INTONATION OF THAI LEARNERS OF ENGLISH FOR MARKING NARROW FOCUS

NISIT KAMPHIKUL

Thesis submitted for the degree of
Doctor of Philosophy
Integrated PhD in Phonetics and Phonology

Newcastle University
School of Education, Communication and Language Sciences

May 2022

Abstract

Languages differ in the ways they mark the focus of information. While English marks focus by means of pitch accent placement and relative acoustic prominence in elements such as f0, intensity and duration, marking focus in Thai can be accomplished by either using a particular word order, changing lexical items or syntactic constructions, using prosody such as extra high tones and phrase-final lengthening to emphasise information, emotions and attitudes, or using combinations of these. In the current study, using prosodic means to express the focus is of special interest. Basically, the prosodic patterns of the two languages involve the use of categorical means such as accent types and gradient means, such as f0, duration and loudness to variable degrees, as well as alignment and scaling. This study thus investigates if there are any differences between Thai learners and native speakers of English in the use of accent choices and the acoustic-phonetic characteristics of focus, in particular via using rising accents.

This study presents the first detailed description covering both phonological and acoustic-phonetic analyses of the prosodic marking of focus in English produced by Thai learners. The aims of this study are threefold: 1) to investigate to what extent Thai learners and native speakers of English mark narrow informative focus prosodically, in particular in terms of the use of accent types and other strategies such as deaccenting; 2) to describe Thai learners' and native English speakers' acoustic cues for marking this kind of focus, as well as tonal alignment and scaling of rising accents associated with marking focus in question; 3) to explore the factors which affect differences in the realisations of focus between Thai learners and native speakers of English, such as L2 English proficiency levels, focus positions, and learners' gender.

The study concentrates on a production experiment conducted on the basis of the Autosegmental-Metrical approach to intonational phonology. Twenty native English speakers and 20 native Thai speakers produced English speech data during question-and-answer tasks recorded in laboratory conditions. Transcriptions of the speech data were based on the Tones and Break Indices system with the help of speech software, Praat. The results showed that native English speakers and Thai learners predominantly used rising accents (L*+H, L+H*, L+<H*) to mark narrow informative focus and they also relied on other different prosodic strategies such as deaccenting and rephrasing. H* can be found in both groups, but L* in Thai group only. Using a high (H*) accent and a rising (L+H*) accent with or without a later peak reflects Thai learners' shared characteristics of how to mark focus using pitch prominence to highlight information. In this case, CAH may be applied for prediction in that shared categories of accents

are easy to produce and ready to use. The use of an accent alongside prosodic strategies such as compressed pitch range and rephrasing can be considered to be due to either non-shared characteristics or the development of a system concerning the characteristics of the L2 intonation patterns produced by Thai learners in expressing the focus. In this case, if CAH was applied, the deaccentuation which is used by native English speakers will be difficult to produce and Thai learners would transfer strategies such as L1 compressed pitch range and rephrasing into L2 English.

In addition, they differed in terms of using acoustic cues, as well as tonal alignment and the scaling of rising accents which they used to mark focus. The analysis of Thai learners' acoustic-phonetic data showed that they tended to employ higher pitch, longer duration and greater intensity and they used rising accents with a delayed peak (L+<H*) resulting in greater pitch slope/size, especially in sentence-initial positions. Factors such as L2 English proficiency level, position of focus, as well as gender still play roles in causing differences such as the misalignment of tonal targets which, it can be speculated, could be a result of the development and use of L2 intonation; whereas an increase in degrees of f0 and duration can be attributed to L1 influence. If Ladd's (1996, 2008) taxonomy of cross-linguistic differences in intonation and Mennen's (2015) L2 intonational learning theory (LILt) are taken into account, this study gives evidence for systematic differences and realisational differences between L1 English and L2 English; that is, in using categorical and gradient prosodic strategies in the expression of focus. Finally, one of the findings has pedagogical implications concerning the use of focus-marking strategies. Therefore, apart from learning where and how native English speakers make words prominent, Thai learners should also learn how native English speakers deaccent out-of-focus material or information. This is because deaccenting mostly co-occurs with an accent placement in the marking of focus in the native English speakers' data.

This study makes contributions to L2 research into intonation and the marking of focus, in particular to Mennen's (2015) L2 intonational learning theory (LILt). It also makes contributions concerning the concepts used to explain tonal alignment and scaling within Ladd's (2008) AM theory by offering empirical evidence supporting an understanding of these tonal phenomena in Thai learners compared to native speakers of English

Keywords: intonation, focus, prosodic strategy, tonal alignment and scaling

Declaration

I hereby declare this thesis is a record of original work done by me except for quotations and citations, which have been acknowledged. I also declare that this work has not been submitted to any other university for the award of any degree or published at any time before.

