Proceedings of ICADSS 2026 Symposium: Interpreting Culture: Context in Art, Literature, and Mass Media
DOI: 10.54254/2753-7064/2026.HT35879

A Study of English /ɪ/ Production among Mandarin ESL Learners Across Different Proficiency Levels

Jiaying Li

Faculty of Education, Universiti Kebangsaan Malaysia, Bangi, Malaysia
18735515335@163.com

Abstract. The English /iː/–/ɪ/ contrast poses persistent learning difficulties for Mandarin Chinese speakers. While perceptual challenges are well documented, systematic acoustic studies on /ɪ/ production across proficiency levels remain scarce. This study examined the acoustic features of English /ɪ/ produced by 16 Mandarin-speaking EFL learners (8 low proficiency, 8 high proficiency) from northern China. Participants read five /ɪ/-containing words (sit, bit, hit, ship, live). F1 and F2 values at vowel midpoint were extracted and compared with native English reference values. Results showed that the low proficiency group produced a numerically lower F1 (382.37 Hz) than the high proficiency group (397.87 Hz), suggesting a higher tongue position, though the difference was not significant ($p = .372$). F2 values were similar across groups (2124.32 Hz vs. 2144.66 Hz, $p = .793$). The non-significant results are attributed to the small sample size and lack of normalization, which led to large individual variation. Pedagogically, the F1 trend suggests that low proficiency learners would benefit from instruction targeting tongue height adjustment to produce a lower /ɪ/ closer to the native target.

Keywords: English /ɪ/ production, Mandarin EFL learners, proficiency level, vowel formants, acoustic analysis

1. Introduction

The contrast between tense and lax vowels /iː/–/ɪ/ is a recognized acquisition difficulty for native Chinese speakers [1, 2]. This difficulty is fundamentally rooted in the phonological differences between the two languages: Mandarin Chinese does not have a phonemic duration contrast and does not use the tense-lax feature in a contrastive way within its vowel system [3]. Therefore, Chinese learners must establish a new phonemic category for /ɪ/ that has no direct equivalent in their native language.

A large number of studies have focused on the perceptual dimension of this acquisition difficulty. Mi et al. [4] demonstrated that native Chinese listeners rely more on duration cues when identifying English vowels than native English listeners, and that their discrimination performance significantly declines when the duration information is balanced. Yang et al. [5] further found that learners have individual differences in their perception of /i/ and /ɪ/, and these are related to learning styles. Kwon and Starr [6] provided an explanation from a phonological perspective, suggesting that learners' perceptual difficulties stem from the absence of the [tense] feature in the mother tongue

DOI: 10.54254/2753-7064/2026.HT35879

phonological system. Yihan [7] extended this exploration to cross-language vowel perception, revealing a consistent native language transfer effect. Meanwhile, Zheng and Wonnacott [8] indicated that explicit phonological teaching can reshape the cue weight strategies in second language learners' perceptions, suggesting that reliance on duration can be changed through training.

The relationship between perception and output in bilingual phonology has always been a topic of continuous attention. Isbell [9] reviewed that there is only a moderate correlation between perception and output, and their development is not synchronous. Empirical studies on Chinese learners support this view. Zhang et al. [10] directly compared the perception and output of English vowels, finding that learners' perception accuracy was higher than their output accuracy; learners could distinguish /iː/–/ɪ/ in perception but often failed to achieve this contrast in output through spectrograms. Wang, Guo and Feng [11] further reported that the gap between perception and output was more prominent among low-level learners. Wang [12] emphasized that progress at the perceptual level, such as adjusting cue weights through training, does not automatically transfer to the output end, and it is necessary to simultaneously examine both modalities in the same group of learners.

2. Literature review

Compared with the rich perception research, systematic acoustic examinations of /ɪ/ production across proficiency levels remain limited. Early studies established that Chinese learners produce /ɪ/ with lower F1 (higher tongue position) and higher F2 (more anterior) than native speakers, along with significantly shorter duration [13, 14]. Munro and Derwing [2] found that L2 vowel production improves in the first two years but may stabilize thereafter, with some non-native features like high tongue position becoming fossilized. Feng and Wang [15] reported duration ratios (/ɪ/–/iː/) of 0.58 for low-level learners, 0.66 for high-level learners, and 0.71 for native speakers, with high-level learners' formants approaching native norms while low-level learners maintained lower F1 and higher F2.

