Perception of Stress in Disyllabic Words in Changsha Xiang: The Effects of Syllable Duration and F0 Contour

ZHOU, Wangqian
The Chinese University of Hong Kong
zhouwangqian@link.cuhk.edu.hk

Lexical stress, when conceived of as a structural position that reflects prosodic headedness in the lexical domain, has been approached from the perspective of neutral/neutralised tones in Chinese [1, 2]. For instance, acoustically, Changsha tones are shortened and flattened in metrically weak positions [3, 4]. Although there have been a few studies on how metrical structures of a tone language influence its native speakers' perception of lexical stress in a second language [5, 6], as well as a number of studies on how Mandarin listeners used acoustic cues to identify and discriminate stress patterns in their native language [7, 8], the perceptual mechanisms underlying stress perception in other Chinese varieties like Changsha Xiang are insofar understudied.

To seek perceptual evidence for the exploitation of acoustic cues (syllable duration and pitch contour) in perceiving disyllabic word stress in Changsha Xiang, an AX discrimination task was used to test listeners' performance in acoustically manipulated conditions. One male and one female speaker of Changsha Xiang were recruited to produce the materials for the AX task. 12 pairs of real disyllabic words contrastive in stress position, like /ˈɕiã34 tɕiɑu33/ ('banana') and /ɕiã34 ˈtɕiɑu34/ ('to intersect'), were produced with two embedding sentences, one providing the contextual cue to its stress placement, the other being a constant metalinguistic statement. In addition, 36 pseudowords with phonotactically illegal structures in Changsha Xiang ([raka], [roko], and [riki]) were constructed. Each pseudoword was elicited following the production of a real disyllabic word that served as a tonal reference. To further clarify stress positions, all pseudowords were also embedded in two declarative carrier sentences as either a verbal or a nominal element.

The word tokens were first segmented and prosodically analysed using ProsodyPro [9]. The acoustic analyses of the stimuli corroborated previous descriptions of the phonetic implementation of Changsha stress (see Figure 1). Three acoustic conditions were then contrived by manipulating the auditory stimuli so that they might retain or obscure the syllable duration and F0 contour as cues to stress. For the both cues condition, all tokens were normalised to 600 ms total duration and 66 dB intensity. For the duration only condition, a new set of tokens was created by resetting the syllables of each right-dominant word to 340 ms and 260 ms (left-dominant tokens' average timing). For the F0 only condition, both syllables were reset to 300 ms for all words. 30 native speakers of Changsha Xiang were involved to indicate if pairs of disyllables had either the same or different stress patterns.

Three-way repeated measures analyses of variance (ANOVAs) were conducted on participants' accuracy rate, reaction time, and sensitivity d' in responding to real words or pseudowords bearing different tones across different acoustically manipulated conditions. The results showed that while both syllable duration and F0 contour provided facilitation to the discerning of trochaic and iambic stress patterns in Changsha, the effect of pitch contour is more important than the effect of syllable duration. Specifically, accuracy and sensitivity were highest when both cues were present, followed by the F0 only condition, and lowest in the duration only condition (Figure 2 and Figure 3). Additionally, the pitch cue was more helpful in processing tones that exhibited larger acoustic differences from their unstressed alternants (Tone 6, Tone 1, Tone 2 and Tone 4), resulting in higher accuracy and sensitivity, as well as shorter reaction times. The pitch cue was also particularly effective in perceiving real words, which is likely due to the benefit from semantic knowledge in top-down processing.

The results challenged the Functional Load Hypothesis of stress marking [10, 11] by showing that pitch information remains paramount in perceiving lexical prominence despite its role in making lexical tone distinctions. This underscores the need to reevaluate prevailing theories regarding stress perception in tone languages. Moreover, the results also called for further acoustic, perceptual, typological and second language research. For example, they provided valid reference for research into Chinese listeners' perceptual learning of stress in other languages, and also the employment of cues in perceiving different levels of prominence.

