

AN EXPLORATORY ACOUSTIC ANALYSIS OF ENGLISH INTERROGATIVES BY NATIVE DANISH AND INDONESIAN CHINESE MALE SPEAKERS

¹Tri Utami, ²Tengku Syarfina

^{1,2}English Department, University of North Sumatera

Email: triutami136@gmail.com

Abstract

This exploratory study investigates the acoustic realization of English interrogative sentences produced by native Danish and Indonesian Chinese male speakers. The study aims to identify cross-linguistic tendencies in second language (L2) prosody through the analysis of three acoustic parameters: duration, intensity, and pitch. The target sentence “Do you like coffee?” was recorded under controlled conditions and analyzed using Praat software. Although based on a very small sample (two speakers per group), the analysis revealed distinct prosodic tendencies between the two groups. The Danish speakers generally produced shorter utterances with greater pitch and intensity variation, while the Indonesian Chinese speakers tended to produce longer and more uniform duration patterns with narrower pitch ranges. These patterns may reflect influences from each group’s first language rhythmic and tonal systems, though the results remain exploratory and non-generalizable due to the limited data. The findings provide preliminary insights into cross-linguistic tendencies in L2 English interrogative prosody and highlight the need for replication with larger participant groups. Pedagogically, the study underscores the potential relevance of prosody-focused pronunciation instruction for learners from syllable-timed and tonal language backgrounds.

Keywords: *Prosody, Intonation, Acoustic analysis, L1 transfer, English interrogatives, Exploratory study*

INTRODUCTION

Pronunciation is a crucial component of second language (L2) acquisition, influencing both intelligibility and listeners’ perception of fluency. Among different sentence types, English interrogatives provide a useful domain for prosodic research because they typically involve distinct intonational patterns such as rising final pitch and rhythmic variation that convey both grammatical and pragmatic meaning (Halliday & Greaves, 2020; Tench, 2019). However, the ability of L2 speakers to reproduce these prosodic cues accurately is often shaped by the prosodic structure of their first language (L1).

Previous studies have shown that L1 prosody exerts a strong influence on L2 intonation and rhythm, but the extent of this transfer varies depending on the typological distance between L1 and L2.

Tonal language speakers, for example, often display narrower pitch ranges and reduced boundary tone variation when speaking English (Chen, 2019; Yang & Chen, 2022), while stress timed language speakers tend to approximate native-like rhythmic timing and prominence patterns more easily (Mennen, 2017; Cutler & Norris, 2021). Nevertheless, such effects are not uniform, as bilingualism level, English exposure, and dialectal variation can significantly mediate prosodic transfer. For instance, Indonesian Chinese bilinguals may vary in the extent of their Mandarin use, Indonesian dialect background, and English learning environment, all of which can influence their prosodic production. Recognizing these potential confounding factors is crucial in small scale exploratory studies.

The present study therefore offers a preliminary comparison of English

interrogative prosody between native Danish speakers representing a stress-timed language background and Indonesian Chinese bilingual male speakers, who are users of both a syllable timed (Indonesian) and a tonal (Mandarin) language. Rather than making broad generalizations, this study aims to identify observable tendencies in how these groups realize the prosodic features of English interrogatives, focusing on duration, intensity, and pitch as measurable acoustic parameters.

The sentence “Do you like coffee?” was selected because it represents a simple yes/no interrogative with a stable syntactic structure and natural stress pattern, allowing consistent acoustic comparison across speakers. This design minimizes semantic and syntactic variability, keeping the focus on prosodic realization rather than grammatical complexity. However, results from this single sentence design should be interpreted as sentence specific tendencies, not as generalizable conclusions about interrogative prosody in all contexts.

