Humanities, Arts and Social Sciences Studies

Research Article

Humanities, Arts and Social Sciences Studies
https://so02.tci-thaijo.org/index.php/hasss
ISSN (Online): 2630-0079

THE INTERPLAY OF SEGMENTALS AND SUPRASEGMENTALS IN THE PERCEPTION OF L2 COMPREHENSIBILITY IN CHINESE-ACCENTED THAI

Peng Hou[*] *and Sarawut Kraisame*

Research Institute for Languages and Cultures of Asia (RILCA), Mahidol University, Thailand

ABSTRACT

Comprehensibility refers to the ease of understanding in second language (L2) speech. This paper examines the relative contribution of segmentals and suprasegmentals on perceived L2 comprehensibility. Fifteen Chinese speakers of Thai were asked to produce Thai segmentals and suprasegmentals in a picture description task. Acoustic analyses demonstrated that Chinese-accented Thai speech was distinct from native Thai speech in both segmentals and suprasegmentals. In addition, native Thai raters (n = 30) also evaluated the degree of comprehensibility using a 9-point Likert scale. A stepwise regression analysis revealed the joint predictiveness of segmentals and suprasegmentals on L2 comprehensibility, which collectively accounted for more than 50% of the variance in comprehensibility ratings. Specifically, suprasegmentals (i.e., filled pauses and speech rate) and segmentals (i.e., duration of diphthongs and quality of back monophthongs) were identified as the most robust predictors for comprehensibility in Chinese-accented Thai. The findings offer direct evidence that comprehensibility is intricately linked to several L2 features from the domains of segmentals and suprasegmentals. Therefore, understanding ~~of~~ the nature of L2 comprehensibility in Thai requires considering the interplay between segmentals and suprasegmentals. In light of this, it is proposed that emphasis is required on both segmentals and suprasegmentals in L2 instruction in order to achieve more comprehensible L2 speech.

Keywords: Comprehensibility; segmentals; suprasegmentals; second language speech; Thai

[*]***Corresponding author:***
Peng Hou
peng.hou@student.mahidol.ac.th

Received: *7 June 2024*
Revised: *3 January 2025*
Accepted: *16 January 2025*
Published: *8 April 2025*

Citation:
Hou, P., & Kraisame, S. (2025). The interplay of segmentals and suprasegmentals in the perception of L2 comprehensibility in Chinese-accented Thai. Humanities, Arts and Social Sciences Studies, 25(1), 169–185. https://doi.org/10.69598/hasss.25.1.270733

1. INTRODUCTION

Adult learners who acquire a second language (L2) often have non-native speech that is typically perceived as accented. This is because acquiring an L2 generally involves the establishment of novel phoneme categories, the reconfiguration of current phoneme boundaries, and the acquisition of unfamiliar suprasegmental elements specific to the target L2. The struggle among adult learners to acquire any of these

domain's results in pronunciations that diverges from the native standard, resulting in accented L2 speech (Edwards & Zampini, 2008). For a considerable period, numerous language learners have made significant efforts to speak an L2 without any noticeable foreign accents. This has emerged as an extremely important objective in the field of L2 education. This pedagogical objective is witnessed in the "nativeness principle" (Levis, 2005), which aims to achieve non-accented L2 speech, or speech devoid of language characteristics that may indicate the speaker as non-native. Nevertheless, it is rare for L2 learners to achieve a level of proficiency in foreign accents that is indistinguishable from that of native speakers (Ortega, 2009). In light of this, a growing number of researchers have emphasized the importance of evaluating L2 speech using criteria relevant to effective L2 communication in practical contexts. As a result, a significant amount of scholarly focus has been placed on fostering L2 speech that is comprehensible, rather than attempting to eradicate a foreign accent. This aligns with the "intelligibility principle" (Levis, 2005), which prioritizes L2 speech that can be understood by the listener, even in the presence of a noticeable accent. The term used to describe this concept is L2 comprehensibility, which refers to the level of ease with which non-native speakers' speech may be understood. It is typically assessed through the use of a rating scale by native listeners.

Empirical studies have also provided evidence to support that a non-native accent does not necessarily hinder effective communication. These studies have shown that accent and comprehensibility, although related, are partially distinct facets. Put simply, speakers who are seen to have a strong accent in their L2 can nevertheless be easily understood, while those who are not comprehensible are always perceived as having a strong accent (Derwing & Munro, 2009). Furthermore, this implies that the linguistic features of L2 speech, which contribute to comprehensibility and accent, exhibit variations (Gass & Varonis, 1984). For instance, Trofimovich and Isaacs (2012) found that different L2 speech features were associated with comprehensibility and accents in native French speakers of English. Accents only correlated with segmentals, whereas comprehensibility was associated with a wider range of L2 features. Saito et al. (2016) replicated similar results pertaining to Japanese native speakers who acquired English as an L2. Comprehensibility was associated with all aspects of L2 language speech, but accents were particularly connected to the precision of segmental accuracy. Comparably, Crowther et al. (2018) discovered a substantial correlation between comprehensibility and several aspects of pronunciation and lexico-grammar, while accents were mostly connected to measures of pronunciation. Based on these factors, it is apparent that comprehensibility is a complex structure potentially associated with a larger number of L2 speech features. However, previous studies have not yet reached a consensus about the L2 speech features that should be highlighted in L2 comprehensibility.

What is the nature of comprehensibility, or which L2 speech features contribute to comprehensibility? The prioritization of this issue is crucial to improving the comprehensibility of L2 speech. Research indicates that both segmental and suprasegmental features contribute to L2 comprehensibility. Segmental features refer to errors made in the articulation of individual consonants and vowels, which can involve either the replacement of one sound with another or the alteration of a sound (Flege & Hillenbrand, 1984; Major & Faudree, 1996). Suprasegmental errors can be categorized into two types: those that affect speech fluency, such as speech rate and pause, and those that affect speech prosody, such as intonation and stress. Certain scholars believe that segmental features significantly contribute to L2 comprehensibility. According to a study conducted by Winters and O'Brien (2013), the significance of segmentals in speech may be determined by manipulating speech samples to blend native suprasegmentals with non-native segments, and vice versa. They also concluded that segmentals appeared to have a greater impact on comprehensibility compared to suprasegmentals. Suzukida and Saito (2021) investigated the impact of specific segmental errors, such as replacing /l/ with /r/ or /v/ with /b/, on the comprehensibility ratings given by native English listeners to Japanese-accented English. Conversely, a different cohort of researchers determined that enhancements in L2 comprehensibility are more likely to take place alongside improvements in suprasegmental features. In their examination of Mandarin-accented English, Munro and Derwing (1995) discovered that more than 70% of the native English judges identified notable correlations between comprehensibility and intonation difficulties. Kang et al. (2010) found that suprasegmentals could explain 50% of the variance in comprehensibility evaluations. Derwing et al. (2004) demonstrated a correlation between various temporal measurements, such as pause and articulation rate, and the level of comprehensibility. Nevertheless, it is important to note that consensus is currently limited regarding the proportional impact of segmentals and suprasegmentals on the perceived comprehensibility of L2. The presence of this discrepancy implies that L2 comprehensibility may be a multifaceted phenomenon that is intricately tied to a range of segmental and suprasegmental features.

It is also necessary to mention that the bulk of the current findings regarding L2 comprehensibility arise from the studies dealing with English. Other languages, especially Asian languages, have not received sufficient attention despite their multi-linguistic contexts. In recent decades, the integration of the ASEAN community and the development of Thailand has increased the number of Thai language learners from other countries, especially from ASEAN countries and China (Plaengsorn, 2017). A large variety of L2 Thai speech has

been documented, and the phonetic features of L2 Thai have also been studied to a certain degree. Among the features under investigation, an overwhelming number of studies target tones and vowels in Thai. For example, Teeranon (2016) conducted an acoustic study of Thai tones produced by Khmer and Vietnamese learners and found both Vietnamese and Khmer speakers acoustically produce clear static and dynamic Thai tones. Still, Vietnamese learners produce clearer contour tones than Khmer speakers when comparing pitch height and pitch contour. Phonphanich and Burusphat (2021) compared the phonetic features of Thai tones produced by Chinese-Zhuang and non-Zhuang students and found that the typological similarities between Thai and Zhuang did not benefit Chinese-Zhuang speakers of Thai in producing accurate Thai tones. In terms of Thai vowels, Lianghiranthaworn and Chapoo (2019) proved that Chinese students in their study mispronounced 71% of vowels. Yi (2017) compared Thai vowels produced by three groups of Chinese speakers, namely, Tai Lue, Naxi, and Yunnanese. She demonstrated that the vowel length distinction of all three groups of Chinese speakers was not as obvious as those of native-Thai speakers, and they produced distorted vowel quality as well. On the other hand, studies on other L2 Thai segmentals and suprasegmentals are rare. Based on existing research, L2 learners of Thai have some distinct pronunciation features that distinguish them from native Thai speakers. Nonetheless, the degree to which these features correlate with comprehensibility, as examined in English, remains ambiguous. Investigating languages beyond English, such as Thai, may yield novel insights into our understanding of L2 communication. It may align with or deviate from research on English, indicating that comprehensibility is fundamentally a cross-linguistic or a more language-specific phenomenon.

