THE ROLE OF INTONATION TYPE AND STRESS POSITION IN NATIVE MANDARIN SPEAKERS' PERCEPTION OF ENGLISH LEXICAL STRESS

Weijing Zhou[1], Francis Nolan[2], Jiawei Jing[3], Jian Gong[4]
[1,3] Yangzhou University, [2] University of Cambridge , [4] Jiangsu University of Science and Technology
[1] zhouwj@yzu.edu.cn, [2] fjnolan@gmail.com, [3] 674581920@qq.com , [4] habaogj@hotmail.com

ABSTRACT

This is an original empirical study investigating the extent to which intonation type (IT) and stress position (SP) influence the perception of L2 English lexical stress by native Mandarin speakers. The stimuli were 16 sets of non-vowel-reducing disyllabic English heteronyms, which differs in stress positions (trochaic vs. iambic) and intonation types (read with falling vs. rising contours). Thirty Mandarin speakers majoring in English completed a lexical stress test using two-alternatives forced-choice identification paradigm, and the accuracy rates (AR) and reaction times (RT) of their lexical stress perception were compared. The results reveal that both IT and SP had significant main effects on subjects' lexical stress judgment, that IT exhibited a more robust role than SP according to effect sizes, and that there are genuinely significant interaction effects between IT and SP signifying a complex interactive network existing in native Mandarin speakers' perception of L2 English lexical stress.

Keywords: intonation type; stress position; English lexical stress perception; native Mandarin speakers.

1. INTRODUCTION

Stress plays a pivotal role in word recognition and speech communication in English, not only because English is a typically 'stress-timed' language in rhythm, but also that stress embodies both segmental and suprasegmental features. It is acoustically manifested in multiple dimensions, especially pitch (f0), duration, intensity, and vowel quality [1] [2] [3] [4] [5] [6] [7] [8]. However, the pattern or weighing of these acoustic cues varies from lexical stress languages (like English) to languages that arguably have no stress or no easily perceived stress in their word phonology (like Mandarin) [6] [9] [10] [11] [12] [13]. Furthermore, variation exists within the family of lexical stress languages [11] [14] [15] [16]. There is also a clear bias in how native English listeners decide that a syllable is stress-bearing with an important cue being a full vowel, but with other features playing an important role [16] [17]. Such complexity is very prone to make it difficult and problematic, directly or indirectly, for non-naïve users to produce and perceive English lexical stress [12] [18] [19] [20] [21].

Research on the perception of L2 English word stress by native French listeners identified so-called 'stress deafness' among native French speakers, as stress (or accent) in French is often regarded as a phrasal phenomenon and consistently falls on the last syllable of words or phrases [19] [20]. Quite a few follow-up cross-linguistic studies supported the claim of the Stress Parameter Model (SPM) [22] [23] that typological stress similarity between L1 and L2 may determine whether the acquisition of L2 word stress is attainable, i.e., if L1 exhibits irregular and contrastive stress patterns, its native speakers tend to have an easy time to encode L2 English lexical stress; otherwise they will encounter obstructions like native French speakers' 'stress deafness'.

However, less research addresses the questions of whether learners whose L1 has a radically different type of lexical prosody (like Mandarin) can encode L2 English stress phonologically and use it for word recognition [20], and whether their L2 English stress perception is influenced by or interacts with some suprasegmental properties of words which constrain lexical access in their L1 (e.g., tones in Mandarin and pitch accents in Japanese). Both issues are worthy of investigation for researchers and practitioners to have a better understanding of native Mandarin speakers' acquisition of English lexical stress as well as the improvement on their English speech performance.

Although vowel quality has been proved to be the most important variable in processing English word stress among both native and non-native English speakers [12] [14] [15] [16] [20] [21], most previous studies have neglected stress contrasts lacking vowel reduction (e.g., PROtest /'prəʊtest/ & proTEST /prə'test/) thus leaving the picture quite incomplete.