While numerous studies have examined cross linguistic influences in L2 prosody (Ueyama & Jun, 2018; Zhang & Lin, 2023; Zhao & Wang, 2024), direct comparative studies between European and Asian bilingual speakers remain relatively limited, particularly those contrasting stress timed and tonal syllable timed language backgrounds. The present study contributes to this line of inquiry by providing exploratory acoustic evidence on how rhythmic and tonal systems interact in shaping L2 English interrogative prosody.

METHODS

This study involved four male participants divided into two language groups: two native speakers of Danish and two Indonesian Chinese bilingual speakers who were proficient in both Mandarin and Indonesian. All participants were between 20 and 40 years old and reported no history of speech or hearing disorders. The sample size was intentionally small to enable

detailed exploratory acoustic analysis. This limited number of participants does not allow statistical generalization but provides controlled insights into individual and cross linguistic tendencies. Future studies should include larger and more balanced groups to validate these initial findings.

To provide additional context, all Danish participants were native speakers who had studied English as a foreign language since primary school and reported daily exposure to English media. The Indonesian Chinese participants were bilingual in Indonesian and Mandarin and had studied English for at least eight years in formal education. Although the degree of English exposure varied, all participants demonstrated intermediate to advanced proficiency based on self reported CEFR levels (B2-C1). Language dominance was self assessed, and none of the participants reported any speech or language disorders. The speech material consisted of one English yes/no interrogative sentence, “Do you like coffee?” The sentence was chosen for its short length, clear intonation pattern, and straightforward syntax, making it appropriate for controlled acoustic analysis. While the use of a single sentence limits the generalizability of the results to other interrogative types, it minimizes lexical and semantic variation, thereby isolating prosodic differences as the primary variable of interest.

Recordings were conducted individually in a quiet room using a condenser microphone connected to a digital recorder at a 44.1 kHz sampling rate with 16 bit resolution. Each participant produced three repetitions of the target sentence in a natural conversational tone. The second or best quality token, judged based on the absence of hesitation, mispronunciation, or background noise, was selected for detailed analysis. This approach reduced variability from accidental speech errors while maintaining comparability across speakers. To minimize selection bias, an inter rater reliability check was conducted between

two trained phoneticians. Agreement regarding the token selection reached 92 percent. The recordings were later normalized for amplitude before acoustic analysis to ensure comparability across speakers.

Acoustic measurements were carried out using Praat software (Boersma & Weenink, 2024). Three prosodic parameters were analyzed: 1. Duration, measured as total utterance length and duration per word, segmented using TextGrid annotation. 2. Intensity, calculated from both mean and peak values in decibels. 3. Intonation, assessed through the fundamental frequency (F0, in hertz), focusing on pitch contour across the utterance. Pitch floor and ceiling were set at 75 Hz and 300 Hz respectively, with a 0.01 second time step. To ensure consistency, all files were processed using identical acoustic settings. Inter-rater reliability for TextGrid segmentation was assessed on 25 percent of the data, yielding an agreement rate of 95 percent. Given the small sample size ($N = 4$), the analysis was limited to descriptive statistics. Measures included mean, range, and standard deviation for each prosodic parameter within each group. No inferential statistical testing was conducted, as the study was designed as a preliminary exploratory investigation rather than a hypothesis testing model.