Beyond proficiency, multiple factors shape L2 vowel production. Input quantity positively affects tense-lax vowel production in young learners [3]. Dialect background interacts with Mandarin to influence English vowel production [16, 17]. The native writing system (Chinese characters vs. pinyin) affects L2 phonological knowledge [18], and systematic phonics instruction improves decoding ability [19].

Gender has also drawn attention in L2 phonology. Casillas [20] incorporated gender in analyzing /i/–/ɪ/ production among early learners, and Wang, Guo and Feng [11] examined its influence on Chinese learners' tense-lax vowel production. Given vocal tract differences between males and females and potential variation in pronunciation strategies, it is necessary to examine whether gender modulates /ɪ/ production and interacts with proficiency.

Despite these contributions, the acoustic trajectory of /ɪ/ as proficiency increases remains underexplored. In particular, it is unclear whether low-proficiency learners rely primarily on duration shortening, while high-proficiency learners approach native spectral quality through formant adjustment, and whether gender moderates this trajectory. This study addresses these gaps by comparing /ɪ/ production across two proficiency groups in both temporal and spectral domains, with the following research question: What are the differences in duration when low- and high-proficiency Chinese learners produce English /ɪ/?

DOI: 10.54254/2753-7064/2026.HT35879

3. Method

3.1. Participants

Sixteen Mandarin-speaking EFL learners (8 males, 8 females) from northern China participated. Their native dialects included Shanxi (13), Northeast (1), Henan (1), and Shandong (1). All used Mandarin as their daily language. Participants were divided into two proficiency groups based on standardized test scores: Low Proficiency (n=8, 4M/4F; CET-4<500 or IELTS<5.5) and High Proficiency (n=8, 4M/4F; CET-4>500 or IELTS≥6). All were undergraduates aged 18-25 with no speech or hearing impairments. A background questionnaire collected age of English onset, daily English use, and prior training [3, 18]. No native English control group was recorded; reference values were obtained from Feng and Wang [15] and Chen et al. [13].

3.2. Materials and procedures

Stimuli included five /ɪ/-containing words (sit, bit, hit, ship, live) embedded in "Say ___ again." Each sentence was read twice, yielding 10 tokens per word (50 total per participant). The recording session lasted approximately 2 minutes per participant.

3.3. Acoustic analysis and statistical procedures

Acoustic analysis was performed using Praat. Vowel onset and offset were manually annotated based on spectrographic and waveform criteria. Duration was calculated as the interval between onset and offset. F1 and F2 were extracted at the vowel midpoint using the Burg LPC algorithm (5 formants; maximum frequency: 5000 Hz for males, 5500 Hz for females). All formants were manually checked and corrected. Since only one vowel was collected per participant, normalization requiring multiple vowel categories (e.g., Lobanov) could not be applied. F1/F2 were retained in Hz and compared separately by gender due to vocal tract length differences, while duration was analyzed across all participants.

Due to the small sample size (n=8 per group), non-parametric Mann-Whitney U tests were used to compare F1, F2, and duration between proficiency groups. All statistical analyses were conducted using SPSS version 27.

4. Results

4.1. Differences between F1 and F2

As shown in Table 1, the F1 value of the low English proficiency group (M= 382.37 Hz) was numerically lower than that of the high English proficiency group (M= 397.87 Hz), and the trend was consistent with the expected direction, suggesting that low English proficiency learners produced /ɪ/ with a higher tongue position. On F2, the values of the low English proficiency group (M= 2124.32 Hz) and the high English proficiency group (M = 2144.66 Hz) were close, and there was no obvious trend of anterior-posterior tongue position difference.

The Mann-Whitney U test showed that the differences in F1 (U= 40.50, p= .372, r= .22) and F2 (U=34.50, p= .793, r= .07) between the two groups did not reach the statistically significant level. The inter-group difference in F1 was a small effect, and the effect of F2 was close to zero.