Chinese speakers of Thai often face challenges due to significant differences in segmental and suprasegmental features between the two languages. Segmentally, Thai possesses nine pairs of duration-contrasted monophthongs, including /i/-/iː/, /e/-/eː/, /ɛ/-/ɛː/, /ɯ/-/ɯː/, /ɤ/-/ɤː/, /a/-/aː/, /u/-/uː/, /ɔ/-/ɔː/, and /o/-/oː/ (Abramson, 1974), whereas Chinese features only six monophthongs—/i/, /y/, /u/, /a/, /ɤ/, and /ə/—without phonemic duration distinctions (Lee & Zee, 2003). Both languages have diphthongs; Thai includes /ia/, /ɯa/, and /ua/, which do not differ by duration, while Chinese has a more extensive diphthong inventory of 11 phonemes, such as /ia/, /iu/, /au/, /ai/, /ou/, /uo/, /ua/, /uə/, /ei/, /ye/, and /ie/ (Li & Zee, 2003). Suprasegmentally, Thai is classified as a stress-timed language, characterized by reduced durations in non-final syllables (Grabe & Low, 2002). In contrast, Mandarin is syllable-timed, exhibiting nearly equal syllable durations in connected speech (Zhang & Lee, 2019). Furthermore, L2 learners generally produce slower speech rates with longer and more frequent pauses than native speakers (Munro & Derwing, 1998; Trofimovich & Baker, 2007). This study examines how Chinese speakers of Thai navigate these segmental and suprasegmental differences in acquiring Thai, and explores the impact of these linguistic features on perceived comprehensibility.

As the preceding discussion suggests, there is currently little understanding of the linguistic features that are relevant to L2 comprehensibility in foreign-accented Thai, especially Chinese-accented Thai. Therefore, the goal of this paper is to extend previous research by examining five different types of L2 speech features, drawn from the domains of segmentals (i.e., vowel duration, vowel quality) and suprasegmentals (i.e., speech rate, pause, stress), to investigate how Chinese speakers of Thai differ in producing these features and further identify how these segmentals and suprasegmentals interplay in the perception of L2 comprehensibility by native Thai listeners.

2. METHODS

2.1 Participants

This study collected speech samples from a group of 15 Chinese speakers of Thai, as well as another group of 15 native Thai speakers. All Chinese learners of Thai were native speakers of Mandarin Chinese. Out of the 15 Chinese speakers of Thai, 12 originated from Shaanxi Province, two were from Henan Province, and one was from Hebei Province. All Chinese speakers of Thai reported being exposed to Mandarin Chinese since birth and identified Mandarin Chinese as their most proficient language. The average age of Chinese speakers of Thai was 22.7 years, with a range of 21 to 23 years. At the time of testing, they had been enrolled in a Thai language major program in their third year at a university in China. All 15 native Thai speakers who provided speech samples were from Bangkok and Nakhon Pathom Province. They expressed that Standard Thai was the most dominant language used in their daily life. They had a mean age of 21.3 years, with a range of 19 to 22 years. At the time of testing, all of them were enrolled in multiple fields of majors at a Thai university. To diminish the influence of gender on acoustic analysis, female participants were recruited exclusively. None of them reported any hearing or listening problems.

In addition, a total of 30 Thai listeners (24 females and 6 males) who were native Thai speakers participated as raters in the task of rating comprehensibility. The average age of the raters was 22.5 years, with a range of 21 to 26. All of the raters were from Bangkok and its surrounding areas. Based on the raters' report,

they assessed their ability to interact with non-native Thai speakers as exceptional. Raters also indicated a limited level of acquaintance with different accented Thai. Considering the raters' limited exposure to different foreign-accented Thai, it can be inferred that they lacked expertise in interacting with non-native Thai speakers and can be considered naïve in recognizing foreign-accented Thai. In addition, none of the raters indicated having received any formal linguistic training. None of them indicated that they had experienced any hearing or speaking difficulties.

2.2 Speech materials

A picture description task was developed to obtain speech samples from both Chinese speakers of Thai and native Thai speakers. The use of picture description would facilitate relatively natural L2 speech production while maintaining control of the speech segmentals (Trofimovich & Baker, 2007). In addition, the task likely imposed minimal demands on the participants to make lexical, syntactic, and pragmatic decisions while producing L2 speech. This task had ten pictures, and each picture was assigned three compulsory-used keywords. Participants were asked to describe the contents of each picture by using the three compulsory keywords occurring together with the picture. Three keywords of each picture were designed to control speech samples' segmental and suprasegmental properties. In the part of segmental, previous studies served as the foundation for selecting potentially problematic vowels for Chinese students. Yi (2017) demonstrated that, when compared to other vowels, Chinese students had more difficulties when pronouncing three diphthongs in Thai, and monophthongs such as /e/, /ɛ/, /ɯ/, /ɤ/, /ɔ/, /o/ and their long counterparts. Additionally, syllable structures were also taken into consideration. The duration of short vowels is believed to be relatively stable in various syllable structures. In contrast, long vowels without any coda have a longer duration than those with a coda. Therefore, this paper examined three types of syllable structures: VS/VN, V:, and V:S/V:N (V, N, and S represent vowels, nasal codas, and stop codas, respectively). To avoid mispronunciation due to unfamiliarity, all keywords were selected from the *Textbooks of Foundation Thai* 1–4 published by Peking University Press in 2011, which were used by the Chinese participants in their learning. A total of 24 keywords fulfilled the above criteria, as listed in Table 1.

Table 1: Selected keywords for the current study

Vowels	VN/VS	V:	V:N/V:S
/ɤ, ɤː/	เยอะ ๆ /jɤ̀ʔ.jɤ́ʔ/ 'abundant/a lot of'	พบเจอ /pʰóp.tɕɤː/ 'encounter'	ทางเดิน /tʰaː.ŋ.dɤː.n/ 'pathway'
/ɯ, ɯː/	ปลาหมึก /plaː.mɯ̀k/ 'squid'	หนังสือ /nǎŋ.sɯ̌ː/ 'book'	แกงจืด /kɛː.ŋ. tɕɯ̀ː.t/ 'soup'
/e, eː/	แม่เหล็ก /mêː.lèk/ 'magnet'	ทะเล /tʰáʔ.leː/ 'sea'	กางเกง /kaː.ŋ.keː.ŋ/ 'pants'
/ɛ, ɛː/	นมแพะ /nom.pʰɛ́ʔ/ 'goat milk'	ตุ๊กแก /túk.kɛː/ 'gecko'	ฝาแฝด /fǎː.fɛ̀ː.t/ 'twins'
/o, oː/	กระจก /kràʔ.tɕòk/ 'glass'	แตงโม /tɛː.ŋ.moː/ 'watermelon'	ข้าวโพด /kʰâː.w.pʰôː.t/ 'corn'
/ɔ, ɔː/	ชาวเกาะ /tɕʰaː.w.kɔ̀ʔ/ 'islander'	ดินสอ /din.sɔ̌ː/ 'pencil'	จิ้งจอก /tɕîŋ.tɕɔ̀ː.k/ 'fox'
/ia/		อ่อนเพลีย /ʔɔ̀ː.n.pʰlia/ 'tired'	นักเรียน /nák.rian/ 'student'
/ua/		ห้องครัว/hɔ̂ː.ŋ.kʰrua/ 'kitchen'	ตำรวจ /tam.rùat/ 'police'
/ɯa/		ผีเสื้อ /pʰīː.sɯ̂a/ 'butterfly'	มะเฟือง /máʔ.fuaŋ/ 'carambola'

This study then assigned the 24 words in Table 1 as keywords to ten pictures for description. Each picture contained three keywords, of which two or three keywords came from the word list (Table 1), and six additional words were complemented to ensure the picture contents were cohesive (see Appendix). The picture description task was designed into three sets by randomizing the order of pictures, and each set occurred on a separate day. This allowed us to collect at least three times for each keyword to facilitate further acoustic analyses. After completing a training and practice session, the actual elicitation procedure was implemented by the researchers in a quiet room. The total number of narratives derived from the recording consisted of 900 narratives (30 speakers × 10 pictures × 3 sets = 900 narratives).