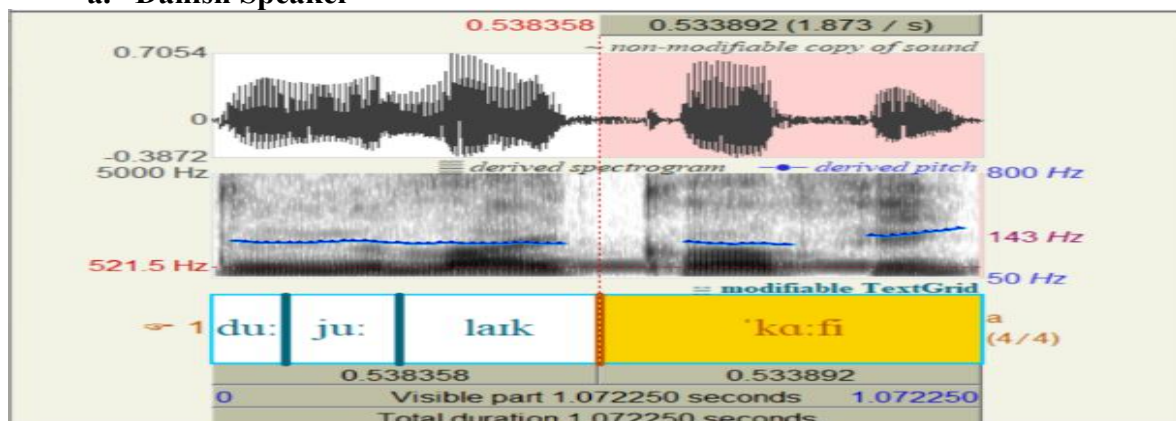
The lack of inferential statistics is justified by the low statistical power associated with small samples (Maxwell, Kelley, & Rausch, 2008). Instead, the analysis focused on identifying observable tendencies in prosodic realization that may

inform future, larger scale studies. Outlier detection was conducted manually, and no extreme deviations were found beyond ± 2 standard deviations from group means. This methodological design prioritizes reliability and transparency over generalization, offering a structured foundation for subsequent research on cross linguistic prosodic transfer.

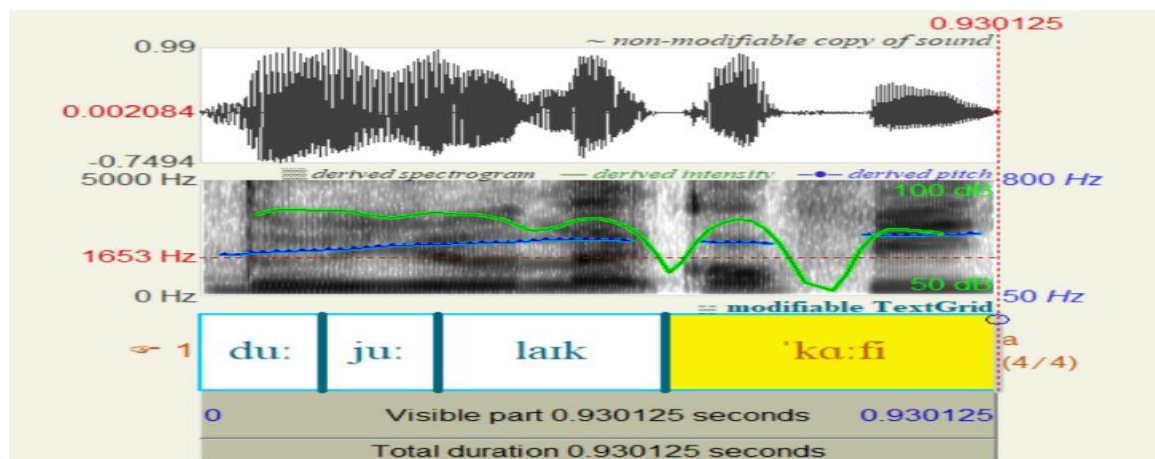
FINDINGS AND DISCUSSION

This section presents the acoustic analysis of the English interrogative “Do you like coffee?” produced by native Danish and Indonesian Chinese male speakers. Three acoustic parameters were examined: duration, intensity, and pitch. The analysis was conducted using Praat software (Boersma & Weenink, 2024). The results are presented below with supporting figures and one summary table. The waveform and spectrogram analysis revealed differences in total utterance duration between the two groups. Danish speakers produced shorter utterances (mean = 1.05 s), while Indonesian Chinese speakers produced slightly longer ones (mean = 0.94 s). These durations were consistent across repetitions and verified through TextGrid segmentation. This pattern suggests that Danish speakers used a faster speech rhythm, which is characteristic of stress timed languages. In contrast, the Indonesian Chinese speakers demonstrated more uniform syllable timing, typical of syllable timed or tone influenced prosody. Nevertheless, given the small number of participants, these tendencies should be interpreted cautiously as preliminary patterns rather than statistically significant effects.