Name: Nisit Kamphikul

Date: May 2022

Acknowledgements

Table of Contents

List of Figures

List of Tables

Chapter 1. Introduction

1.1 Background and Rationale

Over the decades, Thailand's educational system has placed emphasis to a substantial extent on teaching and learning English as a foreign language (EFL) since the reign of King Rama III (1824-1851) according to Durongphan *et al.* (1982) cited in Foley (2005). Thailand's government and educators prioritise English in part because the Association of Southeast Asian Nations (ASEAN) is increasing in size, and its membership has expanded to include Japan, China, and South Korea along with countries in which English is the or one of the official languages such as India, Australia, and New Zealand, representing the ASEAN+3 and ASEAN+6 groups respectively (Kawai and Wignaraja, 2007). The significance of the English language is dramatically growing along with its high social value as a working language for communication among countries in this community (Kirkpatrick, 2008).

All schools and higher educational institutions in Thailand include English as part of formal or informal curricula for learning foreign languages since 1871 (Office of the Education Council, 2017). Thai students normally start to learn English at primary school and continue studying it at high school. The study of English in the compulsory education system lasts for at least twelve years. Despite spending several years in its study, English is still no closer to being at all easier to learn. This is even more discouraging for Thai students, and especially those who have learnt it and need it for use in communication such as in international workplaces (Wiriyachitra, 2001; Hiranburana, 2017).

When considering issues of English pronunciation (for example, English intonation), in instructional settings, possible external sources of problems include little coverage in commercial textbooks, Thai teachers whose English is not native-like and little exposure to English outside the classroom. In this situation, a teacher's pronunciation as well as stress and intonation serve as a model which Thai students listen to and imitate. The chance of using English only arises in English classrooms such as in speaking-and-listening exercises. These issues have been reported by Wiriyachitra (2001), Wongsothorn, Hiranburana and Chinnawongs (2002), and Khamkhien (2010), and mainly occur because Thais normally use their native language for communication in everyday life, including via mass media.

Looking at other possibilities in terms of internal sources, it has long been generally accepted that problems in pronunciation stem partly from first language influence (Lado, 1975) and the critical period for language learning (Lenneberg, 1967). When it comes to intonation, Thai learners have specific problems due to it being a tone language (Abramson, 1962; Iwasaki and Ingkaphirom, 2005). (More details of Thai are provided in section 2.4.) In addition to marking tone lexically, Thai marks phrasal stress and intonation (Henderson, 1949; Abramson, 1979). Another difficulty comes from the choices of intonation patterns that students use to accompany their speech. They may have been taught explicitly about the idea of intonation, along with some classroom practice, but they may make the wrong choices of tonic syllables, as well as of the tones associated with those syllables. Thai learners may lack sufficient knowledge of or exposure to the target language; or otherwise, they may have had implicit exposure to English intonation and then attempt to use it by trial and error. In this case, their English might sound correct if they reproduce exactly the same intonation as their input. If not, their English might sound different. It is likely that their English sounds foreign, monotonous and unintelligible compared to their intentions. The latter case can cause misunderstandings and can lead to communication breakdowns such as an inability to keep a conversation going. The degree of misunderstanding will vary depending on the context in which intonation patterns are used, as well as on the correctness of the pronunciation of segmental phonemes.

Grabe, Kochanski and Coleman (2005, p. 311) described intonation as being "held responsible for numerous instances of miscommunication between native and non-native speakers. In English, these are said to involve primarily the pragmatic impact of utterances and occasions when the wrong intonation causes a difference in grammatical meaning and utterance type." Secondly, Wells (2006, p. 2) pointed out that "almost any intonation pattern is possible in English; but different intonation patterns have different meanings. The difficulty is that the pattern that the learner uses may not have the meaning he or she intends." It is obvious that, without the correct intonation, it may be difficult for the listener to understand the speaker's intentions.

