllabic words where the primary stress falls on medial positions.

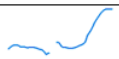
Condition	Native language	Predominantly perceived stress position
 H* L-L% *nf*: **IN**dent	English	**IN**dent (78%)
	Mandarin	**IN**dent (86%)
 H* L-L% *vf*: in**DENT**	English	in**DENT** (93%)
	Mandarin	in**DENT** (84%)
 L* H-H% *nr*: **IN**dent	English	**IN**dent (58%)
	Mandarin	in**DENT** (61%)
 L* H-H% *vr*: in**DENT**	English	in**DENT** (93%)
	Mandarin	in**DENT** (81%)

Table 5: Summary of perceptual results, in contrast with actual stress patterns (illustrated with "indent").

The findings in this study may shed some light on the theoretical debate over the existence of word stress in Mandarin. Duanmu (2007) claims that Mandarin also has lexical stress, which is realized the same way as English lexical stress. He listed some pairs of contrastive words such as *da4yi4* 'general idea' vs. *da4yi4* 'careless', where the two words share lexical tones but differ in patterns of syllable prominence, specifically full-full vs. full-light, to support his claim. According to my native intuitions of Mandarin, the contrast is more of a free variation because producing or perceiving the second syllable of the second word in its full measure of articulation will not trigger any meaning change, not even a sense of oddness. The clear-cut perceptual asymmetry between Mandarin and English speakers on word stress perception observed in this study runs counter to the implication of lexical stress as a property of Mandarin. If Mandarin lexical stress is phonologically real and is realized the same way as in English, the mapping between two stress languages would render Mandarin learners less difficulty and confusion in achieving nativelike patterns in stress perception than the results from the current study have indicated. Similarly, Kijak (2009) proposed that "stress is not operational at the word level" in Mandarin, based on her findings of Mandarin speakers' poor perception of Polish stress. The stress claimed by Duanmu is more similar to phrase stress in English than word stress, since majority of the words he listed are compounds and the "unstressed" monosyllabic words all carry a neutral tone.

5 Conclusion

This study has found that global prosody may affect both native speakers and Mandarin learners in perceiving English word stress, but the effect is more likely to be enhanced in Mandarin speakers' case. It is not too much of a surprise that the strong influence of pitch on Mandarin speakers' stress perception was borne out in an intonational context, considering the tonal status of the language. However, it is puzzling that native speakers also exhibited a decrease in accuracy in the condition of interest, i.e. words with initial stress in rising pitch contours. I suspect that lexical interference may play a role and will leave the puzzle to a future study where nonce words will be used.

References

Archibald, John. 1997. The acquisition of English stress by speakers of nonaccentual languages: Lexical storage versus computation of stress. *Linguistics,* 35:167-181.

Bates, Douglas, Martin Maechler, Ben Bolker, and Steven Walker. 2017. lme4: Linear Mixed-Effects Models using 'Eigen' and S4. R package version 1.1-13. http://CRAN.R-project.org/package=lme4. Accessed 20 August 2017.

Beckman, Mary. 1986. *Stress and Non-stress Accent.* Cambridge, UK: Foris Publications: Dordrecht.

Chao, Yuen Ren. 1980. Chinese tone and English stress. In *The Melody of Language,* ed. L. R. Waugh & C. H. van Schooneveld, 41-44. Baltimore, MD: University Park Press.

Duanmu, San. 2007. *The Phonology of Standard Chinese* (2nd ed.). New York: Oxford University Press.

Fry, Dennis Butler. 1955. Duration and intensity as physical correlates of linguistic stress. *Journal of the Acoustical Society of America,* 27:765-768.

Fry, Dennis Butler. 1958. Experiments in the perception of stress. *Language and Speech,* 1:126-152.

Fry, Dennis Butler. 1965. The dependence of stress judgments on vowel formant structure. *Proceedings of the 5th International Congress on Phonetic Science*, *Münster,* 306–311. Karger: Basel.

Hammond, Michael. 1999. *The phonology of English*. Oxford: Oxford University Press.

Howie, J. M. 1970. The Vowels and Tones of Mandarin Chinese: Acoustical Measurements and Experiments. Doctoral dissertation, Indiana University.

Howie, John Marshall. 1976. *Acoustical Studies of Mandarin Vowels and Tones*. New York: Cambridge University Press.

Jongman, Allard, Yue Wang, Corinne Moore, and Joan Sereno. 2006. Perception and production of Mandarin Chinese tones. In *Handbook of East Asian Psycholinguistics* Vol. 1: Chinese, 209-217, ed. P. Li, L.H. Tan, E. Bates, and O.J.L Tzeng. Cambridge: Cambridge University Press.

Kelly, Michael. 1988. Phonological biases in grammatical category shifts. *Journal of Memory and Language,* 27:343–358.

Kijak, Anna. 2009. How Stressful Is L2 Stress? A Cross-linguistic Study of L2 Perception and Production of Metrical Systems. Utrecht University Dissertation.

Ladefoged, Peter, & Johnson, Keith. 2011. *A Course in Phonetics* (6th ed.). Boston, MA: Wadsworth.

Lehiste, Ilse. 1970. *Suprasegmentals*. Cambridge, MA: M.I.T. Press.

Lieberman, Philip. 1960. Some acoustic correlates of word stress in American English. *Journal of the Acoustical Society of America,* 32:451-454.

Moore, Corinne, and Allard Jongman. 1997. Speaker normalization in the perception of Mandarin Chinese tones. *Journal of the Acoustical Society of America*, 102:1864–1877.

Sluijter, Agaath M. C., and Vincent J. van Heuven. 1996. Spectral balance as an acoustic correlate of linguistic stress. *Journal of the Acoustical Society of America*, 100:2471–2485.

Smith, Jennifer. 2016. Segmental noun/verb phonotactic differences are productive too. In Patrick Farrell (ed.), *Proceedings of the Linguistic Society of America*, vol. 1, article #3717. Washington, DC: Linguistic Society of America.

Wang, Qian. 2008. Perception of English Stress by Mandarin Chinese Learners of English: An Acoustic Study. Ph.D. dissertation, University of Victoria.

Yavaş, Mehmet. 2011. *Applied English phonology* (2nd ed.). Wiley-Blackwell Publishing.

Zhang, Yanhong, and Alexander Francis. 2010. The weighting of vowel quality in native and non-native listeners' perception of English lexical stress. *Journal of Phonetics,* 38:260–271.

Department of Linguistics
Stony Brook University
Stony Brook, NY 11794-4376
yaobin.liu@stonybrook.edu