# An Acoustic and Ultrasound Study Revealing Why Mandarin Speakers Struggle with English Vowel Sounds

Yao-Zhen Zeng and Li-Hsin Ning

*National Taiwan Normal University*

***Background***: Producing the tense/lax vowel contrast in English poses a challenge for second language (L2) learners, especially Mandarin speakers lacking this distinction. According to Speech Learning Model [1], the major difficulty might result from grouping perceptually similar tense/lax vowels into a single category in their native language [2, 3] since the absence of language-specific categories gives rise to a foreign accent in their L2. Previous research studies [4-6] have used ultrasound tongue images to improve the tense/lax vowel differentiation among L2 learners. The findings suggest that these images can aid L2 learners in distinguishing the vowel contrasts, thereby enhancing both their production and perception. Expanding upon this concept, our study aimed to examine two types of vowel categories: (1) **vowel-tenseness** (tense/lax), and (2) **vowel-height** (mid/low), particularly in relation to levels of English proficiency. We sought to determine whether speakers with varying proficiency levels utilize unique articulatory and acoustic cues to produce these two vowel categories in English, with the objective of observing any differences.

***Methods***: Four English L2 learners were recruited at National Taiwan Normal University, with two classified as "advanced" (English majors) and two as "elementary" (freshmen in beginner English classes). Both groups were recorded using an Articulate Instruments Micro ultrasound system with a 4MHz element convex transducer, imaging at a depth of 90 mm and a field of view of 133°. The ultrasound probe was stabilized using an Articulate Instruments headset, and images were recorded at approximately 91 frames per second. Acoustic signals were recorded at 22.05 kHz using a Focusrite Scarlett audio interface. English vowels were elicited in words with and without /h_d/ context. Two types of vowel categories were examined: (1) **vowel-tenseness** for /i/-/ɪ/ and /u/-/ʊ/ (**TENSENESS** × **PROFICIENCY**), and (2) **vowel-height** for /ɑ/-/ʌ/ and /ɑ/-/ɔ/ (**HEIGHT** × **PROFICIENCY**). Three acoustic parameters, including *F1*, *F2*, and *duration*, were analyzed. Formant values were standardized to z-scores for individual comparison. For ultrasound tongue splines, cartesian coordinates were scaled using the minimum y-coordinate values of the participant as the baseline and then converted to polar coordinates. Before conducting the analysis, any incorrectly produced vowels were excluded, leaving a total of 48 tokens under investigation.

***Results***: **Fig. 1.** illustrates that in **vowel-tenseness**, elementary learners (ELs) faced greater difficulty distinguishing the tense/lax contrast compared to advanced learners (ALs). In **vowel-height**, however, both groups effectively distinguished the pairs. The ANOVA results showed that in *F1*, the main effect of **TENSENESS** for the /i/-/ɪ/ pair [$F(1, 11) = 9.48$, $p < .05$] revealed significant differences, suggesting better distinction for high-front vowels compared to high-back vowels. A significant main effect of **HEIGHT** was found for the /ɑ/-/ɔ/ pair [$F(1, 9) = 10.31$, $p < .05$], indicating a clear contrast between low unrounded vowels and mid rounded vowels. Regarding **PROFICIENCY**, significant differences were observed only for the **vowel-tenseness** /u/-/ʊ/ pair [$F(1, 9) = 8.70$, $p < .05$], suggesting varied performance for high-back vowels between the groups. As for *F2*, significant main effects of **HEIGHT** were found for the /ɑ/-/ʌ/ pair [$F(1, 8) = 5.92$, $p < .05$] and the /ɑ/-/ɔ/ pair [$F(1, 9) = 7.72$, $p < .05$], with clear distinction observed for both mid/low back vowel pairs. As for *duration*, a significant main effect of **PROFICIENCY** [$F(1, 24) = 4.54$, $p < .05$] in **vowel-tenseness** was observed, with ALs producing longer durations than ELs (**Fig. 2**). Finally, the Generative Additive Mixed Models (GAMMs) fitted to tongue contours indicated that ELs struggled to differentiate /i/-/ɪ/, /u/-/ʊ/, and /ɑ/-/ʌ/ pairs, with substantial overlap in tongue contours, suggesting challenges in mastering these contrasts of vowel pairs (**Fig. 3**).

***Discussion***: Our acoustic and articulatory findings suggest that the **vowel-height** pair /ɑ/-/ɔ/ was well-performed in both groups, as the participants could distinguish them effectively. However, for the pairs /i/-/ɪ/ and /ɑ/-/ʌ/, despite acoustic analyses revealing no significant differences between the groups, a closer examination of tongue contour images showed that ELs found their distinctions to be challenging, which could be attributed to their absence in Mandarin as well as the less developed English proficiency. Furthermore, this finding also suggests that ultrasonic data might offer valuable insights into subtle articulatory differences that may underlie learning difficulties not evident in the acoustic analyses alone.