Mistakes in the expression of focus or in transitions of meaning between Thai learners and native English speakers can be seen in the following examples (Wells, 2006, p. 139).

a) *She was ˈtrying to lose ˈ<u>weight</u>.*

b) *She was ˈtrying to ˈ<u>lose</u> weight.*

c) *She was ˈ<u>trying</u> to lose weight.*

Explanations of the sources of such mistakes can be drawn from the viewpoints of Grabe, Kochanski and Coleman (2005) and Wells (2006) in the above instance. For instance, Thai learners would utter sentence (a) in response to the question produced by a speaker who asks, 'What was she trying to lose?' by emphasising the final content word. This case is not problematic since it corresponds with making sentence- or phrase-final words prominent in Thai, which primary stress falls on the last syllable of the word or phrase only and it almost always falls on the same syllable or word. When Thai learners are asked, 'What was she trying to do? or 'What was she doing?' they may still utter the sentence (a) which is considered adequate as a response to both questions. However, the point that the researcher would like to raise here is that there is L1 interference that hinders them from producing sentence (b) and (c), or it is still a problem for them to shift the marking of the focus from (a) to (b) and (c) due to L1, resulting in misplacement of the sentence accent. The choice of accents or tones that they employ to highlight information can vary, for example, H*, L+H*, but they can differ from those of native English speakers in terms of misaligning the nuclear accent with the stressed syllable, as in Mennen's (2006, 2007) viewpoint.

As discussed above, it reflects that the differences between L1 and L2 intonation exist. Differences as to L1 phonology cause the characteristics of a person's first language to transfer to the characteristics of the person's second (L2) language or speech. It is getting harder for L2 learners to acquire L2 phonology throughout the course of time, especially when there is an extreme difference between L1 and L2, for example, between individual sounds, or between suprasegmental features. Regarding the latter, intonational theories generally inform about this as a barrier or difficulty in L2 acquisition for L2 learners who want to communicate successfully in L2 and who want to approximate native English speaker standard as closely as possible. Regarding to intonation and lexical tone, they are one of suprasegmental features and they are concerning the use of pitch. The shared characteristics of pitch for intonation and for lexical tone does not prevent a tone language like Thai from using the pitch for an utterance. Therefore, it is likely that Thai learners have Thai-accented English as they transfer lexical tones onto an English word contributing to word meaning, as well as intonation for an entire utterance. And that finally results in the characteristics of L2 intonation produced by Thai learners.

1.2 Definitions of Key Terminology

There are various different technical terms in the literature and studies into prosodic features such as intonation. Researchers usually find one term better than another in the sense of having meanings more appropriate and specific to the context of their studies. This present study is no exception, and so definitions of the key terminology below are provided unless there are explicit definitions given in the literature review.

Native speakers — In this study, the term refers to native speakers of English, or English native speakers, or native English speakers only. Or otherwise, a specific term is used to give different definition, such as Thai native speakers.

Intonation pattern — This term is used interchangeably in this study with the terms 'nuclear contour', or 'nuclear accent contour', or 'nuclear configuration'. All refer to the major pitch movement (nuclear tone) in the British tradition (e.g., fall, rise, rise-fall), or compositions of pitch accents and edge tones in the American school (e.g., H* LL%, L* HH%).

Accent type — This term is restricted to choices of pitch accent types only. In the Autosegmental-Metrical (AM) description, pitch accents and nuclear pitch accents are not distinctive, but the latter is by default the last or rightmost pitch accent associated with most prominent syllable at the right edge of an intermediate or intonational phrase. Types or choices for (nuclear) pitch accents can be monotonal (H*, L*), or bitonal (L*+H), or tritonal (L*HL) pitch accents.

Accent placement — This term is used to refer to the location of (nuclear) pitch accent placement on the accented syllables of the focused words. It is associated with other terms such as 'tonicity', or 'nuclear/tonic syllable placement' and 'nuclear tone' in the British tradition.

4

Nucleus	This term is used to refer to the most prominent syllable of the tone unit or intermediate/intonational phrase. Another name for the nucleus is the tonic (nuclear) syllable. A nucleus or a tonic syllable is one which carries a tone. It has a noticeable degree of prominence.
Tone	This term refers to an identifiable movement or level of pitch that is used in a linguistically contrastive way, such as on the nucleus or tonic syllable in a tone unit in the British school. In this study, tone is used interchangeably with the terms 'tonal targets' or 'pitch targets' (H, L) in the Autosegmental-Metrical theory.
Lexical tone	This term is restricted to 'tones' in tone languages like Thai, which are used only for distinguishing word meanings, or for indicating various aspects of grammar. Usually, lexical tones are identified as being a property of individual syllables, while an 'intonational' or 'prosodic' tone may be spread over many syllables and is associated with a higher level of phrasing (tone groups, or intermediate/intonational phrases).
Deaccentuation	This term is used in instances where the material before or after the focused word is deaccented; otherwise, the specific terms 'pre-focal deaccenting' and 'post-focal deaccenting' are respectively used.
Pitch range compression	This term is used in instances where the material following the focused word is not deaccented. Rather, post-focal materials are still accented but the pitch is lower relative to the preceding one, creating a phenomenon like a downstepped accent.