DOI: 10.54254/2753-7064/2026.HT35879

Table 1. Mean F1 and F2 of English /ɪ/ in each English proficiency group (standard deviation in brackets)

Group	Number	F1 (Hz)	F2 (Hz)
Reference value (male)	—	~400	~1700
Reference value (female)	—	~480	~2000
Low English proficiency	8	382.37 (89.14)	2124.32 (238.69)
High English proficiency	8	397.87 (76.16)	2144.66 (243.75)

Note: The reference values were taken from Feng & Wang [15] and Chen et al. [13]. The reference values for males and females were different because the original resonance peaks were not normalized, and the physiological differences in the length of the vocal tract between males and females would lead to systematic differences in the absolute values of the resonance peaks.

5. Discussion

This study examined the acoustic features of English /ɪ/ produced by Chinese learners at two proficiency levels. Spectral results showed no significant group differences, but the F1 trend aligned with expectations.

5.1. Group comparison in the spectral dimension

Low-proficiency learners produced lower F1 values than high-proficiency learners, consistent with the higher tongue position reported for Chinese learners [13, 18] and with the phonological account that learners may assimilate /ɪ/ to Mandarin /i/ [3]. However, the difference was not significant ($p = .372$, $r = .22$). F2 showed almost no group difference ($r = .07$), suggesting that the front-back tongue dimension may not be a sensitive indicator of proficiency.

The non-significant results likely reflect two factors: the small sample size (n=8 per group) limited statistical power, and the use of raw Hz values without normalization introduced large individual variation (SDs: 76–244 Hz, as shown in Table 1), which may have masked group differences. The fact that F1 showed a trend in the expected direction while F2 showed almost no difference suggests that tongue height adjustment may precede front-back adjustment in L2 development.

5.2. The role of learner background factors

All participants came from the northern Mandarin dialect area, which controlled for dialect-related confounding but limits generalizability. CET/IELTS scores reflect general proficiency rather than pronunciation ability, so future studies should incorporate specialized pronunciation assessments. Input quantity [19], writing system experience [4], and phonetic instruction [14] also influence L2 vowel production and should be examined as covariates in larger samples.

5.3. Teaching implications

The higher tongue position observed in low-proficiency learners suggests that instruction should emphasize tongue height lowering and centralization for /ɪ/, rather than relying solely on duration cues. This aligns with Zheng and Wonnacott's [8] finding that explicit phonetic teaching can reshape perceptual cue weighting.

DOI: 10.54254/2753-7064/2026.HT35879

5.4. Limitations and future directions

Several limitations should be acknowledged. First, the small sample size limited statistical power; the non-significant F1 trend is likely attributable to insufficient participants. Second, without normalization, vocal tract length differences between genders could not be controlled. Future work should include multiple vowel categories (e.g., /iː/, /æ/, /ɑː/, /uː/) to enable normalization. Third, the absence of a native control group and reliance on published reference values may introduce comparability issues due to differing recording conditions. Finally, the homogeneous dialect sample limits generalizability to learners from other regions.

6. Conclusion

This study examined the acoustic production of English /ɪ/ by two groups of Chinese native speakers with different levels of English proficiency. On the spectral dimension, the F1 values of the low English proficiency group were lower than those of the high English proficiency group, suggesting a higher tongue position. This pattern is consistent with the earlier observations of Chen et al. [13] and Liao [14], who reported that Chinese learners tend to produce /ɪ/ with a higher and more fronted articulation compared to native English speakers. It is also in line with the phonological explanation provided by Kwon and Starr [6], who argued that learners assimilate /ɪ/ to the native high front vowel /i/ due to the absence of the tense-lax contrast in Mandarin. However, this difference did not reach statistical significance ($p = .372$), and there was almost no difference in F2 between the two groups ($p = .793$). The non-significant results are mainly attributed to the small sample size (8 participants per group) and the inability to perform normalization processing due to the lack of multiple vowel tokens per speaker, resulting in considerable individual variation that may have masked potential group differences. All participants were recruited from the Northern Mandarin dialect area, which effectively controlled for the confounding effects of dialectal variation, but this also limits the generalizability of the findings to learners from other dialect regions. In terms of pronunciation teaching, the descriptive trend observed in F1 values for low English proficiency learners suggests that instructional efforts should focus on adjusting tongue position, helping learners lower the /ɪ/ tongue position to approximate that of native speakers more closely, in addition to the common emphasis on vowel duration distinctions.

