Acoustic and articulatory correlates of lexical-stress in Mandarin accented English

Boram Kim
CUNY Graduate Center, CUNY, Haskins Laboratories

INTRODUCTION

- L2 sound contrasts that do not exist in the L1 system are often explained in connection with the L1 system. However, this phenomenon has been largely discussed in segmental level rather than suprasegmental level (e.g., lexical stress).
- L2 speakers' articulatory realization of suprasegmental factor has been the focus of less research, and it is, therefore less well understood.
- In American English, both spectral and temporal cues are used to realize tense-lax contrasts [1-2] and lexical-stress [3-7].
- Mandarin L2 speakers exhibited limited use of spectral cues realizing tense-lax distinction [8-10], and lexical-stress [11-14].
- The current study combines these segmental (tense-lax contrasts) and suprasegmental (lexical-stress) factors to examine their interaction.
- It investigated how Mandarin L2 speakers resemble or differ from L1 speakers in using spectral and temporal cues, when tense-lax contrasts and lexical-stress are examined together.

Hypothesis
Mandarin L2 speaker retains the tense-lax contrast that are modulated by lexical stress in a temporal domain but not in spectral domain.

METHOD

- Subset of publicly available Mandarin Accented English Electromagnetic Articulography Corpus was used [24].
- A total 12 speakers' data was analyzed; 6 speakers (3 M; 3 W) from L1 and L2 groups, respectively.
- Target vowels were stressed and unstressed /i - ɪ/ and stressed /u - ʊ/, as unstressed /u - ʊ/ tokens did not occur in sufficient numbers to run statistics (total 3859 tokens).
- Vowel duration (ms) was calculated based on the start and the end of target vowels
- First two formants were extracted using seeding method [25].
- Tongue tip (TT) and tongue dorsum (TD) locations were extracted from the mid-point of each vowel.
- Outliers of both acoustic and articulatory data were removed using elbow method [26] and normalized within each speaker, using a z-score.
- To examine how well the tense-lax pair is separated in spectral domain, the proportion of the overlap area of the vowel ellipses was calculated.

BACKGROUND

Vowel system of Mandarin and English

English [1-2]

- English tense vowels are articulated in more peripheral locations of the vocal tract than lax vowels and feature longer durations than do lax vowels [1-2].

Mandarin [10]

- Mandarin consists of smaller vowel inventory, /a, i, u, y, ə/.
- Mandarin lacks tense-lax distinction.
- In previous studies, when realizing English tense-lax contrast, Mandarin L2 speakers relied more on temporal cue than spectral cues [8-10].

Lexical-stress in Mandarin and English

English

- Stressed syllables are produced with longer **durations**, increased intensities as well as clearer spectral information, relative to unstressed syllables [3-7].
- Articulatory movement becomes larger, longer and faster in prominent speech units by L1 speakers [15-19].
- **Vowel quality** (i.e., full vs reduced) plays crucial role to realize lexical-stress [3,7,12-13].

Mandarin [20]