There have been a number of studies of the second language acquisition of intonation that identify the problems which non-native learners experience when they are learning the intonation system of a new language. These studies, such as those by McGory (1997), Mennen (1999) and Ramiŕez Verdugo (2002), have shown that L2 learners have difficulties with the correct placement of prominence, and that learners use the wrong choice of nuclear tone such

as using rising instead of falling intonation or vice versa (Adams and Munro, 1978; Lepetit, 1989; Wennerstrom, 1994**;** Hewings, 1995; Pickering, 2001). It has also been reported that first language transfer to L2 learners' production of intonation in a second language occurs in terms of intelligibility and foreign accentedness (van Els and de Bot, 1987; Ortega-Llebaria and Colantoni, 2014). In marking focus, Wennerstrom (1994) reported that Thai, Spanish and Japanese L2 learners did not mark new and contrastive information with a higher pitch in the same way as native English speakers. This difference in the use of acoustic cues may be attributed to first language transfer.

To summarise, all of the difficulties discussed above are plausible causes for intonational problems among Thai learners of English as a foreign language. The sources of difficulties that they encounter may range from the phonological structure of intonation, such as nuclear tone choices and the locations of nuclear accent placement, as well as the phonetic implementation of these. In the case of Thai learners, if the nuclear tone that they choose to use is not the problem, it is possible that they not only misplace word stress and phrasal stress but also misalign the nuclear tone with the stressed syllable. Evidence concerning this has been increasingly reported and can be drawn from the relevant literature for the acquisition of various foreign languages in addition to English (Lepetit, 1989; Low and Grabe, 1999; Mennen, 1999; Grabe *et al.,* 2000; Atterer and Ladd, 2004; O'Brien and Gut, 2010).

1.3 Research Gaps and Objectives

In the literature on intonation, however, research that provides evidence in this area is as yet far from sufficient. For the time being, it can be said that a number of studies on L2 prosody, intonation and focus within and across languages have received more attention (e.g., Lepetit, 1989; Hewings, 1995; Wennerstrom, 1994; Pickering, 2001; Xu and Wang, 2001; Grabe, Kochanski and Coleman, 2005; Lange, 2007; Graham and Post, 2018), but these studies are limited to certain languages; for example, English, German, Spanish, Korean and Mandarin Chinese. Chen *et al.* (2001) investigated English sentence stress by Mandarin speakers, whose first language is a tone language similar to Thai and they found that Mandarin speakers produced sentence stress with higher f0 and shorter duration than American English speakers. Even though Wennerstrom (1994) found that Thai participants did not employ higher pitch to signal contrast in meaning, as American English speakers did, issues with respect to tonal alignment and scaling were not investigated. Therefore, this motivation, alongside the need to

fill various other research gaps, led to the current study being established to examine the intonation and focus produced by Thai learners.

In this study, the prosodic features of special interest include choices of nuclear accent types in the expression of focus and acoustic-phonetic correlates of accentual prominence which contribute to the realisation of focus, as well as issues of tonal alignment and scaling. Keeping in mind the fact that the prosodic features in question involve plenty of variability and several of the difficulties mentioned above, it makes sense for the current study to start by examining one particular issue and then letting other researchers interested in this matter investigate other relevant issues.

As far as a language like Thai is concerned compared to others, the situation is different. Apart from its lexical tones, prosodic features such as intonation in Thai can be said to be relatively understudied. There is only a small handful of studies on intonation or even stress in the language (e.g., Luangthongkum, 1977; Luksaneeyanawin, 1983; Potisuk, Gandour and Harper, 1994; Warotamasikkhadit, 2000). Apart from Wennerstrom's (1994) study, little is as yet known about how Thai learners of English mark focus prosodically. In the area regarding the marking of focus prosodically, in particular, issues of the phonetic alignment of focus remain, creating a significant gap in knowledge for research to address in the future. Therefore, the current study sees the advantages of trying to fill this gap, and to attempt to seek more detailed linguistic information by conducting a comparative study of intonation and the marking of focus in L1 and L2 English.

The principal objectives of this study are twofold. Firstly, it aims to investigate the prosodic elements of speech (intonation and focus) produced by Thai learners, whose L1 is a tonal language. Secondly, it is concerned to make a cross-language comparison. Therefore, those prosodic features of native speakers of English were also investigated and taken as a baseline or point of reference. To achieve these objectives, the corresponding research questions described in section 1.5 below were set.