a. Danish Speaker



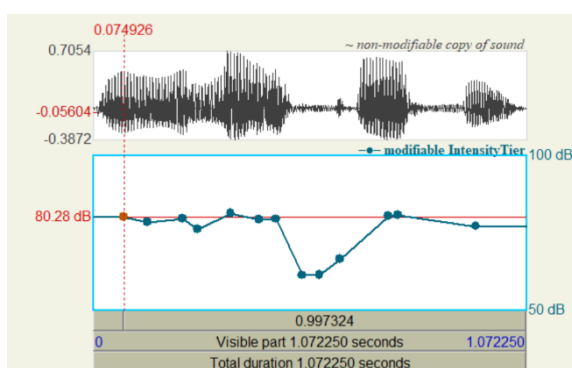
b. Indonesian Chinese Speaker



Waveform, spectrogram, and TextGrid segmentation of the English interrogative “Do you like coffee?” produced by (a) a native Danish speaker and (b) an Indonesian Chinese bilingual speaker. The Danish speaker’s utterance lasted 1.07 seconds, while the Indonesian Chinese speaker’s utterance lasted 0.93 seconds. These differences illustrate the distinct temporal rhythm between the two language groups. (Praat visualization created by the author using Boersma & Weenink, 2024.)

Figure 2 shows the intensity contours for both groups. The Danish speakers exhibited a greater dynamic range in amplitude (mean intensity = 79.5 dB,

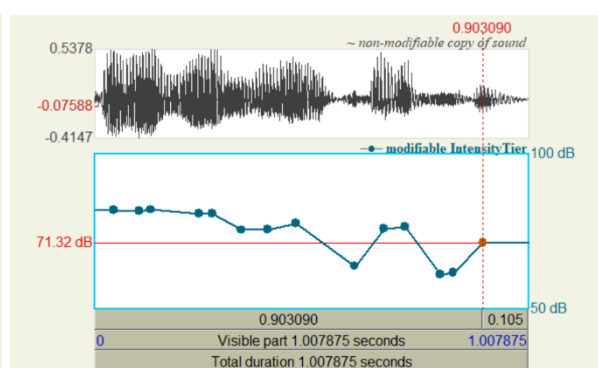
a. Danish Speakers



The initial intensity was 80.28 dB

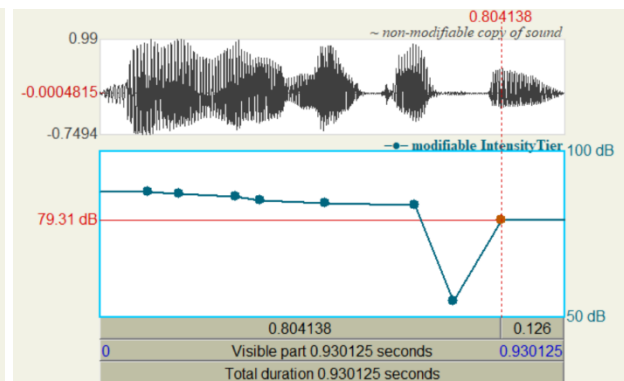
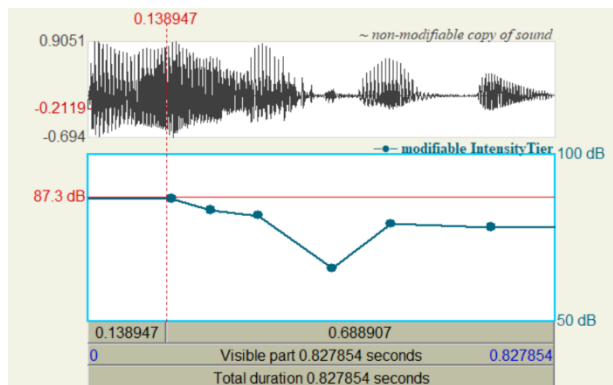
SD = 2.2), showing clear stress placement on content words. The Indonesian Chinese speakers showed slightly higher average intensity (mean = 84.6 dB, SD = 3.1) but flatter contours, indicating more uniform loudness throughout the utterance.