Keywords: perceptual cue, metrical stress, tonal neutralisation, Changsha Xiang

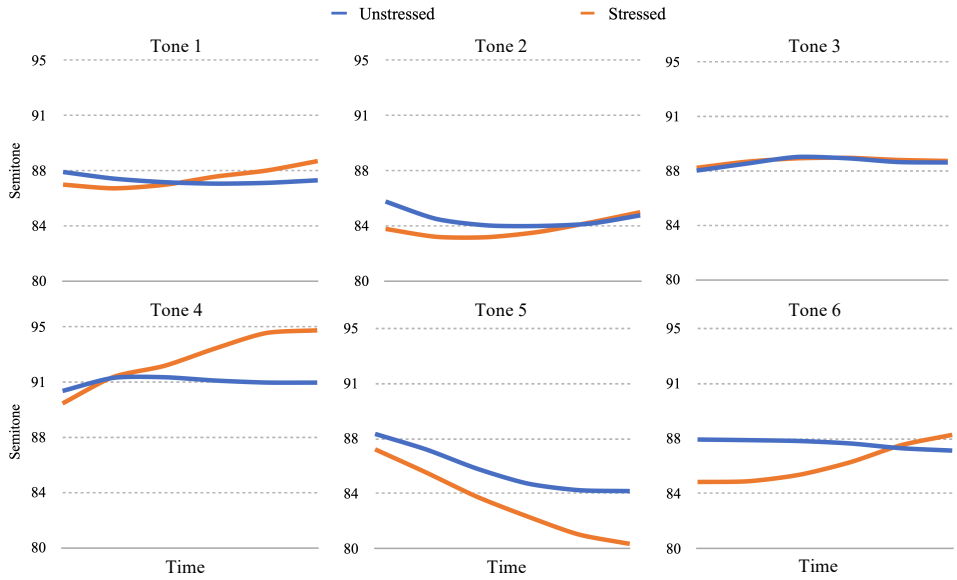

Figure 1. Time-normalised pitch contours of Changsha tones in the second syllable of right- dominant (stressed) and left-dominant (unstressed) disyllables in semitone

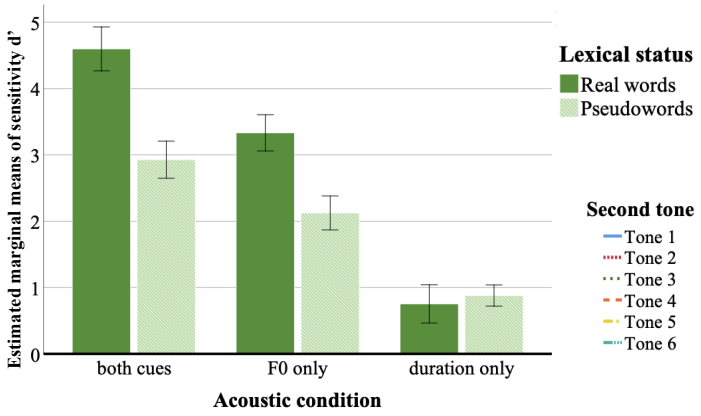

Figure 2. Estimated marginal means of sensitivity d' across lexical statuses and acoustic conditions

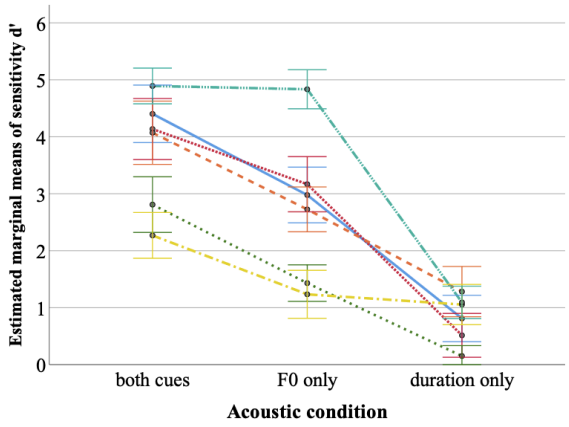

Figure 3. Estimated marginal means of sensitivity d' across acoustic conditions and second tone categories

Selected References

[1] Duanmu, S. (2007). *The Phonology of Standard Chinese*. Oxford University Press.

[2] Feng, S. (2017). *Prosodic Morphology in Mandarin Chinese*. Routledge.

[3] Guo, C., & Chen, F. (2022). Phonetic realizations of metrical structure in tone languages: Evidence from Chinese dialects. *Frontiers in Psychology*, *13*, Article e945973. https://doi.org/10.3389/fpsyg.2022.945973

[4] Zhang, M. (2021). Positional tone sandhi in Changsha Xiang: A production study [conference presentation abstract]. In *1st International Conference on Tone and Intonation (TAI), Sonderborg, Denmark: Abstract book* (pp. 138-139). http://dx.doi.org/10.13140/RG.2.2.31526.86087

[5] Yu, V. Y., & Andruski, J. E. (2010). A Cross-Language Study of Perception of Lexical Stress in English. *Journal of Psycholinguistic Research, 39*(4), 323–344. https://doi.org/10.1007/s10936-009-9142-2

[6] Guo, X., & Chen, X. (2022). Perception of English Stress of Synthesized Words by Three Chinese Dialect Groups. *Frontiers in Psychology*, *13*, 803008–803008. https://doi.org/10.3389/fpsyg.2022.803008

[7] Wang, Y. (2004). The effects of pitch and duration on the perception of the neutral tone in Standard Chinese. *Acta Acustica, 29*(5), 453-461. https://dx.doi.org/10.15949/j.cnki.0371-0025.2004.05.013

[8] Fan, S., Li, A., & Chen, A. (2018). Perception of lexical neutral tone among adults and infants. *Frontiers in Psychology, 9*, Article e00322. https://doi.org/10.3389/fpsyg.2018.00322

[9] Xu, Y. (2013). ProsodyPro — A tool for large-scale systematic prosody analysis. In *Proceedings of Tools and Resources for the Analysis of Speech Prosody (TRASP 2013)*, Aix-en-Provence, France. 7-10.

[10] Berinstein, A. E. (1979). A cross-linguistic study on the perception and production of stress. *UCLA Working Papers in Phonetics, 47*, 1-59.

[11] Van Heuven, V. J., & Turk, A. (2020). Phonetic correlates of word and sentence stress. In C. Gussenhoven, & A. Chen (Eds.), *The Oxford Handbook of Language Prosody* (pp.1501–1565). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780198832232.013.8