1.4 Scope and Limitations of the Study

Since the intonation systems of a language are linked with several factors, it is essential to demarcate the scope of the current study. As clearly stated in the previous sections, this study has the ultimate aim to investigate the intonational features and focus of English as an L1 from native English speakers and as an L2 from native Thai speakers. More specifically, the prosodic features that the current study consider, and which are relevant to the research questions include the following:

Table 1. 1 Scope of the current study

No.	Scope of the study
1	The nuclear pitch accents in intonational phrases used to mark a narrow focus of information in declarative sentences.
2	The acoustic-phonetic characteristics of post-lexical prominence or accentual prominence in the expression of narrow focus.
3	The tonal alignment and scaling of a rising accent associated with the marking of narrow focus.

Several features allow for a cross-language comparison to be conducted in this study. The first point comes from linguistic characteristics specific to Thai. The language itself has lexical tones and also uses intonation, where the latter operates at sentence level to convey meaning such as in statements or questions in a similar way to English. The second point is that the present study investigates the L2 English uttered by Thai learners and L1 English uttered by native English speakers. In this respect, this study assumes that the L2 English characteristics of the Thai learners resembles the L1 English characteristics of native English speakers to a certain extent. Consequently, it is feasible for this study to make comparisons in terms of intonation and the marking of focus between Thai learners and native speakers of English.

1.5 Research Questions

Given the primary objectives, an experimental investigation was designed to examine both the phonological and phonetic realisations of intonational prominence (nuclear pitch accenting) used in marking focus. The phonological study also included a description of other phonological means in the expression of focus, where applicable. The phonetic study also investigated the alignment and scaling of tonal targets in rising accents in the expression of focus. Thus, this study considered two kinds of qualitative and quantitative descriptions in order to address the

following research questions in Table 1.2. The hypotheses are formulated and presented after reviewing the relevant literature in Chapter 2.

Table 1. 2 Research questions for qualitative and quantitative analyses

	Research questions for qualitative analysis	**Chapters**
RQ1	What are the accent types that native speakers of English use to mark narrow informative focus?	Chapter 4
RQ2	What are the accent types that Thai learners of English use to mark narrow informative focus?	Chapter 4
RQ3	To what extent do Thai learners differ from native English speakers in terms of the use of accent types to mark narrow informative focus?	Chapter 4
RQ4	To what extent do the L2 Intonation Learning theory (LILt) and the Contrastive Analysis Hypothesis (CAH) account for the L2 English intonation for focus in this study?	Chapter 7
	Research questions for quantitative analysis	
RQ5	To what extent do the factors such as levels of English proficiency, gender and focus positions affect acoustic characteristics of focus used by Thai learners for the marking of narrow informative focus in terms of the use of f0, intensity and duration?	Chapter 5
RQ6	To what extent do the factors such levels of English proficiency, gender and focus positions affect the temporal relations between the tonal movement in rising pitch accents and segmental strings in narrow-focused words produced by Thai learners of English?	Chapter 6

1.6 Organisation of the Study

The present study investigates one of the most complex but important prosodic domains, which is intonation, and one of the central functions of intonation, which is focus, in the L2 English produced by Thai learners and L1 English produced by native English speakers. A reading-aloud task was designed and administered in laboratory experiments in order to capture and understand the prosodic features in question by means of phonological and phonetic analyses. The two analyses are presented in two separate chapters, while the overall organisation of the study is briefly outlined below.

Chapter 2 presents a literature review of relevant studies on prosody, intonation and focus. This chapter is composed of 6 sections. Section 2.2 introduces the definitions of prosody, intonation and focus specifically used in this study. Sections 2.3 and 2.4 provide detailed accounts of relevant research in English and Thai as well as models of L2 speech acquisition respectively. Section 2.5 then describes previous research on L2 English intonation and focus, as well as factors affecting L2 acquisition. It also introduces English as a foreign language as a model in the English language teaching (ELT) context in Thailand. Section 2.6 elaborates on the theoretical frameworks of intonational phonology and models of intonation transcription with evidence from previous studies. Section 2.7 provides a summary of the chapter. The chapter ends with a restatement of the research questions and hypotheses as formulated according to the results of the literature review.

