PRODUCTION AND PERCEPTION OF ENGLISH LEXICAL STRESS PATTERNS OF THAI AND CHINESE SPEAKERS

LI XINRUI

FACULTY OF LANGUAGES AND LINGUISTICS
UNIVERSITI MALAYA
KUALA LUMPUR

2024

PRODUCTION AND PERCEPTION OF ENGLISH LEXICAL STRESS PATTERNS OF THAI AND CHINESE SPEAKERS

LI XINRUI

DISSERTATION SUBMITTED IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF ARTS (LINGUISTICS)

FACULTY OF LANGUAGES AND LINGUISTICS
UNIVERSITI MALAYA
KUALA LUMPUR

2024

UNIVERSITI MALAYA

ORIGINAL LITERARY WORK DECLARATION

Name of Candidate: Li Xinrui

Matric No: s2016504

Name of Degree: Master of Arts (Linguistics)

Title of Project Paper/Research Report/Dissertation/Thesis ("this Work"):

PRODUCTION AND PERCEPTION OF ENGLISH LEXICAL STRESS PATTERNS OF THAI AND CHINESE SPEAKERS

Field of Study: Phonetics and Phonology

I do solemnly and sincerely declare that:

(1) I am the sole author/writer of this Work;
(2) This Work is original;
(3) Any use of any work in which copyright exists was done by way of fair dealing and for permitted purposes and any excerpt or extract from, or reference to or reproduction of any copyright work has been disclosed expressly and sufficiently and the title of the Work and its authorship have been acknowledged in this Work;
(4) I do not have any actual knowledge nor do I ought reasonably to know that the making of this work constitutes an infringement of any copyright work;
(5) I hereby assign all and every rights in the copyright to this Work to the University of Malaya ("UM"), who henceforth shall be owner of the copyright in this Work and that any reproduction or use in any form or by any means whatsoever is prohibited without the written consent of UM having been first had and obtained;
(6) I am fully aware that if in the course of making this Work I have infringed any copyright whether intentionally or otherwise, I may be subject to legal action or any other action as may be determined by UM.

Candidate's Signature Date: 2024-7-26

Subscribed and solemnly declared before,

Witness's Signature Date: 2024-07-31

Name:

Designation:

PRODUCTION AND PERCEPTION OF ENGLISH LEXICAL STRESS PATTERNS OF THAI AND CHINESE SPEAKERS

ABSTRACT

This dissertation examines Thai and Chinese English, characterized as a variety of English that may exhibit a lack of identifiable lexical stress. Consequently, this might impact the mutual capacity to be recognized in English. There are two main experiments. Firstly, to test the differences in English acoustic features produced by Thai and Chinese L1 speakers, trisyllabic English words with controlled parts of speech obtained from the British National Corpus (BNC) were used. Analyzing the impact of various tonal languages on the lexical stress on English words by studying how Thai and Chinese L1 speakers produce and perceive speech. Praat is used to visualize the acoustic properties of 25 English words generated by four speakers from Thailand and Mainland China. Then, the production results are used as a comparison item for the perception accuracy to study the L1 language effect on L2 lexical stress perception by different language groups. Acoustic analyses of both groups' production indicated they implemented word stress with a longer vowel duration. The stress perception results showed that both Mandarin Chinese and Thai listeners could recognize more than half of the stress positions. A careful examination of the produced acoustic data suggested that average F0 may have been responsible for their stress perception, especially in Chinese groups. The overall findings suggest strong L1 influences on the recognition of L2 English lexical stress. Meanwhile, stress familiarity and word classes also play the great role in stress perception. Further extensive acoustic research on the production and perception of stress in L2 by speakers of tonal languages will enhance our comprehension of how the prosodic system of L1 influences the learning of lexical stress in L2, resulting in a completer and more precise framework for understanding the acquisition of lexical stress across different English varieties. Future studies are needed to do a comparative analysis between the

production of English non-word stimuli and real-world stimuli to fully consider the intonation influence.

Keywords: Lexical stress, Chinese English, Thai English, Production and Perception, Acoustic features

PRODUKSI DAN PERSEPSI POLA TEKANAN LEKSIKAL BAHASA INGGERIS DALAM KALANGAN PENUTUR THAI DAN CINA

ABSTRAK

Disertasi ini mengkaji bahasa Inggeris Thai dan Cina, yang dicirikan sebagai pelbagai bahasa Inggeris yang mungkin menunjukkan kekurangan tekanan leksikal yang boleh dikenal pasti. Akibatnya, ini mungkin memberi kesan kepada keupayaan bersama untuk diiktiraf dalam bahasa Inggeris. Terdapat dua eksperimen utama. Pertama, untuk menguji perbezaan dalam ciri-ciri akustik bahasa Inggeris yang dihasilkan oleh penutur L1 Thai dan Cina, perkataan Inggeris trisuku kata dengan bahagian pertuturan terkawal yang diperoleh daripada British National Corpus (BNC) telah digunakan. Menganalisis kesan pelbagai bahasa tonal terhadap tekanan leksikal pada perkataan Inggeris dengan mengkaji cara penutur L1 Thai dan Cina menghasilkan dan mempersepsikan pertuturan. Praat digunakan untuk menggambarkan sifat akustik 25 perkataan Inggeris yang dijana oleh empat penutur dari Thailand dan Tanah Besar China. Kemudian, hasil produksi digunakan sebagai item perbandingan bagi ketepatan persepsi untuk mengkaji kesan bahasa L1 terhadap persepsi tekanan leksikal L2 oleh kumpulan bahasa yang berbeza. Analisis akustik produksi kedua-dua kumpulan menunjukkan mereka melaksanakan tekanan perkataan dengan tempoh vokal yang lebih lama. Hasil persepsi tekanan menunjukkan bahawa kedua-dua pendengar Cina Mandarin dan Thai dapat mengenali lebih separuh daripada kedudukan tekanan. Pemeriksaan yang teliti terhadap data akustik yang dihasilkan menunjukkan bahawa purata F0 mungkin bertanggungjawab terhadap persepsi tekanan mereka, terutamanya dalam kumpulan Cina. Dapatan keseluruhan mencadangkan pengaruh L1 yang kuat terhadap pengiktirafan tekanan leksikal bahasa Inggeris L2. Sementara itu, kebiasaan tekanan dan kelas perkataan juga memainkan peranan terbesar dalam persepsi tekanan. Penyelidikan akustik yang lebih meluas mengenai produksi dan persepsi tekanan dalam L2 oleh penutur bahasa tonal akan

meningkatkan pemahaman kita tentang bagaimana sistem prosodik L1 mempengaruhi pembelajaran tekanan leksikal dalam L2, menghasilkan rangka kerja yang lebih lengkap dan lebih tepat untuk memahami pemerolehan tekanan leksikal merentasi pelbagai jenis bahasa Inggeris. Kajian masa depan diperlukan untuk melakukan analisis perbandingan antara produksi rangsangan bukan perkataan bahasa Inggeris dan rangsangan dunia sebenar untuk mempertimbangkan sepenuhnya pengaruh intonasi.

Kata kunci: Tekanan leksikal, Bahasa Inggeris Cina, Bahasa Inggeris Thai, Produksi dan Persepsi, Ciri-ciri akustik

ACKNOWLEDGEMENTS

To everyone who has encouraged and helped me along the journey while I have pursued my master's degree, I am eternally grateful.

The most gratitude goes to my amazing supervisor Dr. Chiew Poh Shin, for all of the invaluable support and direction she has given me throughout this research. It would not have been possible for me to complete my studies without her wise counsel, motivating recommendations, and insightful criticism. Her wisdom and extensive knowledge of phonetics and phonology have always impressed me. I am really appreciative of the time and effort she put into this research and the patience with which she reviewed my dissertation.

I would also like to express my gratitude to my penal members: Dr. Roshidah Binti Hassan, Dr. Musaev Talaibek, and Dr. Chew Shin Yi. During my candidature defense, they provided valuable advice on improving my dissertation, pointing out deficiencies I had overlooked.

My great thanks go to my grandparents, parents, and my aunt, who love me unconditionally and give me endless care and financial support. I feel so fortunate to be their daughter.

During the ups and downs of grad school, I received support from these endearing people: my boyfriend Cai An, my bestie Li Jiarou, and Zhang Yu, who were encouraging when I felt down, and their patience with my myriad questions helped me finish the study. I sincerely thank my lovely graduate classmates: Peng Yao, Ed, Wen Jiajia, Hu Yurong, and Liu Xilin; they helped me throughout my postgraduate studies when I lacked clarity about the process. I could not have successfully completed this work without them.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS AND ABBREVIATIONS

AE	:	American English
F0	:	Pitch
μ	:	Mean
Σ	:	Standard Deviation
STR	:	Stressed Syllable
USTR	:	Unstressed Syllable

LIST OF APPENDICES

CHAPTER 1: INTRODUCTION

1.1 Background of the Study

English serves as the lingua franca of the global economy, playing a significant role in public discourse. Moreover, it serves as a gateway to education, prosperity, and other opportunities for many individuals. From less formal and limited kinds to more formal and widespread variations, from globally accepted to recently developed local standards, from linguistic changes influenced by contact to simplified languages and mixed languages, from acquired second languages to varieties learned by non-native speakers, and so forth. English has become so universally spoken that it may be considered "the language that is spoken continuously across the globe" (Seidlhofer, 2002; Jenkins, 2006; Luk & Lin, 2006). It has resulted in language diversification processes that are unmatched elsewhere.

The nature of the language background influences how we interpret speech sounds. Much research has been done on how the first language (L1) affects how speech segments are understood (Archibald, 1992; Pater, 1997; Peperkamp & Dupoux, 2002). According to the main current models of L2 speech, such as the Perceptual Assimilation Model (Best, 1995; Best & Tyler, 2007) and the Stress Typology Model (STM), the level of difficulty in learning L2 sounds is related to the differences between the phonetic and phonological structure of the listeners' L1 and that of the target language. At the same time, in second language learning (L2), production and perception, comprehensibility, and intelligibility are frequently used as assessment instruments. Comprehensibility refers to the ease with which a listener comprehends a speaker's message (Derwing & Munro, 1997). In contrast, intelligibility refers to whether a listener ultimately understands a speaker's intended words (Munro & Derwing, 1995). When evaluating the intelligibility of a speaker, a listener may inaccurately believe they comprehended words they did not understand (Munro & Derwing, 2015).

Munro (1998) defined speech produced by a second language (L2) learner as speech that differs from the phonetic characteristics of L1 speakers of the language in partly systematic ways. That is speech that varies in rhythm, articulation, symbolism, and tone (Hansen & Arslan, 1995). Listeners' linguistic backgrounds and prior exposure to various speech varieties are essential when understanding or evaluating speech with speech that is different from the L1 tongue and the speaker's characteristics. Although somehow accented speech is a significant indicator of an L2-linguistic background, its effects on communication are complicated. Accents can sometimes negatively affect the L2 speaker (Derwing et al., 2014). For instance, it may be challenging for audiences to comprehend speech that deviates from their typical patterns of oral production. Thus, accentuation may result in a loss of intelligibility (Pickering, 2006).

Languages differ not just in their phoneme repertoire but also in their suprasegmental features. The majority of languages tend to employ at least one suprasegmental feature to identify lexical items. Languages divide suprasegmental space in a similar way to how they divide segmental space. Some languages, like Mandarin and Thai, use tones to figure out what something means. Other languages, like Japanese, Norwegian, and Finnish, use pitch accents or duration to distinguish between lexical items. As well as this study's focus language, English, the use of lexical stress to distinguish meaning (Cutler, 2015; Yu & Andruski, 2010; Zhang et al., 2008). Acoustic characteristics, such as pitch, intensity, duration, and formant structure, vary between languages. These distinctions should result in several difficulties for speakers of a particular L1 language attempting to perceive and generate the L2 language clearly. It is generally accepted that L2 speech should be analyzed from diverse perspectives. The work of Isaacs and Trofimovich (2012) demonstrates that speech with accents influenced L1 French English learners' perceptions of recognizability and comprehensibility. The L2 speakers were assigned to describe a brief narrative based on images. Both untrained L1 English observers and TESL-trained

teachers evaluated their speech, with the latter offering comments on the linguistic factors that influenced the rating. One of the most important conclusions from his experiment is that word stress, a suprasegmental feature, was determined to be the most prominent trait for categorizing speakers into comprehensibility levels.

Lexical stress is when the emphasis is placed on a word's syllable. According to Ladefoged (1993), English lexical stress is segmentally related to long or unreduced vowels. Prosodically, fundamental frequency (F0), intensity, and duration serve as cues for lexical stress in English (Lehiste, 1996). According to Cutler and Clifton (1984), about 90% of the lexical stress in English words is placed on the first syllable. Consequently, this makes it simpler for listeners to split words into their component parts (Cutler & Norris, 1988; Donselaar et al., 2005). There are also secondary stressed syllables, typically less prominent than the stressed syllable but more pronounced than an unstressed syllable (Roach, 2009). This study does not examine secondary stress.

However, there appears to be no consensus regarding the significance of lexical stress in English. Jenkins (2002) argued that stress is one characteristic that is not crucial for communication in English perception as an International Language context. In Deterding (2013) 's research, one speaker generated several words with unexpected stress, although these seldom resulted in miscommunications. However, other aspects of his speech caused miscommunications. As a result, Deterding (2013) also reached the same conclusion, saying that "in English as a lingua franca (ELF), variable or unclear lexical stress rarely leads to misunderstandings." Previous research conducted by Kachru has similarly concurred that the recognition of diverse norms within English would not result in a dearth of comprehensibility among its various users. According to Jenkins (2000), lexical stress is a "grey area". It is not considered crucial to acceptability since it is seldom connected to misconceptions because it often happens with other distinct features. On the contrary, Lepage and Busà (2014) demonstrated that vowel reduction and stress impact

the perception and comprehension of speakers of second languages (L2). In reality, Lewis and Deterding (2018) discovered that incorrect stress placements did lead to misconceptions, raising doubts about their absence from EFL instruction. According to Jeong et al. (2020), lexical stress must be emphasized. Then, we will be asked whether this gray area has an essential effect on the perceptibility of L2 in English, which is worthy of consideration by the researchers.

Also, there is still a lot of debate in academia about how to determine word stress. In British English and American English, the stressed syllable is frequently encircled by gentler, faster, unstressed syllables with vowels that have been reduced to more central vowels. In general, unstressed vowels are changed most frequently to schwa /ə/ (schwa) and less regularly to /ɪ/ or /ʊ/ (Collins & Mees, 2013; Cruttenden, 2014; Roach, 2009). However, Roach (2009) notes that not every unstressed syllable contains a reduced vowel. Furthermore, lexical stress is not simply a syllable produced by combining these four features but is more prominent than the surrounding syllables. Ladefoged & Johnson (2014) claim that stressed syllables require more energy and "can always be defined in terms of something a speaker does in one part of an utterance relative to another." Generally, the stressed syllable is louder, longer, altered in pitch, or contains a full vowel (Brazil, 1994), but not equally. For instance, Fry (1955, 1958) used sound manipulation to conclude that loudness had little influence on stress perception and that duration also had an impact. Furthermore, a pitch shift was nearly always seen as a prominent syllable. Morton and Jassem (1965) also discovered that pitch is more critical to listeners than loudness or vowel length. Amplitude did not consistently show itself to be a powerful stress signal (Mol & Uhlenbeck, 1956; Nakatani & Aston, 1978; Cutler, 2015). Similarly, Roach (2009) argues that intensity and vowel quality are less significant than duration and pitch, along with Zhang & Francis (2010) asserting that vowel quality alone appears

to be how English as a native language (ENL) speakers and Mandarin speakers perceive stress.

1.2 Statement of Problem

Prior research on stress perception has mostly examined the impact of the prosodic system on a speaker's L1. These studies have shown a connection between the metrical system of the L1 and the ability to perceive stress qualities in a second language (L2) (Altmann & Vogel, 2002; Peperkamp & Dupoux, 2002). The former assertion has been validated by empirical research, which shows that vowel quality, pitch, length, and intensity are all phonetic indicators of English lexical stress. Additionally, pitch serves as a critical indicator of lexical tone. The acoustic cue hypothesis postulates that speakers of tone languages, such as Thai and Chinese, depend on the pitch to discern the lexical stress in English (Nguyen, 2003; Wang, 2008; Ou, 2010; 2016). In actuality, it is challenging to identify stressed syllables in words generated by non-stress languages, as the positioning of the stress can be movable or ambiguous. It is worth studying whether the same phonetic signal aids L1 lexical tone language speakers to perceive lexical stress, despite background differences between L1 and stress languages (English) (Ou, 2016). On the other hand, the tests mentioned above are the results of assessing perception via researchers' manipulation of L2-produced speech. Furthermore, the materials do not seem to have any direct connection to perception or the fact that they are not natural. To put it another way, the present body of literature study mainly concentrates on the impression of manipulative stress, but it does not take into account speech production. As a result, the purpose of this research is first to investigate the acoustic characteristics of English lexical stress that are created by Thai and Chinese speakers who are L2 English speakers.

On the other hand, there is a relative paucity of literature on the expanded circle of varieties of English used in countries such as Thailand and Mandarin China. Most studies have focused on countries where English is the first language (e.g., Seidlhofer, 2004;

Jenkins, 2000; Zielinski, 2015; Jeong et al., 2017). To illustrate this gap, this paper focuses on the relationship between distinguishability between English variants used in the same intonation background (Thai and Chinese).

1.3 Research Objectives

There are two research objectives of this study:

i. Examine the acoustic features of English lexical stress produced by Thai and Chinese L1 speakers.

ii. Explore the L1 language effect on the English lexical stress perception by Thai and Chinese L1 speakers.

1.4 Research Questions

There are two research questions of this study:

RQ 1. What are the differences in the acoustic features of English lexical stress produced by Thai and Chinese L1 speakers?

RQ 2. To what extent does the L1 influence the English lexical stress perception of Thai and Chinese speakers?

1.5 Significance of the Study

The significance of this study lies in the fact that, in contrast to previous research on stress perception, the studies included the mutual English perception of two L1 languages with tonal backgrounds. This aspect has been largely overlooked in this field of research. Hence, the production study may contribute to developing lexical stress production and perception and world English varieties studies.

The current production results contribute to understanding the articulatory characteristics of L2 English speech produced by Mandarin Chinese and Thai speakers. The production analyses data using the spectrogram, and the phonological description results can be quantified and made more scientific by measuring and calculating the visualized data. Given that previous L2 studies have focused on only one side of Thai or

Mandarin speakers, they have not been carefully put together for comparative studies. With this in mind, the acoustic information extraction findings may be utilized as an L2 English pronunciation teaching guide.

Most importantly, the study's participants engaged in perceptual experiments with the material that they listened to recordings both made by L2 speakers. Considering the outcomes of providing analytical feedback with low perception accuracy, teaching L2 students can benefit from the research findings on how different English variants affect the auditory aspects of lexical stress perception. After all, it is still unclear how linguistic experience influences how L2 lexical stress is perceived. At the same time, it increases the capacity by L2 auditory sessions to recognize lexical stress. It allows instructors to frame the objective of pronunciation and listening training to help students communicate more smoothly.

1.6 Structure of the Thesis

The organization of the dissertation is as follows: Chapter 2 presents a comprehensive summary of the relevant literature. This encompasses prominent models in the field of English L2 perception and production and offers relevant literatures on the production of stress in English L2 and the perception of stress. Chapter 3 provides a detailed account of the current methodology, including information on the participants, stimuli, procedures, and the process of analyzing the collected data. Chapter 4 includes findings from experiments on speech production and perception. Chapter 5 encompasses a comprehensive discussion, conclusion, study limitations, and ideas for further research.

CHAPTER 2: LITERATURE REVIEW

This section provides related theories, which are 1) the Stress Typology Model (STM), 2) the Speech Learning Model (SLM), and 3) the Perceptual Assimilation Model (PAM). Moreover, introduce the World Englishes (WE). Suprasegmental features of Thai English and Chinese English are also explained. Subsequently, previous studies concerning the English word stress of L2 learners are illustrated.

2.1 World Englishes

The widespread dissemination of English as a means of communication has ignited a captivating and contentious discussion on the position of English in its many forms, generally known as World Englishes (Kachru, 1985). Kachru (1997) proposed three circles to divide the English-speaking world. He focused on the historical context of English, its status, and the functions of the language in different regions. According to Kachru, the Inner Circle consists of nations with English-speaking populations, like England, the United States, and Canada. The Outer Circle comprises former colonies such as India, Africa, and Nigeria. In contrast, the Expanding Circle comprises Westernized nations like China, Japan, and Thailand, where English is becoming increasingly important in business, science, and education. Kachru also believed that recognizing a variety of norms would not lead to a lack of intelligibility among different English users. Widdowson (1994) supported Kachru by stating that many bilingual English users acquire the language in educational contexts that emphasize a particular standard and tend to ensure some unifying forms.

The sheer diversity of the WEs opens up more possibilities for English as a world language. Kachru (1985) proposed challenging traditional notions of models, as they are typically associated with Inner-Circle users. The global spread of English has taken an intriguing turn: the L1 speakers of this language appear to have lost their exclusive control over its standardization. The discussion of World Englishes has also raised questions

about teaching and teacher education. In the past two decades, English courses have been offered to train Cambodia, Laos, Myanmar, and Vietnam (CLMV) officers and diplomats, teacher-trainers, liaison officers, interpreters, managers, and administrations (ASEAN Secretariat, 2020c). ASEAN has also established the ASEAN University Network, which comprises over thirty regional universities (Kirkpatrick, 2007). Institutions of higher education are developing more programs with English-medium instruction (EMI) (Bolton & Botha, 2020) to facilitate the successful exchange of faculty, staff, and students across this network. These English interactions in ASEAN are genuine examples of English as ELF in which the interlocutors' primary objective is to communicate across distinct L1s while simultaneously experiencing various varieties of English (Seidlhofer, 2001). Deterding (2013) has studied some of the distinctions in Southeast Asia, using the interlocutors as judges of intelligibility. His research included speakers outside the ASEAN region, including Nigeria, China, and Japan. Nevertheless, in his research, there was no exploration of the phonetic similarities or differences involved. Cross-cultural communication requires exposure to numerous accents and types, which is why WE are referred to as 'English as a Lingua Franca'.

The Thai monarchy saw the strategic importance of the English language from an early stage. They realized its worth as a means to acquire Western knowledge and as a tool for diplomatic purposes. Originally, English proficiency was mainly restricted to the aristocracy and other individuals belonging to the governing elite. Nevertheless, proficiency in the language progressively spread throughout Thai culture, although with various levels of competency. American English (AE) gained increased prominence following World War II. Simultaneously, many Thai individuals acknowledged the significance of acquiring proficiency in the English language to get employment opportunities and establish enterprises that would appeal to the American troops (Snodin & Young, 2015). Currently, the National Education Council of Thailand acknowledges

English as "the predominant global language used for academic and professional reasons" (Noss, 1984). There has been an increasing proliferation of foreign schools that use English as the medium of teaching (Wongsothorn, 2000). However, within Thai society itself, English is still seen as important primarily for international rather than local reasons. Nevertheless, it is not authorized by the Thai government. There are, in reality, significantly more substantial distinctions in Thailand with respect to the status and properties of English; even here, however, the concept of institutionalization is only a broad-brush description that encompasses a multitude of difficulties. Naturally, many nations in Southeast Asia, not just Thailand, struggle in various ways with the difficulties presented by a language that has become widely used and, as a result, is difficult to ignore.

English does not have a privileged position among nations inside the expanding circle, such as China (Crystal, 2003). Due to the widespread use of English and the rapid growth of the English learning sector in recent decades, Jiang (2003) goes as far as to assert that English has become a “Chinese language”. China has the highest number of individuals who are learning and using English globally, with over 400 million people (Wei & Su, 2012). The research on the spread of English in China also made the claim that approximately 400 million individuals had completed junior high school coursework in English (American English) by the late 1990s. Since this information is already 26 years old, one can only guess how many additional millions of individuals have learned English since then and how many people in China at least "know" the language (Wang & Yuan 2013, p. 27). The astounding numbers surrounding the growth of English in China, especially the regular reports of the millions of individuals actively studying the language, are perhaps the most significant aspect of the tale of English in China. Despite this, the figures we do have on the proportion of "English-knowing" individuals in China are, at most, estimates, although some of the estimates have been supported by concrete data. Consequently, regarding English, the significant portion is not how many students study

or have studied the language but rather how many Chinese population really use and utilize the language in their everyday activities. Given these conditions and the fact that English is still considered a non-native language in China, there have been discussions on the ideological implications of English with respect to local languages and Chinese culture (Niu & Wolff, 2007; Pan & Seargeant, 2012). Some studies claim that *Standard English* (inner circle English that population thought) may be spoken with any accent, native or non-native (Crystal, 1999, pp. 10-11; Trudgill & Hannah, 1994; Widdowson, 1994). While there are arguments suggesting that Chinese English is a form of English within the WE paradigm, it is now considered a "performance variety" (Kachru & Nelson, 2006), mostly used for international communication (He & Li, 2009, p. 71). Chinese English, as a performance variation, is based on namely —"standard English." (He & Li, 2009). In other words, it functions as a common language, particularly in the realm of international commerce. These variations are primarily influenced by and rely on the norms of the Inner Circle, particularly British and American English, which are considered the standard varieties. As a result, they are commonly referred to as "norm-dependent varieties" (Kachru, 1985). It is generally agreed with Kirkpatrick (2007b, p.146) that we are not yet able "to claim any distinctive phonological features that are common to all speakers" of "China English," but we also think it is impossible to ignore the fact that the following phonological characteristics of "China English" (among others) are soon to be present: the replacement of the dental fricatives /θ/ and /ð/ with [s] and [d], the insertion of the final [ˊ], a general lack of voiced fricatives, some types of diphthong simplification, avoiding weak forms for function words, a tendency to pronounce multisyllabic words or word groups with syllable-timing, etc. Furthermore, as has been noted, there is a continuum from minimum to maximum acceptability when it comes to the norms of English pronunciation and intonation, i.e. they are are dynamic (Hung, 1992; 2004; Jiang, 2002, p. 11).

Given that English is taught as a subject in all schools and universities and has become widely used in textbooks and English-medium courses (AE) taught in many universities and high schools, it is evident that the field of education is crucial in this context. The only recognized pedagogical models for English language education (ELT) in China up to 2014 were the standard varieties of American English (AE) (Bolton 2003; Adamson 2004). Nevertheless, international English experts have contested this in recent years. This paradigm thus gives birth to the long-running argument about which variation of English should be chosen as the pedagogic model in Outer and Expanding Circle nations (Kachru, 1992; Bamgbos, 2001). According to Cook (1999, p. 185), the dominance of the native speaker in language instruction "has obscured the distinctive nature of the successful L2 user and created an unattainable goal for L2 learners." However, if we can "acknowledge that L2 users have strengths and rights of their own" instead of focusing primarily on the norms of native (inner circle) speakers, we will have a better chance of persuading EFL/ESL students that "they are successful multicompetent speakers, not failed native (inner circle) speakers" (Cook, 1999, p. 204).