Mandarin, as a typical tone language, does not straightforwardly have lexical stress but uses tones to distinguish lexical items, such as 妈 (/mā/ mother), 麻 (/má/ hemp), 马 (/mǎ/ horse), 骂 (/mà/ scold). A number of studies have verified that native Mandarin learners of English as a foreign language (EFL) run into considerable difficulties and pitfalls in perceiving English lexical stress due to the absence of lexical stress in L1 Mandarin word phonology – so-called negative L1 transfer [20] [21] [24] [25] – yet none employed stimuli well controlled in vowel quality, nor were they intended to explore fundamental issues such as whether intonation has a role in native Mandarin speakers' processing of L2 English lexical stress. Due to the theoretical and practical significance of these unattended questions, the present study examines the role of intonation type, stress position, and their interaction in native Mandarin speakers' perception of English lexical stress.

2. METHODS

2.1. Participants

Thirty English majors who speak Mandarin as their L1 were recruited to participate in this study. They were first-year undergraduates majoring in Business English and had just finished a compulsory term course in English Pronunciation before the experiment. They were all born in China and had never been to any English-speaking countries. Their average age was 18.5 years, and the ratio of male to female is 17% (N=5) to 83% (N=25). As their English scores in both the National College Entrance Examination and self-reported overall English proficiency (including listening, speaking, reading, writing and translation) were 'good' and 'excellent', they could be assumed to be near-advanced speakers of English with respect to the whole population of college students. All had normal or corrected-to-normal vision as well as normal hearing.

2.2. Stimuli

Table 1: The stimuli for the perception test of English lexical stress

IT	Falling ↘		Rising ↗	
SP	Trochaic	Iambic	Trochaic	Iambic
1	**DI**gest	di**GEST**	**DI**gest	di**GEST**
2	**IN**crease	in**CREASE**	**IN**crease	in**CREASE**
3	**IN**sult	in**SULT**	**IN**sult	in**SULT**
4	**SUR**vey	sur**VEY**	**SUR**vey	sur**VEY**
5	**IN**cense	in**CENSE**	**IN**cense	in**CENSE**
6	**TOR**ment	tor**MENT**	**TOR**ment	tor**MENT**
7	**IM**port	im**PORT**	**IM**port	im**PORT**
8	**TRANS**form	trans**FORM**	**TRANS**form	trans**FORM**
9	**IM**print	im**PRINT**	**IM**print	im**PRINT**
10	**IN**cline	in**CLINE**	**IN**cline	in**CLINE**
11	**IN**tern	in**TERN**	**IN**tern	in**TERN**
12	**FORE**arm	fore**ARM**	**FORE**arm	fore**ARM**
13	**MIS**fire	mis**FIRE**	**MIS**fire	mis**FIRE**
14	**TRANS**fer	trans**FER**	**TRANS**fer	trans**FER**
15	**MIS**hit	mis**HIT**	**MIS**hit	mis**HIT**
16	**TRANS**plant	trans**PLANT**	**TRANS**plant	trans**PLANT**

Notes: IT=intonation type SP=stress position

The stimuli included 16 sets of disyllabic heteronyms. Each set of words had identical vowels in stressed and unstressed syllables (e.g., /aɪ/ & /e/ in **DI**gest /ˈdaɪdʒest/ vs. di**GEST** / daɪˈdʒest/), but differed in stress positions (trochaic vs. iambic / stress located on the 1st vs. 2nd syllable) and intonation types (read with falling vs. rising contours) (see Figure 1). Therefore, the whole experiment consisted of 64 trials for identification. They were read and recorded by the second author, a native speaker of British English and phonetician. The sampling rate of recording was 44.1kHz.