2.3 Comprehensibility rating task

Before the actual rating, a training session was conducted to ensure that the raters understood the concept of comprehensibility (i.e., how effortless it is to understand foreign speech). They were instructed to

make intuitive judgments and use the scales flexibly (1 = the hardest to understand, 9 = the easiest to understand). Following the training session, the actual rating session was implemented online, and the raters were told to advise the researchers to pause the experiment immediately in the event of any disruptions caused by a poor Internet connection or other issues. The raters evaluated 900 narratives presented in random order for comprehensibility using a 9-point Likert scale. The rating session was conducted across two days, each including two hours, to mitigate fatigue. The raters listened to each narrative once, with a 5-second break between consecutive narratives.

2.4 Acoustic measurements

This study focused on five specific acoustic measurements, encompassing both segmental features (i.e., vowel duration and vowel quality) and suprasegmental features (i.e., speech rate, pause, and stress), as illustrated below. The acoustic analysis of this study applied the digital speech-analysis software Praat (Version 6.1.53) (Boersma & Weenink, 2023). Initially, the acoustic measurements were conducted by the first author. Additionally, this study conducted an inter-rater reliability test to evaluate the reliability of the acoustic data, in which 15% of the measurements were independently validated by a third qualified analyst. The Cronbach's α was .91, demonstrating that the two researchers exhibited internal consistency and responded to the speech samples in a highly comparable manner.

2.4.1 Vowel duration

Regarding vowel duration, there is a basic distinction between Chinese and Thai. Thai contains phonemically short and long vowels, while Chinese does not. This study investigated how Chinese speakers of Thai developed a contrast in vowel duration. Vowel duration encompasses the duration of both monophthongs and diphthongs. The duration of monophthongs was determined by placing cursors on the waveform at the onset of periodicity and the termination of the final pitch period of the monophthongs. In defining the starting and ending points of vowels, the authors employed both the waveform and spectrogram views to ensure more accurate endpoint determination. When necessary, auditory judgment was employed to aid in the segmentation of vowels. For vowels preceding nasals, the endpoint was defined by the shift in formant structure, specifically observing the transitions in F1 and F2 as they approached nasal murmur frequencies, which provided a clearer boundary. In the case of diphthongs, the duration of diphthongs was also measured by dividing it into the 1st vocalic element, the transition, and the 2nd vocalic element in addition to measuring the overall duration of diphthongs (Roengpitya, 2002).

2.4.2 Vowel quality

Pertaining to vowel quality, it denotes the degree of precision with which Chinese pronounced Thai vowels in terms of their specific point of articulation. Chinese and Thai have unique inventories of vowels, comprising both shared and distinctive vowel phonemes. The development of vowel quality invariably entails the generation of novel phoneme categories and the restructuring of preexisting phoneme boundaries. This study examined the first two formant frequencies (F1 and F2) of each vowel. To minimize the impact of initial and final consonants, the quality of monophthongs was assessed using the mean formant frequency values of F1 and F2, calculated from 25% to 75% throughout a normalized duration of 100% for each monophthong. The diphthong quality was assessed by recording the mean formant frequency values of F1 and F2 across a range of 25% to 75%, with each vocalic element normalized to 100%.

2.4.3 Speech rate

As one of the most essential fluency-based aspects of L2 speech, speech rate has been found commonly to be used at a slower rate by learners who speak L2 compared to native speakers (Munro & Derwing, 1995), which could be due to limitations in processing, encoding, and recalling phonological information, or challenges in articulating L2 speech. A slow rate of speech is frequently viewed as an obstacle to understanding (Munro & Derwing, 1998), and an indication of not being a native speaker (Anderson-Hsieh & Koehler, 1988). In this study, speech rate was determined by calculating the ratio of the number of uttered syllables to the overall time of the utterance. If Chinese speakers of Thai were to speak Thai at a speech rate comparable to native speakers, their speech rate ratio would be expected to be equivalent to that of native Thai speakers.

2.4.4 Pause

Pause is another fluency-based suprasegmental feature that is likely associated with comprehensibility since it impacts the overall prosodic structure of L2 speech (Trofimovich & Baker, 2007). Producing pauses in L2 may suggest that the speaker is facing challenges with the task, potentially due to limitations in processing or memory specific to L2 speech. Learners often experience more frequent and extended pauses in their L2 compared to their native language. This study defined pauses as any interruption in the flow of speech that lasted longer than 100 ms. Pauses were categorized into two types: silent pauses and filled pauses. Silent pauses referred to gaps in the speech stream that were devoid of sound, while filled pauses

encompassed any interruptions, such as "Ah" or "Em", which interrupted the flow of speech. In identifying the silent pauses, both the waveform and spectrogram views were carefully examined to distinguish actual pauses from the silent closures of plosives. For true pauses, the endpoint was defined as the beginning of the closure phase for plosives, based on the absence of any formant structure or vocal fold activity preceding the onset of the closure. Both silent and filled pauses were evaluated based on their frequency and duration. The frequency of pauses was determined by taking the average number of pauses for each participant over the 30 pictures, while the duration of pauses was estimated by averaging the pause time for each person throughout the 30 pictures. If Chinese learners of Thai were able to create L2 Thai speech with a level of fluency comparable to native speakers, then their pauses, as evaluated by both frequency and duration, should not exhibit any significant differences when compared to those of native Thai speakers.

2.4.5 Stress

Thai is a stress-timed language that follows a weak-strong pattern in disyllabic words, where stressed syllables are notably longer than unstressed ones (Peyasantiwong, 1986; Potisuk et al., 1996). Chinese is a syllable-timed language that does not have variations in stress levels and has syllables of similar duration. How Chinese speakers of Thai produced the unstressed-stressed pattern of the disyllabic words in the word list (see Table 1) was examined. The duration of stressed and unstressed syllables was assessed using Praat software by placing two cursors at the beginning and end of each syllable. The syllable ratio was calculated by dividing the duration of unstressed syllables by stressed syllables. If Mandarin Chinese speakers adopted Thai-like stress patterns, the ratio of unstressed-to-stressed syllables should be similar to that of native Thai speakers, about 0.5. If Chinese-like syllable timing is utilized, the ratio should be higher, nearing 1.0, as unstressed and stressed syllables have similar duration.

3. RESULTS

Generally speaking, Chinese speakers of Thai differed in terms of both segmentals and suprasegmentals in comparison to native Thai speakers. The native Thai raters also effectively evaluated Thai speech produced by Chinese speakers of Thai to be more difficult to understand in the comprehensibility perception task. Following that, a series of statistical analyses were performed to determine that L2 comprehensibility was a multidimensional concept that contained interplay between L2 segmental and suprasegmental features.

3.1 Acoustic features of segmentals and suprasegmentals produced by Chinese speakers of Thai

3.1.1 Vowel duration

Figure 1 shows the average duration of short and long monophthongs pronounced by Chinese speakers of Thai and native Thai speakers. To make comparisons between the vowel duration of the two groups of speakers, this study performed a z-score normalization for each speaker to control for individual speaking rate variations, ensuring a more standardized basis for comparison between native Thai speakers and Chinese-accented Thai speakers. The normalized dataset was further submitted to a series of *t*-tests to clarify how Chinese speakers differed from native Thai speakers in terms of vowel duration.

Chinese speakers of Thai were able to discern the durational differences between short and long monophthongs, despite the absence of such distinctions in duration in their native language. Nevertheless, it is also apparent that they consistently pronounced Thai monophthongs with a longer duration compared to native Thai speakers. This was particularly apparent in short monophthongs, as indicated by a significant *p*-value of less than .01 for all six short monophthongs examined. Moreover, Chinese speakers exhibited varying levels of proficiency in pronouncing different monophthongs. Out of the six pair monophthongs targeted in this study, three pairs were found to have the highest degree of non-native proficiency, including /ɤ/-/ɤː/ (panel d), /o/-/oː/ (panel e), and /ɔ/-/ɔː/ (panel f). The duration of these monophthongs, as produced by the two groups of speakers, demonstrates a considerable variation compared to native Thai speakers, reaching a level of statistical significance of at least $p < .05$. On the other hand, the most native-sounding performance is observed in the two long monophthongs /ɛː/ (panel b) and /ɯː/ (panel c). This is because there is no significant variation in the duration of these two long monophthongs, whether or not they include a coda.