Academics like Jenkins (2000) and Seidlhofer (2004) have made noteworthy contributions to this topic via their insightful books and papers, particularly Jenkins' "The Phonology of English as an International Language" (2000). Even more, Smith (2005, pp. 58–61) contends that expanding circle learners find English to be unpronounceable, irregular, overly complicated, and sometimes confusing. Given that Thai, Mandarin Chinese and English are topologically very different languages in terms of phonology, lexis, grammar, and discourse pragmatics (e.g., Chinese is a syllable-timed language whereas English is a stress-timed language) (Hung, 2002a; Kirkpatrick, 2006, pp. 73 – 4), the increasingly nativized and acculturated English in expanding circle countries inevitably shows different characteristics to some extent. According to ELF research (Deterding, 2013; Jenkins, 2000), innovative pronunciation is one of the primary causes

of confusion. Lexical stress is one of the most contested aspects of pronunciation. Similarly to this, this explains why the students from the expanding circle (Thai and Mainland Chinese) selected for this research fail to produce and perceive lexical stress in English. Jenkins (2000, 2015a) excludes word stress from features essential to ELF recognizability and intelligibility, whereas Cruttenden (2014) includes it in his suggested features for perceptibility pronunciation in international settings.

2.2 Stress Typology Model

According to word-level prosody, languages are often divided into stress and no stress (Hyman, 2009). According to Vogel (2000), stress positions may either be predictable (i.e., non-contrastive) or utilized to indicate lexical contrast (i.e., contrastive), and he proposed the Stress Typology Model (STM). Based on the characteristics of L1 stress, STM forecasts the success rate of bilingual individuals in perceiving stress patterns in their L2. STM predicts that speakers of L1 predictable stress languages will have difficulty discriminating between different types of stress, but speakers of non-stress languages will do an excellent job of it. STM classifies languages as predictable stress, non-predictable stress, and non-stress languages, which include tone and pitch accent languages. Utilizing the presence or absence of a property regarding L1 lexical stress, it is possible to estimate the L2 speakers' success rate. Speakers of L1 predictable stress languages are expected to have trouble differentiating between various forms of stress. Primary stress is not always predicted in languages that are categorized as having contrastive stress (such as English, Russian, and German). It can happen in several locations within a word, leading to minimal pairings that only vary in stress placement (Vogel, 2000). For instance, the only way to tell certain words apart is by where the emphasis falls. On the other hand, stress does not indicate changes in word meanings in languages with regular stress patterns, such as Arabic, Hungarian, and Finnish (Hayes, 1995; Hulst, 2002). Such languages may have positionally fixed stress, meaning they

always fall at or close to the prosodic word edges (Dupoux et al., 2001; Peperkamp & Dupoux, 2002; Kijak, 2009).

Speakers whose L1 lacks predictable stress patterns (-stress, -predictable) tend to excel in tasks involving stress identification. This is because the absence of stress patterns in their L1 eliminates any impact on their performance. Individuals who speak a first language (L1) that has both a positive stress feature (+stress) and a negative stress feature (-predictable) are expected to have superior performance compared to individuals who speak languages with two positive stress characteristics (+stress and +predictable), which have predictable stress patterns. The model implies that negative parameter settings do not influence the efficacy of L2 stress acquisition. Positive settings, in contrast, are apt to induce interference. Dupoux et al. (2001) conducted tests to assess Spanish and French listeners' memory capacity in accurately recalling the stress location in a series of sequence-recall tasks. The research included participants who were required to listen to sequences of two words with stress contrasts (PIki vs. piKI). The findings indicated that the French speakers had a notably worse ability to distinguish stress differences than the Spanish speakers. Based on these findings (Dupoux et al., 1997 & 2001), the researchers concluded that French listeners had a limited ability to perceive differences in stress, particularly at a higher level of processing known as the phonological level. Dupoux & Peperkamp (2002) argued that French listeners struggle to perceive stress since they do not include stress in their mental representation of words. This is because stress assignment in their native language is consistent and not distinctive.

Altmann (2006) confirmed the validity of this model's predictions. The study found that speakers whose L1 language does not have lexical-level stress, such as Chinese, Japanese, and Thai, had a high performance on an English test that included identifying non-lexical stress. Speakers who speak languages with unpredictable stress patterns, such as Spanish, had the second-highest scores. In contrast, those who are L1 speakers of

languages with predictable stress patterns, such as Arabic, had challenges in identifying the placement of stress in the same test. In contrast to Altmann's study, Lee, Shin, and Garcia (2019) investigated whether native Spanish speakers would outperform Korean speakers in perceiving L2 English stress. Given that Spanish has contrastive lexical prosody while Korean does not. Their research used a strange task that included disyllabic stress pairings that speakers of Southern British English reported. The researchers discovered that Korean individuals learning English had poorer accuracy than Spanish learners on the stress test. This suggests that Korean learners were less adept at recognizing stress positions in comparison to Spanish speakers.

Furthermore, according to STM, it strongly indicates the importance of considering the acoustic properties of stress. These studies argue that second language (L2) listeners with an L1 without stress patterns would rely on the relevant cues in their native language to perceive stress in the target language. This has been supported by studies conducted by Wang (2008) and Wang & Yoon (2008). Equally, since each language has its own special language system. Naturally, L2 lexical stress will behave differently. Thus, the acoustic parameters and performances of non-stress categories of English L2 learners from China and Thailand will be the focus of this study.

2.3 Model of Perception and Production

2.3.1 Speech Learning Model

The Speech Learning Model (SLM) created by Flege (1995) is a model that focuses on the learning of L2 speech. This model encompasses the variations in the acquisition and interpretation of speech sounds during the lifespan. This statement suggests that the cognitive abilities necessary for acquiring speech are still present in adults, even beyond the critical language acquisition time proposed in some models. Adults maintain the capacity to develop novel phonetic categories for speech sounds heard in L2. SLM is widely acknowledged to be the proposed model capable of predicting the success rate of

stress production (Flege, 1995, 2011). SLM is not expressly intended to account for speech perception, and it leverages the correctness and failure of L2 speech perception to explain the learning of L2 production. It provides a broader perspective on L2 speech acquisition, including perception and production, and covers issues such as age of arrival.

In Flege's (2005) work, various propositions were presented, which may be succinctly summarised as follows: 1) When L2 learners are provided with enough input, they are able to recognize the phonetic characteristics of L2 speech sounds appropriately. 2) Similar to the development of L1, the process of L2 involves two key factors: it requires a significant amount of time, and the quality of the information received dramatically impacts it. 3) Similar to the process of L1 development, production in language is influenced by perceptual representations that are kept in long-term memory. 4) The processes and mechanisms that govern the effective learning of L1 speech, including the capacity to develop new phonetic categories, remain functional and available throughout someone's whole life. 5) The phonetic components that comprise the L1 and L2 phonetic subsystems often reside together in a "shared phonological space," exerting a mutual influence on each other.

In sum, achieving a phonetic category that resembles that of an L1 speaker becomes more challenging when one has limited experience in an L2 learning environment. This difficulty arises from the fact that the phonetic systems of the two languages are not entirely distinct (Flege et al., 1992). Namely, producing non-native contrasts with phonetic elements that differ from their L1 language is comparatively more challenging for L2 learners. Archibald (1997) utilized 10 participants representing three language groups to produce 35 test items, and results showed that L1 speakers of non-stress languages did not show sensitivity when completing the production English tasks. He explained that speakers of non-stressful languages only consider stress a lexical phenomenon. Nevertheless, because little is known about how speakers of L1 non-stress

languages produce L2 stress, there is also a considerable limitation in this prediction conclusion, so more empirical research is needed to contribute to this model.

2.3.2 The Perceptual Assimilation Model

Best (1995, 2001) introduced the Perception Assimilation Model (PAM), a model for acquiring second language speech. This approach specifically uses articulatory phonology and asserts that listeners differentiate the speech signal by using information about articulatory gestures (Best, 1995; Fowler et al., 1990). Best (2001) asserts that the active articulator, which includes laryngeal gestures, the position of constriction (place of articulation), and the degree of constriction (manner of articulation), determines these gestures. The Perceptual Assimilation Model (PAM) suggests that the listeners' existing knowledge, whether it is unconscious or conscious, significantly influences how they perceive speech from nonnative speakers. Listeners tend to associate nonnative sounds with a native phoneme or category that shares similar articulatory gestures. Best's (1995 & 2001) studies have provided evidence in support of this conclusion. According to several recent models of nonnative speech processing, learners' initial L2 perceptual system may be equivalent to their L1 language perception (Best & Tyler, 2007; Escudero, 2005). Recent research has looked at the starting point at which learners approach the task of L2 perceptual development. In their study, Escudero & Chládková (2010) compared the perceptual assimilation patterns of Peruvian Spanish listeners for Southern British English (SSBE) and American English (AE) vowels. They discovered that the acoustic similarity between the English varieties and the Spanish vowels was a good predictor of the assimilation patterns. Gilichinskaya & Strange (2010) looked at the PAM framework (Best, 1995) to see if an acoustic comparison of American English (AE) and Russian vowels could predict how listeners would adapt AE vowels to Russian categories. The authors attempted to predict which AE vowel contrasts would be the most challenging to acquire based on the PAM.

The PAM primarily focuses on segments (consonants and vowels) and specifically examines minimal differences between segments. It suggests that nonnative phones are typically perceived based on their similarity to native segments or combinations of segments that are closely related to them in the native phonemic system. The proposed theory says that differences in how far apart two different L2 phones are thought to be from the closest L1 segment(s) could cause differences in how well people can tell the difference between L2 contrasts. The PAM-S, developed by So & Best (2010) and So & Best (2008), is an expanded version of the PAM, first proposed by Best (1995). PAM-S extends the scope of the original PAM by including the perception of nonnative suprasegmental features and its existing emphasis on phones at the segmental level. In the PAM-S, individuals who speak both tonal and non-tonal languages incorporate L2 prosodic categories into their native prosodic categories. These categories include tonal, pitch-accent, and intonational categories. The assimilation patterns are believed to determine how people can distinguish specific contrasts. This prediction is based on the work of So & Best (2010). According to the PAM-S, learners might perceive nonnative prosodic categories as either categorized or uncategorized prosodic categories in their native language systems (So & Best, 2014).

Taking into consideration the fact that not all unexpected sounds are complex for novice listeners to recognize was an advanced suggestion. When it comes to the perception of sounds from different languages, the PAM and PAM-S argued that the degree to which L1 sound contrasts may be recognized is contingent on the manner in which the components of each contrast are perceptually equated to L1 sound categories, respectively. For example, when two nonnative sounds are integrated into a single L1 category, discrimination is projected to be very low. On the other hand, it is predicted to be relatively high when two nonnative sounds are mixed into two distinct L1 categories that are both equally outstanding instances of it. Depending on the degree to which the

category-goodness gap varies, it is anticipated that the situation will experience a shift. According to a number of theories of non-native speech processing, the initial perceptual system of learners of an L2 is equivalent to the language perception of learners of their L1 (Best & Tyler, 2007; Escudero, 2005).

Turn your attention to the focus of this study, the perspective of lexical stress. The learning of Spanish lexical stress by speakers of French, a language without contrastive stress, has proven to be one of the most successful study fields. Dupoux et al. (2008) used several experimental techniques to show that monolingual French speakers were less proficient than native Spanish speakers in differentiating nonce words that differed solely in the stress location. This difficulty was found in cognitively demanding tasks that require more fabulous memory workloads due to the variability of the stimuli but not in more accessible parts such as AXE discrimination tasks, in which listeners must determine whether two tokens are identical or distinct (Dupoux et al., 2001). Additionally, Dupoux et al. (2008) discovered that regardless of L2 proficiency, all learner groups faced an extra challenge with a sequence recall task using cues that varied just in stress location. Dupoux et al., (2008) had concluded that the "stress deafness" was best explained by the incapacity of French listeners to construct distinct phonological representations.

However, their findings do not provide an in-depth discussion of the L2's perception of the acoustic cues. Schwab & Llisterri (2011) tested French listeners with varied levels of exposure to the target language to thoroughly examine L2 perceptions of the auditory signals that convey stress in Spanish. In order to explore the individual and combined impact of pitch, intensity, and duration on accurate identification of Spanish stress, the researchers altered proparoxytone, paroxytone, and oxytone words and pseudo-words. They asked participants to identify the stressed syllable. The authors discovered that experienced Spanish learners could identify stress positions more quickly and accurately than monolingual French listeners, proving that exposure to Spanish enhances sensitivity

to Spanish stress. The current study will also devote itself to this perspective, beginning with the acoustic information of the production experiment and comparing the perception rate in order to investigate the influence of L1 input on L2 perception in greater depth. The investigation of acoustic information has the potential to provide conclusions that are more accurate and convincing.

2.4 Prosody Phonology of Mandarin Chinese

China is a vast and multilingual nation, renowned for its linguistic diversity. In addition to China's lingua franca, Chinese linguists commonly refer to the various languages as regional dialects. Both Mandarin and its dialects extensively utilize a lexical tone system, with lexical tone serving a contrastive function (Fox, 2002; Peng et al., 2005).

2.4.1 Language Profile in China

2.4.1.1 Standard Chinese

There is no internal competition among languages in a single-ethnic nation with one language. Therefore, there is no need to deliberate on the selection of a lingua franca for communication, and the national language may naturally become the lingua franca of the country (Wang, 2014). However, selecting a national lingua franca poses several challenges in multi-ethnic and multilingual nations.

The Han nationality has always been dominant in China, with Chinese as the common language used by over 95% of the population (Wang, 2014, P. 187). Mandarin Chinese was established as the common language of New China in accordance with historical laws of language development (Chen, 1999; Zhou, 2001).

2.4.1.2 Lingua Franca and dialetcs

Chinese scholars widely agree that several regional dialects have consistently coexisted alongside a national lingua franca known as mínzú gòngtóngyǔ, or "nationality-common-language" (Li, 2001; Luke, 2005). This term refers to the linguistic variation used by the predominant group residing in the capital of China. Their authority is derived

from their political influence and economic dominance. According to Li (2001, p. 3), the conventional differentiation between dialects and the Chinese language is based on historical, sociopolitical, and ethnocultural variables rather than just linguistic similarities and variances. Chinese dialectologists use metaphors like "living fossils" to describe specific linguistic features still present in contemporary dialects that are considered essential parts of Archaic Chinese. Examples include syllable-final plosives indicating the entering tone in Cantonese and Min (Di, 2003). Another important evidence is the ability to identify lexico-grammatical correspondences or cognates across various dialects despite differences in current pronunciation.

Due to its linguistic complexity, mainland language is best described as "diglossia with increasing (dialect) bilingualism" (Li, 2006, p. 149; Chen, 1999) or "bidialectism" (Norman, 1988; Chen, 1999). According to Chen (1999, p. 53), "diglossic differentiation" exists between pǔtōnghuà and regional dialects.Pǔtōnghuà is a wide variety of languages utilized in public affairs, education, media, written and broadcast communication, public areas, and interlanguage dialogue. Local dialects are utilized in families and other dialect-speaking groups for daily and interpersonal communication. They are classified as a Low variation.

Chinese linguists generally agree that there are seven primary dialect groupings that Chinese varieties come under, each having its own subdialects (Chao, 1968; Zhou, 2003).

• Mandarin (北方方言, beǐfāng fāngyán; spoken in northern, northwestern),

• Wu (吴语, Wúyǔ, spoken mainly in Shanghai and the provinces Jiangsu and Zhejiang),

• Min (闽语, Mínyǔ, spoken mainly in Fujian province),

• Yue (粤语, Yuèyǔ, spoken mainly in the provinces Guangdong and Guangxi),

• Xiang (湘语, Xiāngyǔ, spoken in Hunan province),

• Gan (贛语, Gànyǔ, spoken in Jiangxi province), and

• Kejia [Hakka] (客家话, Kèjiāhuà, mainly found in small enclaves in different provinces in southern China, notably Guangdong, Guangxi, Fujian and Sichuan)

The distribution of these dialect groupings is quite unequal. Mandarin, commonly referred to as the Northern dialect, is the most geographically widespread variant of the Chinese language. The Mandarin-speaking population constitutes 67% of the overall Chinese population. 75%, communicating with each other without much difficulty. Four primary subcategories of Mandarin have been classified: Northern Mandarin, Northwestern Mandarin, Southwestern Mandarin, and Jiang-Huai (Eastern) Mandarin.

Research on bidialectism has examined language use patterns of pǔtōnghuà and local dialects in three regions (Wú, Mín, Yuè dialects). Studies have shown similar trends, except in Cantonese-speaking areas (Chen, 1999). In a study by Wu & Yin (1984), it was found that 91% of the public understood pǔtōnghuà, compared to just 41% in the early 1950s. Additionally, 50% of people were able to speak pǔtōnghuà. Of those who could speak it, 54% came from Mandarin-speaking regions and 40% from other dialect regions. According to a national poll in 2004, only 53% of the population was proficient in pǔtōnghuà and at least 40% were unable to converse with it (China Daily, 2006). Language interaction and the development of interlanguage are inevitable in bilingual speech studies or diglossic language circumstances. Pǔtōnghuà is rarely spoken in its standard form outside of Beijing, leading to a large degree of language similarity among Chinese mainland populations (Norman,1988, p. 213). Local pǔtōnghuà, also known as "interlanguage," is influenced by local dialects and referred to as "accented Chinese" in phonetic research (Chen, 1999, p. 42). One prevalent characteristic of non-native speakers of the Beijing dialect is the deviation of pitch contours of the four lexical tones (Chen, 1999; Li & Wang, 2003; Li et al., 2006).

2.4.2 Lexical tones

Mandarin Chinese is sometimes assumed to be a nonstress language, in contrast to English, and the critical differentiation between syllables is based on the tone type rather than word level stress (Sluijter & Heuven, 1996). Additionally, there are four tones in Mandarin: tone 1 is high level, tone 2 is mid-rising, tone 3 is low dipping, and tone 4 is high falling. Numerous scholars have looked into the fact that native Mandarin speakers used F0 as their primary cue, particularly the height, and contour of F0 (Howie, 1976), while syllable duration and amplitude contour consistently varied across the lexical tones (Fu et al., 1998; Liu & Samuel, 2004) (see in Fig 2.1).

Gandour (1981), in his study of Mandarin Chinese tones' perception, found that linguistic background impacts the relative emphasis placed on auditory dimensions and the perceptual signal's function interactively to allow for proper tonal identification. English speakers are more concerned with pitch height (e.g., average pitch, extreme endpoint), whereas tonal language listeners (e.g., Chinese, Thai) are more concerned with pitch contour. Tones in Mandarin Chinese are perceived differently based on "height," "direction," and "contour." To separate contour tones (Tones 2, 3, and 4) and the level tone (Tone 1), as well as to differentiate the rising tone (Tone 2), the falling tone (Tone 4), and the falling-rising tone (Tone 3), among these dimensions, the direction of the F0 change is essential. To distinguish between the high tone (Tone 1 and Tone 4), the mid-tone (Tone 2), and the low tone (Tone 3), one uses the pitch height. According to Howie (1976, p. 4), characteristic tones in Mandarin are patterns of voice pitch that align with syllables. However, there is no agreement on the phonetic location of tones in Mandarin for monosyllables. According to Wang (1967) and Chao (1968), the range of tone covers the same area as the vocal part of the syllable, which includes the first consonant when it is voiced. Some other studies argued that the first voiced consonant and nasal ending should not be considered as tone-bearing units (Lin, 1995). Dow (1972, p. 102) asserted

that the tonal patterns exhibit distinctiveness over the vocalic segment of the syllable. Lin (1995) contended that the tonal domain encompasses just the nucleus while eliminating the beginning voiced consonant, the medial, the nasal coda, and the vocalic ending.

Simply put, listeners with different first languages perceive tones using various acoustic signals. A given listener may use several pitch dimensions to recognize tonal differences (depending on the tones in that tonal system). Chang (2010) conducted empirical research on the variations in the expression of the four lexical tones in Mandarin Chinese across different dialects. The study focused on two specific phonetic settings. Analysed were three acoustic properties of tones: f0, rms amplitude, and duration. The impacts of dynamic f0 and rms amplitude movement were investigated using mixed design ANOVAs and independent-samples t-tests. The production study investigated the fundamental frequency (f0), amplitude, and duration, which are the three primary acoustic properties, of the four lexical tones in Beijing Mandarin (BM) and Taiwan Mandarin (TM). The study focused on the isolated and sentence-medial production of these tones by native male and female speakers of these dialects. The findings revealed that, while there were some differences in the form of individual normalized f0 curves within a single tone category, the f0 contours derived from 13 words with different syllable types exhibited virtually identical patterns. Thus, they may be compressed into a unified f0 contour.

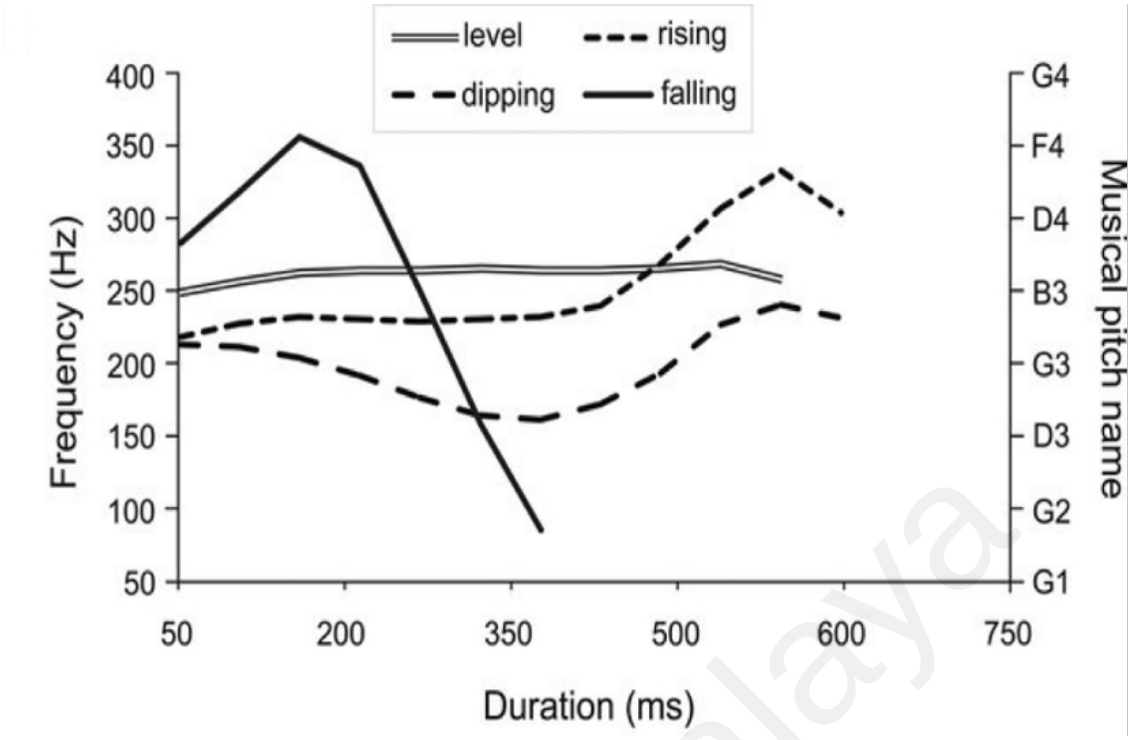

Figure 2.1: Fundamental frequency contours of the four Mandarin Tones

Every pitch contour on the syllable /ma/ corresponds to a distinct lexical unit. The tone "ma1" represents a level tone, "ma2" represents a rising tone, "ma3" represents a dipping tone, and "ma4" represents a falling tone (Peretz, 2002).

2.4.3 Lexical stress and Neutral tone

It is generally recognized that Mandarin Chinese speakers lack awareness about where stress is placed in the language. Mandarin Chinese is a tonal language where tone serves both lexical and phonemic information. The meaning of a word can be altered by applying a new tonal contour to the same segmental sequence in Mandarin Chinese, which has four tones. As a result, the only tonal differences between two words are those that are utterly distinct. For example, CGVX, where C is a consonant, G is a glide, V is a vowel, X is a nasal or an offglide of a diphthong, and VX is the rime, can be used to create a monosyllabic Chinese word (Duanmu, 2000). Monosyllabic words are standalone units. The vowel, therefore, carries tone information most of the time. Syllables with the four typical tones are all stressed from a phonological perspective, as opposed to those with

neutral tones, which, according to phonological norms, lose their tonal identities, such as Tone Sandhi (Duanmu, 2000; Peng et al., 2005). Nonetheless, several studies (Chen et al., 2001; Lai, 2008) argue that Mandarin is a syllable-timed language with no lexical stress.

Chao (1979) believes that Mandarin Chinese is a language with lexical stress. Chinese is thought to have three distinct forms of phonological stress: (1) extra-stressed, (2) normal-stressed, and (3) weak-stressed. Most words in common language are "normally stressed" words, meaning that each syllable has normal stress. On the other hand, Chao (1979) wrote that whether a group of syllables is a word or a phrase, the final syllable always has primary stress, the first syllable always has secondary stress, and the other syllables are all phonetically weak. In other words, even if all of the syllables in a polysyllabic word have normal stress from a phonological perspective, their degrees of syllable stress are not equal. Numerous Mandarin Chinese phonetic investigations back up this assertion (Yan & Lin, 1988; Wang et al., 2003; Yu, 2020). For instance, Yan & Lin (1988) found that the middle syllable is the shortest, the first syllable is the second longest, and the last syllable is the longest in trisyllabic Mandarin Chinese words. While there is some evidence in favor of final stress, it appears too soon to make a definitive judgment based on those phonetic investigations (Lin et al., 1984; Yan and Lin, 1988) since they did not thoroughly regulate the test stimuli to isolate the variables of location and length. Additionally, the lengths of the various tones in Mandarin Chinese vary. The pitch of each of the four Chinese tones is different, with different heights and slopes. Therefore, it is still unknown how much tone, length, or the syllable's placement in the word impact Chinese listeners' preference for last-syllable stress. The experiment also depends heavily on the control of the sample variables.

In Mandarin Chinese, for example, there is a toneless category called neutral tone, which only appears on weak (short) syllables in addition to the four lexical tones that occur on stressed/full syllables. As a result, neutral tone syllables share the acoustic

property of being short syllables (Duanmu, 2014). While the weak syllables cannot carry a full lexical tone but can only carry a neutral tone, the full syllables can carry a full lexical tone and be stressed. Moreover, unstressed syllables are those with a neutral tone. The neutral tone and full tones have different phonetic properties. Chen & Xu (2006) argue that it is shorter in duration than the syllable.

Example (a) displays several F0 realizations of the same neutral-toned morpheme -me in stressed and unstressed words, where F0 of lexical tones is quantitatively specified and ranges from 1 (low) to 5 (high).

(a) (i) duōme duo55 'how'

(ii) shénme shen35 'what'

(iii) zěnme zen 214 'how/what'

(iv) zhème zhe 51 'so/such'

In Example (b), the unstressed (neutral tone) syllable is part of a multisyllabic word that has been fully lexicalized. *lìhai* (to a deep degree, is an adjective or adverb) is a monomorphic word that cannot be further broken down; it has a neutral tone in the second syllable and has a stressed-unstressed pattern throughout. On the other hand, the orthographically identical word produces a stressed-stressed pattern when it comes with two full lexical tones, as in (ii) *lìhài* (of a good or a bad thing; a noun). It is a compound word that can be deconstructed (Duanmu, 2007). The neutral tone syllable in (i) *lìhai* is shorter than the full-toned syllable, and its F0 fluctuates according to the lexical tone that comes before it.

(b) character	Pinyin	meaning
厉害	(i) lìhai	to a deep degree, is an adjective or adverb
	(ii) lìhài	of a good or a bad thing, is a noun

A neutral-toned syllable containing a single mora cannot act individually and must be followed by a full-toned syllable. A trochaic (strong-weak) stress pattern can explain a word like *lìhai*, consisting of a full-toned syllable followed by a neutral-toned syllable. This prosodic analysis matches the fact that morphemes with a neutral-toned syllable and content words never have just neutral-toned syllables. According to this theory, Mandarin has lexical stress and can form a contrastive stress pair with English (e.g., SUBject and subJECT). Nevertheless, stressed-unstressed (lìhai) and stressed-stressed (lìhài) disyllabic words make up the Mandarin stress minimum pair.