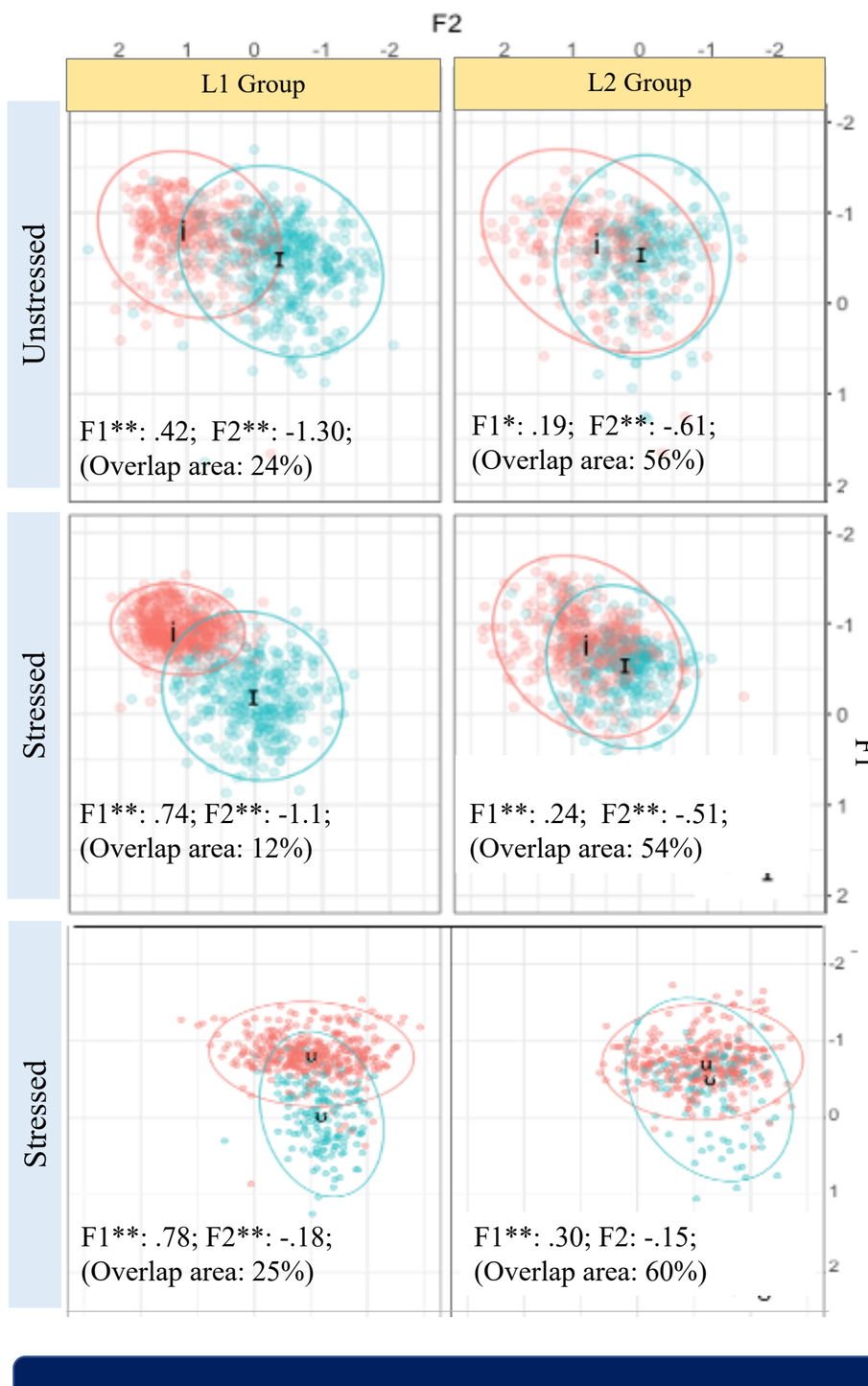
- Lexical tone: full tone *vs* neutral tone
- Stress in Mandarin is encoded in the lexical entry (e.g., *-zi*, *-le*).
- Unstressed morpheme is limited to the syllable that carries a neutral tone.
- Example of contrastive stress (the first syllable must be stressed)

东西	*dōngxī*	'east-west'	(stressed-stressed)
	dōngxi	'stuff'	(stressed-unstressed)