Chapter 3 describes the research methodology used for this study. This includes the design of the experiments, recruitment of the participants, stimuli, and data collection and recording procedures. This study collected data from read speech in laboratory experiments (Cohn *et al.*, 2012). In doing this, it follows the same basic assumption as in previous studies that have suggested that read speech, story retelling, and map tasks as well as spontaneous speech can be utilised to investigate prosodic phenomena (Grabe 1998; Grabe, Post and Nolan, 2001; Fletcher *et al.*, 2002; Grice and Savino, 2003; Grabe, Kochanski and Coleman, 2005; Lickley, Schepman and Ladd, 2005). Furthermore, experiments in the laboratory, such as those recording speech, can be considered very beneficial in enabling the capture of phonological and phonetic phenomena (Lickley, Schepman and Ladd, 2005; Xu 2010).

In addition, the chapter gives more detailed accounts of the annotation and transcription of intonational patterns, as well as of the analysis of data using R software. In the section on data measurement and analysis, the phonological analysis of intonation is dealt with in terms of pitch events alongside the marking of focus with nuclear pitch accents in accordance with the Autosegmental-Metrical theory and the ToBI system. Then, the acoustic-phonetic analysis is described which is used in examining the acoustic correlates of intonation (nuclear pitch accents used for marking focus) and their tonal alignment and scaling.

However, this study is still limited in the design of speech materials in English. For example, not all of the stimuli could be constructed with fully voiced words or put into carrier sentences of the same length. In addition, there is yet no work on the development of Thai-ToBI, which would be useful when making cross-language comparisons. The development of such a language-specific system for prosodic transcription is far beyond the scope of the current study.

Chapter 4 presents the results of the study concerning the phonological analysis of intonation in terms of using nuclear pitch accents for the marking of narrow focus. To answer the research questions, the results are reported and organised in the following order. Firstly, the pitch accent types for marking focus as used by native English speakers are reported, and secondly those used by Thai learners of English are described. Furthermore, similarities and differences in the marking of narrow focus between L1 and L2 English are illustrated, particularly in terms of the choices of accent types and deaccenting post-focal materials. Graphic visualisations in R are presented where relevant.

Chapter 5 presents the results of the present study regarding the acoustic characteristics of intonation and focus. To answer the research questions, the results are reported and organised in the following order. The acoustic correlates involved in the realisation of intonation and focus (fundamental frequency, duration and intensity) used by native English speakers are first reported, followed by those used by Thai learners of English. Similarities and differences in those acoustic-phonetic correlates of intonation and focus between L1 English and L2 English are illustrated. Graphic visualisations in R are presented where relevant.

Chapter 6 presents the results of the present study regarding the phonetic characteristics of intonation and focus. To answer the research questions, the results are reported and organised in the following order. The phonetic characteristics in terms of the alignment and scaling of tonal targets in the rising pitch accent used in marking focus is examined. The results of tonal alignment and scaling produced by native English speakers were first reported, followed by those used by Thai learners of English. Similarities and differences in those phonetic characteristics between L1 English and L2 English are illustrated. Graphic visualisations in R are presented where relevant.

Chapter 7 presents discussions derived from the results of the findings concerning the phonological analysis of intonation in terms of the use of nuclear pitch accents in marking narrow focus (Chapter 4). Then findings regarding the acoustic-phonetic analysis of intonation and focus (Chapter 5) are considered. These discussions are also linked to the literature review in Chapter 2, considering how learning models (the L2 Intonation Learning theory (LILt) and the Contrastive Analysis Hypothesis (CAH)) can account for the results of the present study, and how the findings can make contributions to existing learning models.

Chapter 8 provides a summary of the main findings corresponding to the research questions, leading to a statement of the final conclusions of the study. The chapter concludes with discussions of the contributions of the present study to several existing fields and enterprises, including research into L2 prosody, intonation and focus, and implications for the design of syllabuses concerning the teaching and learning of L2 English. In addition, suggestions are given concerning investigations of these prosodic features in new areas of research; for instance, in the context of English as a *lingua franca* in the ASEAN community. References and appendices are provided following this chapter.

Chapter 2. Literature Review

2.1 Introduction

This chapter gives descriptive accounts of the literature and previous studies regarding prosodic features as well as other relevant phenomena which contribute to the realisation of focus. The chapter starts with an introduction to and the definitions of prosody, intonation and focus specific to the current study. Sections 2.3 and 2.4 then present language-specific descriptions of prosody, intonation and focus in English and Thai. Section 2.5 considers previous cross-linguistic research and studies of L2 prosody, intonation and focus. Section 2.6 gives accounts of the theoretical framework of intonation in terms of the American and British approaches, and the models used in the analysis of intonation: ToBI and IViE. Finally, section 2.7 ends the chapter with a summary and restatements of the research questions and hypotheses.