2.5 Prosody Phonology of Thai)

2.5.1 Segmental System

There are nine monophthongs and three diphthongs in standard Thai. In Thai, vowel length is phonemic. Nine short vowels and nine long vowels may be formed from the nine monophthongs, totaling eighteen vowel sounds. There are twenty-one vowel phonemes in total in Thai. Relative duration in Thai is the physical equivalent of the differential in phonemic vowel length.

Meanwhile, according to Tuaycharoen (1990), the 21 consonant phonemes in Thai are as follows:

Table 2.1: Thai Consonantal Sounds

	Bilabial	Labio-dental	Alveolar	Lamio-prepalatal	Palatal	Velar	Glottal
Plosive	p b p^h		t d t^h			k k^h	ʔ
Nasal	m		n			ŋ	
Fricative		f	s				h
Affricate				tɕ $tɕ^h$			
Tap			ɾ				
Lateral			l				
Semivowel	(w)				j	w	

2.5.2 Suprasegmental System

2.5.2.1 Lexical tones

Five lexical tones are phonetically distinct in Thai: mid (- or unmarked), low (ˋ), falling (ˆ), high (ˊ), and rising (ˇ). These five tones can be distinguished by their voice quality, pitch height, and pitch contour (see Table 2.2). (Comrie, 1987; Wayland, 1997; Hirst, 1998). All syllables can carry one of the five tones, although not all syllables can. Syllabic structure influences how these lexical tones are distributed. (Wayland et al., 2006; Moren & Zsiga, 2006). All five tones can occur on syllables ending in a long vowel or a sonorant segment (m, n, ŋ, w, j). However, only the low and falling tones are permitted in syllables with a long vowel followed by a non-sonorant segment (p, t, k), and just the low and high tones are permitted in syllables with a short vowel and a non-sonorant segment. The lexical tones of English words borrowed from Thai (Gandour, 1979) are also limited in the same way. Futhermore, prior studies (Gandour, 1983) have shown that three distinct acoustic characteristics, which are average pitch level, pitch direction, and pitch slope, act as perception signals for distinguishing the five tones in Thai. In addition, Abramson (1978) discovered that although a distinct pitch level is enough to recognize static tones, a quick pitch movement is required to identify dynamic tones.

Compared to Mandarin Chinese tones, the essential perceptual signals of Mandarin tones are F0 height and F0 contour (as mentioned in section 2.4.2). Perception studies have shown that Mandarin listeners tend to prioritize the contour dimension above the height dimension (Gandour, 1984; Massaro et al., 1985). Further investigation into the F0 contour has shown that the F0 turning point, which is the specific point when the direction of the F0 contour transitions from descending to ascending, serves as the perception cue that distinguishes between Tone 2 and Tone 3 (Moore & Jongman, 1997; Shen & Lin, 1991). Furthermore, research has shown that altering vowel duration may impact how Tone 2 and Tone 3 are perceived (Garding et al., 1986). Increasing the

duration of vowels consistently led to a greater number of Tone 3 answers. In a cross-linguistic study on tone perception, Gandour (1983) demonstrated that speakers of tone languages, such as Mandarin Chinese and Thai, frequently prioritize the direction dimension of F0 over its height dimension.

Table 2.2: Characteristics of Thai tones

Tone	Tone mark	Pitch contour	Pitch height	Voice quality
Mid	unmarked	Level	Medium	non-glottalized
Low	`	Level	Low	non-glottalized
Falling	ˆ	contour	High to Low	glottalized
High	´	Level	High	glottalized
Rising	ˇ	contour	Low to high	non-glottalized

2.5.2.2 Syllables and Stress

Another suprasegmental factor is stress. Thai stress works phonetically (Luksaneeyanawin, 1998) or primarily for contrast and emphasis, which differs from English (Saengsuriya, 1989). Thai is sometimes said to be monosyllabic (Hemakom et al., 2021). While it is true that many Thai words only have one syllable, there are also a large number of polysyllabic terms. These words with more than one syllable can be loanwords, combinations of Thai words with one syllable, combinations of native Thai words, or even combinations of two or more loanwords. The vast majority of loanwords in Thai are of Indian origin. These terms are roughly as common in Thai as words with Latin or Greek roots are in English (Gedney, 1947), and others are borrowed from other languages, such as Chinese, English and Pali words. Despite being adjusted in various ways to fit Thai speech and writing, these terms have grown to play a significant role in the Thai language.

According to some Thai stress studies, the syllable in the word-final position bears intense stress (Luksaneeyanawin, 1998; Saengsuriya, 1989; Peyasantiwong, 1986). Hass (1964), Hiranburana (1971), and Gandour (1976) are other linguists who have also explored the function of stress in Thai. Even though they all have different approaches and points of view, they all seem to agree that the syllable at the end of the word is the most important or has the most stress. Surintramont (1973), who contends that a tone-neutralization rule must apply from left to right, also raised specific issues that have not yet received a suitable response. However, Gandour (1979) contends that this explanation is flawed and is based on an inadequate collection of facts. He provides examples of the opposite, including "poet,"/cintàʔkàwii/- /cintàʔkawii/ or /cintaʔkawii/, also /wannáʔkháʔdii/- /wannáʔkhaʔdii/or /wannaʔkhaʔdii/, which means 'literature.'

According to some Thai stress studies, the syllable in the word-final position is the most noticeable or bears the strong stress (Luksaneeyanawin, 1998; Saengsuriya, 1989; Peyasantiwong, 1986). Further investigation showed that Thai is a language with a fixed accent (Luksaneeyanawin, 1998). In the stress system of monosyllabic Thai words, the stress is placed on content terms, while grammatical words remain unstressed. The major emphasis is always placed on the word's last syllable in polysyllabic words, while the secondary accent is placed according to a set of rules.

Luksaneeyanawin (1998) says that "accent" is "the potential for the syllable or syllables in a word to be stressed either when the word is spoken on its own or when it is used with other words in an expression" (p. 10). The final syllable of trisyllabic and tetrasyllabic words receives the major stress. The syllable farthest from the primary stress will get secondary stress if the two remaining syllables are of the same kind. The secondary emphasis will be placed on the remaining syllable if it is a non-linker syllable (Luksaneeyanawin, 1998). Furthermore, Cutler (1986) proposed that stress is not employed for lexical purposes but for word boundary detection if it can occur in only one

location for a phonetic segment. Stress has an impact on Thai vowel length realization. Unlike a long unaccented syllable, which is pronounced as a relatively short unstressed syllable, a long-emphasized syllable is always spoken as a noticeably long stressed syllable (Luksaneeyanawin, 1998). The absence of stress does not impact the quality of vowels in Thai, since there is no counterpart to the mid-central unstressed vowel, often known as schwa, in English (Panlay, 1997). Peyasantiwong (1986) also provided several phonological principles to distinguish between weakly stressed and typically stressed syllables.

Similarly, Wayland et al. (2006) and Gandour (1983) have acknowledged that native Thai speakers will be attentive to vowel length in their English production and perception of lexical stress. In contrast, Thai does not permit consonant clusters in syllable ending places. So, native Thai speakers may become less sensitive to this aspect of syllable structure and what it means for stress patterns. However, the participants in the previous study were relatively inexperienced English learners. Also, in this study, the Thai students who were good at English controlled the participants who were chosen to control the experimental variables. This means that different control groups may lead to different experimental results.

2.6 Previous studies on English lexical stress by L2 learners

2.6.1 The difficulty of Learning Lexical Stress by L2 speakers

Since English is a lexical stress language, the syllables in each word with more than one syllable will have different levels of importance. Some syllables may act as the focal point for emphasis, giving stress, others are never stressed (Culter, 2015; Tremblay & Owens, 2010; Ou, 2010). In other words, English tries to avoid having two stressed syllables in a row, and its speech is made up of alternating stressed and unstressed syllables. This propensity for stress alternation has an apparent implication. English words can not only have one stressed syllable but have three, four, or even more. This

means that there are different levels of stress. Due to stress's function in derivational morphology, English does change stress across words in several ways (Ou, 2010). We can either adMIRE a BAron as a PERson who is aristoCRATic or express our admiRAtion for his baRONial skill to perSONify the arisTOCracy by adding a derivational prefix to an English word, which results in the creation of a morphologically similar word of a different grammatical class.

Vowels in English can be full or reduced. Full vowels can be diphthongs (e.g., oil) or monophthongs (e.g., ill), but they all have the same property. Reduced vowels are centered, with the most prevalent one being the schwa. In English, every stressed syllable must have a full vowel (e.g., the first vowel in the word). Stress may not be placed on any syllable with a reduced vowel, such as the second syllable of a word. Simplified, the phonology of English varies from that of lexical stress languages without reduced vowels, where the only way to indicate stress is through suprasegmental differences (Ladefoged, 2006). This segmental property is essential to how stress works in phonology and how people understand words and sentences. Since it is involved in speech perception, a slip of the tongue or a non-native speaker's bad pronunciation that changes the patterning of full and reduced vowels makes it very hard to figure out what word was meant.

Learners of English as a second language have the most challenging time with the first and second lexical stress locations. Chomsky & Halle (1968) say that the rules for giving primary emphasis to words depend on the type of word (compounds, prefixes and suffixes words, or simple words) and the weak or strong syllables. The rules for determining the location of major emphasis on basic words are as follows: Disyllabic words such as "MASter" and "CREAtor" have stress on the second-to-last syllable. Trisyllabic or longer words like "exAMple" and "ratiO" have stress on either the second-to-last or third-to-last syllable (Fudge, 1948; Karjo, 2016).

Since the position of stress is flexible in English, a word may be stressed to the left or right of where it usually is while still first seeming valid. Any English word stress error thus entails two crucial factors: the degree to which the given word stress error generates simultaneous changes in vowel quality (i.e., full vs. reduced stress exchanges and the direction of stress shift). English word stress changed leftward is less harmful than stress shifted rightward, according to researchers like Cutler & Clifton (1984), who studied L1 English listeners, and Field (2005), who studied L1 and a variety of L2 English listeners. "Increasing the number of vowel errors may DEcrease comprehensibility" is an example of how Field says this may be partially explained by how contrastive stress in English permits leftward emphasis shifts.

At the same time, the connection between segmental and suprasegmental seems to have sparked much debate among academics. According to Zielinski (2015), it is not easy to distinguish between segments and suprasegmentals entirely. Segmental and lexical stress may not necessarily be unconnected if vowel quality affects perceived stress. Deterding's (2006) example of a Mandarin speaker pronouncing the word simply with an added vowel at the end, [dʒʌstə], is used by Zielinski (2015) to illustrate how the addition of a vowel (expected suprasegmental) resulted in an extra syllable being misunderstood and a different stress pattern. Suprasegmental inventions may also be segmental versions.

2.6.2 Production studies of English L2 lexical stress

There is evidence that L2 English stress production is challenging for Mandarin speakers. The present findings indicate that the observed problem may be attributed to the interference caused by the Mandarin tonal system. Juffs (1990) noted Chinese speakers who were enrolled in college but had little to no exposure to spoken English outside of the classroom. Many of these utterances had different stress placements than inner circle and outer circle utterances, indicating that they simply did not know which syllables needed stress in the utterances they were supposed to make.

Nevertheless, they showed signs of difficulties with the manipulation of certain correlates of stress, even when stress was applied to the proper syllable. For instance, some speakers tend to adopt a falling tone to indicate an English-stressed syllable. It is possible that these speakers were overextending the English tendency to use a sharply falling F0 contour for strongly emphatic stress rather than being aware of the general association between English stress and a higher average F0, as evidenced by the use of a falling tone with its overall lower F0 for a stressed syllable (Chao, 1972). This might indicate that these speakers simply covered the stressed English syllable with the Mandarin high tone's characteristics.

Similarly, data from George's (2019) duration measurement experiments also indicate that there may be an L1 impact on the greater degree of ultimate duration; however, there may also be another possibility, which is that they just paused more often. In addition, their research reveals that the pitch contours of unstressed syllables do seem to have a goal comparable to neutral tone syllables. As a result, some ambiguous data still supports the native Mandarin stress system's role. However, the above-mentioned are manipulation experiments, not materials produced under natural conditions. For example, George's (2019) CVC syllables in the English wordlist for Experiment 3 may have artificially shortened the vowels, making comparing duration across languages problematic.

Previous research on English word stress in Thai contexts focused on the stress production of English words with different syllables by L1 Thai university learners (Khamkhien, 2010), while Aungcharoen (2006) investigated the stress produced by L1 Thai secondary learners. It was shown that L1 Thai learners had difficulties assigning English word stress to disyllabic and polysyllabic words. The same study that found Thai speakers to have problems with the distribution of stressed syllables in English words was conducted by Jangjamras (2011). Thai speakers have difficulty in producing stress on CVVO and CVO syllables due to the influence of Thai tone distribution laws. Assigning

stress to the syllable CVVO in Thai proved to be the most difficult due to its limited ability to handle low tone and falling tone. The syllable CVO was the second most complex since it can only accommodate low tone and high tone in Thai. Then, Thai speakers create final stress more precisely than initial stress as Thai is claimed to have a fixed-final stress pattern.

It is crucial to keep in mind that there is an ongoing debate on the function of an acoustic parameter that has phonemic significance in a person's L1 and how it relates to the production and perception of L2 stress. Jangjamras (2011) discovered that vowel duration was the main acoustic factor associated with the development of final stress in non-tonal languages. This discovery aligns with the research conducted by Nguyễn et al. (2008) and implies that acoustic characteristics that have phonemic significance in one's L1 might serve as an acoustic indicator in the production of stress in L2. In addition, the use of F0, which is employed in a contrasting way in Thai lexical tones was identified as one of the most prominent features in the development of initial stress in Thai L2, along with F0 range and intensity. Nevertheless, despite the efforts made to closely resemble genuine words, the English non-word wordlist used in the experiment nevertheless had a noticeable artificial quality. The generated material seems to lack naturalness.

2.6.3 Perception studies of English L2 lexical stress

Numerous scholars have pointed out that stress may be measured in terms of its F0, duration, and intensity (Fry, 1954 & 1958; Ladefoged & Johnson, 2014). The four psychological dimensions that a listener picks up on are duration, loudness, pitch, and the quality of the vowels. In contrast, non-centered vowels, glottal stops, and aspirated plosives are secondary signals of a stressed syllable that have more to do with segmental features. (Kuhlen, 1986). Since perception depends on these acoustic qualities, investigations have been conducted to determine if acoustic cues have a substantial impact on the difficulties that listeners may have in perceiving stress in an L2.

Various research studies investigating the impact of acoustic signals have identified two factors that could influence the perception of lexical stress. The initial aspect to consider is whether the relative importance assigned to auditory cues for stress perception is similar to that employed for L1 prosody. Different ways of weighing auditory signals and prosody in L1 and L2 can cause problems with perception (Braun & Mani, 2011; Chrabaszcz et al., 2014). Alternatively, the comprehension of suprasegmental information is influenced by a person's linguistic background, which determines their preferred acoustic signals for understanding (Jongman & Moore, 2000; Yu & Andruski, 2010). This is in addition to the debate about which syllable in a Mandarin Chinese word should be stressed first (the first or the last). For example, Russian and English have similar stress languages from a typological point of view, but Russian speakers have trouble understanding English stress. This was because Russian and English stress were pronounced differently (Chrabaszcz et al., 2014). On the other hand, the second element is connected to L1's lexical usage of one or more phonetic stress correlates (Peperkamp et al., 2010). When L2 signals serve as indicators of linguistic differences in their respective L1s, listeners exhibit heightened sensitivity towards them (Wang, 2008; Ou, 2010; Lengeris, 2009; Schertz, 2020). For example, the work of Kijak (2009) demonstrates that the Polish language includes phonetic signals to identify stress that are distinct from those used to designate L1 tones, and the Chinese listeners were indifferent to these differences. According to Yu & Andruski (2010), it was found that Chinese listeners primarily relied on pitch to determine the location of stress in bisyllabic English words, whereas English listeners utilized pitch, duration, and intensity. These findings align with Mandarin Chinese tone analyses, demonstrating that listeners from various linguistic cultures process suprasegmental information using various acoustic cues. Kijak asserted that the speakers from tone languages, such as Chinese speakers, are only able to perceive stress in L2 when the acoustic properties of L2 are equivalent to those of tone

in L1, especially F0. Fry (1954, 1958) was one of the first to explain how duration and intensity can be used to judge stress. Despite focusing on lexical stress in limited noun/verb minimal pairs in English, Fry made voice stimuli with different physical dimensions that could change on their own and in a planned way. For instance, F0 changed in 16 steps, whereas duration and intensity altered separately within 5 steps. When these two things were explored, listeners' ratings showed that the durational manipulation had a more considerable effect than the intensity manipulation. Yet other researchers have claimed that F0 serves as an indicator of accent rather than stress recognition (Van, 2002; Chrabaszcz, 2014; Zahner, 2019).

Wang's (2008) study showed that length and intensity manipulations only predict a small number of Chinese participants' stress judgments, but F0 makes it easier to predict each Chinese participant's stress judgment. Accordingly, the individual logistic analysis indicates that how Chinese people feel the stress depends only on how F0 changes, not its length or intensity. When it came to the F0 cue, Chinese people relied on it much more than English L1 speakers. Comparing the two groups at each stage of the F0 modification provides more support. When the second syllable was higher in F0, fewer Chinese people than English L1 speakers thought the first syllable was more critical. Conversely, when the first syllable in F0 is higher, Chinese speakers perceive the stress on the first syllable as substantially more excellent than theirs. In a similar way, this study's results show that F0 is the most crucial predictor of stress among English L1 speakers. In their group, F0 exhibits the highest dependency score in comparison to duration and intensity. Little difference in the results about the acoustic signals of F0 may be seen in the study by Lai (2004), because of their perceptual weight in Mandarin Chinese as L1, novice learners of Mandarin as L2 may not exhibit sensitivity toward variations in the maximum F0 range. In contrast, they place greater emphasis on duration. Moreover, once proficient learners acquire awareness of F0 signals as a significant perceptual cue in the English language,

they are able to utilize them extensively. This is shown through the distinct perception observed in relation to different maximum F0 ratios. The role of F0 in Mandarin acquisition greatly influences the proficiency of L2 learners. Novice students consistently employ the same mean and maximum F0 values across verb and noun readings in their output. He claims that since F0 is a phonemic characteristic in the learners' L1, they are compelled to apply it consistently. L1 language and age were employed as between-subjects factors in the statistical analysis, while stressed and unstressed were used as within-subjects factors. Additionally, the results suggested that the English speech of Mandarin speakers was caused by their L1 language exposure to Mandarin lexical tones.

Furthermore, Jangjamras (2011) examined the potential impact of Thai learners' tonal backgrounds on their ability to perceive English stress patterns at the lexical level. Jangjamras (2011) revealed that native Thai individuals generally recognized the accurate placement of lexical stress in disyllabic non-words is on par with that of Americans. This outcome was anticipated for Americans, given that English tends to stress the first consonant, especially in disyllables (Cutler, 1986). However, native Thai speakers also distinguished between initial and final stress more effectively than American participants (Jangjamras, 2011, pp. 160-170). Considering the consistent presence of stable final stress in multisyllabic words in Thai, this discovery does not provide evidence to support the notion that Thai speakers would perceive initial stress as less preferable than final stress (Luksaneeyanawin, 1988). The fact that Thai participants in this study's final stress judgment needed greater reaction time further supported their initial stress preference, which was about 200 ms longer than the initial stress judgment. Thai likely displays the phrase "final stress" rather than "final word stress." The fact that Thais have trouble with final stress shows that L1 stress patterns at the phrase level do not always make it easier to understand L2 stress at the word level. More investigation is required to determine the exact causes, which may or may not make NT more susceptible to final stress. On the

other hand, research by Wayland et al. (2006) shows that independent of syllabic structure or lexical class, Thais favor final stress assignment over initial stress. The researcher showed the participants the target word, along with four different non-words embedded in the carrier sentences "I'd like to", "I'd like a." (The final "to" and "a" are produced in the form of schwa [ə]). However, the methodologies employed in Wayland et al.'s research and the present study exhibit dissimilarities, which could have led to a different pattern of stress preference. Then, the participants were asked to choose which statement they liked best without being asked directly whether they heard the first or last stress. Based on the syllable structure and word class of the target non-existent words, a prediction was made regarding the preference for stress placement at either the beginning or end. Although, it is also possible that the Thai participants instinctively associated the default stress position of the Thai phrase due to the target stimulus being positioned at the end of the carrier sentence, that is, the final stress of the phrase. The default position for the stress of the Thai phrase, that is, the final stress of the phrase.

For F0, analyzing the acoustic parts of English stress showed that F0, which is used in a different way in Thai lexical tones, was among the most distinctive parts of Thai L2 initial stress generation, along with F0 range and intensity. Also, since average F0 has a bigger effect size than intensity, both groups thought that average F0 has the biggest effect on how people perceive stress (Jangjamras, 2011). Fry (1958) examined how different types of variations in fundamental frequency (F0) can be used to assess stress levels, addressed a different agreement. His stimulus words used three different F0 patterns while varying duration: level (-), a linear change (\,), and a curvilinear change (-\). His research revealed that the assessment of stress position was not significantly affected by the combinations of F0 patterns on stressed and unstressed syllables. Listeners are able to properly detect stress as long as a frequency shift has been recognized.

The literature review mentioned above revealed conflicting findings on the influence of L1 on L2 stress perception. However, the majority of prior research on the perception of lexical stress in English has been conducted by individuals whose L1 is English. Little is known about how other listeners (from the outer and expanded circles) would react to incorrect lexical stress production, such as L2 production. Without further investigation, it is impossible to determine whether their limited mastery of language rules causes the failure of L2 students to perceive lexical stress, specific stress placement patterns, or their failure to recognize syllables exhibiting suitable acoustic characteristics (Altmann, 2006). It is worth noting that most research changed the sounds of common words in synthetic speech in order to test how acoustic signals affect how L2 stress is perceived. This would fail to accurately depict the authentic correlation between syllables that are stressed and those that are not, which might lead to an unrealistic reproduction of human speech. Such as, in Altmann's (2006) research, 10 participants in each language group were shown non-words written in English orthography. The length of the 46 non-word test questions ranged from two to four syllables and were all open syllables. She could eliminate the familiarity effect, but there were still issues; for example, it was impossible to do a statistical analysis since so many terms were removed owing to coding issues. Furthermore, researchers employed actual words as stimuli in the scant studies that examined the impact of L2 stress generation using naturally recorded utterances; researchers used actual lexical items as stimuli, a practice that may potentially result in untrustworthy conclusions (Yu, 2008; Edmunds, 2009; Karjo, 2016). More precisely, any variations in performance might be attributed to the participant's acquaintance with the stimuli rather than their L1. Generally speaking, further investigation is required to see if the performance of those who use non-stressed language is affected by how they distribute or manage stress.

2.7 Studies of Lexical Familiarity

There is a suggestion that word familiarity, which is a subjective assessment based on familiarity ratings, would be a more precise measure. Gernsbacher (1984) conducted a series of experiments that showed how word familiarity explained the inconsistent relationships between word frequency and several other lexical factors (such as concreteness, polysemy and bigram frequency) that were previously reported in studies on lexical access. This was particularly observed for low-frequency words. This theory was examined via a series of tests with low-frequency words that exhibited variations in subjective frequency evaluations and either bigram frequency, concreteness, or polysemy. In addition, Gernsbacher (1984) discovered that familiarity ratings were more reliable in predicting reaction times in a lexical judgement task compared to concreteness evaluations. Gernsbacher (1984) proposed that subjective judgements of experience lexical familiarity might provide a more complete measure than written word frequency. This is because it takes into consideration all instances of encountering a certain lexical item. For instance, measurements of familiarity may indicate the extent to which one has been exposed to words via language production, as well as their experience with aural, visual, and written representations.

Pitt & Samuel (1993) also suggested that the impact of the lexicon on the recognition of speech sounds might be affected by differences among people's perceptions of word familiarity. Flege et al. (1995) examined how subjective knowledge of vocabulary affected the ability of both experienced and novice Japanese participants to recognize English words. Here, three noteworthy discoveries must be succinctly summarized: 1) The "positive" set of words, which are more familiar than their minimal pairs, would provide a more accurate identification of liquids compared to the "negative" set of words that are less familiar than their minimal pairs. 2) Inexperienced Japanese (IJ) subjects would experience a greater impact from lexical familiarity compared to experienced

Japanese (EJ) subjects due to the higher phonetic ambiguity of English liquids for IJ subjects. 3) Considering the fact that English /l/ sounds are generally less recognized by Japanese adults compared to word-initial /ɹ/ sounds, it can be inferred that the impact of lexical familiarity would be more pronounced for /l/ than /ɹ/.

Cynthia & John (1990) also conducted the experiments on lexical familiarity, using visual and auditory lexical decision tasks to determine the effect of familiarity, and added a variable to their experiment, which is the effect of frequency. The researchers used the lexical decision task in their first tests because of its widespread usage in the visual domain. The experiments analysis was interested in establishing effects of familiarity and frequency with a lexical decision task for visually (Experiment 1) and auditorily presented words (Experiment 2). The subsequent experiments extend these findings to the naming task (Experiments 3 and 4). The data showed that the pattern of frequency and familiarity effects diverged along two distinct dimensions. The impact size of the familiarity varied depending on the task, with an effect seen for the lexical choice task. Furthermore, the magnitude of the familiarity effect was significantly high for every naming delay. Familiarity scores may indicate a post lexical component, unlike frequency. considers presumably influenced by the ease of creation. The statistics given here support the idea that familiarity may be used as a measure across other senses, including how easily something can be produced. The cross-modal character of familiarity is further confirmed by the consistent familiarity effects seen in both the auditory and visual modalities. Additionally, Nusbaum & Dedina (1985) have previously shown familiarity effects for stimuli that were equalized for frequency using visual lexical decisions. In the field of auditory perception, the application of lexical decision has not been as common. However, Luce (1986) have previously used auditory lexical decisions to illustrate the impact of lexical frequency.

As previously stated, it may be inferred that the ability of L2 learners to recognize English segmental features will be impacted by their level of familiarity with the objects being assessed in the experiment. Several academics have presented empirical data corroborating this presumption (Yoshida, 1995; Yamada & Tohkura, 1992). The investigation of familiarity effects across different sensory modalities is of relevance, since Gernsbacher (1984) has proposed that subjective familiarity assessments reflect the cumulative experience with a word. If this is true, one would anticipate familiarity effects irrespective of medium. Additionally, empirical evidence appears to confirm that native Chinese and Thai participants learn the pronunciation of textbooks from an early age in American English (Hinkel, 1999; Pollap, 2010; Rawlings, 2013), which can also be referred to in section 2.1. Familiarity might provide a useful gauge of lexical stress learning. Robust familiarity effects are seen across tasks and domains, indicating a significant role for subjective familiarity.

2.8 Summary of Literature Review

In this chapter, I have presented a summary of research studies that have investigated the production and perception of stress in L2 students. The literature reviewed in this chapter offers valuable insights into how the phonological and phonetic characteristics of L1 influence the way L2 learners produce and perceive lexical stress, particularly when examining different variations of world English. However, there is still a need for further extensive exploration into L2 stress production and perception.

CHAPTER 3: METHODOLOGY

This section provides details of the research design.