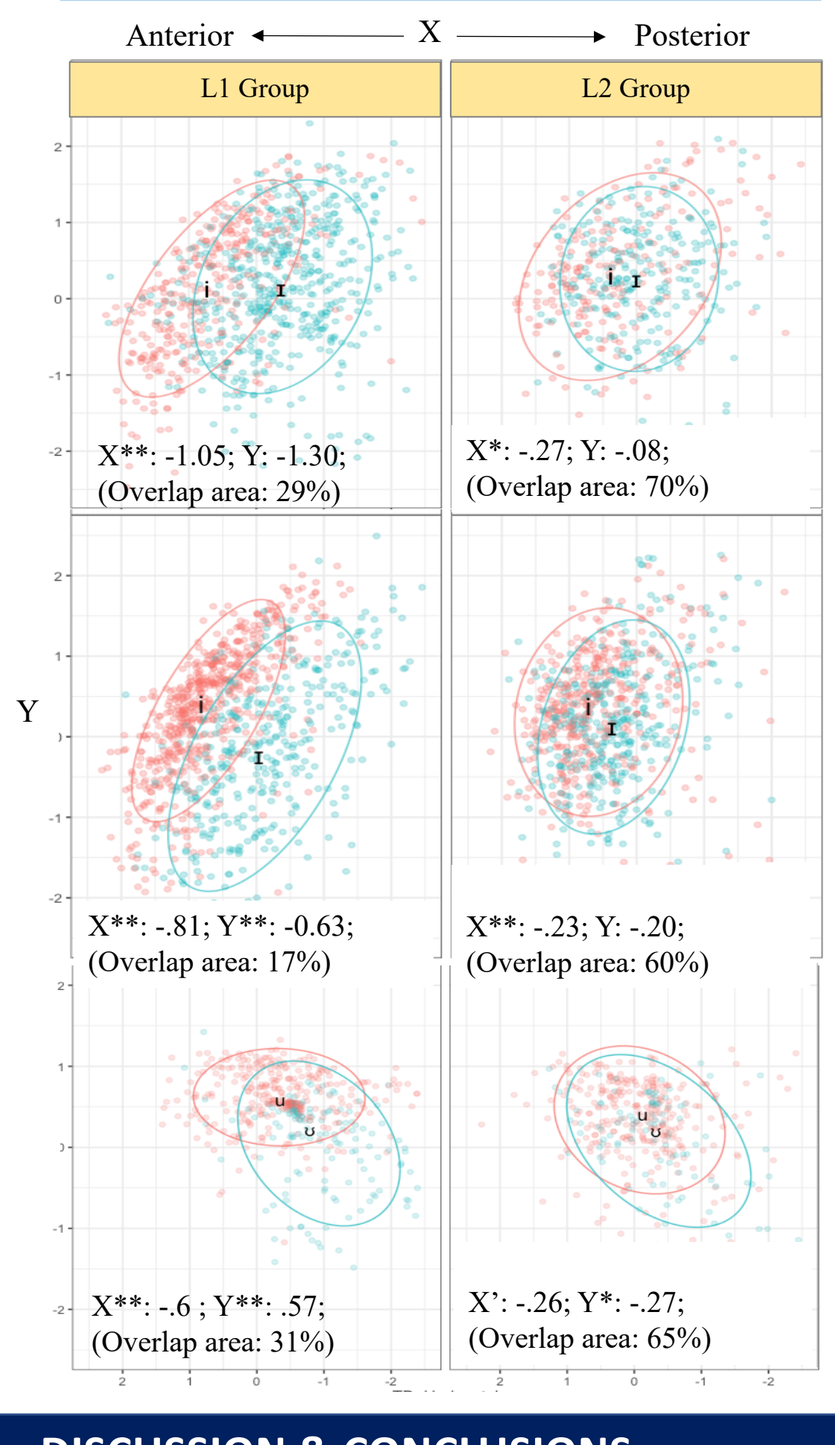
- **Duration** is the primary cue to the stress contrast [21-22].
- F0 is the secondary cue: F0 of an unstressed syllable carrying a neutral tone depends on the lexical tone of the preceding syllable.
- **Vowel quality** is not considered important as a cue to Mandarin stress, at least as compared to English stress [21,23].