Numerous prior studies have demonstrated that the duration of long vowels is roughly double that of short vowels (Abramson, 1974; Misra et al., 2010; Tsukada, 2009). The results of the present study reveal a durational ratio of short vowels, long vowels with a coda, and long vowels without a coda as 1: 1.69: 2.43, closely aligning with established duration ratios from prior research. Meanwhile, Chinese speakers of Thai also distinguish vowel duration in these three syllable patterns with a ratio of 1: 1.23: 1.72. However, the distinctions between short and long vowels are less evident than those articulated by native Thai speakers.

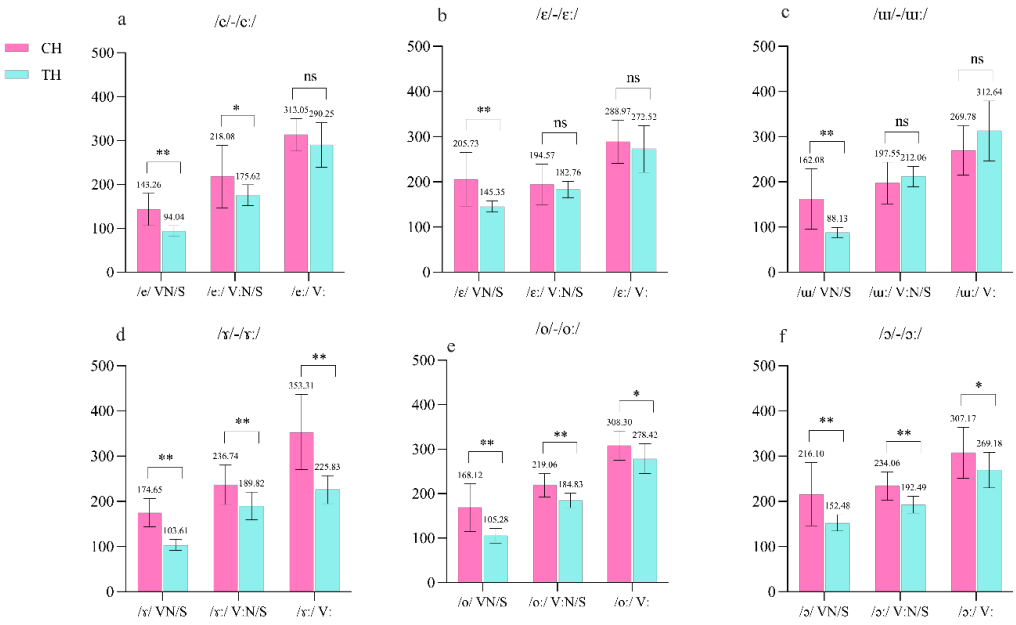

Figure 1: Monophthong duration

Chinese speakers of Thai extended the duration of Thai diphthongs, similar to monophthongs, in contrast to native Thai speakers (Figure 2). In addition, Chinese speakers of Thai also exhibited distinct realization compared to native Thai speakers when the diphthongs were divided into the 1st vocalic element, the transition, and the 2nd vocalic element. The greatest significant disparities were found in the duration of the 1st and 2nd vocalic elements. Among native Thai speakers, diphthong pronunciation is characterized by the 1st vocalic element occupying a larger percentage, generally 49.50% of the overall duration (panel d), whereas the 2nd vocalic element has a significantly smaller portion (28.46% in panel d). Nevertheless, Chinese speakers of Thai displayed a significant disparity in the length of diphthongs when compared to native Thai speakers. Although the 1st vocalic element has the greatest duration (43.33% in panel d), its prominence is significantly diminished by the 2nd vocalic element (35.95% in panel d). Thus, it can be deduced that the diphthongs in Chinese speakers of Thai are differentiated by a shortened 1st vocalic element (i.e., /i/, /ɯ/, and /u/) and a lengthened 2nd vocalic element (i.e., /a/).

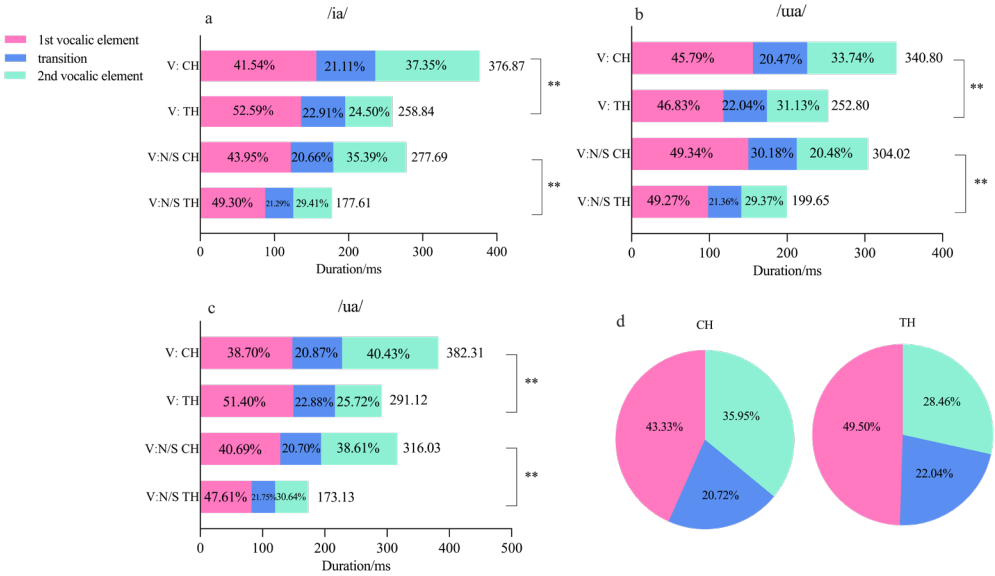

Figure 2: Diphthong duration

3.1.2 Vowel quality

In general, Chinese speakers of Thai produced distinct monophthong quality in comparison to native Thai speakers, as seen by the significant differences in F1 or F2 values depicted in Figure 3. In addition, Chinese speakers of Thai also demonstrated different degrees of proficiency in pronouncing various Thai monophthongs. More precisely, the monophthong /ɛː/ (panel b) achieves the highest level of nativeness compared to the other monophthongs investigated, as both its F1 and F2 values are not significantly distinct from those of native Thai speakers. On the other hand, the monophthongs /ɔ/-/ɔː/ (panel f) are likely the least native vowel sounds, as indicated by the significantly distinct F1 and F2 values of these two monophthongs.

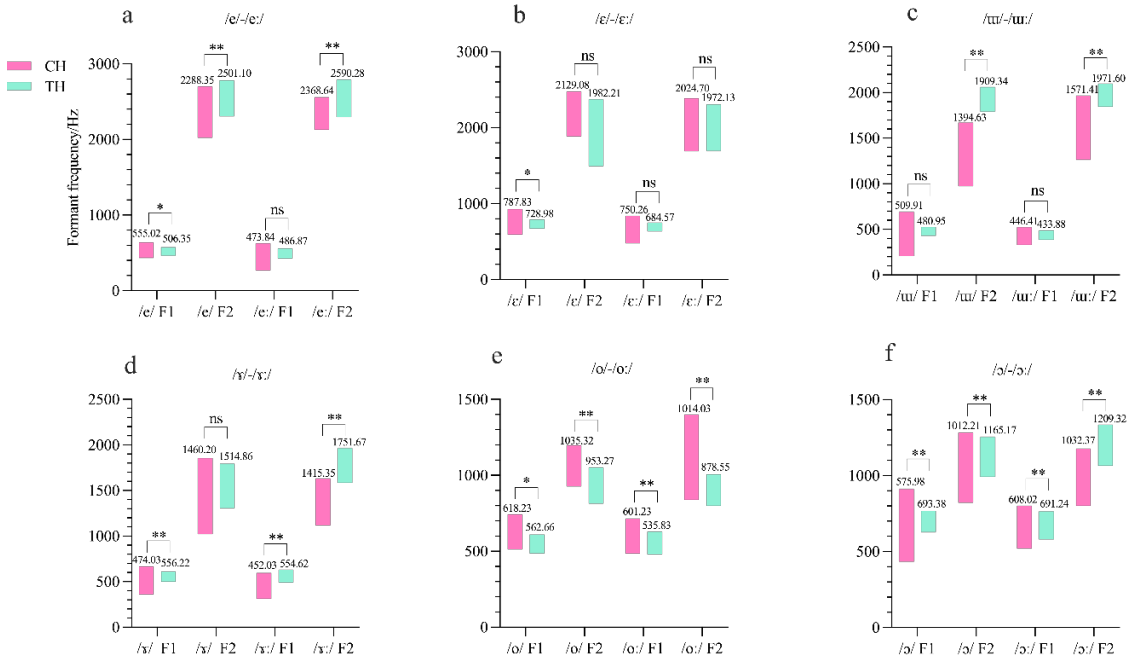

Figure 3: Monophthong quality

Overall, the diphthongs produced by Chinese speakers of Thai are similar in quality to the diphthongs produced by native Thai speakers, in contrast to the distorted qualities of monophthongs. The two groups of participants showed similar spectral trajectories in their pronunciation of the three diphthongs (see Figure 4). Chinese speakers of Thai have achieved a high level of competence in producing Thai diphthongs in terms of vowel quality.