3.1 Research Design

The Chinese and Thai students each got a stress production and perception test as part of this research's basic design. The theories involved in this experiment's design refered to sections 2.2 — 2.3. The production experiment was based on STM to explain whether the production of English L2 stress is predictable. Additionally, the PAM suggested that the listeners' existing knowledge significantly influences how they perceived speech from L2 speakers. The PAM also served as the basis for the interpretation of the results of perception experiments to determined how the perception of L2 lexical stress was affected. On the other hand, the conclusion of the experiments can also be verified the correctness of this model.

In the production experiment, the acoustic parameters were measured vowels by using Praat (Version 6.4.13) (Boersma & Weenik, 2024): average F0 (Hz); duration (ms); and intensity (dB), to examine the acoustic characteristics used in the implementation of stress. It was determined that these particular target cues were selected not only because they were crucial signals for lexical stress, as noted earlier (refer to section 2.4 & 2.5), but also because of the multiple functions they play in the overall realization of stress. The significance of F0 in the perception of stress by Mandarin and Thai L2 English learners was particularly relevant given their unique phonemic peculiarities. Duration and intensity were important factors in lexical stress of English production and perception (Wang, 2010; Ou, 2010).

Regarding the perception experiment, the entire set of recognition tasks was devised based on Ernie et al.'s (2020) work. The listeners were tasked with identifying the stress placement in trisyllabic words, and subsequently, the researcher analyzed mutual stress

location identification and accuracy scores to be investigated the transfer effect of L1 onto English L2 prosodic speech perception.

3.2 Participants

As Michael et al. (2011) pointed out the experimental sample size of twenty to fifty individuals can be used as a basic sample for quantifying the results at resolution. Based on this idea, a total of 34 adult students (participants details refer to 3.2.1 & 3.2.2) with normal hearing participated in the experiments: two recording groups and two listening groups. Meanwhile, the participants all had at least a bachelor's degree. Before beginning the recording process, both of them signed a consent form to participate.

It was impossible to be balanced the speaker sample by gender since the ratio of men and women learning languages was unbalanced. Additionally, the outcomes of this perceptual experiment revealed that gender differences did not exert a significant influence on variations in perceptual results (see sections 4.1 & 4.2). Consequently, the researcher posited that this variable had a negligible impact on the overall study's reliability. Two separate stages made up this study: a production task in the first stage and a perception task in the second stage.

3.2.1 Participants in the Production Stage

Stage 1 consisted of 4 speakers (2 Thai females and 2 Chinese females). The recruitment of speakers used judgmental sampling. The The researcher showed that Mainland China can be roughly divided into north and south, and the selected speakers were from one south and one north; their L1 is Mandarin Chinese, so they were somewhat representative, while Thailand's geographical division was not so complicated (Dutt, 1996, pp. 234-266), and the two recorded speakers selected were also from Bangkok. Their L1 is standard Thai (Bangkok Thai language), so they could be also showed representativeness. Thai speakers are all 21 years old, while Chinese speakers are all 28 years old. Is because the Thais who was selected happens to be the undergraduate

graduate student, the age was slightly on the small side, and the Chinese speakers was the master student (refers to Table. 3.1).

The 4 students who participated in the stimulus recording were chosen based on certain language profiles. About the L1, in order to ensure the consistency of the L1 language variable in the production sample, Mandarin Chinese was used as the L1 for two Chinese speakers, while standard Thai was used for Thai speakers. Specifically, two Chinese speakers explicitly stated that their L1 was Mandarin Chinese, and they had been communicating with their parents, siblings, teachers, classmates, and friends in Mandarin since birth. But it also showed that since the family was born and raised in Yangzhou and Shiyan, when communicated with their parents, dialects were also used. Notably, their parents employed the Yangzhou and Shiyan dialects more frequently than they speak dialects to their parents. Meanwhile, it was simple to compared and matched their average English proficiency values, including passed the IELTS test with a score of 6.5 and above. IELTS scores were a prerequisite for students from specific countries seeking visas to enrolled in universities. Typically, achieved an 'Overall Band Score' between 6.0 and 7.0 was widely recognized as evidence of English language proficiency for tertiary education institutions globally (Ciccarelli, 2001). Additionally, considered that the experimental population must be students with a certain level of English proficiency, in order to reduce the cost of time, they were all drawn from a random Faculty of Language and Linguistics sample to ensure sample representativeness. It was because there was a core course in Language and Linguistics and an understanding of some introductory phonetics and phonology knowledge that can quickly understand the experimental concepts and processes.

Selected speakers are willing to participate in multiple rounds of interviews and shared the recordings under supervision (to ensure the reliability of the recordings). They also pointed out that they had previously participated in similar acoustic experiments, shown

varying activity levels, and were familiar with the recording process, which reduced the possibility of mistakes.

Table 3.1: Summary of language backgrounds of speakers

Main language used as L1	English proficiency	Female	Male	Age range	Total
Standard Thai	IELTs: (6.5)	2	0	21	2
southwest mandarin	IELTs: (7)	1	0	28	1
Lower Yangtze	IELTs: (6.5)	1	0	28	1

(Note: The English Proficiency cell contains the mean value in parentheses)

3.2.2 Participants in the Perception stage

Besides, another 30 participants (15 Thais: 5 males, 10 females, and 15 Chinese: 2 males, 13 females) were in stage 2. All of them used a common snowball sampling design to be adopted. The age range of the Thai listeners was generally between 20 and 26, while the age range of the Chinese listeners appeared to be one or two years older than that of the Thais (see Table 3.2, broken down by region of China), but the actual difference was not large). It is convenient to control the listeners and, at the same time, increased the possibility of the researcher contacting the experimental survey population, which was more feasible.

Thai listeners were second-language English learners from Rangsit University. English was their major. Their English listening, speaking, reading, and writing skills were got a fluent level (the IELTS scores about 6.5 and above). There were two types of internationally recognized English proficiency tests in Thailand: Tofel and IELTS. (See

Appendix E for the Equivalency scores of the two tests). 6 Thai students took the Thai TOEIC exam (test score 605 and above) (μ=680), and 9 of the Thai students took the IELTS exam with a score of 5.5 or above (μ=6.5). The questionnaire indicated that Standard Thai was also the listeners' native language, as per their background information. They said that they used standard Thai this language both at home and at school. Other than that, English is seldom utilized.

Meanwhile, the native Chinese listeners were students from the Faculty of Language and Linguistics at the Universiti Malaya. In the context of ELF, the number of Mandarin Chinese students studying in Malaysia has increased in recent years. The 60th Universiti Malaya in the QS University Rankings (2024) has also attracted more Chinese students to the university. Additionally, each listener has to get an IELTS score of at least 6.5 or an equivalent. This score indicated that the listeners were able to make at least a "fluent judgment of English pronunciation as heard," as described by the IELTS band descriptions. The questionnaires also revealed that Chinese listeners commenced their English studies as early as six years old. The following table (Table 3.2) details the background information of the listeners.

Table 3.2: Summary of language backgrounds of listeners

Main language used as L1	**English proficiency**	**Female**	**Male**	**Age range**	**Total**
Standard Thai	Tofel: (608) IELTs: (6.5)	10	5	20-26	15
Mandarin	(6.5)	6	1	25-30	7
southwest mandarin	(6.5)	3	1	25-27	4

Hakka	(6.5)	1	0	29	1
Lower Yangtze	(7)	2	0	28-31	2
Wu	(7.5)	1	0	23	1

(Note: The English Proficiency cell contains the mean value in parentheses)

The language background of the Thai participant was much simpler than that of the Chinese participant, excepted for one who studied in Portugal for two years. He was born in Thailand and lives there. From the language background questionnaire found that he can be skilled in fundamental skills of English, so it appears that his international exposure had minimal impacted on his perception of English lexical stress from speakers with different L1 pronunciations.

It is worth noting that since there are 34 different provinces in the Chinese mainland, each province belongs to a different dialect region (see section 2.4.1.2 of the literature review), the Chinese audience had a different L1 language environment. Also, as we mentioned before, Mandarin has been categorized into four main subcategories: Northern Mandarin, Northwestern Mandarin, Southwestern Mandarin, and Jiang-Huai (Eastern) Mandarin. The dialect spoken in Beijing, which has distinct linguistic characteristics, may be considered a subdialect of Northern Mandarin. The six remaining primary dialect groups are often known as the "Southern dialects" due to their speakers, who make up around 30% of the Han population, residing in the region located south of the Yangtze River (excluding the southwest) (Chen, 1999).

Almost all of the participants had received education from countries in the outer circle (Malaysia) at some point during their academic career, typically for over a year and a half. However, one participant completed both undergraduate (3 years) and graduate (4 years) studies in Malaysia. Additionally, three Chinese participants had experience studying and working in other countries, their diverse L1 language and education backgrounds were

reflected in the language they most frequently used at home and elsewhere. (refer to Table 3.3 for more details). These three students used English much more frequently than other Chinese participants, engaging in communication with family and friends outside of class, and achieving an IELTS score of 7.7 or higher. Meanwhile, all of them were fluent Mandarin Chinese users.

Table 3.3: Diverse Participants' Linguistic Environments

Country	**Gender**	**Main Language used in L1**	**English Proficiency（IELTS）**	**Countries differ from others**
Thai	Male	Standard Thai	7	Portugal (2 years)
Chinese	Female	Mandarin	8.5	Singapore (since primary school); UK (bachelor); Malaysia (1.5 years) Australia (1 year); Malaysia (4 years)
	Female	southwest mandarin	7	
	Female	Mandarin	7	Nepal (1 year); Mexico (1.5 years); Ecuador (2 years); Malaysia (3 years)

3.3 Stimuli

One thing, in particular, is that different from previous studies in which stimuli were presented with nonce words so that it was more likely to avoid the familiarity effect (Jangjamras, 2021). A total of 20 real words of English in Appendix A and another 5 target words with carrier sentences in Appendix B were drawn from the British National

Corpus (BNC), a frequently used corpus in second language studies. In this study, the target words listed in Appendix A and B were sorted based on the sequence in which they appeared in the recording procedures. Also, to be considered is that English typically does not contain words exceeding three syllables without internal morphological composition; it has an effect on the distribution of stress (Albadar, 2021). Thus, based on Albadar's (2021) theory, the target words listed were relatively frequently used in second language studies, and all of these were trisyllabic words, so the side effects of syllables that may impeded the interpretation of the data were reduced. Bisyllabic words were also avoided because forced two-choice identification tasks may be biased against them. Moreover, the arrangement of stress positions systematically differed among the three syllables in various word categories, including noun class, verb class, and adjectives, which maximizes comprehensiveness.

Meanwhile, five target words be placed in sentences (see Appendix B) in which a context is provided, and all word classes used in isolation were contained. There was a distinction worth mentioning in the design of stimuli. In prior research, fixed carrier sentences were utilized to present target words. Sentential prominences in English were not dispersed uniformly throughout all sentence locations. It frequently exhibited prosodic prominence, characterized by pitch emphasis and occasional elongation, as a result of its positional impact. This elongation was, in part, a result of domain initial strengthening, an articulatory phenomenon in addition to contemporaneous lengthening in lexical items resulted from accentuation (Cho, 2001; Fougeron & Keating, 1997). A prosodic phrase's edge speech unit was enhanced spatially and temporally. Therefore, the target words were placed in the different middle of the carrier sentence to be controlled the sentential intonation constraints influence, but not at the beginning or end. However, it was difficult for the selection of 20 independent target words to fully conform to the phonological syllable structure. According to a pilot study by the researcher, the listeners

had trouble distinguishing the coda /m/ from /n/ and had trouble hearing speech production cues such the less audible onsets like /f/ or /s/. Crucially, segmentation work was also facilitated by having a single start consonant. Thus, the choice of stimuli also tried to took these factors into account. Only word types that had been identified as being suitable in prior research (Jangjamras, 2008; Albadar, 2021) were chosen. This was carried out to reduce testing time and fatigue.

3.4 Data Collection

Prior to conducting the experiment, participants were required to provide their consent by signing a form (See Appendix F). An initial questionnaire (see Appendix C) was used to gather background (details refers to 3.2.1 and 3.2.2) data on the participants (recording personnel and perception subjects), including their age, level of English ability, etc.

3.4.1 Production Task

As a result of the limitations of time and space caused by COVID-19, the reding materials and guidelines were sent to the speakers with a PDF, the recordings were conducted in a booth designed to reduced sound interference to make sure their recordings fulfilled the requirements for analysis. The recording equipment included using a high-quality cassette recorder, the microphone was positioned on a stand, maintained a consistent distance of around 20 cm from the listener's mouth. Speakers were urged to produce in a manner that was authentic, maintained a standard pace and volume.

Two sets of recordings were planned (refer to section 3.3). The researchers added a word and a sentence (not a target filter word or sentence) at the beginning and end of the two sets of experimental materials to ensure the speech production of the target tokens was clear and stable. Then two Thai speakers and another two Chinese speakers separately read each word in isolation and each sentence that contained the relevant target word. Recorded the same recording process three times and selected the one with the most

stable sound quality (moderate volume, less noise, and clear pronunciation) as the materials for the production experiment for analysis.

3.4.2 Perception Task

Thai and Mandarin Chinese were tonal languages that had distinct prosodic properties, which differed from those of English. This perception experiment aimed to be examined the variations in perceiving stress patterns in English among speakers with different L1 prosodic systems. Based on previous literature (Ingvalson et al., 2011), tone language speakers should be possessed the ability to detected stress. In other words, listeners whose L1 incorporates tone distinctions at the word level utilized suprasegmental cues of tone when perceived stress. The anticipated performance of the tonal speakers in perceived English stress was expected to be favorable. However, since the focus of this dissertation was on the recordings produced by speakers whose L1 was not English, the researcher investigated the potential transmission of L1 acoustic properties as a whole by compared the accuracy of the perception of stimuli produced by different groups of participants to stimuli produced by groups of speakers who were similar but different from their own language. Specifically, the perceptual experiment was the mutual experiment of the two language groups (Chinese and Thai). That was, Thai listeners listened to the audio of Chinese recorders, and the Chinese listened to Thai recordings. They were informed that the session aimed to gauge how effectively Chinese and Thai English speakers could perceive one another rather than measure their English proficiency.

The perception task consisted of two components. They were given access to an online questionnaire to complete the perceptual tested through a provided Google link. It was carried out to avoid the cost of invalid travel time for the researcher as well as the participants. Meanwhile, to monitor the progress of the listeners' experiment, the listeners needed to share the screen in Google Meet, and the researcher recorded the entire experimental process to facilitate the analysis of the experiment. The recording was

focused on questions that the listener may had at the end of the listening task, in the open-ended question "Please specify the reason for the perception problem", such as: "Do I need to focus on the phonetic and phonology aspect or other linguistic influences?" or "What are the implications of suprasegmental?", etc. The problems they generated may also affect the analysis of RQ2 results.

To ensure that all the listeners were familiar with the experimental procedures, a concise introduction of the experiment was provided at the outset. Through a Google Form, they could be accessed an online questionnaire and completed the perception tests. The perception tasks started by explaining the implication of English lexical stress, which was intended to guarantee that all listeners understand what lexical stress was. The minimal stress pairings "present (noun) vs. present (verb)" and "record (noun) vs. record (verb)" were displayed as examples in addition to the implication description on the questionnaire (see Appendix D). Before the first block began, the listeners received five warm-up practice trials with non-test items. After checked the sound volume of the headset with a sample sentence, made sure that both the researcher and the listeners had a good internet connection and did not cause voice quality issues or sound lagging. Then, listeners were asked to click the audio link only once based on their own pace (which might involve the perceived the task) and then completed the task of identified the appropriate stress, then clicked on the following audio, and so on (see Appendix D).

The test was divided into three tasks. The first section was to wrote out the word orthographically according to what they heard and capitalized the stressed syllables after heard the target word presented separately. The second section, the listeners proceeded to the second set of recordings. Each sentence contained a target word was played this time, after which the listeners selected the number of stressed syllables of the target word based on what they had heard. The final part was a listener's self-assessment, which included the difficulty level of selecting the two perception tasks. The difficulty level was divided

into five categories. It also included an open-ended question about what learners think was caused them to had difficulty perceiving lexical stress.

3.5 Data Analysis

The researcher carefully reviewed all speaker recordings to verified the natural pronunciation of target words, ensured clear segments and actual stress placement, before selected stimuli for production and perception tasks. After the recording was finished and then imported to Praat (Version 6.4.13) (Boersma & Weenik, 2024), Praat was used to performed noise elimination operations on each recording and evaluated vowel quality, duration(ms), pitch (HZ) levels, and intensity (dB) for each token. (These measurements are discussed further in the next section).

3.5.1 Production Task

3.5.1.1 Data Annotation

Once again, the process involved reading each word three times, resulting in a total of 300 words for this elicitation procedure (25 target words* 3 repetitions * 4 speakers). According to the guidelines provided in the Praat manual, a pitch range of 100 – 500 Hz was utilized for measuring F0 in female speakers. However, when analyzed creaky voice (which accounted for approximately 3% of all tokens), a specific token across all talkers had its pitch floor set at 50 Hz. Due to various disturbances or technical issues, a total of 233 tokens were excluded from the analysis. Consequently, only 67 words remained for further examination as each production was presumed to reflect the speaker's utmost effort in producing stress appropriately.

According to Ladefoged (2001), segmentation criteria were based on audio cues in waveform, spectrogram, and auditory perception. Vowel boundaries (onset and offset of stressed and unstressed vowels) were segmented according to the following criteria (Table. 3.4):

Table 3.4: Annotation in spectrogram

Tier	Boundaries
Tier1: Target Words	Entire target words
Tier 2: Syllable	Entire target syllables Boundaries: Start from the start point of the plosive, voicing bar, or vowel formant. End at the disappearance of F2.
Tier 3: Vowel	The vowel onset was defined as the first upward zero crossing at the beginning of voicing in the waveform, usually accompanied by the onset of the F1. On the spectrogram, it was the visible cessation of aspiration of the onset of voicing. The vowel offset was marked at the downward zero crossing immediately following the final glottal pulse in the waveform (Zhang et al. 2008); On the spectrogram, this was usually the cessation of all formants; If extra oscillation or cycles were present, a sudden drop in F1 was used as a cue (de Jong 2004);

	The offset of a syllable with an obstruent coda, downward zero crossing, was defined as the end of F2; In a syllable with a bilabial nasal coda (e.g., /n/), the visual cues to the end of the vowel included a decrease or drop in F1 on the spectrogram and rather faint higher formants compared to that of the vowel portion (See Fig. 2.1). When the vowels were nasalized, there was usually an extra peak of energy, generally between the oral and nasalized portions of the vowel. (The nasalized portion of the vowel was included in the vowel duration.) Extra aperiodic cycles or creaky voice after the regular offset (for both open and closed syllables) seen on the waveform were excluded (Ladefoged, 2001).
Tier 4: Remark	The advice came after a doctor with extensive experience in spectrogram observation and annotation and a Ph.D. specializing in phonetics and phonology double-checked the annotation. The Vowel boundary —— onset and offset will be at the nearest zero crossing. The transient will be included.

See Fig 3.1 below. In many cases, when the vowels underwent nasalization, there tended to be an additional peak of energy typically situated between the oral and nasalized segments of the vowel. (The duration of the nasalized segment was considered part of the overall vowel duration.) Any supplementary irregular cycles or creaky voice occurred after the regular cessation (for both open and closed syllables) observed on the waveform were disregarded.

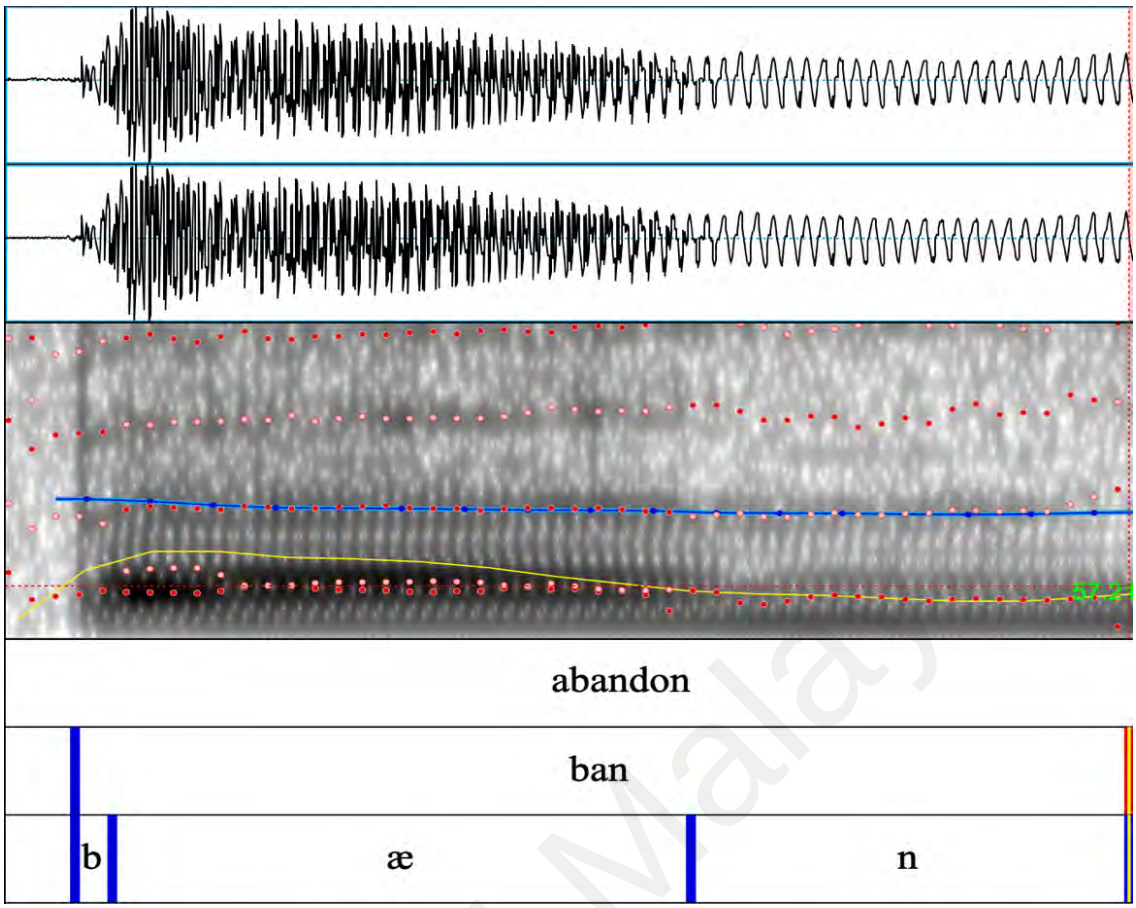

Figure 3.1: A spectrogram of the word "abandon" with vowel boundaries.

3.5.1.2 Acoustic Measurement

Acoustic analysis Software Praat (Version 6.4.13) be used to be measured the acoustic parameters duration (ms), average intensity (dB), and F0 (Hz) for each token. All recorded utterances were divided into segments and acoustically evaluated. To ensure consistency, the present study employed acoustic measurements of individual vowels instead of syllables. This choice was made because stress effects in English primarily manifest in vowels rather than consonants (Fry, 1995). Followed the guidelines provided by the Praat manual, a pitch range of 100-50 Hz was set for female speakers and 75-300 Hz for male speakers when assessed F0. The mean of F0, duration, and intensity of vowels were measured by average measurements across each vowel in Praat.

The acoustic data that included F0, vowel duration, and intensity was first extracted manually; due to manual extraction may be due to bias in the selection of the average

value. Therefore, this data lacks accuracy and scientific validity. Then, the researcher rechecked it with the script. Furthermore, in cases where the F0 values of a target word could not be determined by the pitch tracker in Praat, individual examination of those tokens was conducted. If this scrutiny resulted in obtaining F0 values, the token was included for analysis. Tokens without determinable F0 values were excluded from analysis based on F0. Additionally, tokens lacking a pitch accent on the target word were also considered for exclusion. It is worth mentioning that only 3 out of 67 tokens (equivalent to 0.9% of the data) were excluded due to any of these aforementioned reasons during this dissertation study.

3.5.1.3 Production analysis

In production experiments, the most important task is to identify stressed syllables. The researcher began by examining the raw values by comparing the mean values of f0 (Hz), vowel duration (ms), and intensity (dB) of the two groups in the two language groups of participants. The average value of the three acoustic signals in a syllable is higher than the average value of the other two syllables, establishing that syllable as the stressed syllable of the word. There are prominent syllables in the three acoustic parameters (the raw values are obviously larger than the other two syllables). Therefore, word-by-word analysis should be carried out according to different part-of-speech classifications in order to demonstrate the tendency of lexical stress. This is also attributed to the variation in emphasis on acoustic parameters across different lexical stress studies, as well as in experiments on word stress production in various languages. Therefore, a word-by-word analysis by classification is deemed more reliable in the present study.

Due to the limitations of the participants, the production experiment was unable to complete the parametric test. The participants only had four recorders available to experiment with. A straightforward descriptive-quantitative study was carried out by the researcher in order to compare the acoustic characteristics of the lexical stress of English

words generated by two distinct language groups (Thai and Chinese). The researcher analyzed the 64 tokens separately and then classified them by part of speech in order to summarize the findings.

3.5.2 Perception analysis

Firstly, the respondents' written answers to what they heard and the detected stressed syllables in their responses were evaluated, and these results were compared to the actual production for different syllable stresses. The participants' replied to each word were classified as either accurate or wrong, or in other words, called different or same depending on which syllable the participants picked as the stressed syllable based on actual utterance. Based on whether the stress identified matches the intended stress as spoken by the speaker (as recorded previously, where each word had prominent stressed syllables), an evaluation was conducted to determine differentiation. The four participant groups' (two Thai and two Chinese groups) respective accuracy means were calculated for each token. In addition to stress similarities, different responses, that is, participants were selected a syllable that does not contained the primary stress in words, were also analyzed. If there were a typo situation, the researcher would mark it simultaneously, extracted the marker word individually until the end of the experiment, asked the participants during the retrospective session, and determined the correct syllable they perceived. If the entire word was completely misspelled (unable to understand the meaning of syllables and utterly unaware of syllable division rules), this data should be removed. Since the target words required for this experiment were simple high-frequency words, the researcher expected that there were not this kind of typographical errors. After the experiment, the researchers collated the participants' answers and found that only two mistakes (ratio - ratial; allocate – allegate) were made by the same person.

It was worth noting that part 1 of the first experiment required participants to orthographically wrote out the words they heard. This orthography task involved how to

be established the division of syllables. Johnson (2017) pointed out that the syllable served as the primary unit for explaining the combinations of phonemes in a language. Additionally, they demonstrated that numerous rules governing allophones of sounds can solely be expressed in terms of the syllable. For instance, in English, /h/ can only occured at the beginning of a syllable, /bɪ,heɪv/ (Redford & Randall, 2005). Therefore, it was crucial for listeners to understand the concept of syllable.

The syllable prominence theory was based on auditory perceptions. Syllables were associated with peaks in prominence, typically aligning with the number of vowels present (Johnson, 2017). Although this principle holds true in most instances, there exist some exceptions. For example, the word 'believes/bɪli:v/' contained two vowels and two peaks of sonority. However, the term 'spy /spaɪ/' possessed only one vowel; nevertheless, the sonority level of '/s/' surpasses that of '/p/', resulting in two sonority peaks. It was important to note that when analyzed syllable structure, we focused on phonetic sounds rather than their written representation (Derwing & Eddington, 2014). The only thing on which we needed to focus in this study was trisyllabic words. In the context of syllable rules, variations can be observed in the division of triphthongs, such as the /aɪə/ sound in BrE fire /faɪə/, which may be perceived as either one or two syllables. Sequences involved sounds like /iə/ could potentially be analyzed as having one or two syllables. For instance, a word like lenient /li:niənt/ might be subject to different interpretations.