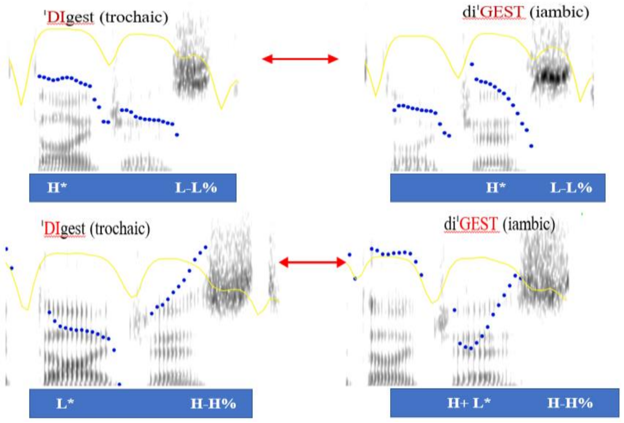

Figure 1:The Intonation patterns of *digest* with stress position on the 1st vs. 2nd syllable.

2.3. Data collection and analysis

The experiment was conducted in Yangzhou University Laboratory of Phonetics, Hearing & Cognitive Science, with the participants seated in a comfortable chair at around 70cms from an LCD monitor. The stimulus presentation program was written with E-Prime (Version 3.0). The auditory stimuli were presented binaurally through headphones at 70dB SPL. And the test follows a two-alternatives forced-choice paradigm.

The experiment consisted of 4 steps: **Step 1**: ***Introduction***. Prior to the experiment, the instructions on how to do the test and requirements for doing the task were delivered on the screen. **Step 2**: ***Rehearsal***. A practice test of eight self-paced trials plus a response after each trial was carried out so as to ensure all the participants knew exactly the procedure and the task. **Step 3**: ***Main Test***. The test consisted of 64 trials, with each trial presented only once. The presentational order of the trials was randomized across participants so that they were unlikely to be able to anticipate the upcoming trial. Each trial began with a red cross on the screen and an auditory instruction "ding" sound for 1000 milliseconds. Then a target word was binaurally delivered over the headphones and the participants were required to identify the word with perceived pattern of intonation type (IT) or stress position (SP) shown on the screen as quickly as possible without sacrificing accuracy. The data of accuracy rates (AR) and reaction time（RT）were generated and stored onto the computer right after each trial. **Step 4**: ***Ending***.

A thank-you note was presented on the screen when all the trials were over.

The data of AR and RT were then checked and calculated via Excel (Version 2017) before being statistically analysed via Jamovi (Version 1.2).

3. RESULTS

3.1. Effects of intonation type

As shown in Table 2 and Figure 2, the average RTs to perceive all the target lexical stresses in falling vs. rising intonation are 2989ms and 3260ms respectively with a gap of 271ms, indicating that the participants took longer or more effort to discriminate lexical stress embedded in rising intonation than in falling intonation, though the RT differences are not statistically significant ($F_{(29)}$ = -1.76, P=.090, Effect Size = -0.320).

Table 2: Paired-sample T-test results: The AR & RT means of the perception of target lexical stress read in falling vs. rising intonation.

	N	M	SD	T	df	P	Effect Size
F-RT (ms)	30	2989	928	-1.76	29	.090	-0.320
R-RT (ms)	30	3260	926				
F-AR (%)	30	70.71	15	4.66		.000	0.851
R-AR (%)	30	54.17	6				
OvrM-AR(%)	30	62.44	11				

Notes: F=falling intonation; R=rising intonation; M=mean
RT=reaction time; AR=accuracy rate; OvrM-AR=overall Mean of accuracy rate.

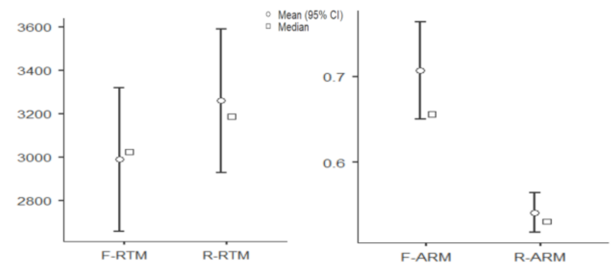

Figure 2:The AR & RT means of perception of target lexical stress read in falling (F) vs. rising (R) intonation.