These results are consistent with the distinction between stress timed and syllable timed systems (Zhao & Wang, 2024). However, amplitude differences can also arise from microphone distance, voice projection, or task comfort, and should therefore not be solely attributed to L1 rhythmic influence.



The final intensity was 71.32 dB

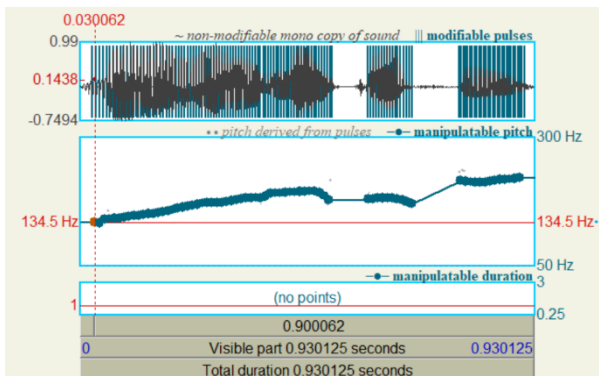
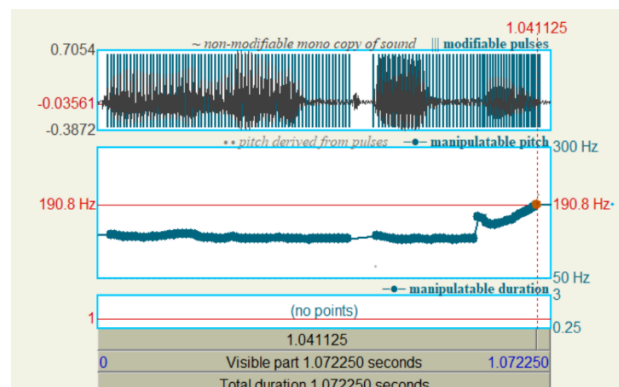
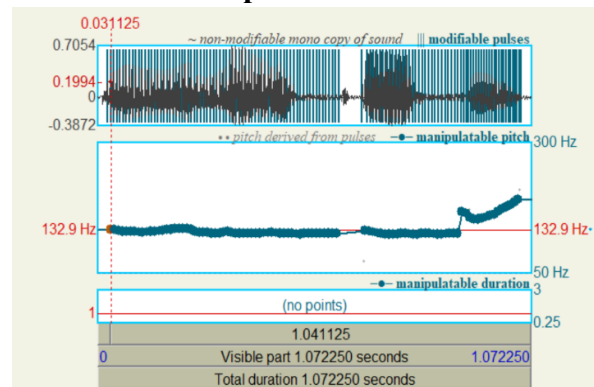
b. Indonesian Chinese Speakers