However, in this experiment, there was only one confusing target word, volunteer/, vɒlən'tɪə (ɾ)/ that defaulted to a three-syllable word. As a result, there were almost no syllable divisions that can confuse the listeners.

In the third part of the perception experiment, which was the part where the listeners self-evaluated the difficulty of the above two sections of the perception tasks. After the researchers counted the perceived choices of different L1 listeners about the difficulty of presented the target words independently or accompanied carrier sentences, we focused

more on examined the listeners' narratives as well as terminology that they found difficult to grasp to supplement the test results. At the same time, the researcher compared the perception results of different listeners with the lexical stress annotation in the dictionary so as to be conducted a triangulation analysis, which determined the extent to which the listener's L1 affected the lexical stress perception of Chinese and Thai students and make the experiments completer and more reliable.

3.6 Summary of Methodology

This chapter outlines the methodologies employed in two experiments. The production experiment involved the participation of 2 Chinese and 2 Thai speakers, while the perception experiment included 15 Chinese and 15 Thai listeners. English words sourced from the British National Corpus (BNC) were utilized to examine how different L1 speakers (Chinese and Thai) produce lexical stress in English, as well as investigate any potential influence of their L1 on the perception of English lexical stress. A detailed acoustic analysis was conducted on selected speech samples to identify the acoustic features associated with lexical stress for both language groups. In the speech perception tasks, listeners were tasked with identifying the location of stress in English words under various conditions (words isolated or within sentences), with accuracy scores being recorded. By analyzed participants' stress production locations and perception accuracy, this study aimed to explore how factors such as stress position, acoustic features, and language group impact both stress production and perception.

CHAPTER 4: FINDINGS

This section showcases the findings derived from the experiments conducted in Chapter 3, focusing on stress production and perception. The outcomes are also divided into two segments: stress production and stress perception.

4.1 Production Results

This section contained a detailed presentation of the recording production's outcomes. The acoustic parameters F0, duration, and intensity were examined in the findings to determine if participants employed these acoustic correlates to indicated stress and to be assessed the variations in the acoustic features of English lexical stress created by Thai and Chinese speakers.

First, the researchers listed 25 target words (see Table 4.1-4.25), including words in isolation and words with sentences to separated analysis. The researchers explained the results of each word under three different parts of speech (noun, verb, and adjective).

4.1.1 Noun class

As shown in the table below (Tables 4.1–4.7), in the seven noun target words "japanese", "opponent", "ratio", "volunteer", "transistor," "basketball," and "musician," the difference between the mean and standard deviation of the duration (ms) of both the words produced by the Chinese group and the Thai group was more significant than the difference in F0 (Hz) and intensity (dB).

When we saw F0 and duration, the difference was relatively large; when the Chinese group spoke of these 7 target words, the pitch level was relatively high, about 216 Hz, and the gap was not broad. Nevertheless, in the pronunciation of Thais, the F0 value in each syllable of the words had a comparatively large gap. There was almost no difference in intensity values; Chinese and Thai pronounced these 7 words at average intensity without great fluctuations. It can be seen that these 7 nouns, Chinese and Thai

pronunciators, mainly relied on duration to discriminate lexical stress since the numerical difference of each syllable was the largest.

Table 4.1: The Mean Value of Three Acoustic Cues for "japanese"

	Production	Acoustic pattern	Syllable 1	Syllable 2	Syllable 3	Standard Division
japanese	Chinese	F0 (Hz)	219	221	234.5	8.43
		Intensity (dB)	67.5	63.5	66.5	2.08
		Duration (ms)	50	41	195.5	86.72
	Thai	F0 (Hz)	146	118.5	198	40.37
		Intensity (dB)	62.5	64	67.5	2.57
		Duration (ms)	80.5	37.5	140.5	51.73

(Note: Std.D. = Standard deviation for each syllable mean rating)

Word 'japanese' - The Chinese speakers put the stress on syllable 3. The duration of the syllable 3 vowel was significantly higher than that of the initial and second syllables, about 195.5 ms; Thai speakers also placed stress on syllable 3. The gaps between F0 and duration on syllable 3 compared with the first and second syllables were relatively large. The F0 of the third syllable is 198 Hz, which was larger than both the first two syllables, which were about 70 Hz, and in terms of vowel duration, the difference between the third syllable and the first two syllables was about 80 ms. Therefore, both F0 and vowel duration were vital parameters in Thai production. And in this noun word, the stressed syllable both placed at syllable 3.

Table 4.2: The Mean Value of Three Acoustic Cues for "opponent"

	Production	Acoustic pattern	Syllable 1	Syllable 2	Syllable 3	Standard Division
opponent	Chinese	F0 (Hz)	215.5	244.5	204	20.87
		Intensity (dB)	61	65.5	64	2.29
		Duration (ms)	33	129.5	106	50.32
	Thai	F0 (Hz)	182.5	238.5	191.5	30.07
		Intensity (dB)	63.5	69	67	2.78
		Duration (ms)	38.5	150	66	58.09

(Note: Std.D. = Standard deviation for each syllable mean rating)

In the word 'opponent', Chinese speakers place stress on syllable 2 (130 ms), which results in a noticeable difference in vowel duration compared to the initial syllables (97 ms) and a difference of 24 ms from syllable 3. Thai speakers also place stress on syllable 2, with a difference in F0 of about 52 Hz between the first and third syllables, as well as a significant discrepancy in vowel duration, approximately 100 ms compared to the first and third syllables. The stressed syllable is consistently placed at the second position for both languages.

Table 4.3: The Mean Value of Three Acoustic Cues for "ratio"

	Production	Acoustic pattern	Syllable 1	Syllable 2	Syllable 3	Standard Division
ratio	Chinese	F0 (Hz)	242	221	178.5	32.35

		Intensity (dB)	68	64	62	3.06
		Duration (ms)	101.5	122.5	203.5	53.86
	Thai	F0 (Hz)	243	195.5	156	43.56
		Intensity (dB)	70	66.5	62.5	3.75
		Duration (ms)	130	54	165	56.75

(Note: Std.D. = Standard deviation for each syllable mean rating)

The word "ratio" was stressed on the third syllable in Chinese, with a small difference in vowel duration between syllables 1 and 2, but a significant difference of 90 ms with syllable 3. Similarly, Thais also placed stress on the third syllable, demonstrating a considerable difference in vowel duration of 111 ms from syllable 2 and 35 ms from the initial syllable. Therefore, it can be concluded that the stressed syllable of this word is consistently placed at the third position in both language groups.

Table 4.4: The Mean Value of Three Acoustic Cues for "volunteer"

	Production	Acoustic pattern	Syllable 1	Syllable 2	Syllable 3	Standard Division
volunteer	Chinese	F0 (Hz)	220.5	213	217.5	3.77
		Intensity (dB)	69	65.5	67	1.76
		Duration (ms)	99	48	254	107.29
	Thai	F0 (Hz)	170.5	164	201	19.75

		Intensity (dB)	66.5	66.5	69	1.44
		Duration (ms)	62	80	263	111.22

(Note: Std.D. = Standard deviation for each syllable mean rating)

Word 'volunteer' - Chinese speakers placed stress on the syllable 3, the vowel duration was significantly different from the length of the first two syllables. The difference in duration from the first syllable is 155 ms, and the difference with the second syllable was even larger, about 206 ms. But since the third syllable was a diphthong, it is to be expected that the duration is significantly higher than the first two syllables. Thais also showed the stress on the syllable 3, with a huge difference in vowel duration. The difference between the initial and second syllables of the Thai speakers was not large, but the difference between them and the third syllable was big, at about 200 ms. However, since the third syllable was a diphthong, the duration was significantly longer than the first two syllables.

Table 4.5: The Mean Value of Three Acoustic Cues for "transistor"

	Production	Acoustic pattern	Syllable 1	Syllable 2	Syllable 3	Standard Division
transistor	Chinese	F0 (Hz)	227	233.5	191	22.89
		Intensity (dB)	63	67.5	63	2.60
		Duration (ms)	50	110	103	32.81
	Thai	F0 (Hz)	183.5	233.5	133.5	50.00
		Intensity (dB)	65	66.5	64.5	1.04

		Duration (ms)	86	71	81.5	7.70

(Note: Std.D. = Standard deviation for each syllable mean rating)

Word ‘transistor’ – Chinese speakers put stress at syllable 2, different in vowel duration with the initial syllable to 60 ms, and 7 ms difference with the last syllable; While Thais showed stress at syllable 1, with huge different in F0, the Std. D. showed values of 50 between syllables. However, for this word, the stressed syllable placed on the two language groups is quite different. Chinese speakers are stressed in syllable 2, while Thais are stressed in syllable 1. Moreover, considering the provided information, researchers found it essential to establish a consistent set of criteria for determining lexical stress. Thus, we continued to depend on the duration as a factor in determining the lexical stress in the pronunciation of "transistor" by Thai speakers.

Table 4.6: The Mean Value of Three Acoustic Cues for “basketball”

	Production	Acoustic pattern	Syllable 1	Syllable 2	Syllable 3	Standard Division
basketball	Chinese	F0 (Hz)	238	212	140	50.77
		Intensity (dB)	65.5	60	64.5	2.93
		Duration (ms)	92	55.5	127	35.75
	Thai	F0 (Hz)	241.5	218.5	172.5	35.13
		Intensity (dB)	71	67	68	2.08
		Duration (ms)	138.5	37	58.5	53.49

(Note: Std.D. = Standard deviation for each syllable mean rating)

Word 'basketball' - Chinese speakers tend to place stress on syllable 3, leading to a noticeable difference in vowel duration of about 53 ms compared to syllable 2 and a difference of about 35 ms compared to syllable 1. For Thais, stress was placed on syllable 1, resulting in a significant difference in F0 in syllable 3 of about 196 Hz compared to the other two syllables, as well as a substantial difference in vowel duration for syllable 1 of about 91 ms compared to the other syllables.

Table 4.7: The Mean Value of Three Acoustic Cues for "musician"

	Production	Acoustic pattern	Syllable 1	Syllable 2	Syllable 3	Standard Division
musician	Chinese	F0 (Hz)	220.5	230	217.5	6.53
		Intensity (dB)	66	66	63.5	1.44
		Duration (ms)	52	80.5	50.5	16.90
	Thai	F0 (Hz)	138	237.5	170.5	50.74
		Intensity (dB)	66.5	70	67	1.89
		Duration (ms)	76	100	117.5	20.83

(Note: Std.D. = Standard deviation for each syllable mean rating)

The word 'musician' was pronounced differently by Chinese and Thai speakers. Chinese speakers stressed the second syllable, resulting in a 30 ms difference in vowel duration compared to the initial and last syllables. On the other hand, Thai speakers placed stress on the third syllable, with a greater difference in pitch (F0) than vowel duration. In addition, the F0 of the second syllable appeared higher than that of the other two syllables. Furthermore, based on the given information above, it was necessary for researchers to

standardize the elements used to determine lexical stress. Therefore, we still relied on duration as a determinant for lexical stress in Thai speakers' pronunciation of "musician".

4.1.2 Verb class

Among the six verbs mentioned in Table 4.8-Table 4.13, it was observed that, except for the word "modify" pronounced by Chinese speakers and the word "recommend" pronounced by Thais, there was a greater difference in F0 values between syllables than in duration. This suggests that speakers placed more emphasis on F0 when pronouncing these two words' syllables. On the other hand, for the remaining four words - "allocate," "understand," "referee," and "abandon" - compared to F0 and intensity, there was a considerable difference in the duration of word syllables regardless of whether they were spoken by Chinese or Thai speakers.

The analysis also revealed that the duration difference between these four words was significantly higher when pronounced by Chinese speakers compared to their Thai counterparts."

Table 4.8: The Mean Value of Three Acoustic Cues for "modify"

	Production	Acoustic pattern	Syllable 1	Syllable 2	Syllable 3	Standard Division
modify	Chinese	F0 (Hz)	243	239	195.5	26.35
		Intensity (dB)	68	61.5	62	3.62
		Duration (ms)	117.5	71.5	88.5	23.26
	Thai	F0 (Hz)	229.5	213	168.5	31.55
		Intensity (dB)	72.5	70.5	67	2.78

		Duration (ms)	157	50.5	170	65.56

(Note: Std.D. = Standard deviation for each syllable mean rating)

Word ‘modify’ - Chinese speakers put the stress at syllable 1, little difference showed in F0 and vowel duration; For Thais at syllable 3, evident different in vowel duration about 170 ms, obvious longer than syllable 2 to 120 ms, and the gap with syllable 1 was not so big, about 13 ms.

Table 4.9: The Mean Value of Three Acoustic Cues for “allocate”

	Production	Acoustic pattern	Syllable 1	Syllable 2	Syllable 3	Standard Division
allocate	Chinese	F0 (Hz)	233.5	219	201.5	16.02
		Intensity (dB)	63.5	67.5	61.5	3.06
		Duration (ms)	44	116.5	102	38.36
	Thai	F0 (Hz)	220.5	207	160.5	31.48
		Intensity (dB)	70.5	70	67	1.89
		Duration (ms)	70.5	67	151.5	47.81

(Note: Std.D. = Standard deviation for each syllable mean rating)

Word ‘allocate’ - Chinese speaker placed the stress on syllable 2. The duration of the syllable 2 is longer than that of the initial and last syllables; However, Thais placed stress on syllable 3, clear different in vowel duration also, about 152 in syllable 3, which was longer than that of initial and second syllables to 82 ms.

Table 4.10: The Mean Value of Three Acoustic Cues for “understand”

	Production	Acoustic pattern	Syllable 1	Syllable 2	Syllable 3	Standard Division
understand	Chinese	F0 (Hz)	224.5	226.5	218.5	4.16
		Intensity (dB)	63	65.5	64	1.26
		Duration (ms)	43.5	37.5	83	24.72
	Thai	F0 (Hz)	181	168	204	18.23
		Intensity (dB)	66	65.5	71.5	3.33
		Duration (ms)	48	36.5	110	39.54

(Note: Std.D. = Standard deviation for each syllable mean rating)

Word 'understand' – The deposition of the stress for Chinese speakers were at syllable 3, huge different in vowel duration, about 45 ms; Also, for Thais at syllable 3, different same with phenomenon of Chinese, but the gap between stressed and unstressed syllables was greater, about 68 ms.

Table 4.11: The Mean Value of Three Acoustic Cues for "recommend"

	Production	Acoustic pattern	Syllable 1	Syllable 2	Syllable 3	Standard Division
recommend	Chinese	F0 (Hz)	228	230	203.5	14.76
		Intensity (dB)	64.5	62.5	63.5	1.00
		Duration (ms)	82	27	114.5	44.23

	Thai	F0 (Hz)	230.5	212	169	31.55
		Intensity (dB)	72	70.5	66.5	2.84
		Duration (ms)	60	46.5	42.5	9.17

(Note: Std.D. = Standard deviation for each syllable mean rating)

Word 'recommend' – The placement of the stress for Chinese speakers at syllable 3, different in vowel duration about 115 ms, obvious longer than other syllables; Thais place stress on syllable 1, with different in F0 and vowel duration. The difference in F0 looked huge than duration in Thai speakers, yet the largest number still appeared on the initial syllable.

Table 4.12: The Mean Value of Three Acoustic Cues for "referee"

	Production	Acoustic pattern	Syllable 1	Syllable 2	Syllable 3	Standard Division
referee	Chinese	F0 (Hz)	222	215.5	217	3.40
		Intensity (dB)	64	64	65	0.58
		Duration (ms)	48.5	31.5	263	129.03
	Thai	F0 (Hz)	221.5	194.5	136.5	43.43
		Intensity (dB)	73.5	69.5	67	3.28
		Duration (ms)	101	40	150	55.11

(Note: Std.D. = Standard deviation for each syllable mean rating)

Word ‘referee’ – The placement of stress of Chinese speakers were at syllable 3, huge different in vowel duration, about 224 ms; Same position for Thais, also with different in F0 and vowel duration looked in Standard Devision. The longer duration still appeared on the syllable 3, about 150 ms, much higher than syllable 2 to 110 ms, and higher than syllable 1 to 50 ms.

Table 4.13: The Mean Value of Three Acoustic Cues for “abandon”

	Production	Acoustic pattern	Syllable 1	Syllable 2	Syllable 3	Standard Division
abandon	Chinese	F0 (Hz)	221.5	222.5	206	9.25
		Intensity (dB)	60.5	66	64.5	2.84
		Duration (ms)	57	111.5	70	28.46
	Thai	F0 (Hz)	187	209.5	172.5	18.64
		Intensity (dB)	64.5	73	68.5	4.25
		Duration (ms)	45	111	108.5	37.40

(Note: Std.D. = Standard deviation for each syllable mean rating)

Word ‘abandon’ - Chinese speakers put the stress at syllable 2, huge different in vowel duration. The vowel duration of the syllable 2 is quietly longer than that of syllable 1 and syllable 3; Also, Thai speakers put the stress on syllable 2, with different in vowel duration and F0.

4.1.3 Adjective class

When looked at the 7 adjectives, we found that Chinese seem to paied more attention to the proportion of F0 when produced adjectives, and it showed this pattern in words "positive"; "rational"; "consistent" and "cultural".

However, Thais still seemed to be insisted on pronunciation based on vowel duration; whether it was these four words or the remained three words, "absolute," "fantastic," and "delicate" are the same.

An interesting finding was that among adjectives, the difference in the acoustic parameter of F0 and duration was not as significant as in pronouncing nouns or verbs between Chinese and Thai pronounced different syllables. Of course, intensity was still the most negligible difference, whether a noun, verb, or adjective.

Table 4.14: The Mean Value of Three Acoustic Cues for "positive"

	Production	Acoustic pattern	Syllable 1	Syllable 2	Syllable 3	Standard Division
positive	Chinese	F0 (Hz)	268	227.5	201	33.74
		Intensity (dB)	67	64.5	61	3.01
		Duration (ms)	100.5	78.5	118	19.79
	Thai	F0 (Hz)	209	175	181	18.15
		Intensity (dB)	74	64.5	68.5	4.77
		Duration (ms)	86.5	46	81	21.97

(Note: Std.D. = Standard deviation for each syllable mean rating)

Word ‘positive’ – The stress deposition of Chinese speakers was at syllable 3, different in F0 and vowel duration; While Thais were at syllable 1, with difference in F0 and vowel duration. However, there is not much difference between the stressed and unstressed syllables of word “positive” in Chinese or Thai speakers.

Table 4.15: The Mean Value of Three Acoustic Cues for “rational”

	Production	Acoustic pattern	Syllable 1	Syllable 2	Syllable 3	Standard Division
rational	Chinese	F0 (Hz)	242	224	188.5	27.22
		Intensity (dB)	66.5	66	61.5	2.75
		Duration (ms)	103.5	75	99.5	15.43
	Thai	F0 (Hz)	235	203	185	25.32
		Intensity (dB)	73	69	66	3.51
		Duration (ms)	87	25.5	44	31.55

(Note: Std.D. = Standard deviation for each syllable mean rating)

Word ‘rational’ – The placement of stress of Chinese placed the stress at syllable 1, difference shows in F0 and vowel duration; For Thais also at syllable 1, also with different in F0 and vowel duration, with obvious different in vowel duration with the second syllables to 61 ms, the difference from the syllable 3 is 55 ms;

Table 4.16: The Mean Value of Three Acoustic Cues for “absolute”

	Production	Acoustic pattern	Syllable 1	Syllable 2	Syllable 3	Standard Division

absolute	Chinese	F0 (Hz)	230.5	241.5	212	14.91
		Intensity (dB)	60	62.5	63.5	1.80
		Duration (ms)	83.5	63	173	58.50
	Thai	F0 (Hz)	188.5	176.5	181	6.06
		Intensity (dB)	71	67	67	2.31
		Duration (ms)	94	33	175	71.23

(Note: Std.D. = Standard deviation for each syllable mean rating)

Word 'absolute' - Chinese speakers put the stress at syllable 3, clear different in vowel duration compared with unstressed syllable, about 100 ms; Thais also put the stress on the same place, different in vowel duration. The obvious difference showed about 142 ms with syllable 2, and 81 ms with syllable 1.

Table 4.17: The Mean Value of Three Acoustic Cues for "consistent"

	Production	Acoustic pattern	Syllable 1	Syllable 2	Syllable 3	Standard Division
consistent	Chinese	F0 (Hz)	221	248	217.5	16.69
		Intensity (dB)	62	65	61	2.08
		Duration (ms)	98.5	90	81	8.75
	Thai	F0 (Hz)	190.5	187	176	7.57

		Intensity (dB)	67	68.5	68.5	0.87
		Duration (ms)	31.5	80	66	24.96

(Note: Std.D. = Standard deviation for each syllable mean rating)

Word 'consistent' – The stress placement of Chinese speakers was at syllable 1, different in vowel duration, the gap between syllables was not significant; The stress placement of Thai speakers was at syllable 2, difference showed in vowel duration, the duration in syllable 2 was 80 ms, the difference from the first syllable (31.5 ms) was greater than the difference from the third syllable (66 ms).

Table 4.18: The Mean Value of Three Acoustic Cues for "fantastic"

	Production	Acoustic pattern	Syllable 1	Syllable 2	Syllable 3	Standard Division
fantastic	Chinese	F0 (Hz)	227	235.5	208	14.08
		Intensity (dB)	64.5	60.5	59	2.84
		Duration (ms)	60	84.5	112.5	26.27
	Thai	F0 (Hz)	177.5	209	155	27.12
		Intensity (dB)	68.5	68.5	63.5	2.89
		Duration (ms)	53	134.5	84.5	41.10

(Note: Std.D. = Standard deviation for each syllable mean rating)

Word 'fantastic' - Chinese speakers place the stress at syllable 3, clear different in F0 and vowel duration; While Thais' stress placement was at syllable 2, different in F0 and

vowel duration. The vowel duration was obviously the element that both language group speakers produced lexical stress.

Table 4.19: The Mean Value of Three Acoustic Cues for "delicate"

	Production	Acoustic pattern	Syllable 1	Syllable 2	Syllable 3	Standard Division
delicate	Chinese	F0 (Hz)	251	224.5	190.5	30.33
		Intensity (dB)	65	63.5	60	2.57
		Duration (ms)	131	59.5	71	38.39
	Thai	F0 (Hz)	230.5	224	164.5	36.37
		Intensity (dB)	72	69.5	68	2.02
		Duration (ms)	67	54	144	48.64

(Note: Std.D. = Standard deviation for each syllable mean rating)

Word 'delicate' – The deposition of stress for Chinese speakers were at syllable 1, the difference shows in F0 and greater in vowel duration, the vowel duration in syllable 1 was 131 ms, longer than syllable 2 to 71 ms; than syllable 3 to 60 ms ; While for Thais were put on syllable 3, greater difference in vowel duration, the syllable 3 was about 144 ms, the initial and second syllable only has 67 ms and 54 ms.

Table 4.20: The Mean Value of Three Acoustic Cues for "cultural"

	Production	Acoustic pattern	Syllable 1	Syllable 2	Syllable 3	Standard Division
cultural	Chinese	F0 (Hz)	247	228.5	187.5	30.45

		Intensity (dB)	65.5	62	63	1.80
		Duration (ms)	95	65.5	81.5	14.77
	Thai	F0 (Hz)	227.5	189.5	162.5	32.65
		Intensity (dB)	70	68.5	63.5	3.40
		Duration (ms)	100	49.5	31.5	35.51

(Note: Std.D. = Standard deviation for each syllable mean rating)

Word ‘cultural’ – The stress placement for Chinese speakers were at syllable 1, clear different in F0 than vowel duration; For Thais were at syllable 1, the difference shows in vowel duration, Thais still seemed to be insisted on pronunciation focused on duration.

4.1.4 Words in sentence

As for the 5 target words in the carrier sentence, the words "decided" and "important" were similar to the previous target words in the "noun classes" category that occurred when the target words occurred separately, and both Chinese and Thais relied more on duration to pronounce syllable stress. The words "analyze" and "citizen" pronounced by Chinese also showed the enormous difference in duration, but Thais showed the dissimilarity. The values of these two words showed that the difference between F0 and duration in syllables spoken by Thais seems to be similar, with F0 slightly higher than duration.

The acoustic parameter of F0 in the word "example" was essential, and the difference between syllables was greater than duration in F0, whether Thai or Chinese pronounced (refer to Table 4.22). The only consistent observation was a slight variation found in intensity values.

Table 4.21: The Mean Value of Three Acoustic Cues for "decided"

	Production	Acoustic pattern	Syllable 1	Syllable 2	Syllable 3	Standard Division
decided	Chinese	F0 (Hz)	227	259	225.5	18.92
		Intensity (dB)	66	66.5	67	0.50
		Duration (ms)	64	153.5	72.5	49.40
	Thai	F0 (Hz)	132	135.5	184.5	29.35
		Intensity (dB)	63	66	66.5	1.89
		Duration (ms)	56.5	156	96	50.10

(Note: Std.D. = Standard deviation for each syllable mean rating)

Word 'decided' – Chinese speakers put the stress at syllable 2, distinct in duration, the duration of syllable 2 was 153.5 ms, while syllable 1 and syllable 3 only had 64 ms and 72.5 ms, the gap can be seen to be wide; Thai speakers also placed it at syllable 2, the biggest difference in duration, the duration of syllable 2 has 156 ms, the syllable 1 had 56.5 ms, the difference was about 100 ms; the syllable 3 had 96 ms, the difference was much smaller than the difference between the unstressed syllable 1.

Table 4.22: The Mean Value of Three Acoustic Cues for "example"

	Production	Acoustic pattern	Syllable 1	Syllable 2	Syllable 3	Standard Division
example	Chinese	F0 (Hz)	240	224.5	289.5	33.95

		Intensity (dB)	68	68.5	67.5	0.50
		Duration (ms)	71	86.5	61.5	12.62
	Thai	F0 (Hz)	171.5	246	188	39.13
		Intensity (dB)	59	65.5	65	3.62
		Duration (ms)	42.5	75	57	16.28

(Note: Std.D. = Standard deviation for each syllable mean rating)

Word 'example' – The stress deposition for Chinese speakers were at syllable 2, clear different in F0 than vowel duration; For Thais were at syllable 2, the difference also showed in F0 than vowel duration. In this word, Chinese and Thai speakers mainly used F0 to produce stressed syllable.

Table 4.23: The Mean Value of Three Acoustic Cues for "important"

	Production	Acoustic pattern	Syllable 1	Syllable 2	Syllable 3	Standard Division
important	Chinese	F0 (Hz)	216	236.5	238.5	12.45
		Intensity (dB)	63	67.5	66	2.29
		Duration (ms)	60.5	105.5	59.5	26.27
	Thai	F0 (Hz)	177	258	196	42.36
		Intensity (dB)	63.5	67	69.5	3.01

		Duration (ms)	31	123	63.5	46.66

(Note: Std.D. = Standard deviation for each syllable mean rating)

Word 'important' – Chinese speakers placed the stress at syllable 2, obvious different in duration, the stressed syllable had 105.5 ms, while unstressed syllable only had 60.5 ms and 59.5 ms; Also, for Thais at syllable 2, the huge difference in F0 and vowel duration. For duration, the difference between stressed and unstressed syllables was as large as 100 ms (syllable 2: 123 ms; syllable 1 31 ms; syllable 3: 63.5 ms).