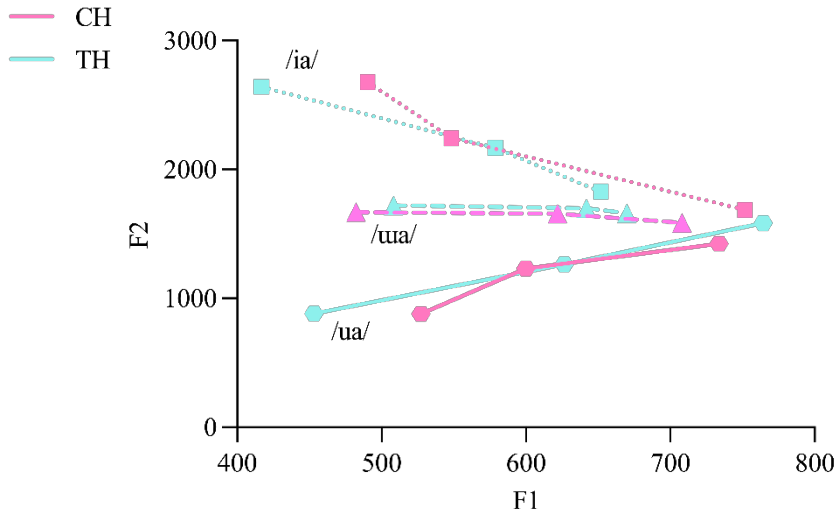

Figure 4: Diphthong quality

3.1.3 Speech rate

Regarding the speech rate shown in Figure 5, it was observed that native Thai speakers articulated 3.46 syllables per second, but Chinese speakers of Thai only produced 1.95 syllables per second. The speech rate ratios were subjected to an independent sample *t*-test to compare the two groups of speakers. The investigation demonstrated a significant disparity, as Chinese speakers of Thai produced a significantly fewer number of syllables compared to native Thai speakers, indicating that Chinese speakers of Thai had a lower level of fluency in speaking Thai as they articulated a smaller number of syllables and likely conveyed less information to raters.

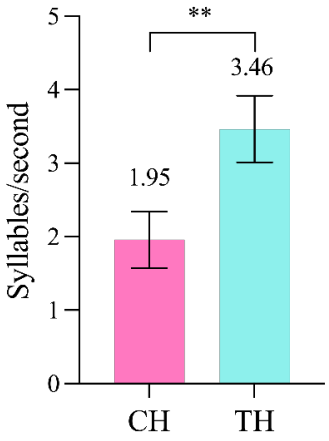

Figure 5: Speech rate

3.1.4 Pauses

Figure 6 depicts the number (panel a) and duration (panel b) of silent and filled pauses, respectively. Chinese speakers of Thai had a greater frequency of silent and filled pauses when producing Thai speech flow. However, only the number of filled pauses demonstrates a significant difference. Moreover, they also showed longer durations for both filled and silent pauses compared to native Thai speakers. Yet only the difference in filled pauses is statistically significant. Hence, the primary distinction between the two groups of speakers resides in their use of filled pauses. Increased frequency and duration of filled pauses of Chinese speakers may result in the awkward and fragmented flow of L2 Thai speech, making it perceptually challenging to comprehend.

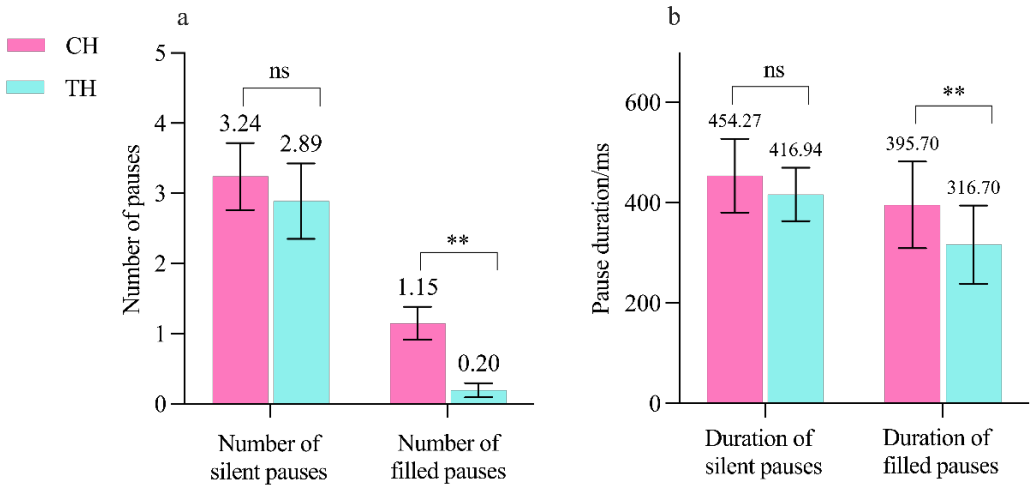

Figure 6: Pauses

3.1.5 Stress

The ratio between the duration of stressed and unstressed syllables is contrasted in Figure 7. The ratios obtained for each speaker were submitted to an independent sample *t*-test to compare the two speaker

groups. The results demonstrated that Chinese speakers of Thai showed a significantly increased stress ratio compared to native Thai speakers. As previously stated, Mandarin Chinese is a language that follows a syllable-timed pattern, while Thai is a stress-timed language. The stress ratio derived from Chinese speakers of Thai is expected to approximate 1.0, whereas that of native Thai speakers is anticipated to approximate 0.5. Figure 7 indicates that Chinese speakers of Thai were negatively impacted by their native language when it came to producing stress in Thai.

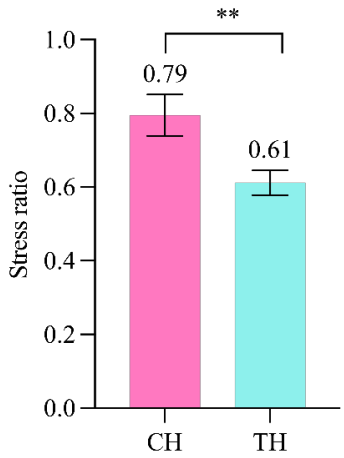

Figure 7: Stress

3.2 Rating of comprehensibility

Mean comprehensibility scores for each speaker were determined by averaging the rating scores provided by 30 native raters for the 30 picture narratives, as illustrated in Figure 8. Reliability was calculated to evaluate the consistency of the ratings. Native raters showed a high degree of agreement in their comprehensibility rating, demonstrated by a high Cronbach's α coefficient of .85. Native Thai speakers were rated as being remarkably more comprehensible by native raters, with an average score of 7.52. Chinese speakers of Thai received an average comprehensibility score of 5.73, indicating that their L2 Thai speech was viewed as more difficult to understand. A significant difference was found in the comprehensibility rating between the two speaker groups, $t(28) = -5.29$, $p < .01$. The comprehensibility perception experiment effectively allowed native raters to distinguish between non-native speakers and native speakers based on their level of comprehensibility.