2.2 Introduction to Prosody, Intonation and Focus

2.2.1 Definition of prosody

A stream of speech in any language is viewed as being composed of basic strings of consonant and vowel sounds. Descriptions of such individual segmental phonemes are already known to be the subject of segmental phonology. However, there are still several other aspects of speech other than the segments that occur and are involved in speech in a language. For example, in a tone language like Thai, lexical tones are used jointly with a string of consonants and vowels to distinguish between words with different lexical meanings. For example, the word 'māː' with a mid-tone refers to a verb meaning 'to come', whereas the word 'máː' with a high tone refers to a noun meaning 'a horse'. In this sense, the choices of tones of this kind alter the lexical identity of syllables or words; therefore, lexical tones are considered to be properties of syllables or words in Thai. In English, it can be said that the language does have tones, but they are not used in the same way as in Thai. English makes use of tones for intonation, and such a use of tones does not influence the lexical meaning of the word (Wells, 2006). Furthermore, stress in English is another instance of a factor involved in speech. English makes use of stress to make a syllable more prominent than its neighbours, and such uses of stress are considered to be among the properties of a syllable. Therefore, lexical tones in Thai as well as stress and

intonational tones in English are matters of suprasegmental or prosodic structure in each language's phonology.

The terms suprasegmental, and prosody are used interchangeably in the literature and in research into aspects of speech that typically occur beyond the level of phonemes or segments (Cruttenden, 1997; Roach, 2009). Roach (2009) explains that the term 'suprasegmental' is likely to be used most often in American work, whereas the term 'prosody' is used more often by the British. The relevant domains are not limited to lexical tones, intonation, and stress, but also include rhythm and speech rate, and other factors. Meanwhile, Cruttenden (1997, p. 1) characterises prosody as follows:

> Prosodic features may extend over varying domains: sometimes over relatively short stretch of utterances, like one syllable or one morpheme or one word … sometimes over relatively longer stretch of utterances, like one phrase, or one clause, or one sentence …

He further points out that the former domain is associated with the lexical tones of tone languages, and the latter with intonational tones of intonation languages. However, he accepts that there is still no clear distinction between the two. This is because a single word, in reality, can itself become a complete utterance. In addition, it is not surprising that the term 'prosody' itself is quite often used interchangeably with the term 'intonation' to refer to the same concept. This is because the two terms cover a wide range of similar phenomena, as discussed in the following paragraphs and in section 2.2.2.

Prosodic features refer to a number of larger units of speech which encompass individual consonants and vowels, according to Lehiste (1970) and Cruttenden (1997). These prosodic features include, for example, vocal pitch, length and loudness. They are also traditionally considered to be three underlying prosodic elements of speech, all of which affect speech beyond groups of segments and individual segments and which are regularly used for linguistic purposes. For instance, variations in the pitch of a syllable or word are studied in relation to lexical tones and lexical meaning, and pitch, length and loudness in relation to word stress or lexical prominence. It is not surprising that linguists typically exploit those prosodic features to examine the linguistic characteristics and functions of intonation, if they view intonation as involving sentence stress or post-lexical prominence. In this case, linguists investigate intonation as one specific prosodic domain by examining its perceptual correlates. Apart from this, linguists can investigate intonation at the phonetic level by examining its acoustic-phonetic

correlates at the same time. These acoustic parameters of intonation include fundamental frequency, duration, and intensity (Cutler, Dahan and van Donselaar, 1997).

To conclude, it is quite clear from the paragraphs above that the definitions and applications of the term 'prosody' vary. For research into the prosody of a given language, as in the current study, it is good practice as a starting point to understand such terminological distinctions before embarking on any investigation of prosodic domains and features. It is also important for this study to take a stance and to make the clear point that the term 'prosody' in this study will be used as an umbrella term, covering the concept of intonation under investigation.

2.2.2 Definition of intonation

Subsumed under prosody, intonation constitutes one of the most crucial and complex elements of speech in intonational languages like English. Intonation may be defined in various different ways from one scholar to another (Cruttenden, 1997; Hirst and Di Cristo, 1998; Wells, 2006; Grice and Baumann, 2007; Ladd, 2008; Roach, 2009). Some definitions of intonation are exemplified as follows.