Table 4.24: The Mean Value of Three Acoustic Cues for "analyze"

	Production	Acoustic pattern	Syllable 1	Syllable 2	Syllable 3	Standard Division
analyze	Chinese	F0 (Hz)	224.5	212	181.5	22.12
		Intensity (dB)	63	64.5	63	0.87
		Duration (ms)	90	82.5	156.5	40.73
	Thai	F0 (Hz)	237	239	177.5	34.94
		Intensity (dB)	72	71	68.5	1.80
		Duration (ms)	107.5	67.5	99.5	21.17

(Note: Std.D. = Standard deviation for each syllable mean rating)

Word 'analyze' – The stress placement for Chinese speakers were at syllable 3, clear different in vowel duration between syllable 3 with 156.5 ms and syllable 2 about 82.5 ms; Thais were put it at syllable 1, however, the large difference showed in F0 more than duration.

Table 4.25: The Mean Value of Three Acoustic Cues for "citizen"

	Production	Acoustic pattern	Syllable 1	Syllable 2	Syllable 3	Standard Division
citizen	Chinese	F0 (Hz)	258.5	246	224.5	17.20
		Intensity (dB)	67	60	69	4.73
		Duration (ms)	82	45.5	103	29.10
	Thai	F0 (Hz)	211	234	178.5	27.89
		Intensity (dB)	66	62.5	67	2.36
		Duration (ms)	84	47	93.5	24.57

(Note: Std.D. = Standard deviation for each syllable mean rating)

Word 'citizen' – The placement of stress for Chinese speakers were at syllable 3, clear different in vowel duration; Thai speakers also place stress at syllable 3, and the huge different in F0 and vowel duration.

4.1.5 Summary of Production Results

The production tests showed that when target words were spoken in isolation in noun classes and verb classes, vowel duration significantly affected the stress of L2 English words spoken by students whose first language was Chinese or Thai. Unlike noun and verb classes, there was not much difference in vowel duration and F0 between stressed and unstressed syllables in adjective classes. In terms of target words spoken with sentences, regardless of the target word class, the vowel duration and F0 still affected the production performance of the speakers in both language groups. In general, the situation

was the same as when the target word was read separately, and the stress performance was different from that of different language groups.

Thai and Chinese production had the same performance of verbs and nouns in the isolation of target words, and the stressed syllables being distinguished from the unstressed syllables mainly depended on the vowel duration. Chinese speakers seem to rely more on vowel duration than Thai speakers when produced verbs. Thai speakers still relied mainly on vowel duration when produced adjectives, but Chinese relied more on F0 when produced adjective lexical stress. At the same time, we found that the intensity value was nearly the same in each syllable, whether viewed in terms of different language groups or three parts of speech.

4.2 Perception Results

The results of the perception experiment are presented in detail in this section. Firstly, Section 4.2.1 reports the results showing the identification accuracy of stressed syllables for both language groups. Then, the introduction of American English pronunciation as a reference for contrast first, and the specific low perception accuracy is explained through the utilization of three cues by language groups to perceive stressed syllables is shown more in Section 4.2.2. Finally, presents the difficulty self-evaluation of participants in language groups respectively.

4.2.1 Perception Accuracy

The term "Accuracy", as shown in Table 4.26 below, refers to the ratio of listeners who perceived the same (as correct) lexical stress as produced by the different language groups compared to the total number of listeners (n=15). Mentioned the same (correct) criterion was to compare the stress of syllabic words produced by the speaker. Whether the stressed syllables perceived by the participant were consistent, and if consistent, they were considered correct.

For ease of observation and subsequent difference assessment, the researcher had bolded the target words' results with poor perception accuracy for Thai perception and italicized the terms with low perception accuracy for Chinese perception. After the results were presented, they were extracted separately for comparative analysis.

Table 4.26 demonstrates that when compared to the data generated by two language groups, the researcher found that among the three different parts of speech, Thai perceived verbs with the highest accuracy of 83%, followed by nouns, about 71%, and the lowest one was adjective words with 57% accuracy. Thais properly detect the isolation target words with an overall listening accuracy rate of 70% (n = 20) and correctly perceived the target word embedded in the carrier sentence with an overall listening accuracy rate of 40% (n = 5).

Looking at the data comparison of Chinese, it seems that Chinese listeners were better at perceiving adjectives, with an accuracy of 72%, while verbs were only half of it, and less than half of the noun classes perception accuracy (43%) showed that Chinese seem to be not god at perceiving the noun words stress pronounced by Thais' speakers. Chinese recognized target words in isolation with a 60% (n=20) accuracy rate and when the part of a sentence with a 80% (n=5) accuracy rate.

Overall, the accuracy of lexical stress in the two stimulus sets (spoken in isolation and with sentences) produced by the Thai and Chinese speakers in perceiving each other's language groups was almost identical, and the accuracy was basically above 50%.

Table 4.26: Results of lexical stress perception evaluation of production by two language groups

	Thai perception			Chinese perception		
Target word	Produced syllable	Selected syllable	Accuracy (n)	Produced syllable	Selected syllable	Accuracy (n)
isolation						

Noun	japanese	3	3	80% (n=12)	3	3	87% (n=13)
	opponent	2	2	80% (n=12)	2	2	80% (n=12)
	ratio	**3**	**1**	7% (n=1)	*3*	*1*	*7% (n=1)*
	volunteer	3	3	67% (n=10)	3	3	73% (n=11)

Table 4.26: Results of lexical stress perception evaluation of production by two language groups, continued

	transistor	2	2	87% (n=13)	*1*	*2*	*33% (n=5)*
	basketball	**3**	**1**	**7% (n=1)**	1	1	87% (n=13)
	musician	2	2	60% (n=9)	*3*	*2*	*13% (n=3)*
Verb	modify	1	1	87% (n=13)	*3*	*1*	*13% (n=2)*
	allocate	2	2	67% (n=10)	*3*	*1*	*7% (n=1)*
	understand	3	3	67% (n=10)	3	3	87% (n=13)
	recommend	**3**	**1**	**33% (n=5)**	1	1	60% (n=9)
	referee	3	3	53% (n=8)	*3*	*1*	*20% (n=3)*

	abandon	2	2	80% (n=12)	2	2	80% (n=12)
Adjective	**positive**	**3**	**1**	**20% (n=3)**	1	1	60% (n=9)
	rational	1	1	73% (n=11)	1	1	80% (n=12)
	absolute	**3**	**1**	**13% (n=2)**	3	3	53% (n=8)

Table 4.26: Results of lexical stress perception evaluation of production by two language groups, continued

	consistent	2	2	87% (n=13)	*2*	*1*	*33% (n=5)*
	fantastic	**3**	**2**	**7% (n=1)**	2	2	93% (n=14)
	delicate	1	1	73% (n=11)	*3*	*1*	*13% (n=2)*
	cultural	1	1	60% (n=9)	1	1	60% (n=9)
With sentence							
	decided	2	2	53% (n=8)	2	2	60% (n=9)
	example	**2**	**3**	**27% (n=4)**	2	2	73% (n=11)
	important	**1**	**3**	**20% (n=3)**	2	2	80% (n=12)

	analyze	3	3	67% (n=10)	1	1	*67%* (n=10)
	citizen	3	1	40% (n=6)	3	1	20% (n=3)

(Note: Accuracy = Proportion of same identification.)

4.2.2 Low Perception Accuracy Analysis

Knowing the perception rates of distinct Thai and Chinese language groups, the researchers focused on target words created by cross groups with low perceived rates. In order to more clearly prove the cause of target words with low perception rate (Accuracy $\leqslant$ 50%), the author introduced American English as a comparative reference. Chinese and Thai participants learn the pronunciation of American English from textbooks at an early age (Hinkel, 1999; Pollap, 2010; Rawlings, 2013) (refers to Section 2.1 World Englishes), their exposure to English variants was arguably limited, many English tutors may only be familiar with a single variant of English. That is why the study involved different variants of English between Thailand and Mainland China and had to mention the situations of "low perception accuracy" when participants perceive lexical stress.

The researcher also compared three significant acoustic parameters in terms of average F0, vowel duration, and average intensity in target words with low perception. By compared the mean value difference and standard deviation of the three cues perceived by the listeners of the stressed syllables and the unstressed syllables, their perceptual dependence can be seen.

4.2.2.1 Thai listeners

One particular finding appeared to be that the low perception accuracy situation happened more obviously when the speech production was different from the textbook stress patterns, which also indicated AE. The stress production of 7 of those terms: "ratio," "basketball," "positive," "absolute," "fantastic," "important," and "citizen" are different

from AE stress placement. Among them, excepted for the target word "important," the perceived stress syllables were different from the actual production and AE stress placement; the other six words with low perception accuracy were consistent with AE. This was especially obvious when the target word was produced in isolation.

On the one hand, it was observed that Thai listeners exerted more effort in accurately perceiving the target words when the third syllable was stressed. They always perceived that the word stress was at the initial syllable (refers to Table 4.26), which matched with AE placement. Therefore, it can be guessed that it was due to the influence of familiarity, as listeners had been formally exposed to AE since elementary school. On the other hand, referred to sections 4.1.1—4.1.4, where three important acoustic parameters also provided the basis. With the exception of the target word "absolute," the F0 value of the stressed syllables perceived by Thai listeners was higher than the other two syllables, in other words, Thai listeners relied heavily on the F0 when perceiving English L2 stress from the acoustic information of target words with low listening resolution.

And seeing that the other two stress production words match AE stress placement words, the first syllable of the word "recommend" is considered stressed. With Chinese pronunciation and American English stress appeared on the tail syllable, while 60% of Thais perceived the stress on the first syllable, and only 33% of Thais perceived it correctly. Interestingly, from the acoustic information of the produced word, the intensity of the first syllable (65 dB) was higher than others, more than 3 dB (refer to Table. 4.11), and it was guessed that the possible influencing factor was intensity. In another particular word, "example," the third syllable was considered a stressed syllable. Table 4.22 showed that the values of F0 were higher than the other two syllables. The values of the Std.D between syllables were 33.95. So, it can also be proved that the acoustic features of F0 play an important role in the perception of Thai listeners.

Table 4.27: Comparison of Low Perception by Thai participants

Target Words		Produced syllable	Selected syllable	AE	Accuracy
isolation	ratio	3	1	1	6%
	basketball	3	1	1	6%
	recommend	3	1	3	33%
	positive	3	1	1	20%
	absolute	3	1	1	3%
	fantastic	3	2	2	6%
with sentence	example	2	3	2	27%
	important	1	3	2	20%
	citizen	3	1	1	40%

4.2.2.2 Chinese listeners

The fact that low perception accuracy occur more obviously when speech production was different from AE stress patterns can still be demonstrated in Chinese listeners' perception results. Among the 10 target words with low perception accuracy, 9 target words had different stress patterns from AE stress placement, and they were "ratio," "transistor," "modify," "allocate," "referee," "delicate," "musician," "analyze," and "citizen". The stressed syllables perceived by Chinese listeners for all of these 9 target words were consistent with AE stress placement, no matter the target word appeared in isolation or with a sentence. The researchers observed that with the exception of the word "musician," which matched AE stress placement and stress appeared in syllable 2, the stress perception of other words matched AE, and the stress appeared in the initial syllable.

There were exceptions in the target word "consistent." Chinese listeners perceived the lexical stress in the initial syllable, but the stress of AE and the actual Thai production were shown on the second syllable. Table 4.17 showed the stressed syllables perceived by Chinese listeners, with the value of F0 higher than that of the other two syllables, and

although the difference in the F0 values between syllables was not large, it can also be guessed that Chinese students mainly relied on F0 to perceive English lexical stress.

Another interesting phenomenon was that regardless of whether the actual production matched the AE stress placement, the value of F0 in the stress syllable perceived by Chinese was higher than that of other syllables, indicated that they relied heavily on F0 to perceived English lexical stress. That also implied that even when F0 was not employed and vowel duration was favored in the production of stress (as is the case with Chinese and Thai stimuli) (see Table 4.1—4.25), Chinese listeners nonetheless used pitch to assessed stress sites.

Meanwhile, by compared the mean value difference and standard deviation of the other two cues (refer to 4.1.1-4.1.4) perceived by the Chinese listeners of the stressed syllables and the unstressed syllables, their perceptual dependence can be seen. All three acoustic parameters depended, but in addition to F0, followed by vowel duration. Compared to the Chinese group, the Std.D. vowel duration values for the Thai perception group stand out. It demonstrated how Thai listeners perceived word stress concerned vowel duration. Target words with low Chinese perception rates had stressed syllables comparable to unstressed syllables, and their Std.D. values were strongly lower than those of the Thai group, demonstrated that vowel duration had little effect on Chinese perception. Additionally, the mean intensity values for both groups were slightly higher in perceived stressed syllables than in unstressed syllables, as seen in the low-perception words. In contrast, the average values of the Chinese population were all somewhat higher than those of the Thai group. It may also be stated that the linguistic background was mostly unaffected.

Table 4.28: Comparison of Low Perception by Chinese participants

Target Words	Produced syllable	Selected syllable	AE	Accuracy

isolation	ratio	3	1	1	6%
	transistor	1	2	2	33%
	modify	3	1	1	13%
	allocate	3	1	1	7%
	referee	3	1	1	20%
	consistent	2	1	2	33%
	delicate	3	1	1	13%
	musician	3	2	2	13%
with sentence	analyze	2	1	1	13%
	citizen	1	1	1	20%

4.2.2.3 Self Evaluation

The self-evaluation from listeners at the end of the perceptual test was also an essential part of the low perceptual accuracy analysis. Table 4.29 illustrated the reasons behind the difficulties that the listeners faced while completing the perception tasks. Since it was an open-ended Q&A, there may be no limit to one reason for the expression.

Nearly half of the listeners in both language groups agreed that the main cause of the difficulty was the suprasegmental and segmental differences. Meanwhile, 7% (n = 1) of listeners expressed in both language groups stated the difficulties were caused by the unclear audio, although it should be noted that the quality of recordings was checked for clarity before being sent to the listeners.

The Chinese listeners explained that the segmental difference was reflected in changes in schwa, such as /pɒzətɪv/-/pɒzɪtɪv/; or /bɑːskɪtbɔːl/ - /bɑːskətbɔːl/; or difficulty for Chinese to distinguish between Thai pronounced consonant of /ɾ/ and /v/, resulting in perceptual difficulties. Meanwhile, 7% of Chinese listeners admitted that they had poor proficiency in the language, which was why they had difficulty making out what was being said by the Thai speakers, while 7% of Chinese believe they were unfamiliar with

Thai accents. One of the Chinese listeners did not state the reasons for their troubles in finished the task since she really could not analyze the specific reason.

In Thai groups, a total of 33% of the listeners thought that it was difficult for them to recognize the words due to the difficult accent used in the recording. In addition, 13% of Thais agreed that the accent was unfamiliar to them. 7% (n = 1) of the Thai listeners reported confusion about the target word, i.e. homonyms.

Table 4.29: Reasons for the Difficulties in Completing the Perception Task

Groups	Reasons	%
Thai	Word confusion	7%
	Segmental difference	40%
	Suprasegmental difference	40%
	Unfamiliar words	7%
	Difficult accent	33%
	Unfamiliar accent	13%
	Unclear audio	7%
Chinese	Segmental difference	47%
	Suprasegmental difference	47%
	Unfamiliar accent	7%
	Unclear audio	7%
	Poor proficiency	7%
	Difficult accent	13%
	Not stated	7%

4.2.3 Summary of Perception Results

In sum, perception experiments are observed in terms of accuracy. In terms of the three parts of speech, the perceived accuracy of the three classes was completely different, also according to the language groups.

The Thai listeners perceived that the verbs had the highest accuracy, followed by nouns, while adjectives were found to have the lowest accuracy. In terms of Chinese listeners, they were good at perceived the lexical stress of adjectives; the verbs and adjectives were about 50% correct and perform slightly better on verbs than on nouns. When the target word appear in a sentence, the accuracy of Chinese (80%) is significantly higher than that of Thai (40%). It proved that Chinese were better at perceived lexical stress when the target word was embedded in a sentence, while Thai listeners were better at perceived stress when the target word appeared independently.

When extracting words with a perception accuracy of less than 50%, the researcher found that both the lexical stress familiarity and L1 input had a large impact on the accuracy. When speech production differed from the AE stress patterns, also means low stress familiarities, the poor perception accuracy problem became more apparent. At the same time, due to the influence of the tonal language structure of Thai and Mandarin Chinese, F0 was the acoustic parameter that the two language groups relied most on to distinguish the stress of L2 words. However, when comparing the two language groups, Chinese listeners seem to be relied more on F0 to discern L2 stress than Thai listeners. Lastly, intensity is an acoustic feature that played almost no role.

4.3 Summary of Findings

The findings of the two main experiments are summarized below. The production tests showed that of the three main acoustic features, vowel duration had the most significant effect on the stress of L2 English words spoken by students whose first language was Chinese or Thai. This was followed by F0, and the intensity value was nearly the same in each syllable. At the same time, the researchers classified the position of the stress in the syllables. They discovered that when the stress fell on the first and second syllables, the values of the three acoustic features tended to be high. In contrast, when the stress was placed on the final syllable, only vowel duration exhibited a significant value.

On the other hand, the results of the perception experiment showed that the perception accuracy of Chinese and Thai students was almost the same. However, Thai students performed well in the case of target word independence (about 70%), and Chinese students had a slightly lower score (about 60%). When the target word appeared along with the sentence, the perception rate of Thai students was much lower than that of Chinese students. (Accuracy: Chinese 80%, Thai 40%). When both language groups' speech production deviated from the AE stress patterns, the poor perception accuracy problem became more evident, which may be related to the familiarity effect. In addition, F0 had a substantial impact on the L2 lexical stress perceptions of Chinese and Thai students, followed by vowel duration and intensity. The detailed discussion of the implications arising from the specific outcomes of each experiment presented in the subsequent chapter.

CHAPTER 5: GENERAL DISCUSSION AND CONCLUSIONS

The analyses in Chapter 5 have provided answers to the two main research questions raised in the Introduction:

- What are the differences in the acoustic features of English lexical stress produced by Thai and Chinese L1 speakers?
- To what extent does the L1 influence the English lexical stress perception of Thai and Chinese speakers?

The production and perception of lexical stress are discussed independently in the sections that follow. After that, all of the findings will be combined with consideration of existing lexical stress models. The last part concludes by outlining the current study's limitations and making recommendations for further investigation.

5.1 Stress Production

First of all, it is assumed that due to the influence of the tonal language rules, the acoustic signals relied on by Thai and Chinese pronunciation to produce lexical stress did not show a big difference; Although they are both tonal languages, they have their own tonal systems.

5.1.1 Difference of Intensity

Overall, the acoustic studies revealed that both Mandarin Chinese and Thai students used the acoustic features of fundamental frequency (F0), intensity, and length in a comparable fashion to generate English lexical stress. In comparison to unstressed syllables, both groups generated stressed syllables with increased F0, longer duration, and greater intensity. This explanation aligns with the overall observations that L2 English cross-language speakers stressed syllables (Ou, 2007; George, 2019; Adams & Munro, 1978).

The overall difference in intensity between Mandarin speakers and Thai speakers was no higher than 5 dB; it is improbable that this discrepancy had a substantial impact

on the perception of their non-naiveness in producing the English stress contrast. With the results of intensity, there was no difference of the language background for both the stressed vowels and unstressed vowels. The mean intensity of stressed vowels for Chinese speakers was 65 dB, and 69 dB for Thai speakers. The value of unstressed vowels for Chinese speakers was 63 dB, and 67 dB for Thai speakers. This result indicated that Chinese and Thai speakers depend less on intensity when producing English lexical stress. Generally, stressed vowels were typically greater in intensity than unstressed vowels.

In a similar vein, research suggests that duration and pitch differences are more efficient in indicating stress than intensity differences (Mattys, 2000; Morton & Jassem, 1965; Zhang, 2005). There seems to be no great disagreement about this acoustic information, either in previous studies or in the current study.

5.1.2 Difference of F0

According to previous studies, F0 was predicted to significantly affect cue in accents of L2 English words spoken by Chinese speakers (Chao, 1972; Altmann, 2002); Mandarin speakers produced English syllables with significantly higher pitch range than Thai speakers, no matter whether stressed or not or which part of speech. This kind of findings align with Chen et al. (2001) and Zhang et al.'s (2008) findings and conclusions regarding the average F0 performance of Mandarin Chinese speakers. Both of these researchers have contended that the reason for this phenomenon is due to the tendency of Chinese speakers to use a larger frequency range. In contrast to English stress, Mandarin high tones are generated with an F0 at an even greater proportion of the speaker's whole pitch range (Shen, 1989; Adams & Munro, 1978). The L1 tone domain, which Mandarin speakers may use to manipulate the F0 feature, is still having a significant impact. Looking at another possible interpretation concerns F0 as a strongly affected stressed syllable cue in Chinese L2 English speakers. Chomsky &

Halle (1968) stated that the principles for placing words in the major stress rely on the word type and the syllable weight. In this study, there were three-word categories (noun, verb, adjective) and different situations of target word placement (in isolation and with sentence).

However, Chinese speakers not only achieved a higher average F0 in stressed syllables but also excessively elongated these syllables beyond necessity. This suggests that Chinese speakers incorporated the characteristics of the Mandarin high tone into the stressed syllable in English, including its association with an extended duration of the syllable, rather than simply producing a generally elevated average F0. These findings indicate that Mandarin speakers, although being aware of which syllable to stress, may rely on their tonal patterns when producing stress.

In addition, F0 was also one of the most differentiated parameters in Thai L2 stress production; it was the primary cue used contrastively in Thai lexical tones (Nguyễn et al., 2008). The significant finding occurred when the target word was introduced within a sentence. The situation appeared in the word "example"; the difference between the values of F0 was larger than the vowel duration (seen in the mean values in Table 4.22) that appeared in both Thai and Chinese production. Meanwhile, the results showed that when the Thai speakers pronounced "analyze" and "citizen," the standard deviation between the stressed and unstressed syllables of F0 and vowel duration was tiny, even F0 was slightly higher than the duration. The researcher believed that this was the effect of the target words in the sentence, which made the results seem to be different from previous studies. This explanation suggests that the implementation stress may be influenced more by the word type rather than language background.

5.1.3 Difference of Vowel Duration

From the production results, we can see that the largest acoustic cue that Thai and Chinese speakers used to produce stressed and unstressed syllables in the noun and verb classes is vowel duration rather than F0. Only except for the verb 'modify'.

Vowel duration is the most relied upon acoustic signal by Thai speaker (Khamkhien, 2010; Jangjamras, 2011). With respect to the results of mean duration, the present results are consistent with the results of Jangjamras (2011), vowel duration is the most relied upon acoustic signal by Thai speaker (Khamkhien, 2010; Jangjamras, 2011). Corroborative results can also be found in Satravaha's (2002) and Nguyễn et al. (2008)'s study. Both of their results pointed out that Thai speakers preferred vowel duration when highlighting stressed English syllables. They have attributed such conclusions to the effect of language background's syllable structure.

It is also assumed that Thai speakers produce more target words with final stress than initial stress since Thai was reported to have a fixed pattern of final stress (Jaiprasong &Pongpairoj, 2020; Albadar, 2021). If we solely consider the quantity, it appears that Thai speakers tended to exhibit a higher occurrence of final stress compared to initial and second position stress, thereby could supporting the hypothesis. There are 25 target words in total; 11 words pronounced by Thai speakers have the stress on the final syllable, 7 are located in the initial syllable of the word, and 7 are located in the middle syllable. It seems to support the well-known fact that Thai-speaking English strongly tends to stress word-finally. However, it is more challenging to account for the fact that due to Thai speakers having a set pattern of final stress, they have to produce more target words with final stress than with initial stress. One explanation could be that due to the small number of speakers who completed the production experiment in this study, it lacks a certain validity, thus it is not sufficient to prove the accuracy of the hypothesis that Thai speakers produce more final stress than the initial or second position stress.

Interestingly, the vowel duration difference of the three syllables of different words produced by the Chinese group was more significant than that of the Thai group. This was probably a consequence of vowel duration being the acoustic signal that has the greatest impact on Chinese L2 English speakers, not F0, at least in noun and verb target words. This seems to be the original hypothesis that does not hold water. However, when we looked at the target words of the adjectives, we found that the acoustic features of 57% (n=7) of the adjectives showed that the F0 difference between stressed and unstressed syllables produced by Chinese speakers appeared to be slightly greater than the difference in vowel duration. Therefore, it is concluded that when the Chinese produced adjectives, they relied more on the acoustic cue of F0 than vowel duration. It appears that Chinese speakers encounter difficulties in strictly controlling F0, possibly due to interference caused by the part of speech.

5.2 Stress Perception

The analysis and discussion of the stress responses observed in Chapter 4 are presented in this chapter. In the previous chapter, intriguing findings were uncovered regarding how listeners' perception of stress is influenced by their L1s. In general, Thai and Mandarin Chinese identified the lexical stress location of trisyllabic as having a relatively high accuracy. However, speakers from different expanding circle countries pronounced the materials. This outcome in the task of identifying stress is comparable to the results obtained by Altmann (2006), who observed that individuals speaking languages such as Chinese, Thai, and Korean, which lack lexical-level stress patterns, achieved flawless scores when it came to identifying stress in English words. That is, Thai listeners identified stress in the two stimulus sets (isolation and with sentences) almost as well as Chinese participants.

5.2.1 Effect of L1 on Part of Speech

Firstly, the lexical class has been proven to considerably affect stress placement patterns on real words and independently across Thai and Mandarin-Chinese English listeners. That is, these tonal language listeners, like late Spanish-English bilinguals (Guion, 2005), possess stress placement depending on lexical classes. This view was also consistent with previous results by Kelly & Block (1988) and the fact that late learners, such as Arcuili & Cupples (2003), displayed awareness of English stress patterns based on lexical classes.

In this study, Mandarin Chinese listeners of English obviously had a lower mean of correct response for noun words (43% accuracy) in their perception, whether those words were produced in isolation or with carrier sentences. In the adjective class (72% accuracy), Mandarin Chinese listeners performed much better than in the noun or verb class (50%). The importance of lexical classes for stress perception can be seen from the disparity in perception.

Different parts of speech also had an impact on Thai English language learners, as shown in Table 4.26. On the contrary, they obviously had a lower mean of correct response for adjective class (57% accuracy) in their recognition of whether those words were produced in which situation. However, as for the noun and verb classes, their performance seemed to be relatively good. Regarding perceived verbs, the perception accuracy reached 83%, and the noun category had a correct accuracy of 71%. Compared with Chinese listeners, their perception of the consistency of these two speech parts can be considered excellent.

5.2.2 Effect of L1 on Employment of Lexical Familiarity

In addition to the impact of part of speech on Chinese and Thai listeners' perceptions of English word stress, the researcher found that form-extracted words with low

perception accuracy from the previous chapter (see Tables 4.27 and Table 4.28) and the listeners' familiarity with lexical words also affect the perception of stress position.

Except for the word "fantastic," where the perceived stress was on the second syllable, all of the words with low perception in Thai listener groups had the stress in the initial syllable. Interestingly, when compared to American English (AE), we found that the position of the stressed syllables perceived by the Thai listeners was precisely the same as that of the AE words. However, it was different from the stressed position of the production position, and this was the case in all target words with low perception by Thai listeners, including the word 'fantastic,' where the stress is perceived to be in the second syllable. However, the findings of this study contradict previous research suggesting that native Thai speakers have a better perception of final stress compared to initial stress. This is considering that Thai language has fixed final stress in English words (Luksaneeyanawin, 1983). Additionally, Altmann's (2006) research supports the preference for initial stress among Thai listeners as it took them approximately 200 ms longer to judge final stress compared to initial stress. It is possible that the presence of fixed final stress in Thai language reduces sensitivity to perceptual cues for stress in this position rather than enhancing sensitivity towards final stress. This is due to the fact that neither its location nor its acoustic correlates are recorded in the NT long-term representation. Meanwhile, Thai individuals have encountered greater difficulty with final stress, which may imply that L1 phrase-level stress patterns do not always promote the detection of L2 word stress. Similarly, it may be possible to suggest further validation of Altmann's research, which is the familiarity effect mentioned in this section. Since the textbooks that Thai students were exposed to from an early age and the critical teaching standards guided by teachers in schools were American English, it can be said that Thai students' fundamental English learning was based on American English (Datesman, 1997; Hinkel, 1999; Pollap, 2010).