RESULTS

Spectral domain: formants

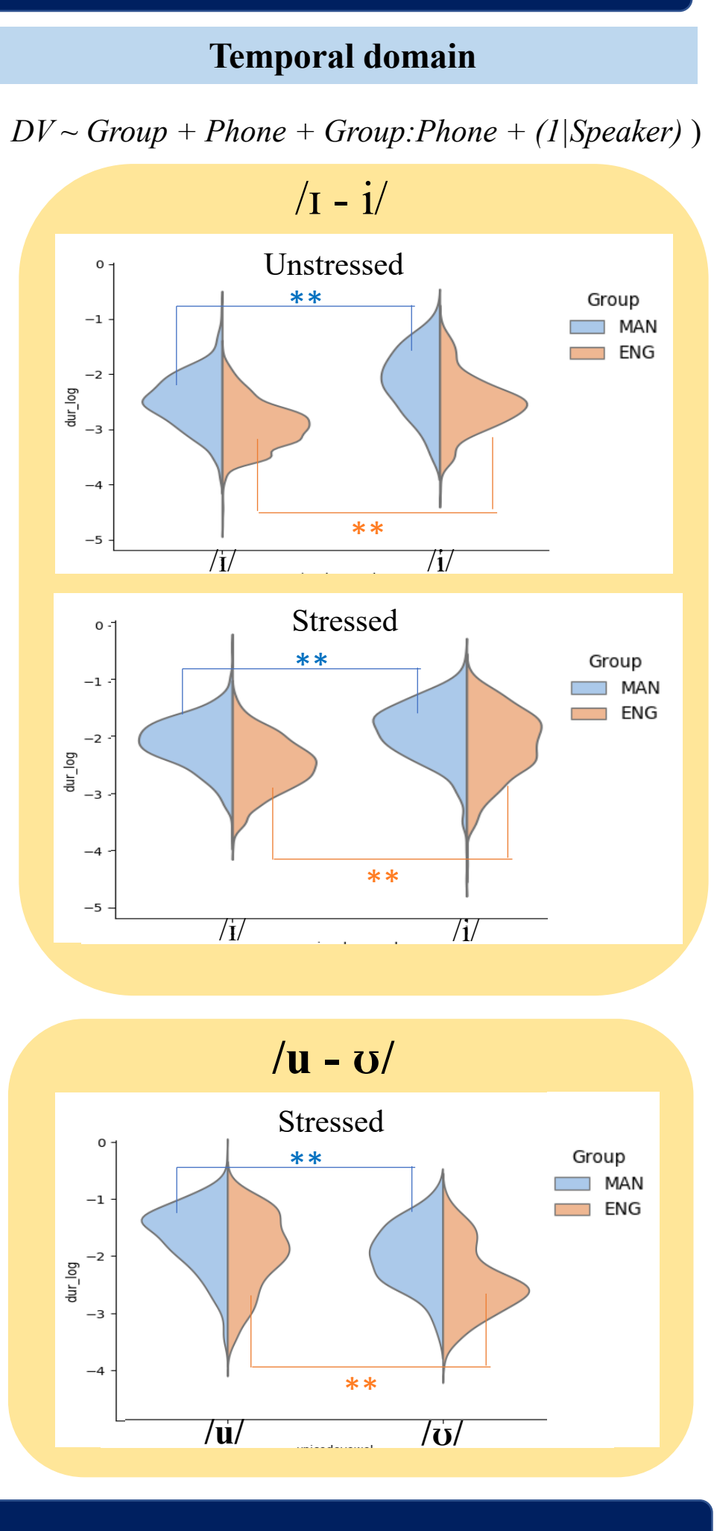


Spectral domain: TT (/i - ɪ/) and TD (/ʊ- u/)



Temporal domain

DV ~ Group + Phone + Group:Phone + (1|Speaker))



DISCUSSION & CONCLUSIONS

- Mandarin L2 speakers merged tense-lax contrasts in the spectral domains showing larger overlap in vowel ellipses, regardless of stress conditions.
- Mandarin L2 speakers retained the contrasts in temporal domain in both stressed and unstressed conditions.
- Examining F0 would be interesting as F0 is a secondary cue to Mandarin lexical-stress, while it is a cue to phrasal prominence (i.e., sentential stress) in English.
- Future study is needed to examine other articulatory gestures such as lips aperture, and/or jaw movements to better understand the articulatory correlates of lexical-stress and the kinematics of L2 vowel production.