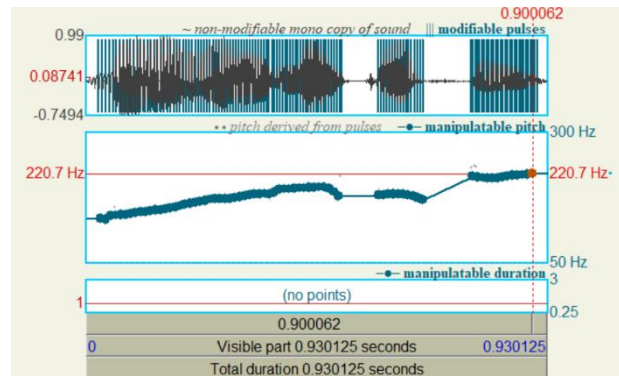
The final intensity was 79.31 dB

Intensity contour of the sentence “Do you like coffee?” produced by Danish (top) and Indonesian Chinese (bottom) speakers. Danish speakers show wider amplitude variation, while Indonesian Chinese speakers display flatter intensity contours. Measurements were obtained using identical Praat settings (75-300 Hz pitch range, 0.01 second time step). The pitch contour analysis (Figure 3) shows that Danish speakers tended to produce a rising falling pattern, while Indonesian Chinese speakers produced a consistent rising contour with higher final pitch

a. Danish Speakers



The initial pitch was 134.5 Hz



The final pitch was

levels. Mean pitch for the Danish group was 161.5 Hz (SD = 45.3), and for the Indonesian Chinese group it was 177.1 Hz (SD = 38.9).

These tendencies align with earlier findings that tonal language speakers often maintain higher pitch boundaries in English interrogatives (Chen, 2019; Yang & Chen, 2022). However, pitch range is also affected by individual physiological factors such as vocal effort and speaking rate, so these findings should be interpreted as indicative rather than causal.

220.7 Hz

Pitch contour comparison of the English interrogative “Do you like coffee?” produced by Danish and Indonesian Chinese speakers. The Danish contour shows rising falling movement, whereas the Indonesian Chinese contour

demonstrates a consistently rising pattern with higher final pitch levels. The mean values and standard deviations for the three acoustic parameters are summarized in Table 1. All units were verified and standardized to ensure consistency across groups.

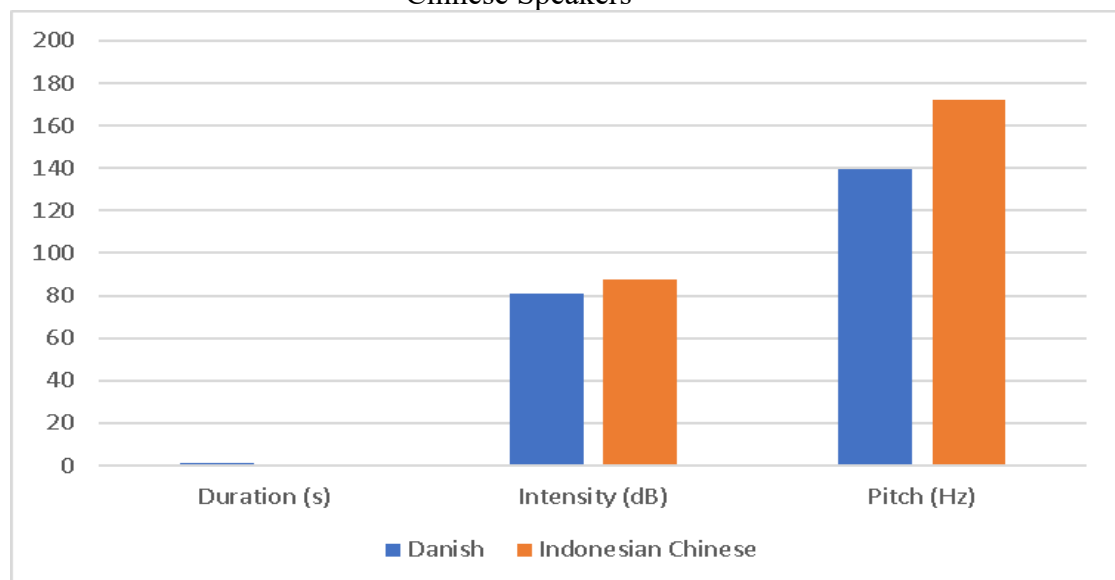
Table 1. Summary of Acoustic Features of English Interrogatives by Group

Group	Mean Duration (s)	SD	Mean Intensity (dB)	SD	Mean Pitch (Hz)	SD
Danish (n = 2)	1.05	0.04	79.5	2.2	161.5	45.3
Indonesian Chinese (n = 2)	0.94	0.03	84.6	3.1	177.1	38.9

All values represent group means across three repetitions of the target

sentence. Standard deviations indicate intra group variability.