As for the correct perception rates, the overall mean of correct perception rate is 62.44%. Specifically, the average perceptual AR in falling intonation is 70.71%, whereas that in rising intonation is 54.17%. The gap is 17.54%, and the AR differences are overwhelmingly significant ($F_{(29)}$ = 4.66, P< .001, Effect Size = 0.851). This proves that, although the overall average AR range is fairly moderate, the participants' perception of target lexical stress embedded in falling intonation significantly outperformed that in rising intonation, and that there seem to be a falling-intonation bias in Mandarin speakers' distinguishing lexical stress in L2 English. Also, it could be assumed that the lower AR of perceived English lexical stress in rising intonation corroborates the longer RT during the perception.

3.2. Effects of stress position

Table 3 and Figure 3 show the data of participants' perception of target lexical stress located on the 1st vs. 2nd syllable for falls, rises, and for falls and rises combined respectively. The overall RTs to discriminate the target lexical stress located on both 1st and 2nd syllables are 3106ms and 3143ms respectively with a gap of 37ms, and the RTs to judge lexical stress in sub-categorical 1st vs. 2nd syllables read in falling and rising intonation contours are 2981ms vs. 2997ms embedded in falling intonation and 3230ms vs. 3290ms in rising intonation with gaps of 16ms and 60ms respectively. The overall and sub-categorical RT differences are not statistically significant, indicating that the time the participants employed in discriminating the lexical stress is not significantly different for stress either located on the 1st syllable or on the 2nd syllable. As for the correct perception rates,

Table 3 Paired-sample T-test results: The AR & RT means of the perception of target lexical stress located on the 1st vs. 2nd syllable.

	N	M	SD	T	df	P	Effect Size
S1F-RT (ms)	30	2981	1232	-0.082	29	0.935	-0.015
S2F-RT (ms)	30	2997	869				
S1R-RT (ms)	30	3230	886	-0.416	29	0.680	-0.076
S2R-RT (ms)	30	3290	1112				
S1-RT (ms)	30	3106	928	-0.451	29	0.655	-0.082
S2-RT (ms)	30	3143	776				
S1F-AR (%)	30	72.71	23	0.558	29	0.581	0.102
S2F-AR (%)	30	68.75	26				
S1R-AR (%)	30	42.29	32	-2.269	29	**0.031**	-0.414
S2R-AR (%)	30	66.04	26				
S1-AR(%)	30	57.50	11	-2.030	29	**0.042**	-0.371
S2-AR(%)	30	67.40	18				

Notes:S1F/S2F= stress located on the 1st / 2nd syllable with falling intonation
S1R/S2R= stress located on the 1st / 2nd syllable with rising intonation
RT=reaction time AR=accuracy rate M=mean `

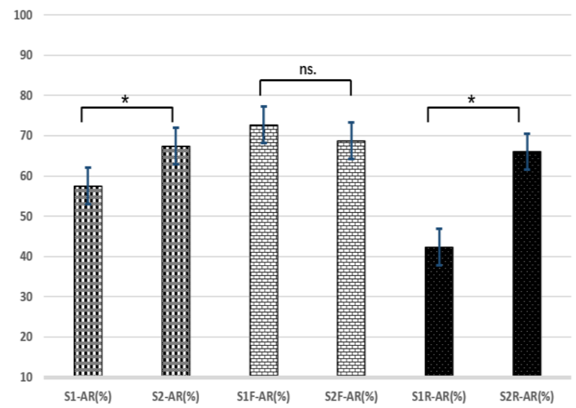

Figure 3: Paired-sample T-test results: The AR means of the perception of target lexical stress located on the 1st vs. 2nd syllable.

the average ARs of lexical stress located on the 1st vs. 2nd syllable are 57.50% and 67.40% respectively, and the gap between the two reaches significance ($F_{(29)}$ = -2.030, P= .042, Effective Size= -0.371), exhibiting a fairly certain iambic over trochaic advantage for Mandarin speakers when they do English word stress judgement. In addition, a further comparison between the perceptual ARs of word stress located on the 1st and 2nd syllable with falling intonation (72.71% & 68.75%) and those with rising intonation (42.29% & 66.04%) reveals that the AR differences in falling intonation is not significant ($F_{(29)}$ = 0.558, P= .581), whereas the difference in rising intonation is truly significant ($F_{(29)}$ = -2.269, P= .031). This not only provides robust evidence for the falling-intonation advantage in Mandarin speakers' doing English lexical stress judgement task, but also suggests there could be a potential interaction effect between intonation type and stress position in Mandarin speakers' perception of English lexical stress such that low pitch on a stressed first syllable of a disyllable causes problems.