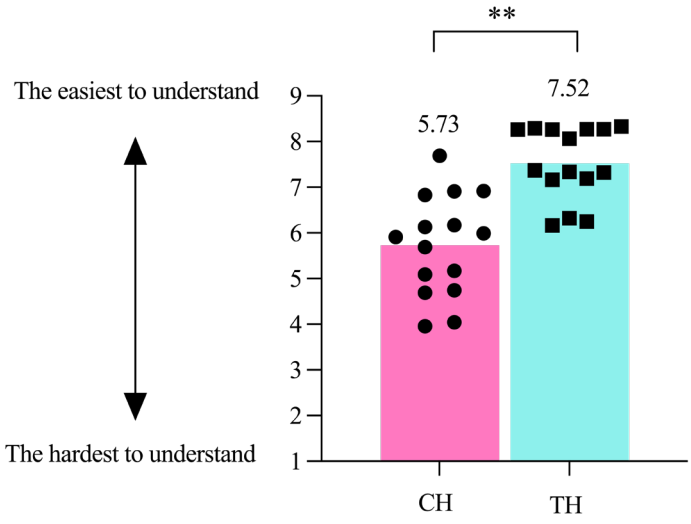

Figure 8: Comprehensibility rating results

3.3 Segmentals and suprasegmentals contributed to L2 comprehensibility

The previous analyses show that Chinese speakers of Thai had varying levels of proficiency in producing L2 segmentals and suprasegmentals. They were also perceived as more difficult to comprehend by native raters. The following analysis aimed to identify the particular segmentals and suprasegmentals that influenced the perception of L2 speech comprehensibility and to assess their impact. Correlation analysis, principal component analysis, and regression analysis were performed to achieve this objective. The results of these three analyses will be discussed in turn.

3.3.1 Correlation analysis

A preliminary correlation analysis examined the bivariate correlation between comprehensibility and all segmentals or suprasegmentals. Table 2 reports statistically significant correlations. Firstly, it is crucial to highlight the importance of suprasegmentals, which had a considerably significant correlation with comprehensibility. A negative correlation was found between stress and comprehensibility, showing that higher stress ratios were associated with lower levels of comprehensibility. Chinese speakers of Thai did not show a notable decrease in the length of the initial syllable in disyllabic words, in contrast to native speakers. This led to an increased stress ratio, which hindered the L2 speech's comprehensibility. A negative correlation exists between the number and duration of filled pauses and comprehensibility. More frequent and longer filled pauses would lead to reduced comprehensibility. Conversely, a direct correlation was found between speech rate and comprehensibility. The native raters demonstrated greater comprehension when the speaker used Thai speech with more syllables. In terms of segmentals, the duration of diphthongs was found to be linked to how understandable they were perceived. Chinese speakers of Thai have shown a distinct pronunciation of Thai diphthongs, as mentioned before. The 1st vocalic element was shortened, and the 2nd vocalic element was lengthened. Thus, Thai raters considered this specific pronunciation difficult to understand. Besides this, the duration of short monophthongs /e/, /ɯ/, /ɤ/, and long monophthong /ɔː/ followed by a coda also negatively correlated to comprehensibility, indicating a longer duration of vowels made L2 Thai speech harder to comprehend. Besides these, six quality features of vowels were also significantly correlated to comprehensibility, namely F1 of /ɔ/ and the 1st vocalic element of /ia/, as well as F2 of /ɯ/, /ɯː/, /ɔː/, and the 2nd vocalic element of /ia/. The remaining segmentals and suprasegmentals did not exhibit a significant correlation with comprehensibility and were not submitted for further analyses.

Table 2: Significantly correlated L2 features with comprehensibility

L2 features	*r*	*p*
Number filled pauses	-.70	.000**
Speech rate	.69	.000**
Duration of the 2nd vocalic element of /ua/ V:N/S	-.65	.000**
F2 of /ɔː/	.64	.000**
Duration of /ua/ V:N/S	.62	.000**
F1 of the 1st vocalic element of /ia/	-.61	.000**
Duration of /ɯ/	-.56	.001**
Duration of the 2nd vocalic element of /ia/ V:	-.55	.002**
Duration of the 2nd vocalic element of /ua/ V:	-.55	.002**
Duration of /ɤ/	-.54	.002**
Stress	-.54	.002**
Duration of the 1st vocalic element of /ua/ V:	.54	.002**
F2 of /ɯ/	.52	.003**
Duration of the 1st vocalic element of /ia/ V:	.52	.004**
F2 of the 2nd vocalic element of /ia/	.51	.004**
Duration of /ɔː/ V:N/S	-.50	.005**
Duration of /ua/ V:N/S	-.50	.005**
Duration of /e/	-.50	.005**
Duration of the 2nd vocalic element of /ia/ V:N/S	-.49	.007**
Length of filled pauses	-.47	.008**
F2 of /ɯː/	.47	.009**
F1 of /ɔ/	.46	.010**

(** $p < .01$)

3.3.2 Principal component analysis

The correlation analysis above indicated that 22 segmentals and suprasegmentals were strongly correlated with L2 comprehensibility. This study then performed a principal component analysis to combine

these variables into groups and enhance the interpretability of the dataset. Table 3 shows that the principal component analysis effectively decreased the number of segmentals and suprasegmentals from 22 to four components while preserving the maximum information. The four components explained 79.42% of the variance in the initial 22 segmentals and suprasegmentals. Component 1 consisted solely of segmentals. Two vowel quality features are involved, namely F1 of /ɔ/ and F2 of /ɔː/. Additionally, this component included several vowel durational features, encompassing the duration of both monophthongs and diphthongs. The duration of /ɯ/ and /e/, and the duration of /ɔː/ when followed by a coda, were included in Component 1. Three durational features of the diphthong /ua/ followed by a coda, specifically. The total duration and the duration of its 1st and 2nd vocalic elements were also taken into account. Component 2 included both segmentals and suprasegmentals. It measured the duration of the 1st and 2nd vocalic element in both /ia/ and /ua/ across different syllable structures, as well as the duration of /ɤ/ and F2 of /ɯː/. In addition to these, stress and duration of filled pauses, as two suprasegmentals are incorporated into this component, rendering Component 2 multidimensional. Component 3 exclusively concentrated on the quality features of vowels, particularly analyzing F2 of /ɯ/, the 2nd vocalic element of /ia/, and F1 of the 1st vocalic element of /ia/. Component 4 is a purely suprasegmental element encompassing two suprasegmental properties, namely the number of filled pauses and speech rate.

Table 3: Principal component loadings of segmentals and suprasegmentals that were correlated to comprehensibility

L2 features	Components			
	1	2	3	4
Duration of /ɯ/	.84			
Duration of /e/	.79			
Duration of the 1st vocalic element of /ua/ V:N/S	-.77			
Duration of the 2nd vocalic element of /ua/ V:N/S	.75			
F1 of /ɔ/	-.61			
Duration of /ɔː/ V:N/S	.60			
F2 of /ɔː/	-.59			
Duration of /ua/ V:N/S	-.59			
Duration of the 1st vocalic element of /ia/ V:		-.75		
Duration of the 2nd vocalic element of /ia/ V:N/S		.73		
Duration of /ɤ/		.72		
Duration of the 2nd vocalic element of /ia/ V:		.72		
Length of filled pauses		.70		
Duration of the 2nd vocalic element of /ua/ V:		.65		
Duration of the 1st vocalic element of /ua/ V:		-.64		
F2 of /ɯː/		-.59		
Stress		.57		
F2 of the 2nd vocalic element of /ia/			.82	
F2 of /ɯ/			.62	
F1 of the 1st vocalic element of /ia/			-.55	
Number of filled pauses				.86
Speech rate				.84

3.3.3 Regression analysis

The four identified principal components (i.e., predictor variables) and the comprehensibility scores (i.e., predicted variable) were subjected to a stepwise regression analysis to investigate the segmentals and suprasegmentals that mostly contributed to comprehensibility. Table 4 is the regression model generated from the regression analysis, consisting of all four components. This regression model has achieved the ability to account for 54.4% of the variance in comprehensibility according to the Adjusted R^2 value. All the components discovered in the principal component analysis were significant predictors of comprehensibility. Nevertheless, their predictability exhibited dissimilarities, as evidenced by the standardized coefficients presented in Table 4. The standardized coefficient with the maximum value was -.552, corresponding to Component 4. It suggests that a one-unit rise in Component 4 resulted in a drop of .552 units in the comprehensibility of connected speech. The standardized coefficients for Component 1 were -.488, indicating that a one-unit increase in Component 1 led to a corresponding drop of .488 units in comprehensibility. The standardized coefficients for Component 2 were reported as -.331. An increase of one unit in Component 2 resulted in a decrease of -.331 units in comprehensibility. The predictability of Component 3 was the lowest, with standardized coefficients of .272. This suggests that a one-unit change in Component 3 would lead to a .272 change in comprehensibility.