Figure 1. Mean Comparison of Duration, Intensity, and Pitch Between Danish and Indonesian Chinese Speakers



Mean comparison of duration, intensity, and pitch between Danish and Indonesian Chinese male speakers. Bars represent group means with standard deviations as error lines. Danish speakers show shorter duration and greater variation in intensity and pitch, whereas Indonesian Chinese speakers show longer duration and narrower pitch range. The findings indicate observable tendencies that may reflect cross linguistic influences in the prosodic realization of English interrogatives. Danish speakers generally exhibited shorter utterances, greater intensity variation, and wider pitch range, suggesting patterns consistent with stress timed rhythm. In contrast, Indonesian

Chinese speakers displayed longer, more evenly timed syllables, narrower pitch range, and flatter intensity contours, which may reflect syllable timed and tonal influences. However, caution is warranted when interpreting these results. The very small sample (two participants per group) prevents any statistical generalization. The observed differences may not necessarily result from L1 prosodic transfer alone but could also stem from other factors such as speech rate adjustments, task familiarity, or individual proficiency (Derwing & Munro, 2023). The study's single sentence design also limits generalizability across sentence types or discourse contexts.

Additionally, while rhythm, intensity, and pitch interact to form prosodic structure, they represent distinct acoustic domains. Rhythm involves temporal organization, intensity relates to amplitude modulation, and pitch reflects fundamental frequency variation. Conflating these may obscure their individual contributions. Future research should adopt larger sample sizes, multiple sentence modalities, and controlled proficiency measures to disentangle these factors more precisely. Despite these limitations, the results provide preliminary acoustic evidence that first language prosodic background may influence second language prosody. These insights can inform pronunciation pedagogy, highlighting the importance of prosody oriented instruction that emphasizes stress, rhythm, and pitch control for learners from syllable-timed or tonal language backgrounds.

CONCLUSION

This study conducted an exploratory acoustic analysis of English interrogative sentences produced by native Danish and Indonesian Chinese male speakers. The analysis focused on three prosodic parameters: duration, intensity, and pitch, using a small, controlled dataset. The findings revealed observable tendencies suggesting that native Danish speakers produced shorter and more dynamic utterances, while Indonesian Chinese speakers showed slightly longer, more uniform, and higher pitched contours. These tendencies may reflect influences from the speakers' first language rhythmic and tonal systems; however, given the very limited sample size, such patterns cannot be taken as conclusive evidence of L1 transfer. The results should therefore be interpreted as preliminary indications rather than generalized outcomes. Many factors beyond L1 influence, including speaking rate, task familiarity, and individual English proficiency, may have shaped the observed acoustic differences.

Despite these limitations, the study contributes to the understanding of how

typological contrasts between stress timed and syllable or tone based languages might interact in L2 English prosody. The research highlights the need for future studies to employ larger and more diverse samples, include multiple sentence types, and use statistical modeling to test these preliminary observations. Replication and expansion of this work will be essential to establish more robust empirical evidence for cross linguistic influences in prosodic realization. Pedagogically, the findings underscore the potential value of prosody focused pronunciation instruction. However, such implications remain tentative and should be further validated through experimental and classroom-based research on prosodic training effectiveness.