The same familiarity effect seems to have occurred in the listening results of Mandarin Chinese listeners (Rawlings, 2013). As shown in Table 4.28, when all the words with low perception accuracy for Chinese listeners were compared with the AE word stress, the position of word stress was exactly the same. This is probably a consequence of the familiarity effect of word stress in AE. After all, the English knowledge that the Mandarin Chinese have been exposed to since childhood is also based on AE.

Lexical familiarity plays a pivotal role in the perception of English lexical stress. The familiarity effect leads to increased effort in accurately perceiving the stress patterns in words produced by Chinese speakers, who are not accustomed to this particular variation of English. Consequently, achieving high accuracy in their perception of lexical stress becomes challenging for them.

5.2.3 Effect of L1 on Difference in Acoustic Properties

Nonetheless, the acoustic properties of the test words seem to present another possible interpretation for stress identification by both groups. It has been previously mentioned that the perception of L2 stress is heightened among listeners when it is accompanied by informative acoustic cues in their native languages (Ingvalson et al., 2011). According to what was described in the section on the literature review, previous studies on the recognition of stress in L2 have yielded inconclusive results regarding the ability of individuals who speak a non-stress variant of their native language, which is also a tonal language, to perceive stress in L2.

5.2.3.1 L1 Effects on F0

The researcher's detailed analysis of the acoustic properties of perception stimuli revealed that there was a consistent and significant difference in the average F0 (Hz) between stressed and unstressed syllables across various syllabic structures. The correlation results on acoustic ratio values shed light on how listeners perceive stressed syllables. It is possible that the dissimilar stress identification observed among Thai and

Mandarin Chinese speakers compared to their production could be attributed to the average F0 differences. Based on noticeable differences, it can be speculated that average F0 plays a crucial role in stress perception for both language groups, surpassing duration and intensity differences in terms of effect size. This suggests that both intonation language groups might subconsciously exhibit sensitivity towards prominent variations in F0 at stressed positions, as F0 serves as the primary acoustic indicator of English stress system and intonation patterns. The studies conducted by Adams & Munro (1978), Chen et al. (2001a), and Jangjamras (2011) on acoustic characteristics have been compared with the conclusion, leading to the determination that they are essentially identical. According to other research (Beckman, 1986), English listeners also used F0 as the primary signal to emphasize perception in a stressed position.

5.2.3.2 L1 Effects on Vowel Duration

It is also noteworthy that, despite a greater variance in vowel duration, this characteristic has no impact on how accurately stress is recognized in Thai and Mandarin Chinese. This discovery suggests that, despite vowel duration's greater degree of saliency, the significance of duration in stress perception is relatively lower compared to F0. This could potentially be attributed to the elongation effect observed in Thai speakers' production of final vowels. In simpler terms, as a longer duration is expected for a final vowel compared to an initial vowel, the role of differences in vowel duration as an acoustic indicator of stress becomes diminished. Therefore, Thai listeners had different results when they perceived the stress at the end of words. In like manner, the result from Jangjamras (2011) suggests that despite its higher degree of saliency, duration is relatively less important than F0 in final stress perception. This is due, perhaps, to the final vowel lengthening effect. On the other hand, Yu & Andruski (2010) support my speculation that average F0 contrast would be the primary cue in perception of Chinese participants, and duration is relatively less important than F0 in stress perception.

5.2.3.3 L1 Effects on Intensity

In addition, Turk & Sawusch (1996) proposed that time and intensity were evaluated in a holistic manner and that listeners were unable to utilize these factors alone when evaluating stress judgments. The perceived results of the Chinese listener's group also showed striking agreement with the previously discussed Wang's (2008) study, where duration and intensity manipulation could only predict stress judgments in a small number of Chinese participants, but F0 could more easily predict stress judgments in each Chinese participant. Based on these findings from intensity and vowel duration, I propose that the weight of these two perceptual cues is parallel in their perception. Similar to the investigation of Wang (2008). He points out that when pitch is not present, individuals who speak tone languages may have challenges identifying L2 stress. Also, as shown in the literature review (2.6.3), contrary to the findings of Lai (2004), she suggested that L2 learners were not sensitive to changes in the maximum F0 range because of their perceptual weight in Mandarin L1.

In contrast, Jangjamras (2011) pointed out that the relationship between the acoustic data of the perception stimuli and stress identification accuracy was investigated through a stepwise regression analysis. The results revealed that intensity (15%) was one of the strongest predictors of native Thai participants final stress perception scores, followed by increased vowel duration (5%) and average F0 (4%), respectively.

5.3 The Current Findings Based on Stress Learning Model

The Speech Learning Model (SLM), created by Flege (1995; 2005), is a contemporary model that explains the process of acquiring L2 speech. The main objective of the SLM is to explain the variability in individuals' ability to properly acquire or not acquire phonetic components in an L2. Given that stress placement may carry varying degrees of importance in differentiating lexical categories across Thai and Chinese English learners, the theory of SLM might potentially provide explanations for some observed occurrences

in the trials. Furthermore, this work aims to investigate the applicability of SLM in the context of English word stress learning. It is worth noting that prior research on SLM has mostly concentrated on the acquisition of segments.

The duration facts do support that Thai speakers' use of vowel duration is influenced by their L1 phonology, given that 56% (n=14) of all target words produced by Thai speakers place the stress on the third syllable, similarly to the long vowel tendency normally at the end of the words in their L1 language (Wayland, 2006). However, it is a more common phenomenon observed when the target word appears in isolation situations. It is possible that this final lengthening is instead just a prosodic boundary effect. Since the target stimuli were placed in the sentence final position, Thai participants automatically made references to the default position of phrasal stress in Thai, phrase final stress.

For Mandarin Chinese, there isn't a clear comparison between their L2 English performance and their L1 performance. The vowel duration measurements also do not clearly indicate anything about tone, though Chinese speakers did the same as Thai speakers on vowel duration. Meanwhile, if we take the features of neutral tone as indicative of a toneless unstressed syllable, there is an indication that unstressed English syllables are being treated similarly. In the initial syllables, we see that they are somewhat comparable to neutral tones, and although no comparison was possible between English unstressed syllables and neutral tones in the first syllable, that contour does appear to have the kind of mid-contour one might expect for a neutral tone syllable in that position. But, in this study, there is no clear evidence for the influence of Mandarin tones except for the influence of F0. Only this acoustic feature can support the L1 tonal influence explanation of SLM. To some extent, this is consistent with Altmann's (2006) findings, stress patterns exhibited by speakers of L1 non-stress languages such as Chinese, Japanese, and Thai do not exhibit a common grouping.

5.4 The Current Findings Based on Speech Typology Model

According to Altmann and Vogel (2006), the Stress Typology Model (STM) is the model that accurately predicts the success rate of both perception and production by English L2 learners. The facts presented here seem to both support and dispute the theory of STM. In the stress identification test, Thai listeners showed competence (accuracy rate of stressed syllable perception), the same as Chinese listeners in the stress identification task. According to what is expressed in the literature review section of this study, Thai and Chinese were classified as tone languages; the prediction made by STM that speakers of L1 non-stress language (-stress, -predictable) would do very well in recognizing stress tests proved accurate.

As an alternative, Thai may also be considered a language that falls under the category of fixed stress language, which is characterized by high levels of stress and predictability (+stress, +predictable). The prediction made by the STM that speakers who receive stress that is predictable would have trouble detecting the location of stress does not seem to be supported by the data. Thai listeners in this research were distinct from the participants in Altmann (2006) who were exposed to predicted stress in the sense that they did not have any obvious perceptual difficulties while attempting to determine the location of the stress. Providing a detailed description of the stress category is one approach that may be used to determine typological. It may be necessary for STM to indicate that suprasegmental characteristics (such as tone and pitch accent) with phonemic status should be ranked higher than features with phonetic status, such as phrasal stress in Thai (Altmann, 2006). This is in addition to the categorization of languages based on the surface stress patterns that may be seen. Strictly speaking, the contrastive tone system in Thai argues that Thai should be classified as a non-stress language rather than a stress language. This is despite the fact that Thai and other languages have the same fixed final tension stress.

On the other hand, the results have shown that Chinese listeners had a slightly lower perception accuracy when it came to perceiving English stress in trisyllabic words in isolation situations compared to Thais. However, the accuracy in carrier sentences is higher than that of Thai listeners. But in fact, the difference between the two language groups is actually not large on the whole. Therefore, it can be argued that their proficiency in using lexical tones in their native language also contributes to their advantage. These findings align with previous research indicating that exposure to tonal languages enhances listeners' ability to identify lexical stress in L2.

Another underlying concern is that the STM did not provide a comprehensive explanation for why participants of languages that do not contain stress performed so well in the stress detection task. The STM just make the assumption that the positive settings in L1 stress were likely to induce interference that was seen. In this assumption, the stress location in L1 is the only thing that is taken into consideration; sensitivity to acoustic characteristics to stress perception is not take into account. The current study indicates that the stress position of the L1 (final in standard Thai) did not have a significant impact on the stress perception. This is due to the fact that both groups of speakers had a well preference for initial stress. It is quite probable that the sensitivity to acoustic elements of stress is, in fact, a significant component that contributes to the performance of Thai speakers in identifying stress. It is also speculated that the low average F0 value of the final stress may be one of the reasons why the Thai listener did not successfully perceive the final stress. It is thus possible that variations in degrees of sensitivity to acoustic aspects of L2 stress, rather than only surface stress patterns, might give an improved rationale for the superior stress identification that Thai and Chinese listeners are able to achieve. In other words, the utilization of suprasegmental dependence in L1 may explain a portion of the effectiveness in English L2 stress identification. And suprasegmental signals are frequently used by Thai and Chinese students.

Although the Speech Learning Model (SLM) is the first thing that comes to mind when it comes to stress production, it seems that Altmann & Vogel (2006) also proposed the STM prediction section on stress production. It is predicted by STM that speakers who lack stress in their own language would create stress placement methods that are not target-like. It is important to note that this prediction does not hold true for the current discovery since the capacity to perceive stress was evaluated while the location of stress was controlled. Based on the findings of the production, it seems that speakers of Thai and Chinese are able to recognize the intended stress placement of the other language groups. These production findings demonstrate that stress might be articulated by speakers of non-stressing languages, including Mandarin Chinese, and Thai.

5.5 Conclusion

When Thai and Chinese L1 speakers produced the acoustic features of English lexical stress, there were differences depending on the part of speech and whether the target word had context. After visual analysis of the data, it was found that vowel duration is the most relied upon acoustic signal by Thai speakers, which is consistent with previous studies. Chinese speakers also mainly use vowel duration to differentiate stressed syllables, followed by F0, while the intensity was nearly the same in each syllable, with no effect on differentiating stressed to unstressed syllables. Above are the main findings of the first research question in this article.

In this dissertation, the second research inquiry aimed to investigate if there exists a correlation between the listening abilities of different tonal L1 groups and the L1 of the speakers who created the stimuli. Data from perception experiments indicate that the average F0 contrast is the main indicator, whereas duration and intensity play a relatively minor role compared to F0 in stress perception for both language groups. Put differently, tone languages speakers transmit the L1 tonal information, i.e. the contrastive phonological use of tone promotes stress perception, and as a result, they depend on F0

as a meaningful phonetic acoustic feature in their sense of stress. This kind of finding seems to corroborate the previous scholar's data (for example, Kijak, 2009; Andruski, 2010). The results of this research hold significance because they provide further proof that the impressive abilities of tonal students were not simply attributed to their dependence on the pitch for recognizing English L2 lexical stress. Additionally, word stress placements in all the varieties had some similarities with AE. Therefore, they argue that more than just L1 input contributes to variations in L2 English word stress. The familiarity of the speech production also plays a role. Perception accuracy decreases noticeably when there is a mismatch between speech production and the stress patterns of AE. This phenomenon could be seen in the perception findings of both language groups. The results of the study demonstrate that Chinese listeners consistently aligned with Thai listeners in accurately identifying word stress. This indicates that even when speakers of different languages produced the stimuli, the tonal language listeners performed well in perceiving stress at the phonetic level. Specifically, the study examined stress perception and found that despite significant differences in the English word stress produced by Thai English L2 speakers and Chinese English L2 speakers, the listeners were still able to successfully complete the perceptual experiment and accurately determine the intended position of the word stress. In instances where F0 did not play a role in indicating stress (as the primary cue was vowel duration), the majority of Chinese listeners demonstrated accurate identification of stress placement.

5.6 Limitations of the Current Study

This research examined two aspects of stress acquisition, specifically stress perception and production. Given the emphasis on acoustic elements within the suprasegmental domain, it is recommended to conduct additional comparative investigations pertaining to segmental data. After the researcher visualized the data in Praat, they found that there was also a difference attention in use of segmental cues, that is, vowel quality. Scholars

like Boram (2021)'s research, a notable disparity was observed in the F1 and F2 of Mandarin L2 speakers compared to those from inner-circle nations with regards to their approach towards cue weighting for lexical stress. Research on lexical stress indicates that focusing just on suprasegmental signals for lexical stress does not adequately describe how lexical stress is realized in L2 speech. This is supported by several studies (Connell et al., 2018).

Secondly, only 4 speakers (two Thai and two Chinese) were adopted for comparison in the production stage, and the post-experimental research analysis could only perform simple descriptive analysis and could not intuitively see significant differences. As we all know, parametric tests such as t-tests or p value require a large sample size (generally greater than 10) (Neslihan, 2022). Although the conclusions of these statistical analyses are not necessarily the best and most accurate, they can also be used as a reference for data presentation.

Thirdly, the chosen English wordlist was not strictly controlled in terms of its segments and syllabic structures. The focus was solely on the frequency of this target word list in L2 studies to ensure it is as common as possible. However, it should be noted that the wordlist utilized in this study exhibited two distinct syllabic structures in stressed and unstressed positions, along with varying vowels. Consequently, there might have been some variability (such as segmental errors) introduced during the detailed acoustic analysis. If all unstressed syllables had similar vowels and syllabic structures, comparing raw values across the acoustic parameters of stressed syllables would lead to more statistically significant findings.

Eventually, the participants who had adopted from Mainland China had different L1(Chinese dialect & pǔtōnghuà). Although, *pǔtōnghuà* is the official language of instruction in educational institutions. The intention behind the propagation of pǔtōnghuà was not to eradicate Chinese dialects in an unnatural manner but rather to gradually

diminish the extent of their use. In the 1990s, linguists began reevaluating the cohabitation of pǔtōnghuà and Chinese dialect, seeing them not as mutually incompatible but rather as complementing (Guo, 2004; Li, 2006). According to phonetic research on several regional variants of contemporary Standard Chinese (SC), conducted by Li et al. (2006) and Deng et al. (2006), bilingual speakers of the official standard Chinese (pǔtōnghuà) sometimes include elements from their local dialects while speaking. However, in the field of speech production and perception, the majority of research on Mandarin Chinese use the word "Mandarin" in a general sense, without explicitly stating which specific geographical variant of Mandarin was investigated. On the other hand, due to the diverse linguistic environments, most Mandarin Chinese populations speak more than one Chinese language as their L1. The heterogeneity of L1 in Mandarin Chinese may be an unstable variable in perception experiments, but the effect is almost non-existent based on the above results.

5.7 Future Direction of lexical stress' research

This research is a thorough investigation that examines the perception and production of stress among speakers of tonal languages. It explores the influence of the speakers' native tonal background and applies it to stress perception and production, focusing on three fundamental acoustic characteristics. It has enhanced our comprehension of how prosodic characteristics influence the production and perception of lexical stress in English. Although lexical stress acting a crucial role in speech perception, there is a need for more research that investigates the impact of stress across different tonal languages. Therefore, the purpose of this dissertation's experimental research was to investigate how students from native languages with various topological characteristics, specifically in terms of word-level stress use in their first language, produce and perceive language. In addition, the data used in the present study included spontaneously generated words from various speakers of diverse first languages, in contrast to prior studies that employed

stimuli generated by speakers from inner circle nations (such as Americans or British) or acoustically manipulated speech. It is advisable for future studies to do a comparative analysis between the outcomes of English non-word stimuli and English real-world stimuli. Since English non-words can better control the influence of syllable structure, and at the same time, the influence of intonation can be fully considered.

Further topics of investigation might include the comparison of students' performance across various stress typologies in relation to the same task. Initially, it is intriguing to investigate if speakers who do not use stress in their L1 would have any similarities or discrepancies with those who speak a language with unpredictable stress patterns when it comes to perceiving stress in an L2. Second, doing the mutual perception tasks can be the same in these studies. Given the fact that both non-stress language students and non-predictable stress students do not use stress contrastively, it is possible that non-stress language users who excel in stress recognition tasks may not perform well in stress discrimination. If this is the case, this study will be further validated in support.

In addition to cross perceiving each other's English word stress, the comparative analysis of the prosodic systems of the different L1 languages themselves has not been directly tested. Although Thai and Chinese are both tonal languages, there are still different language systems, and whether there is lexical stress in Chinese is still very controversial (Ladd, 2008; Xu, 2021; Feng, 2023). So, this comparison can highlight the difference in perceived accuracy scores and present groundbreaking discoveries about the perception of stress in L2 by students with tonal backgrounds. The forthcoming discoveries will form a substantial collection of cross-linguistic data when combined with the current results. This research direction has the potential to serve as a valuable resource for presenting a continuity approach to the learning of L2 stress perception and production, with a particular pay attention on acoustic evidence.

REFERENCES

Adams, C., & Munro, R. R. (1978). In search of the acoustic correlates of stress: fundamental frequency, amplitude, and duration in the connected utterance of some native and non-native speakers of English. *Phonetica, 35*(3), 125-156. https://doi.org/10.1159/000259926

Adamson, B. (2004). *China's English: A History of English in Chinese Education*. https://eprints.qut.edu.au/21267/

Albadar, I. B. (2021). *The effect of native prosodic properties on the perceptibility of lexical stress: a cross-linguistic experimental study*. https://udspace.udel.edu/handle/19716/30271

Altmann, H., & Vogel, I. (2002). L2 acquisition of stress: The role of L1. In DGfS. *Annual Meeting "Multilingualism Today" in Mannheim, Germany*.

Altmann, H. (2006). *The perception and production of second language stress: a cross-linguistic experimental study*. [Doctoral Dissertation, University of Delaware].

Abramson, A. S. (1978). Static and dynamic acoustic cues in distinctive tones. *Language and Speech, 21*, 319–325. https://doi.org/10.1177/002383097802100406

Archibald, J. (1992). Transfer of L1 parameter settings: some empirical evidence from Polish metrics. *Canadian Journal of Linguistics, 37*, 301–339. https://doi.org/10.1017/S0008413100019903

Archibald, J. (1997). *The acquisition of English stress by speakers of nonaccentual languages: Lexical storage versus computation of stress.* The Acquisition of English Stress by Speakers of Nonaccentual Languages: Lexical Storage Versus Computation of Stress. https://doi.org/10.1515/ling.1997.35.1.167

Bamgbos.e, A. (2001). World Englishes and globalisation. *World Englishes 20*, 357–64. https://doi.org/10.1111/1467-971X.t01-1-00220

Bender, C. (1988). *Chinese phonology*. [B.A Thesis], Swarthmore College.

Beckman, M. E., & Pierrehumbert, J. B. (1986). Intonational structure in Japanese and English. *Phonology, 3,* 255-309. https://doi.org/10.1017/S095267570000066X

Best, C. T. (1995). *A direct realist view of cross-language speech perception.*" In W. Strange (Ed.), Speech perception and linguistic experience: Issues in cross language research (pp. 171–203).

Best, C. T., & Tyler, M. D. (2007). Nonnative and second-language speech perception: Commonalities and complementarities. In M. J. Munro & O.-S. Bohn (Eds.), *Language experience in second language speech learning: In honor of James Emil Flege* (pp. 13-34). John Benjamins.

Boersma, P., & Weenink, D. (2021). Praat: doing phonetics by computer. [Computer program]. Version 6.0.48. Retrieved from http://www.praat.org/ [17.02.2019].

Bolton, K. (2003). *Chinese Englishes: A Sociolinguistic History*. Cambridge: Cambridge University Press.

Bolton, K., Botha, W., & Kirkpatrick, A. (Eds.). (2020). *The Handbook of Asian Englishes*. John Wiley & Sons. https://doi.org/10.1002/9781118791882

Braun, B., Lemhöfer, K., & Mani, N. (2011). *Perceiving unstressed vowels in foreign accented English*. Acoustical Society of America. https://doi.org/10.1121/1.3500688

Brazil, D. (1994). *Pronunciation for advanced learners of English student's book*. Cambridge University Press.

Brown, A. (1989). Some thoughts on intelligibility. *The English Teacher, 8*, 1-11.

Chang, C. Y. (2010). *Dialect differences in the production and perception of Mandarin Chinese tones* [Doctoral dissertation, The Ohio State University]. http://rave.ohiolink.edu/etdc/view?acc_num=osu1282159730

Chao, Y. R. (1968). *A grammar of spoken Chinese.* Berkeley: University of California Press.

Chao, Y. R. 1968. *A Grammar of Spoken Chinese*. University of California Press.

Chen, P. (1999). *Modern Chinese. History and sociolinguistics.* Cambridge: Cambridge University Press.

Chen, P. (1999). *Modern Chinese: history and sociolinguistics.* Cambridge University Press.

Chen, Y., Robb, M., Gilbert, H., & Lerman, J. (2001). Vowel production by Mandarin speakers of English. *Clinical Linguistics & Phonetics, 15*(6), 427-440. https://doi.org/10.1080/02699200110044804

Chen, Y., & Xu, Y. (2006). Production of weak elements in speech–evidence from f_0 patterns of neutral tone in Standard Chinese. *Phonetica, 63*(1), 47-75. https://doi.org/10.1159/000091406

Cho, T. (2001). Effects of morpheme boundaries on intergestural timing: Evidence from Korean. *Phonetica, 58*(3), 129-162. https://doi.org/10.1159/000056196

Chomsky, N., & Halle, M. (1968). The sound pattern of English.

Chrabaszcz, A., Winn, M., Lin, C. Y., & Idsardi, W. J. (2014). Acoustic cues to perception of word stress by English, Mandarin, and Russian speakers. *Journal of Speech, Language, and Hearing Research, 57*(4), 1468-1479. https://doi.org/10.1044/2014_JSLHR-L-13-0279

Collins, B., & Mees, I. M. (2013). *Practical phonetics and phonology: A resource book for students.* Routledge.

Comrie, B. (Ed.). (1987). *The world's major languages.* Routledge.

Cook, V. (1999). Going beyond the native speaker in language teaching. *TESOL Quarterly 22*, 185–209.

Crystal, D (1999) The future of Englishes. *English Today 58,* 10–20. https://doi.org/10.1017/S0266078400010816

Crystal, D. (2003). *English as a global language.* Cambridge university press. https://doi.org/10.1017/CBO9780511486999

Cruttenden, A. (2014). *Gimson's pronunciation of English* (8th ed.). Routledge. https://doi.org/10.4324/9780203784969

Cutler, A. (1986). Forbear is a homophone: Lexical prosody does not constrain lexical access. *Language and speech, 29*(3), 201-220. https://doi.org/10.1177/002383098602900302

Cutler, A. (2015). Lexical Stress in English Pronunciation. In M. Reed & J. M. Levis (Eds.), *The Handbook of English Pronunciation* (pp. 106–124). John Wiley & Sons, Inc. http://onlinelibrary.wiley.com/doi/10.1002/9781118346952.ch6/summary

Cutler, A., & Clifton, C. E. (1984). The use of prosodic information in word recognition. *Attention and Performance X: Control of Language Processes,* 183–196.

Cutler, A., & Norris, D. (1988). The role of strong syllables in segmentation for lexical access. *Journal of Experimental Psychology: Human perception and performance, 14*(1), 113. https://doi.org/10.1037/0096-1523.14.1.113

Cynthia. M. C., & John. M. (1990). Word Familarity and Frequency in Visual and Auditory Word Recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition 1990, 16*(6), 1084-1096. https://doi.org/10.1037/0278-7393.16.6.1084

Deng, D., Feng, S., and Lu, S. (2006). The contrast on tone between Putonghua and Taiwan Mandarin. *Sheng Xue Xue Bao [Acta Acustica], 31*(6), 536-541. https:// 10.15949/j.cnki.0371-0025.2006.06.011

Deterding, D. (2006). The pronunciation of English by speakers from China. *English Worldwide, 27*(2), 175-162. https://doi.org/10.1075/eww.27.2.04det

Deterding, D. (2013). *Misunderstandings in English as a lingua franca: An analysis of ELF interactions in Southeast Asia.* De Gruyter Mouton. https://doi.org/10.1515/9783110288599

Derwing, B. L., & Eddington, D. (2014). The experimental investigation of syllable structure. *The Mental Lexicon, 9*(2), 170-195. https://doi.org/10.1075/ml.9.2.02der

Derwing, T. M., & Munro, M. J. (1997). Accent, intelligibility, and comprehensibility: Evidence from four L1s. *Studies in Second Language Acquisition, 19*(1), 1-16. https://doi.org/10.1017/S0272263197001010

Derwing, T. M., Fraser, H., Kang, O., & Thomson, R. I. (2014). L2 accent and ethics: Issues that merit attention. *English in multilingual contexts: Language variation and education*, 63-80.

Di, Shiyu. (2003). *Hànyu fangyánxué* ('The study of Chinese dialectology'). Chongqing: Xinán Shifàn Dàxué Chubanshè.

Dow, F. D. M. (1972). *An outline of Mandarin phonetics.* Australian National University Press.

Duanmu, S. (2007). *The phonology of standard Chinese*. OUP Oxford. https://doi.org/10.1017/S0025100316000359

Duanmu, S. (2014). Syllable structure and stress. *The handbook of Chinese linguistics*, 422-442.

Dupoux, E., Peperkamp, S., & Sebastián-Gallés, N. (2001). A robust method to study stress "deafness". *The Journal of the Acoustical Society of America, 110*(3), 1606-1618. https://doi.org/10.1121/1.1380437

Dupoux, E., Sebastián-Gallés, N., Navarrete, E., & Peperkamp, S. (2008). Persistent stress 'deafness': The case of French learners of Spanish. *Cognition, 106*(2), 682-706. https://doi.org/10.1016/j.cognition.2007.04.001

Escudero, P. (2005). *Linguistic perception and second language acquisition: Explaining the attainment of optimal phonological categorization*. Netherlands Graduate School of Linguistics. http://wwwlot.let.uu.nl/

Escudero, P., & Chládková, K. (2010). Spanish listeners' perception of American and Southern British English vowels. *The Journal of the Acoustical Society of America, 128*(5), EL254-EL260. https://doi.org/10.1121/1.3488794

Field, J. (2005). Intelligibility and the Listener: The Role of Lexical Stress. *TESOL Quarterly, 39*(3), 399–423. https://doi.org/10.2307/3588487

Flege, J. E. (1995). Second language speech learning theory, findings, and problems. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 233–77). Baltimore: York Press.