3.3 Interaction of intonation type & stress position

Table 4: GLM repeated measures results of IT, SP and their interaction in the perception of target lexical stress.

Tests of Within-Subjects Effects						
Source	Type III Sum of Squares	df	Mean Square	F	P	η²p
IT	37749	1	37749	643.651	.000	.936
Error(IT)	1759	29	58			
SP	1841.	1	1841	18.464	.000	.302
Error (SP)	2992	29	99			
IT * SP	2430	1	2430	25.553	.000	.460
Error (IT*SP)	2853	29	95			

Notes: IT=intonation type SP=stress position

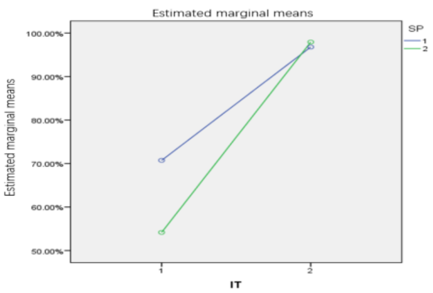

Figure 4:Within-Subjects Effects of IT & SP in the perception of target lexical stress.

To verify the possible interaction effects between intonation type (IT) and stress position (SP) in participants' lexical stress perception, GLM repeated measures were conducted to test the within-subjects effects of 2 types of IT (falling vs. rising) and 2 patterns of stress position (on the 1st vs. 2nd syllable) on the perception of target English lexical stress.

As displayed in Table 4 and Figure 4, both intonation type and stress position display significant main effects on the participants' perception of target lexical stress (IT: $F_{(1,29)}$ = 643.65, P=.000, $\eta^2 p$= .936; SP: $F_{(1, 29)}$ =18.464, P= .001, $\eta^2 p$= .302), and the effect of intonation pattern is more robust than that of stress position according to their effect sizes. This could be considered as confirmative evidence for the effective roles of both factors in Mandarin speakers' perception of English lexical stress. Moreover, there is a genuinely significant interaction effect between intonational type and stress position ($F_{(1, 29)}$ =25.553, P= .000, $\eta^2 p$= .460), demonstrating a complex interactive network existing in native Mandarin speakers' perception of English lexical stress.

4. DISCUSSION

The implications of these findings are several. First, the overall perception AR of English lexical stress embedded in 16 sets of English heteronyms (OvrM-AR=62.44%) proves that native Mandarin speakers, whose L1 does not have lexical stress, can encode L2 English stress phonologically and use it for word recognition to a great extent. This encouraging outcome could be accounted for partly by the participants' near-advanced overall English proficiency and partly by their exposure to one term of systematic instruction on English pronunciation, which helped develop their theoretical knowledge and practical competence in perceiving and producing English lexical stress. Secondly, when judging L2 English lexical stress in falling vs. rising intonation, the participants did significantly better with falling intonation than with rising intonation. Such a falling-intonation bias demonstrates the influence of L1 falling intonation in L2 lexical stress perception. Thirdly, when judging L2 English lexical stress located on the 1st vs. 2nd syllable, the participants' performance did not match native speakers' clear bias towards the initial syllable. Instead, they did better both in the 1st syllable with falling intonation and in the 2nd syllable with rising intonation, indicating native speakers' stress bias is mediated by L2 learners' suprasegmental properties of their L1 word phonology, such as lexical tones in Mandarin. Specifically, given that the learners' weakest performance occurred when stress position was on a low-pitch first syllable followed by a high-pitch terminal (S1R), Mandarin EFL learners are doomed to find it hard to associate stress with a syllable lacking high pitch and to adapt to their perception by seeking help from other parameters, such as duration and intensity. Given the predominance of falling accents in English, an association of higher pitch with stress is a reasonable generalisation and a lack of high pitch on the stressed first syllable of a disyllable causes problems for Mandarin EFL learners. That's exactly why their S1R-AR% was so strikingly low.