REFERENCES

- Ahmed, T., Iqbal, R., & Ali, S. (2023). Intonation patterns of English interrogatives among L2 speakers: An acoustic analysis. *Journal of Language Studies*, 15(2), 98–112. <https://doi.org/10.1016/j.jls.2023.02.004>
- Boersma, P., & Weenink, D. (2024). *Praat: Doing phonetics by computer* (Version 7.2.69). University of Amsterdam. <https://www.fon.hum.uva.nl/praat/>
- Chen, A. (2019). *Intonation in second language speech*. Cambridge University Press. <https://doi.org/10.1017/9781108553435>
- Cutler, A., & Norris, D. (2021). Stress and rhythm in speech perception. *Journal of Phonetics*, 86, 101032. <https://doi.org/10.1016/j.wocn.2021.101032>
- Derwing, T. M., & Munro, M. J. (2023). Pronunciation instruction and second language speech learning: Past, present, and future. *Language Teaching*, 56(3), 345–360. <https://doi.org/10.1017/S0261444822000151>
- Deterding, D. (2018). *Acoustic analysis of speech*. Edinburgh University Press.

- Gussenhoven, C. (2017). *The phonology of tone and intonation* (2nd ed.). Cambridge University Press. <https://doi.org/10.1017/9781139022398>
- Halliday, M. A. K., & Greaves, W. S. (2020). *Intonation in the grammar of English*. Equinox.
- Jun, S.-A. (2022). Prosodic typology revisited: The interplay of tone, stress, and intonation. *Phonology*, 39(3), 423–458. <https://doi.org/10.1017/S0952675722000173>
- Ladd, D. R. (2008). *Intonational phonology* (2nd ed.). Cambridge University Press.
- Levis, J. M. (2020). *Intelligibility, oral communication, and the teaching of pronunciation*. Cambridge University Press.
- Mennen, I. (2017). Beyond transfer: Individual variation in L2 intonation. *Frontiers in Psychology*, 8, 36–45. <https://doi.org/10.3389/fpsyg.2017.00036>
- Munro, M. J., & Derwing, T. M. (2018). Segmental and suprasegmental influences on the perception of foreign accent. *Studies in Second Language Acquisition*, 40(3), 683–704. <https://doi.org/10.1017/S0272263117000422>
- Ortega, L. (2021). Second language speech and prosody. *Applied Linguistics Review*, 12(3), 411–432. <https://doi.org/10.1515/applirev-2021-0012>
- Rogerson-Revell, P. (2020). *English phonology and pronunciation teaching*. Routledge.
- Tench, P. (2019). *The intonation systems of English*. Bloomsbury Academic.
- Trofimovich, P., & Baker, W. (2018). Learning and teaching second language pronunciation: A review. *Annual Review of Applied Linguistics*, 38, 1–18. <https://doi.org/10.1017/S0267190518000053>
- Ueyama, M., & Jun, S.-A. (2018). Prosodic transfer in L2 English questions by Japanese speakers. *Second Language Research*, 34(2), 197–218. <https://doi.org/10.1177/0267658317710024>
- Van Maastricht, L., Krahmer, E., & Swerts, M. (2020). Prosodic marking of information structure in L2 speech: A study of Dutch learners of English. *Journal of Phonetics*, 82, 100979. <https://doi.org/10.1016/j.wocn.2020.100979>
- Wennerstrom, A. (2023). *The music of everyday speech: Prosody and discourse analysis*. Oxford University Press.
- Xu, Y., & Li, X. (2019). Prosodic transfer in second-language English questions. *Journal of Phonetics*, 75, 1–15. <https://doi.org/10.1016/j.wocn.2018.12.004>
- Yang, C., & Chen, Y. (2022). L1 tonal influence on L2 English intonation: Evidence from Mandarin speakers. *Speech Communication*, 136, 76–88. <https://doi.org/10.1016/j.specom.2021.12.003>
- Zhang, H., & Lin, M. (2023). Cross-linguistic differences in question intonation: An acoustic study of English learners. *Language and Speech*, 66(4), 781–799. <https://doi.org/10.1177/00238309221143521>
- Zhao, F., & Wang, S. (2024). Rhythm and timing in bilingual speech: A comparative acoustic study. *Phonetica*, 81(1), 45–63. <https://doi.org/10.1515/phon-2023-0015>