References

[1] Cebrian, J., Gorba, C., & Gavaldà, N. (2021). When the easy becomes difficult: Factors affecting the acquisition of the English /iː/-/ɪ/ contrast. *Frontiers in Communication, 6*, 660917. https://doi.org/10.3389/fcomm.2021.660917

[2] Munro, M. J., & Derwing, T. M. (2008). Segmental acquisition in adult ESL learners: A longitudinal study of vowel production. *Language Learning, 58*(3), 479-502.

[3] Li, M., Chen, Y., & Cui, J. (2020). The effect of input on the production of English tense and lax vowels by Chinese learners: Evidence from an elementary school in China. In *INTERSPEECH* (pp. 2372-2376).

[4] Mi, L., Tao, S., Wang, W., Dong, Q., Guan, J., & Liu, C. (2016). English vowel identification and vowel formant discrimination by native Mandarin Chinese-and native English-speaking listeners: The effect of vowel duration dependence. *Hearing Research, 333*, 58-65.

[5] Yang, X., Shi, F., Liu, X., & Zhao, Y. (2016). Learning styles and perceptual patterns for English /i/ and /ɪ/ among Chinese college students. *Applied Psycholinguistics, 37*(3), 673-701. https://doi.org/10.1017/S014271641500020X

[6] Kwon, J., & Starr, G. (2023). How L1-Chinese L2-English learners perceive English front vowels: A phonological account. *Glossa: A Journal of General Linguistics, 8*(1). https://doi.org/10.16995/glossa.9282

[7] Yihan, L. (2024). *Perception and categorisation of vowels in German and English by Chinese university learners* [Master's thesis]. University of Oxford.

DOI: 10.54254/2753-7064/2026.HT35879

[8] Zheng, Z., & Wonnacott, E. A. (2026). Effects of explicit phonetic instruction in shaping L2 perceptual cue weighting: Evidence from English /i:/-/ɪ/. *Journal of Speech, Language, and Hearing Research*.
[9] Isbell, D. R. (2016). The perception-production link in L2 phonology. *MSU Working Papers in Second Language Studies, 7*(1).
[10] Zhang, A., Feng, H., Wang, S., & Dang, J. (2016). Relationship between perception and production of English vowels by Chinese English learners. *2016 10th International Symposium on Chinese Spoken Language Processing (ISCSLP)*. IEEE, October, pp. 1-5.
[11] Wang, L. J., Guo, J. W., & Feng, H. (2023). Perception and production of English tense and lax vowels by Chinese EFL learners.
[12] Wang, X. (2020). L2 speech learning: Perception, production & training. *Cadernos de Linguistica, 1*(1), 1-22.
[13] Chen, Y., Robb, M., Gilbert, H., & Lerman, J. (2001). Vowel production by Mandarin speakers of English. *Clinical Linguistics & Phonetics, 15*(6), 427-440.
[14] Liao, J. S. (2006). *English vowel production of Mandarin speakers* [Doctoral dissertation, The University of Texas at Austin].
[15] Feng, H., & Wang, L. (2024). Acoustic analysis of English tense and lax vowels: Comparing the production between Mandarin Chinese learners and native English speakers. *The Journal of the Acoustical Society of America, 155*(5), 3071-3089.
[16] Wen, X., & Jia, Y. (2016). Joint effect of dialect and Mandarin on English vowel production: A case study in Changsha EFL learners. In *Interspeech*, 185–189.
[17] Zeng, J., Pillai, S., & Chew, S. Y. (2025). The production of English monophthongs by Chinese Yi and Han speakers: A comparative acoustic analysis. *English Today*, 1–8.
[18] Martin, K. I. (2017). The impact of L1 writing system on ESL knowledge of vowel and consonant spellings. *Read Writ, 30*, 279–298. https://doi.org/10.1007/s11145-016-9673-5
[19] Li, S., & Woore, R. (2023). How Chinese learners decode L2 English words: Evidence from a phonics instruction program. *Read Res Q, 58*, 584–600. https://doi.org/10.1002/rrq.515
[20] Casillas, J. (2015). Production and perception of the /i/-/I/ vowel contrast: The case of L2-dominant early learners of English. *Phonetica, 72*(2–3), 182–205.