5. REFERENCES

[1].Beckman, M. E. 1986. *Stress and non-stress accent*. Dordrecht: Foris.

[2]. Fry, D. B. 1955. Duration and intensity as physical correlates of linguistic stress. *Journal of Acoustic Society of America, 27*(4), 155-158.

[3] Fry, D. B. 1958. Experiments in the perception of stress. *Language and Speech, 1*, 126- 152.

[4] Ladefoged, P. 2016. *A Course in Phonetics*. Boston, MA: Thomson/Wadsworth.

[5] Lehiste, I., and Fox, R. A. 1992. Perception of prominence by Estonian and English listeners. *Language and Speech, 35*(4), 419-434.

[6] Lehiste, I. 1970. *Suprasegmentals*. Cambridge: MA: MIT Press.

[7] Sluijter, A. M.C., & van Heuven, V. J. 1996a. Spectral balance as an acoustic correlate of linguistic stress. *Journal of the Acoustical Society of America*, 100(4), 2471–2485.

[8] Sluijter, A. M. C., & van Heuven, V. J. 1996b. Acoustic correlates of linguistic stress and accent in Dutch and American English. In *Proceedings of the fourth international conference on spoken language processing* 2, Philadelphia, 630–633.

[9] Xu, Y. 1997. Contextual tonal variations in Mandarin. *Journal of Phonetics*, 25, 61–83.

[10] Xu, Y. 1998. Consistency of tone-syllable alignment across different syllable structures and speaking rates. *Phonetica*, 55, 179–203.

[11] Cutler, A.2005. Lexical tress. In: *The Handbook of Speech Perception,* D. B. Pisoni and R. E. Remez (eds.), 246-289, Orford: Blackwell.

[12] Cutler, A. 2015. Lexical stress in English pronunciation. In: *The Handbook of English Pronunciation.* R. Marnie and J. M. Levis (eds.),106-124, Oxford: Blackwell.

[13] Duanmu, S. 2000. *The Phonology of Standard Chinese*. Oxford: Oxford University Press.

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

[15] Slowiaczek, L. M. 1990. Effects of lexical stress in auditory word recognition. *Language and Speech* 33: 47-68.

[16] Slowiaczek, L. M.1991. Stress and context in auditory word recognition. *Journal of Psycholinguistic Research 20*: 465-481.

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

[18] Archibald, J. 1997. The acquisition of English stress by speakers of non-accentual languages: lexical storage versus computation of stress. *Linguistics*, 35, 167-181.

[19] Dupoux, E., Pallier, C., Sebastian, N., and Mehler, J. 1997. A destressing ‘‘deafness’’ in French?. *Journal of Memory and Language*, 36, 406–421.

[20] Lee, G., Shin, D, and Garcia, M.T.M., 2019. Perception of lexical stress and sentence focus by Korean-speaking and Spanish-speaking L2 learners of English. *Language Sciences*, 72: 36-49.

[21] Zhang, Y., Nissen, S. L., and Francis, A. L. 2008. Acoustic characteristics of English lexical stress produced by native Mandarin speakers. *Journal of the Acoustical Society of America*, 123 (6), 4498-4513.

[22] Peperkamp, S.A., 2004. Lexical exceptions in stress systems: arguments from early language acquisition and adult speech perception. *Language*, 80 (1), 98– 126.

[23] Peperkamp, S., Dupoux, E., 2002. A typological study of stress ‘deafness’, *Laboratory phonology*, 7, 203–240.

[24] Zhang, Y., and Francis, A. 2010. The weighting of vowel quality in native and non-native listeners' perception of English lexical stress. *Journal of Phonetics*, 38 (2), 260– 271.

[25] Lado, R. 1957. *Linguistic Across Cultures*, Ann Arbor: University of Michgan Press.