[1] Ladefoged, P. (2005). *Vowels and Consonants*; [2] Ladefoged, P., & Maddieson, I. (1996). *The sounds of the world's languages* (Vol. 1012). Blackwell Oxford. [3] Cambier-Langeveld, T., & Turk, A. E. (1999). A cross-linguistic study of accentual lengthening: Dutch vs. English. *Journal of Phonetics*, *27*(3), 255–280. [4] Cole, J., Mo, Y., & Hasegawa-Johnson, M. (2010). Signal-based and expectation-based factors in the perception of prosodic prominence. *Laboratory Phonology*, 1(2), 425–452. [5] Fry, D. B. (1958). Experiments in the perception of stress. *Language and speech*, 1(2), 126-152. [6] Kochanski, G., Grabe, E., Coleman, J., & Rosner, B. (2005). Loudness predicts prominence: Fundamental frequency lends little. *The Journal of the Acoustical Society of America*, 118(2), 1038–1054. [7] Sluijter, A. M., & Van Heuven, V. J. (1996). Spectral balance as an acoustic correlate of linguistic stress. *The Journal of the Acoustical Society of America*, 100(4), 2471–2485. [8] Chen, Y. (2006). Production of Tense-Lax Contrast by Mandarin Speakers of English. *Folia Phoniatrica et Logopaedica*, 58(4), 240–249. [9] Jiang, H. (2008). Effect of L2 phonetic learning on L1 vowels. Dept. of Linguistics-Simon Fraser University. [10] Lin, T., & Wang, L. (2001). Yuyinxue Jiaocheng (A course in phonetics). Daxue Chubanshe (Peking University Press). [11] Campbell, N., & Beckman, M. (1997). Stress, prominence, and spectral tilt. *Intonation: Theory, Models and Applications*. [12] Cutler, A., Mehler, J., Norris, D., & Segui, J. (1986). The syllable's differing role in the segmentation of French and English. Journal of Memory and Language, 25(4), 385–400. [13] Cutler, A., & Van Donselaar, W. (2001). Voornaam is not (really) a homophone: Lexical prosody and lexical access in Dutch. Language and Speech, 44(2), 171–195. [14] Sluijter, A. M., & Van Heuven, V. J. (1996). Spectral balance as an acoustic correlate of linguistic stress. The Journal of the Acoustical Society of America, 100(4), 2471–2485. [15] Beckman, M. E. (1992). Prosodic structure and tempo in a sonority model of articulatory dynamics. *Papers in Laboratory Phonology II: Segment, Gesture, Prosody*. [16] Beckman, M. E., & Edwards, J. (1994). Articulatory evidence for differentiating stress categories. Papers in Laboratory Phonology III: Phonological Structure and Phonetic Form, 7–33. [17] De Jong, K. J. (1995). The supraglottal articulation of prominence in English: Linguistic stress as localized hyperarticulation. The Journal of the Acoustical Society of America, 97(1), 491–504. [18] Harrington, J., Fletcher, J., & Beckman, M. (2000). Manner and place conflicts in the articulation of accent in Australian English. Papers in Laboratory Phonology V, 40–51. [19] Cho, T. (2006). Manifestation of prosodic structure in articulatory variation: Evidence from lip kinematics in English. *Laboratory Phonology*, *8*, 519–548. [20] Peng, S., Chan, M. K., Tseng, C., Huang, T., Lee, O. J., & Beckman, M. E. (2005). Towards a Pan-Mandarin system for prosodic transcription. Prosodic Typology: The Phonology of Intonation and Phrasing, 230–270. [21] Lin, T. (1985). Tantao Beijinghua qingyin xingzhi de chubu shiyan [On neutral tone in Beijing Mandarin]. In S. Hu (Ed.), Beijing yuyin shiyanlu (Working papers in experimental phonetics, pp. 1–26). Beijing: Beijing Daxue chubanshe [Peking University Press]. [22] Qin, Z., Chien, Y.-F., & Tremblay, A. (2017). Processing of word-level stress by Mandarin-speaking second language learners of English. *Applied Psycholinguistics*, 38(3), 541–570. [23] Zhang, Y., & 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. [24] Ji, A., Berry, J., & Johnson, M. T. (2014). The Electromagnetic Articulography Mandarin Accented English (EMA-MAE) corpus of acoustic and 3D articulatory kinematic data. 7719–7723. [25] Chen, W., Whalen, D. H., & Shadle, C. H. (2019). F0-induced formant measurement errors result in biased variabilities. The Journal of the Acoustical Society of America, 145(5), EL360–EL366. [26] Whalen, D. H., Chen, W.-R., Tiede, M. K., & Nam, H. (2018). Variability of articulator positions and formants across nine English vowels. *Journal of Phonetics*, *68*, 1–14.