Table 4: Results of regression analysis using segmentals and suprasegmentals as predictors for comprehensibility

Predictor variables	R^2	Adjusted R^2	p	Standardized coefficients	Sig
Component 4	.607	.544	.000	-.522	.000
Component 1				-.488	.005
Component 2				-.331	.014
Component 3				.272	.040

Hence, it can be seen there is interplay between the segmentals and suprasegmentals in terms of the perception of L2 comprehensibility. The two suprasegmentals encompassed in Component 4, the number of filled pauses and speech rate, played the most crucial role in how native Thai raters perceived comprehensibility. Additionally, the segmentals in Component 1 are also fairly reliable predictors of comprehensibility. For instance, the duration of two short monophthongs is particularly outstanding, including the duration of /ɯ/ and /e/. Moreover, the durational features of diphthongs are also emphasized, particularly the duration of the 1st and 2nd vocalic elements in /ua/. The quality features of some back monophthongs are also relatively reliable predictors of comprehensibility. The segmental and suprasegmental features contained in Components 2 and 3 show a relatively limited level of predictability in predicting comprehensibility.

4. DISCUSSION AND CONCLUSIONS

This study investigated the segmentals and suprasegmentals that contribute to the perceived comprehensibility of L2 Thai speech by Chinese-speaking learners of Thai. Models predicting comprehensibility based on specific segmental and suprasegmental parameters showed that suprasegmentals, especially filled pauses and speech rate, are the primary factors influencing perceived L2 comprehensibility. Moreover, the control of segmentals, such as the duration of diphthongs and the quality features of back monophthongs, also impacts the comprehensibility of L2 Thai speech. The results of this study offer solid evidence that L2 comprehensibility is complex and interconnected with several L2 features in the areas of segmentals and suprasegmentals. Hence, it is necessary to examine the interplay between segmentals and suprasegmentals to better understand the nature of L2 comprehensibility in Thai.

It is also crucial to acknowledge that specific L2 features related to suprasegmentals appear to significantly influence how comprehensible L2 speech is perceived by native raters. This is consistent with previous research that examined the impact of L2 phonological features on the degree of comprehensibility (Bergeron & Trofimovich, 2017; Kang et al., 2010; Pongprairat & Luksaneeyanawin, 2013; van Maastricht et al., 2021). Suprasegmentals such as speech rate, stress, pause, and intonation are linked to the comprehensibility of L2 speech. Disagreement exists over the importance of segmentals compared to suprasegmentals in improving the comprehensibility of L2 speech. The results of the present study demonstrate the critical significance of suprasegmentals in L2 comprehensibility. However, it must be acknowledged that suprasegmentals are not adequately addressed in L2 instruction because the current emphasis in L2 teaching is mainly on segmentals, which include vowels and consonants. It would be beneficial for L2 learners to allocate instructional time more effectively by incorporating training on suprasegmentals. Since the primary goal of L2 instruction is to attain comprehensible L2 speech, highlighting the vital significance of suprasegmentals would undoubtedly help achieve this goal. On the other hand, emphasizing suprasegmentals does not inevitably diminish the significance of segmentals. The current study emphasizes the importance of integrating both suprasegmentals and segmentals in instruction in L2 instead of viewing them as a competitive dichotomy, as reflected in the interplay between segmentals and suprasegmentals in L2 comprehensibility.

The results of the regression analysis showed that a mix of suprasegmentals and segmentals explained around 50% of the variance in comprehensibility. What could explain the remaining 50% of the variance that was not addressed in the current study? The remaining variance could be attributed to other unstudied suprasegmentals or segmentals in the current study, such as intonation, tones, consonants, etc. Additionally, L2 linguistic features from the domains of lexico-grammar or discourse could also serve as reliable predictors for comprehensibility. Future studies should focus on more measurable L2 features to obtain a more thorough understanding of what constitutes a comprehensible L2 speech. In addition to including more L2 features, the current results must be validated with different groups of L2 speakers and raters to identify whether factors affecting comprehensibility are unique to a particular group of participants. It has been proven that the linguistic and social backgrounds of speakers and raters play an important part in L2 speech perception (Kang, 2012; Piske et al., 2001; Saito et al., 2019). The present study focused primarily on monolingual Thai speakers and raters who were unfamiliar with accented Thai speech and lacked linguistic training. Subsequent research could include a broader range of speakers and raters with diverse language and social backgrounds.

As an exploratory study, this study specifically targeted naturally produced L2 speech samples, which refers to the speech samples that had not been manipulated in terms of their acoustic properties to assess the specific impact size of any selected segmentals or suprasegmentals. Thus, these L2 segmentals or suprasegmentals that were evaluated exhibited inherent covariance, making it impossible to separate the influences of these properties and individually determine their impact on comprehensibility. In future studies, L2 speech samples can be modified or manipulated to examine the independent impact of certain suprasegmental and segmental elements. Furthermore, one approach is to extract segmental or suprasegmental properties from the speech samples of one group and overlay them over the speech samples of another group (Jilka, 2000; van Maastricht et al., 2021; Winters & O'Brien, 2013). Further research efforts will enhance our understanding of L2 comprehensibility phenomena in a broader sense, and contribute specifically to the ongoing discussion over the comparative significance of suprasegmentals and segmentals in determining perceived L2 comprehensibility.

This study provides important insights into the acoustic properties of L2 Thai speech articulated by Chinese learners. However, certain methodological limitations must be acknowledged. The implementation of a picture description task, although efficient in eliciting naturalistic speech, introduced variability in speech rate that may have affected the reliability of acoustic measurements, such as vowel duration. Further research could benefit from adopting speaking tasks with more controlled speech contexts or implementing various normalization methods to address this variability. Additionally, the criteria for identifying segmental boundaries, particularly in syllables with nasal finals, could be improved by incorporating other acoustic cues, such as spectrum or energy levels, to enhance accuracy. The differentiation between silent pauses and the closure phases of stop consonants presents a potential source of inaccuracy, underscoring the need for stricter segmentation techniques or more sophisticated acoustic analysis tools. Furthermore, it should be noted that some acoustic demarcations, including those employed in the present study, may involve a degree of arbitrariness. It is essential to adopt a well-defined arbitrary measure and apply it consistently. Addressing these challenges in future research would enhance the reliability and generalizability of findings in the study of L2 speech.

REFERENCES

Abramson, A. (1974). Experimental phonetics in phonology: Vowel duration in Thai. *PASAA*, *4*(1), 71–90. https://doi.org/10.58837/CHULA.PASAA.4.1.3

Anderson-Hsieh, J., & Koehler, K. (1988). The effect of foreign accent and speaking rate on native speaker comprehension. *Language Learning, 38*, 561–613. https://doi.org/10.1111/j.1467-1770.1988.tb00167.x

Bergeron, A., & Trofimovich, P. (2017). Linguistic dimensions of accentedness and comprehensibility: Exploring task and listener effects in 2nd language French. *Foreign Language Annals*, *50*(3), 547–566. https://doi.org/10.1111/flan.12285

Boersma, P., & Weenink, D. (2023). *Praat: doing phonetics by computer* (*Version 6.3.09)* [Computer software]. http://www.praat.org/

Crowther, D., Trofimovich, P., Saito, K., & Isaacs, T. (2018). Linguistic dimensions of L2 accentedness and comprehensibility vary across speaking tasks. *Studies in Second Language Acquisition*, *40*(2), 443–457. https://doi.org/10.1017/S027226311700016X

Derwing, T. M., & Munro, M. J. (2009). Putting accent in its place: Rethinking obstacles to communication. *Language Teaching*, *42*, 476–490. https://doi.org/10.1017/S026144480800551X

Derwing, T. M., Rossiter, M. J., Munro, M. J., & Thomson, R. I. (2004). Second language fluency: Judgments on different tasks. *Language Learning*, *54*(4), 655–679. https://doi.org/10.1111/j.1467-9922.2004.00282.x

Edwards, J., & Zampini, M. (2008). *Phonology and second language acquisition*. John Benjamins.