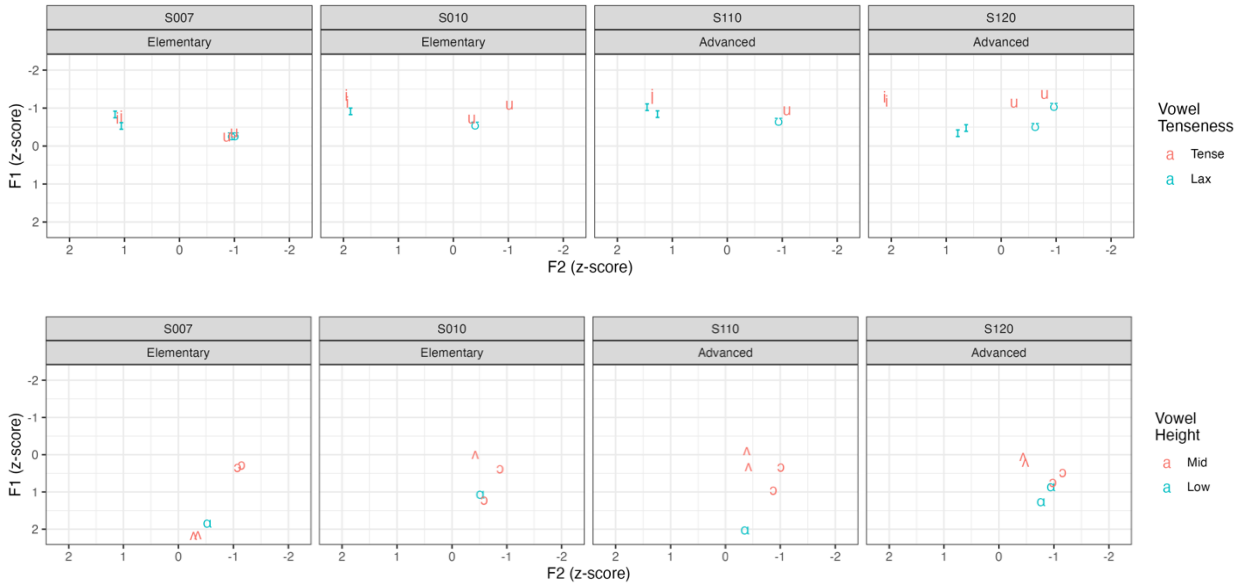


**Fig. 1.** IPA symbol for **vowel-tenseness** and **vowel-height** plotted on a vowel chart for each participant

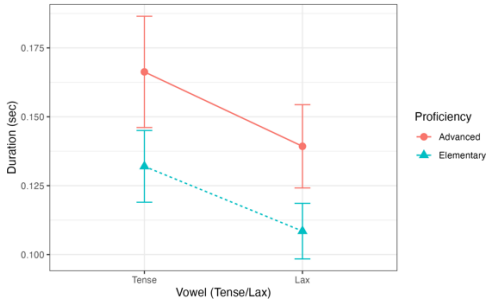


**Fig. 2.** The interaction between **TENSENESS** and **PROFICIENCY** for *duration*

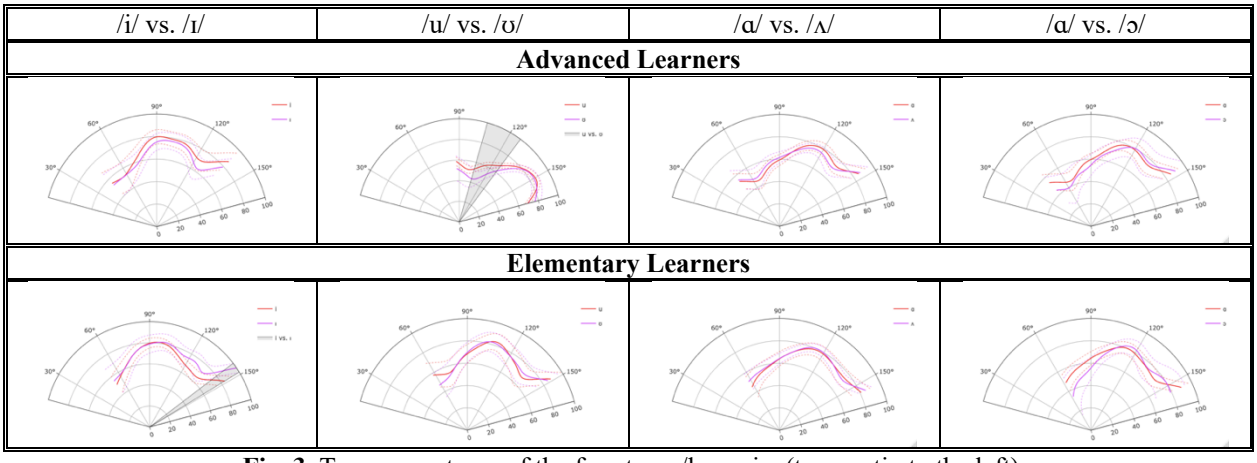


**Fig. 3.** Tongue contours of the four tense/lax pairs (tongue tip to the left) for advanced learners (top) and elementary learners (bottom)

## References


[1] J. E. Flege, "Second language speech learning: Theory, findings and problems," in *Speech Perception and Linguistic Experience: Theoretical and methodological issues in cross-language speech research*, W. Strange Ed.: Baltimore: York Press, 1995, pp. 233-277.

[2] J. E. Flege, "The production of "new" and "similar" phones in a foreign language: evidence for the effect of equivalence classification," *Journal of Phonetics,* vol. 15, no. 1, pp. 47-65, 1987.

[3] J. E. Flege, O.-S. Bohn, and S. Jang, "Effects of experience on non-native speakers' production and perception of English vowels," *Journal of Phonetics,* vol. 25, no. 4, pp. 437-470, 1997.

[4] Y.-h. S. Chang, "Effects of Production Training With Ultrasound Biofeedback on Production and Perception of Second-Language English Tense–Lax Vowel Contrasts," *Journal of Speech, Language, and Hearing Research,* vol. 66, no. 5, pp. 1479-1495, 2023.

[5] B. Kühnert and C. Pillot-Loiseau, "Teaching pronunciation with direct visual articulatory feedback: pedagogical considerations for the use of ultrasound in the classroom," *Recherches Anglaises et Nord Americaines,* vol. 55, pp. 9-24, 2022.

[6] J. Y. Song and F. Eckman, "Using ultrasound tongue imaging to study covert contrasts in second-language learners' acquisition of English vowels," *Language Acquisition,* vol. 28, no. 4, 2020.