Flege, J. E., Munro, M. J., & MacKay, I. R. (1995). Factors affecting strength of perceived foreign accent in a second language. *The Journal of the Acoustical Society of America, 97*(5), 3125-3134. https://doi.org/10.1121/1.413041

Fougeron, C., & Keating, P. A. (1997). Articulatory strengthening at edges of prosodic domains. *The journal of the acoustical society of America, 101*(6), 3728-3740. https://doi.org/10.1121/1.418332

Fox, A. (2002). *Prosodic Features and Prosodic Structure: The Phonology of'Suprasegmentals'.* OUP Oxford. https://doi.org/10.1093/oso/9780198237853.001.0001

Fry, D. (1955). Duration and Intensity as Physical Correlates of Linguistic Stress. *The Journal of The Acoustical Society of America, 26*(1), 138-138. https://doi.org/10.1121/1.1908022

Fry, D. (1958). Experiments in the Perception of Stress. *Language And Speech, 1*(2), 126-152. https://doi.org/10.1177/002383095800100207

Fu, Q.-J., Zeng, F.-G., Shannon, R. V., & Soli, S. D. (1998). Importance of tonal envelope cues in Chinese speech recognition. *The Journal of the Acoustical Society of America, 104*(1), 505–510. https://doi.org/10.1121/1.423251

Gandour, J. T. (1979). *Tonal rules for English loanwords in Thai.* Theraphan et al. Studies in Tai and Mon-Khmer phonetics and phonology in honour of Eugénie JA Henderson, Bangkok, Chulalongkorn University Press.

Gandour, J. (1983). Tone perception in Far Eastern languages. *Journal of phonetics, 11*(2), 149-175. https://doi.org/10.1016/S0095-4470(19)30813-7

Gandour, J. (1984). Tone dissimilarity judgments by Chinese listeners. *Journal of Chinese Linguistics, 12,* 235–261.

Gårding, E., Kratochvil, P., Svantesson, J. O., & Zhang, J. (1986). Tone 4 and Tone 3 discrimination in modern standard Chinese. *Language and Speech, 29*(3), 281-293. https://doi.org/10.1177/002383098602900307

Gedney, W. J. (1947). *Indic loanwords in spoken Thai*. Yale University.

Gernsbacher, M. A. (1984). Resolving 20 years of inconsistent interactions between lexical familiarity and orthography, concreteness and polysemy. *Journal of*

Experimental Psychology: General, 113, 256-281. https://doi.org/10.1037/0096-3445.113.2.256

Ghosh M., & Levis J. M. (2021). Vowel quality and direction of stress shift in a predictive model explaining the varying impact of misplaced word stress: Evidence from English. *Frontiers in Communication, 6*. https://doi.org/10.3389/fcomm.2021.628780

Gilichinskaya, Y. D., & Strange, W. (2010). Perceptual assimilation of American English vowels by inexperienced Russian listeners. *The Journal of the Acoustical Society of America, 128*(2), EL80-EL85. https://doi.org/10.1121/1.3462988

Görlach, M. (1998). Varieties of English world-wide: Where we stand. *Links & letters,* 13-36.

Guo, L. (2004). The relationship between Putonghua and Chinese dialects. In *Language Policy in the People's Republic of China: Theory and Practice since 1949*. M. Zhou (Ed.). Boston: Kluwer Academic Publishers, 45-53.

Hansen, J. H. L. & Arslan, L. (1995). Foreign accent classification using source generator-based features. *In 1995 International Conference on Acoustics, Speech, and Signal Processing* (Vol. 1, pp. 836-839). https://10.1109/ICASSP.1995.479824

Hayes, B. (1995). *Metrical stress theory: Principles and case studies.* University of Chicago Press.

He, D., & Li, D. C. S. (2009). Language attitudes and linguistic features in the 'China English' debate. *World Englishes, 28*(1), 70–89. https://doi.org/10.1111/j.1467-971X.2008.01570.x

Hemakom, A., Jitwiriyanontb, S., Rugchatjaroena, A., & Israsena, P. (2021). The development of Thai monosyllabic word and picture lists applicable to interactive speech audiometry in preschoolers. *Clinical linguistics & phonetics, 35*(9), 809–828. https://doi.org/10.1080/02699206.2020.1830301

Hinkel, E. (Ed.). (1999). *Culture in second language teaching and learning.* Cambridge University Press.

Hiranburana, S. (1971). *The role of accent in Thai grammar.* University of London, School of Oriental and African Studies.

Hirst, D., & Di Cristo, A. (1998). A survey of intonation systems. *Intonation systems: A survey of twenty languages, 144*, 152-166.

Howie, J. M. (1976). *Acoustical studies of Mandarin vowels and tones* (Vol. 18). Cambridge University Press.

Hung, T. T. N. (1992) Some problems and issues in the representation of pronunciation in a Singapore English dictionary. In A. Pakir (ed.), *Words in a Cultural Context* (pp. 30–36). UniPress.

Hyman, L. M. (2009). How (not) to do phonological typology: the case of pitch-accent. *Language sciences, 31*(2-3), 213-238. https://doi.org/10.1016/j.langsci.2008.12.007

Isaacs, T., & Trofimovich, P. (2012). Deconstructing comprehensibility: Identifying the linguistic influences on listeners' L2 comprehensibility ratings. *Studies in Second Language Acquisition, 34*(3), 475-505. https://doi.org/10.1017/S0272263112000150

Jangjamras, J. (2011). *Perception and production of English lexical stress by Thai speakers.* University of Florida. https://doi.org/ 10.1121/1.4783772

Jenkins, J. (2000). *The Phonology of English as an International Language*. OUP Oxford.

Jenkins, J. (2002). A Sociolinguistically Based, Empirically Researched Pronunciation Syllabus for English as an International Language. *Applied Linguistics, 23*(1), 83–103. https://doi.org/10.1093/applin/23.1.83

Jenkins, J. (2015). Repositioning English and multilingualism in English as a Lingua Franca. *Englishes in Practice, 2*(3), 49-85. https://doi.org/10.1515/eip-2015-0003

Jeong, H., Thorén, B., & Othman, J. (2017). Mutual intelligibility of Malay-and Swedish-accented English: An experimental study. *Indonesian Journal of Applied Linguistics, 7*(1), 43-53.

Jeong, H., Thorén, B., & Othman, J. (2020). Effect of altering three phonetic features on intelligibility of English as a lingua franca: a Malaysian speaker and Swedish listeners. *Asian Englishes, 22*(1), 2-19. https://doi.org/10.1080/13488678.2018.1536817

Jiang, Y. (2003). English as a Chinese language. *English Today, 19*(2), 3–8. https://doi.org/10.1017/S0266078403002013

Johnson, S. O. (2017). *Phonological correlates of socio-economic background of English speech of secondary school students in oyo and ogun states*, [Doctoral dissertation]. UISpace.

Jones, D. (2003). *English pronouncing dictionary.* P. Roach, J. Hartman, & J. Setter (Eds.). Cambridge University Press.

Jongman, A., & Moore, C. B. (2000). The role of language experience in speaker and rate normalization processes. *In Sixth International Conference on Spoken Language Processing.*

Kachru, B. B. (1985). The bilinguals' creativity. *Annual review of applied linguistics, 6*, 20-33. https://doi.org/10.1017/S0267190500003032

Kachru, B. B. (1997). World Englishes and English-using communities. *Annual review of applied linguistics, 17*, 66-87. https://doi.org/10.1017/S0267190500003287

Kachru, Y. & Nelson, C.L. (2006). *World Englishes in Asian Contexts*. Hong Kong: Hong Kong University Press.

Kanokpermpoon, M. (2007). Thai and English consonantal sounds: A problem or a potential for EFL learning?. *ABAC journal, 27*(1).

Karjo, C. H. (2016). Accounting for L2 learners' errors in word stress placement. *Indonesian Journal of Applied Linguistics, 5*(2), 199-208.

Kijak, A. (2009). How stressful is L2 stress? A cross linguistic study of L2 perception and production of metrical systems [Doctoral dissertation]. Universiteit Utrecht.

Kirkpatrick, A (2007b). *World Englishes: Implications for International Communication and English Language Teaching*. Cambridge: Cambridge University Press.

Kirkpatrick, A. (2010). *English as a Lingua Franca in ASEAN: A Multilingual Model.* Hong Kong: Hong Kong University Press. https://www.jstor.org/stable/j.ctt1xcs49

Kruatrachue, F. (1960). *Thai and English: A comparative study of phonology for pedagogical applications*. Indiana University. http://wals.info/refdb/record/471

Ladefoged, P., & Johnson, K. (2014). *A course in phonetics.* Cengage learning.

Lai, Y. (2008). *Acoustic realization and perception of English lexical stress by Mandarin learners* [Doctoral dissertation, University of Kansas].

Lengeris, A. (2009). Perceptual assimilation and L2 learning: Evidence from the perception of Southern British English vowels by native speakers of Greek and Japanese. *Phonetica, 66*(3), 169-187. https://doi.org/10.1159/000235659

Lepage, A., & Busà, M. G. (2014). Intelligibility of English L2: The effects of incorrect word stress placement and incorrect vowel reduction in the speech of French and Italian learners of English. *In Proceedings of the International Symposium on the Acquisition of Second Language Speech Concordia Working Papers in Applied Linguistics, 5*(2014), 387-400.

Lewis, C., & Deterding, D. (2018). Word stress and pronunciation teaching in English as a lingua franca contexts. *CATESOL Journal, 30*(1), 161-176.

Li A. and Wang, X. (2003). A contrastive investigation of standard Mandarin and accented Mandarin. In *Proceedings of the 5th International Symposium on Chinese Spoken Language 2006*. 2345-2348. https://doi.org/10.21437/Eurospeech.2003-647

Li, A., Xiong, Z., & Wang, X. (2006). Contrastive study on tonal patterns between accented and standard Chinese. *In Proceedings of the 5th International Symposium on Chinese Spoken Language 2006* (pp. 157-168).

Li, D. C.S. (2006). Chinese as a lingua franca in greater China. *Annual Review of Applied Linguistics, 26*, 149-176. https://doi.org/10.1017/S0267190506000080

Li, R. L. (2001). *Hànyu fangyánxué* [The study of Chinese dialectology]. Beijing: Gaodang Education Press.

Lin, C. Y., Wang, M., Idsardi, W. J., & Xu, Y. I. (2014). *Stress processing in mandarin and Korean second language learners of English.* Cambridge University Press. https://doi.org/ 10.1017/S1366728913000333

Lin, M.-C. (1995). A perceptual study on the domain of tones in Standard Chiense. *Chinese Journal of Acoustics, 14* (4), 350-357.

Lin, M., & Yan, J. (1988). The characteristic features of the final reduction in the neutral-tone syllable of Beijing Mandarin. *Phonetic laboratory annual report of phonetic research, 37*, 51.

Liu, S., & Samuel, A. G. (2004). Perception of Mandarin lexical tones when F0 information is neutralized. *Language and Speech, 47*(2), 109–138. https://doi.org/10.1177/00238309040470020101

Luce, P. A. (1986). *Neighborhoods of words in the mental lexicon.* Unpublished doctoral dissertation, Indiana University.

Luke, K. K. (2005). Cóng shuangyu shuangyán kàn Xianggang shèhuì yuyán biànqian [Bilingualism and diglossia: Sociolinguistic changes in Hong Kong]. *Zhongguó shèhuì yuyánxué [The Journal of Chinese Sociolinguistics], 1*, 82–88.

Luksaneeyanawin, S. (1998). *Intonation systems: a survey of twenty languages* (pp. 345-359). Cambridge University Press. https://doi.org/10.1353/lan.2000.0088

Massaro, D. W., Cohen, M. M., & Tseng, C. (1985). The evaluation and integration of pitch height and pitch contour in lexical tone perception in Mandarin Chinese. *Journal of Chinese Linguistics, 13,* 267–290. https://www.jstor.org/stable/23767517

Mattys, S., & Samuel, A. G. (2000). Implications of stress-pattern differences in spoken-word recognition. *Journal of Memory and Language, 42*(4), 571-596. https://doi.org/10.1006/jmla.1999.2696

Michael, S. M., & Anton, K., & Joseph, G. (2011). *A Handbook of Process Tracing Methods for Decision Research: A Critical Review and User's Guide*. Psychology Press. https://doi.org/10.4324/9780203875292

Mol, H., & Uhlenbeck, E. M. (1956). The correlation between interpretation and production of speech sounds. *Lingua, 6,* 333-353. https://doi.org/10.1016/0024-3841(56)90029-8

Moore, C. B., & Jongman, A. (1997). Speaker normalization in the perception of Mandarin Chinese tones. *Journal of the Acoustical Society of America, 102,* 1864–1877. https://doi.org/10.1121/1.420092

Morén, B., & Zsiga, E. (2006). The lexical and post-lexical phonology of Thai tones. *Natural Language & Linguistic Theory, 24*(1), 113-178. https://doi.org/10.1007/s11049-004-5454-y

Morton, J., & Jassem, W. (1965). Acoustic correlates of stress. *Language and speech, 8*(3), 159-181. https://doi.org/10.1177/002383096500800303

Munro, M. J. (1998). The effects of noise on the intelligibility of foreign-accented speech. *Studies in Second Language Acquisition, 20*(2), 139-154. https://doi.org/10.1017/S0272263198002022

Munro, M. J., & Derwing, T. M. (1995). Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language Learning, 45*(1), 73-97. https://doi.org/10.1111/j.1467-1770.1995.tb00963.x

Munro, M. J., & Derwing, T. M. (2015). Intelligibility in research and practice: Teaching priorities. In M. Reed & J. M. Levis (Eds.), *The handbook of English pronunciation* (pp. 377-396). John Wiley & Sons, Inc.

Nakatani, L. H., & Aston, C. H. (1978). Perceiving the stress pattern of words in sentences. *The Journal of the Acoustical Society of America, 63*(S1), S55-S55. https://doi.org/10.1121/1.2016717

Nguyen, T. A. T. (2003). *Prosodic Transfer: The Tonal Constraints on Vietnamese Acquisition of English Stress and Rhythm.* [Unpublished dissertation]. University of Queensland. https://doi.org/10.14264/107093

Nguyễn, T. A.-T., Ingram, C. L. J., & Pensalfini, J. R. (2008). Prosodic transfer in Vietnamese acquisition of English contrastive stress patterns. *Journal of Phonetics, 36*, 158-190. https://doi.org/10.1016/j.wocn.2007.09.001

Niu, Q. & Wolff, M. (2007). Linguistic failures. *English Today, 23*(1), 61–4. https://doi.org/10.1017/S0266078407001125

Norman, J. (1988). *Chinese.* Cambridge, UK: Cambridge University Press.

Noss, R. B. 1984. *An Overview of Language Issues in South-East Asia 1950–1980*. Oxford University Press.

Nusbaum, H., & Dedina, M. (1985). *The ejects of word frequency and subjective familiarity on visuallexical decisions.* (Research on Speech Perception, Progress Resport No. 11). Indiana University.

Ou, S. (2010). Taiwanese EFL learners' perception of English word stress. *Concentric: Studies in Linguistics, 36*(1), 1-23.

Ou, S. (2016). Perception of English lexical stress with a marked pitch accent by native speakers of mandarin. *Taiwan Journal of Linguistics, 14*(2).

Panlay, S. (1997). *The effect of English loanwords on the pronunciation of Thai*. [Master's. Thesis], Michigan State University.

Pan, L. & Seargeant, P. (2012). Is English a threat to Chinese language and culture? *English Today, 28*(3), 60–6. https://doi.org/10.1017/S0266078412000302

Pater, J. (1997). Minimal violation and phonological development. *Language Acquisition, 6*(3), 201-253. https://doi.org/10.1207/s15327817la0603_2

Peng, G., & Wang, W. S. Y. (2005). Tone recognition of continuous Cantonese speech based on support vector machines. *Speech Communication, 45*(1), 49-62. https://doi.org/10.1016/j.specom.2004.09.004

Peperkamp, S., & Dupoux, E. (2002). *A typological study of stress 'deafness'. In: C. Gussenhoven & N. Warner* (eds.) Papers in Laboratory Phonology 7. Mouton de Gruyter, 203-240.

Peperkamp, S., Vendelin, I., & Dupoux, E. (2010). Perception of predictable stress: A cross-linguistic investigation. *Journal of Phonetics, 38*(3), 422-430. https://doi.org/10.1016/j.wocn.2010.04.001

Peyasantiwong, P. (1986). Stress in Thai. *Thai Studies in Honor of William J. Gedney.* 211-230.

Phetkla, C. (2020). *Vowel adaptation in English loanwords in Thai* [Doctoral dissertation, Newcastle University]. http://theses.ncl.ac.uk/jspui/handle/10443/4968

Pickering, L. (2006). Current research on intelligibility in English as a lingua franca. *Annual Review of Applied Linguistics, 26,* 219-233. https://doi.org/10.1017/S0267190506000110

Pitt, M. A., & Samuel, A. G. (1993). An empirical and meta-analytic evaluation of the phoneme identification task. *Journal of Experimental Psychology: Human Perception and Performance, 19*(4), 699.

Pollap, T. U. N. T. I. G. A. (2010). *The relationship between American cultural background knowledge and the English proficiency of Thai students majoring in English.* [Unpublished masters' thesis]. Srinakharinwirot University.

Rawlings, M., & Sue, E. (2013). Preparedness of Chinese students for American culture and communicating in English. *Journal of International Students, 3*(1), 29-40.

Redford, M. A., & Randall, P. (2005). The role of juncture cues and phonological knowledge in English syllabification judgments. *Journal of Phonetics, 33*(1), 27-46. https://doi.org/10.1016/j.wocn.2004.05.003

Richards, M. G. (2016). *Not all word stress errors are created equal: Validating an English word stress error gravity hierarchy.* [Unpublished doctoral dissertation). Iowa State University.

Roach, P. (2009). *English phonetics and phonology paperback with audio CDs (2): A practical course.* Cambridge university press.

Saengsuriya, N. (1989). *Variations in duration, pitch and amplitude of Thai words in different utterance conditions.* [Doctoral dissertation]. University of Kansas.

Satravaha, N. (2002). *Tone classification of syllable-segmented Thai speech based on multilayer perceptron*. [Doctoral dissertation]. West Virginia University. https://doi.org/10.33915/etd.1611

Schertz, J., & Clare, E. J. (2020). Phonetic cue weighting in perception and production. *Wiley Interdisciplinary Reviews: Cognitive Science, 11*(2), e1521. https://doi.org/10.1002/wcs.1521

Schwab, S., & Llisterri, J. (2011). Are French speakers able to learn to perceive lexical stress contrasts?. In *ICPhS* (pp. 1774-1777).

Seidlhofer, B. (1999). Double standards: teacher education in the Expanding Circle. *World Englishes 18,* 233–45. https://doi.org/10.1111/1467-971X.00136

Seidlhofer, B. (2001). Closing a conceptual gap: The case for a description of English as a lingua franca. *International journal of applied linguistics, 11*(2), 133-158. https://doi.org/10.1111/1473-4192.00011

Seidlhofer, B. (2002). Pedagogy and local learner corpora. *Computer learner corpora, second language acquisition and foreign language teaching,* 213-234.

Seidlhofer, B. (2004). Research perspectives on teaching English as a lingua franca. *Annual review of applied linguistics, 24*, 209-239. https://doi.org/10.1017/S0267190504000145

Shen, X. S., & Lin, M. C. (1991). A perceptual study of Mandarin tones 2 and 3. *Language and Speech, 34*, 145–156. https://doi.org/10.1177/002383099103400202

Sluijter, A. M., & Van Heuven, V. J. (1996). Spectral balance as an acoustic correlate of linguistic stress. *The Journal of the Acoustical Society of America, 100*(4), 2471-2485. https://doi.org/10.1121/1.417955

Smith, L. E., & Nelson, C. L. (1985). International intelligibility of English: Directions and resources. *World Englishes, 4*(3), 333-342. https://doi.org/10.1111/j.1467-971X.1985.tb00423.x

Snodin, N. S., & Young, T. J. (2015). Native-speaker varieties of English: Thai perceptions and attitudes. *Asian Englishes, 17*(3), 248-260. https://doi.org/10.1080/13488678.2015.1083354

So, C. K., & Best, C. T. (2008). Do English speakers assimilate Mandarin tones to English prosodic categories? *Ninth Annual Conference of the International Speech Communication Association.* http://handle.uws.edu.au:8081/1959.7/45242

So, C. K., & Best, C. T. (2010). Cross-language perception of non-native tonal contrasts: Effects of native phonological and phonetic influences. *Language and Speech, 53*(2), 273–293. https://doi.org/10.1177/0023830909357156

So, C. K., & Best, C. T. (2014). Phonetic influences on English and French Listeners' assimilation of Mandarin tones to native prosodic categories. *Studies in Second Language Acquisition, 36*(2), 195–221. https://doi.org/10.1017/S0272263114000047

Surintramont, A. (1973). Some aspects of underlying syllable structure in Thai: evidence from Khamphuan--a Thai word game. *Studies in the Linguistic Sciences, 3*(1), 121-142.

Suzukida, Y., & Saito, K. (2019). Which segmental features matter for successful L2 comprehensibility? Revisiting and generalizing the pedagogical value of the functional load principle. *Language Teaching Research,* 1-20. https://doi.org/10.1177/1362168819858246

Tremblay, A., & Owens, N. (2010). The role of acoustic cues in the development of non-target-like second-language prosodic representations. *Canadian Journal of Linguistics/Revue canadienne de linguistique, 55*(1), 85-114. https://doi.org/10.1017/S0008413100001389

Trudgill, Peter, & Hannah, J (1994). *International English*. London: Arnold.

Tsukada, K. (2008). An acoustic comparison of English monophthongs and diphthongs produced by Australian and Thai speakers. *English world-wide, 29*(2), 194-211. https://doi.org/10.1075/eww.29.2.05tsu

Tuaycharoen, P. (1990). Phonetics and practical phonetics. https://doi.org/10.1017/9781108289849

Turk, A. E., & Sawusch, J. R. (1996). The processing of duration and intensity cues to prominence. *The Journal of the Acoustical Society of America, 99*(6), 3782-3790. https://doi.org/10.1121/1.414995

Van der Hulst, H. (2002). Stress and accent. *Group*, *4*(4), 246, 254.

Van Donselaar, W., Koster, M., & Cutler, A. (2005). Exploring the role of lexical stress in lexical recognition. *The Quarterly Journal of Experimental Psychology Section A, 58*(2), 251-273. https://doi.org/10.1080/02724980343000927

Vogel, P. (2000). Grammaticalization and part of speech systems. *Approaches to the typology of word classes,* 259-284. https://doi.org/10.1515/9783110806120.259

Walker, R., & Zoghbor, W. (2015). *The pronunciation of English as a lingua franca. In M. Reed & J. M. Levis* (Eds.), The handbook of English pronunciation (pp. 433-453). Wiley Blackwell.

Wang, W. S-Y. (1967). Phonological features of tones. *International Journal of American Linguistics, 33* (2), 93-105.

Wang, X. Y. (2014). The establishment and selection of China's lingua franca from the perspective of national sovereignty. *Hebei Academic Journal, 34*(3): 187-190.

https://kns.cnki.net/kcms2/article/abstract?v=MTbc36RhFpSvzO2AXRqNBpoSjkffnww8-wDSmKZYmgyw6ooZSmBozx0JwHSoW1emUyHOGBbCHfwqs1IbdXf-pPI9eJyOUJVMp3vpOKpFMY40klxzzsUZmszTTNjIIty00A7sKjJXubQ=&uniplatform=NZKPT&language=CHS

Wang, Y., Jongman, A., & Sereno, J. A. (2003). Acoustic and perceptual evaluation of Mandarin tone productions before and after perceptual training. *The Journal of the Acoustical Society of America, 113*(2), 1033-1043. https://doi.org/10.1121/1.1531176

Wang, Q. (2008). *Perception of English stress by Mandarin Chinese learners of English: An acoustic study.* [Doctoral dissertation]. University of Victoria. http://hdl.handle.net/1828/1282

Wang, Q., & Yoon, T. J. (2008). The transfer of L1 acoustic cues in the perception of L2 Lexical stress. *Canadian Acoustics, 36*(3), 126-127.

Wang, H. & Yuan Z. (2013). The promotion of Putonghua (Mandarin Chinese): An overview. In Y. Li & W. Li (Eds.), *The language situation in China* (vol. 1, 27–40). Mouton De Gruyter.

Wayland, R. (1997). Non-native production of Thai: acoustic measurements and accentedness ratings. *Applied Linguistics, 18*(3), 345-373. https://doi.org/10.1093/applin/18.3.345

Wayland, R., Guion, S.G., Landfair, D. and Li, B. (2006). Native Thai speakers' acquisition of English word stress patterns. *Journal of Psycholinguistic Research 35*, 285-304. https://doi.org/10.1007/s10936-006-9016-9

Wei, R. & Su., J. (2012). The statistics of English in China. *English Today, 28*(3), 10–14. https://doi.org/10.1017/S0266078412000235

Wells, J. C. (2008). *Longman pronunciation dictionary* (3rd ed.). Pearson Longman. https://doi.org/10.1017/S0025100309003934

Widdowson, H. G. (1994). The ownership of English. *TESOL quarterly, 28*(2), 377-389.

Wongsothorn, A. (2000). Thailand. In Wah Kum Ho and Ruth Y. L. Wong, eds., Language Policies and Language Education. *Singapore: Times Academic Press*, 307–320.

Wutiwiwatchai, C., & Furui, S. (2007). Thai speech processing technology: A review. *Speech communication, 49*(1), 8-27. https://doi.org/10.1016/j.specom.2006.10.004

Yamada, R. A., & Tohkura, Y. I. (1992). The effects of experimental variables on the perception of American English/r/and/l/by Japanese listeners. *Perception & psychophysics, 52*, 376-392. https://doi.org/10.3758/BF03206698

Yoshida, Y. (1995). *On pitch accent phenomena in Standard Japanese.* University of London, School of Oriental and African Studies.

Yu, V. Y., & Andruski, J. E. (2010). A cross-language study of perception of lexical stress in English. *Journal of psycholinguistic research, 39*, 323-344.

Zahner, K., Kutscheid, S., & Braun, B. (2019). Alignment of f0 peak in different pitch accent types affects perception of metrical stress. *Journal of Phonetics, 74,* 75-95. https://doi.org/10.1016/j.wocn.2019.02.004

Zhang, Y., & Francis, A. (2010). *The weighting of vowel quality in native and nonnative listeners' perception of English lexical stress.* Elsevier. https://doi.org/10.1016/j.wocn.2009.11.002

Zhang, Y., Nissen, S. L., & Francis, A. L. (2008). Acoustic characteristics of English lexical stress produced by native Mandarin speakers. *The Journal of the Acoustical Society of America, 123*(6), 4498–4513. https://doi.org/10.1121/1.2902165

Zhou, M. L. (2001). The spread of Putonghua and language attitude changes in Shanghai and Guangzhou, China. *Journal of Asian Pacific Communication, 11*, 231–253. https://doi.org/10.1075/japc.11.2.07zho

Zhou, Q. S. (2003). Zhongguó yuyán zhèngcè yuyán guihuà yánjiù [The language policy and language planning in China]. In Q.S. Zhou (ed.), *Guójia, mínzú yuyán – yuyán zhèngcè guóbié yánjiù* [Nation, ethnic group and language—language policy by nations] (pp.250–275). Y wén Chubanshè.

Zielinski, B. (2015). *The segmental/suprasegmental debate.* In M. Reed & J. M. Levis (Eds.), The handbook of English pronunciation (pp. 397-412). Wiley-Blackwell. https://doi.org/10.1002/9781118346952.ch22