Flege, J., & Hillenbrand, J. (1984). Limits on phonetic accuracy in foreign language speech production. *Journal of the Acoustical Society of America, 76*, 708–721. https://doi.org/10.1121/1.391257

Gass, S., & Varonis, E. (1984). The effect of familiarity on the comprehensibility of nonnative speech. *Language Learning, 34*, 65–89. https://doi.org/10.1111/j.1467-1770.1984.tb00996.x

Grabe, E., & Low, E. (2002). Durational variability in speech and the Rhythm Class Hypothesis. In C. Gussenhoven & N. Warner (Ed.), *Laboratory phonology 7* (pp. 515–546). De Gruyter Mouton. https://doi.org/10.1515/9783110197105.2.515

Jilka, M. (2000). *The contribution of intonation to the perception of foreign accent: Identifying intonational deviations by means of F0 generation and resynthesis* [Doctoral dissertation, Universitat Stuttgart, Germany]. CiteSeer.

Kang, O. (2012). Impact of rater characteristics and prosodic features of speaker accentedness on ratings of

international teaching assistants' oral performance. *Language Assessment Quarterly*, *9*(3), 249–269. https://doi.org/10.1080/15434303.2011.642631

Kang, O., Rubin, D. O. N., & Pickering, L. (2010). Suprasegmental measures of accentedness and judgments of language learner proficiency in oral English. *The Modern Language Journal*, *94*(4), 554–566. https://doi.org/10.1111/j.1540-4781.2010.01091.x

Lee, W. S., & Zee, E. (2003). Standard Chinese (Beijing). *Journal of the International Phonetic Association*, *33*(1), 109–112. https://doi.org/10.1017/S0025100303001208

Levis, J. (2005). Changing contexts and shifting paradigms in pronunciation teaching. *TESOL Quarterly*, *39*, 367–377. https://doi.org/10.2307/3588485

Lianghiranthaworn, S., & Chapoo, L. (2019). Thai language pronunciation problems of Chinese overseas students of Maejo University, Chiang Mai, Thailand. *Journal of Liberal Arts, Maejo University*, *7*(2), 114–129.

Major, R., & Faudree, M. (1996). Markedness universals and the acquisition of voicing contrasts by Korean speakers of English. *Studies in Second Language Acquisition, 18*, 69–90. https://doi.org/10.1017/S0272263100014686

Misra, D., Yadav, R., Punyayodhin, S., & Narang, V. (2010). Durational contrast and centralization of vowels in Hindi and Thai. In M. Dong, G. Zhou, H. Qi, & M. Zhang (Eds.), *2010 International Conference on Asian Language Processing* (pp. 339–342). IEEE. https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=5681541&isnumber=5681540

Munro, M. J., & Derwing, T. M. (1995). Processing time, accent, and comprehensibility in the perception of foreign-accented speech. *Language and Speech, 38*, 289–306. https://doi.org/10.1177/002383099503800305

Munro, M. J., & Derwing, T. M. (1998). The effects of speaking rate on listener evaluations of native and foreign-accented speech. *Language Learning*, *48*, 159–182. https://doi.org/10.1111/1467-9922.00038

Ortega, L. (2009). *Understanding second language acquisition*. Hodder Arnold.

Peyasantiwong, P. (1986). Stress in Thai. In R. J. Bickner, T. J. Hudak, & P. Peyasantiwong (Eds.), *Papers from a Conference on Thai Studies in Honor of William J. Gedney* (pp. 211–230). Michigan papers on South and Southeast Asia, Center for South and Southeast Asian Studies, University of Michigan, Ann Arbor.

Phonphanich, S. H., & Burusphat, S. (2021). The positive and negative effects of tonal L1 on the acquisition of tonal L2: A case of the standard Thai tones and the Chinese Zhuang learners. *LEARN Journal: Language Education and Acquisition Research Network*, *14*(2), 28–58.

Piske, T., MacKay, I. R., & Flege, J. E. (2001). Factors affecting degree of foreign accent in an L2: A review. *Journal of Phonetics*, *29*(2), 191–215. https://doi.org/10.1006/jpho.2001.0134

Plaengsorn, R. (2017). *The science of teaching Thai as a foreign language*. Chulalongkorn University Printing House. [in Thai]

Pongprairat, R., & Luksaneeyanawin, S. (2013). Degree of comprehensibility in Thai accented English rated by native speaker judges. *Journal of Language Teaching and Learning (PASAA)*, *45*, 35–55. https://doi.org/10.58837/CHULA.PASAA.45.1.3

Potisuk, S., Gandour, J., & Harper, M. P. (1996). Acoustic correlates of stress in Thai. *Phonetica*, *53*(4), 200–220. https://doi.org/10.1159/000262201

Roengpitya, R. (2002, June). Different durations of diphthongs in Thai: A new finding. In P. Chew (Ed.), *Annual Meeting of the Berkeley Linguistics Society* (pp. 43–54). Berkeley Linguistics Society.

Saito, K., Tran, M., Suzukida, Y., Sun, H., Magne, V., & Ilkan, M. (2019). How do second language listeners perceive the comprehensibility of foreign-accented speech?: Roles of first language profiles, second language proficiency, age, experience, familiarity, and metacognition. *Studies in Second Language Acquisition*, *41*(5), 1133–1149. https://doi.org/10.1017/S0272263119000226

Saito, K., Trofimovich, P., Isaacs, T., & Webb, S. (2016). Re-examining phonological and lexical correlates of second language comprehensibility: The role of rater experience. In T. Isaacs & P. Trofimovich (Eds.), *Interfaces in second language pronunciation assessment: Interdisciplinary perspectives* (pp. 141–156). Multilingual Matters.

Suzukida, Y., & Saito, K. (2021). Which segmental features matter for successful L2 comprehensibility? Revisiting and generalizing the pedagogical value of the functional load principle. *Language Teaching Research*, 25(3), 431–450. https://doi.org/10.1177/1362168819858246

Teeranon, P. (2016). An acoustic analysis and perceptual study of Thai tones produced by Vietnamese and Khmer speakers. *Journal of Language and Culture*, *35*(1), 81–100. [in Thai]

Trofimovich, P., & Baker, W. (2007). Learning prosody and fluency characteristics of second language speech: The effect of experience on child learners' acquisition of five suprasegmentals. *Applied Psycholinguistics*, *28*(2), 251–276. https://doi.org/10.1017/S0142716407070130

Trofimovich, P., & Isaacs, T. (2012). Disentangling accent from comprehensibility. *Bilingualism: Language and Cognition*, *15*(4), 905–916. https://doi.org/10.1017/S1366728912000168

Tsukada, K. (2009). An acoustic comparison of vowel length contrasts in standard Arabic, Japanese and Thai. In M. Zhang, H. Li, K. Lua, & M. Dong (Eds.), *2009 International Conference on Asian Language Processing* (pp. 76–79). IEEE. https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=5380774&isnumber=5380723

van Maastricht, L., Zee, T., Krahmer, E., & Swerts, M. (2021). The interplay of prosodic cues in the L2: How intonation, rhythm, and speech rate in speech by Spanish learners of Dutch contribute to L1 Dutch perceptions of accentedness and comprehensibility. *Speech Communication*, *133*, 81–90. https://doi.org/10.1016/j.specom.2020.04.003

Winters, S., & O'Brien, M. G. (2013). Perceived accentedness and intelligibility: The relative contributions of F0 and duration. *Speech Communication*, *55*(3), 486–507. https://doi.org/10.1016/j.specom.2012.12.006

Yi, L. (2017). A Comparison of the Acoustic Characteristics of Thai Vowels produced by Chinese and Thai Participants. *Humanities Journal*, *24*(1), 302–326. [in Thai]

Zhang, J., & Lee, S. (2019). Acquisition of English speech rhythm by Chinese learners of English at different English proficiency levels. *Phonetics and Speech Sciences*, *11*(4), 71–79. https://doi.org/10.13064/ksss.2019.11.4.071

APPENDIX


ชาวเกาะ ทะเล ปลาหมึก
islander sea squid
milk
Goat milk
นมแพะ ตุ๊กแก ลิ้นชัก
goat milk gecko drawer
หนังสือ ดินสอ นักเรียน
book pencil student
จิ้งจอก แตงโม ทุ่งนา
fox watermelon field
ห้องครัว แกงจืด ข้าวโพด
kitchen soup corn
ฝาแฝด แม่เหล็ก ของเล่น
twins magnet toy
ผีเสื้อ มะเฟือง ดอกไม้
butterfly carambola flower
กระจก กางเกง เด็กหญิง
mirror pants girl
ทางเดิน เยอะๆ ต้นไม้
pathway abundant/a lot of tree
ตำรวจ อ่อนเพลีย พบเจอ
police tired encounter