Starting from the definition that seems to be most easily understood and typically heard in English language classrooms, intonation is about how we say things, rather than what we say. It concerns the melody of speech (Wells, 2006), and it involves complex unconscious mechanisms in pronunciation. These are likely to be among several reasons why intonation lessons are neglected in pronunciation classes. However, the term 'intonation' becomes more complex when linguists refer to its broader and narrower meanings involving issues of levels of representation and analysis when describing intonation. Defined in terms of a narrow meaning, intonation in this sense is simply viewed as a matter of variations in pitch patterns over the whole sentence that speakers use to convey meaning or to communicate with listeners (Xu, 2017). In this case, speakers modulate the patterns of the pitch of the voice, resulting in various patterns of intonation corresponding to their intended meaning throughout the course of a conversation. Similarly, Cruttenden (1997, p. 7) viewed intonation as a phenomenon involving "the occurrence of recurring pitch patterns, each of which is used with a set of relatively consistent meanings, either on single words or on groups of words of varying length". Figure 2.1 below illustrates how intonation is defined related to levels of representation and analysis in describing intonation according to Hirst and Di Cristo (1998, p. 7).
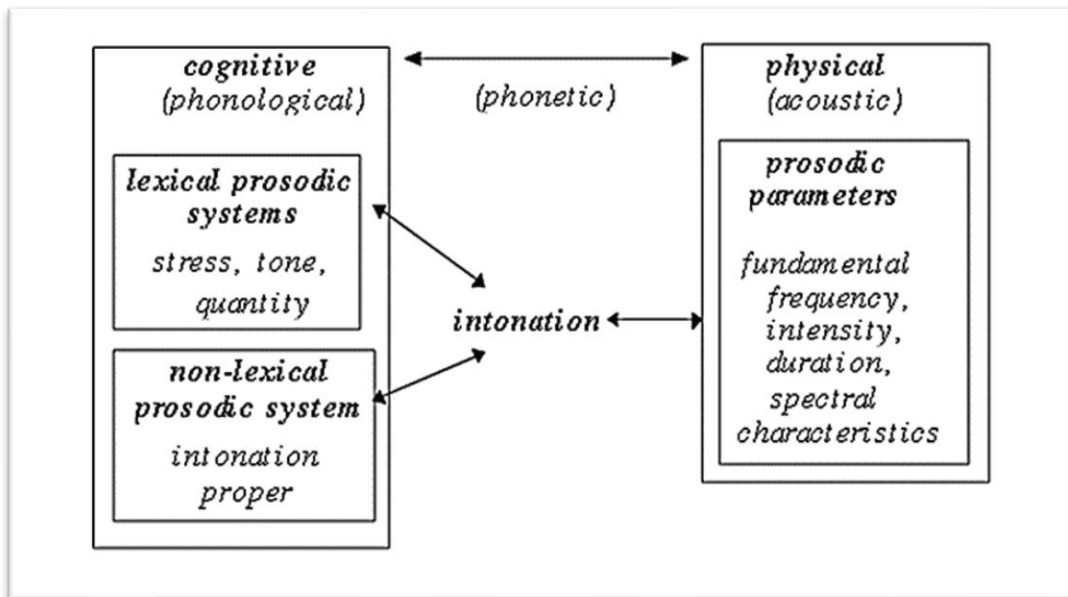
Figure 2. 1 General prosodic characteristics of languages (Hirst and Di Cristo, 1998, p. 7)

However, intonation can also be defined in terms of having a broader meaning. Intonation in this sense is viewed as a mixture of several prosodic features. According to Roach (2009) and Xu (2017), intonation is here used in an equivalent manner to prosody such that it covers other prosodic features such as duration, intensity, voice quality, and the like. This viewpoint is in line with Crystal's (1969, p. 195) statement that "intonation is viewed, not as a single system of contours, levels, etc., but as a complex of features from different prosodic systems … the most central are tone, pitch-range and loudness, with rhythmicality and tempo closely related". Similarly, Cole (2015) defines intonation as "the systematic use of suprasegmental properties (or for some authors, just pitch) at the phrase or utterance level to mark linguistic information beyond word identity … post-lexical information" (p. 2). Using intonation with a broader meaning is similar to its use with a narrow meaning in that both kinds of meanings involve the use of pitch operating at the level of the utterance to convey post-lexical meanings or information, as shown in Figure 2.1.

Having said that, research into intonation can be divided into two groups when classified according to which definition is used. The first investigates intonation by mainly examining the role of vocal pitch (fundamental frequency), and the second investigates intonation by examining vocal pitch together with duration, intensity, and other variables. To date, there is a tendency for more recent studies to approach the intonation of a language by means of using a

broader definition. This is because it seems to reflect the general use of intonation in actual speech to a greater extent. In the present study, 'intonation' is defined as part of the umbrella concept of 'prosody', so that duration and intensity, besides fundamental frequency, are included in the investigation, following a number of previous studies (e.g. Cutler, Dahan and van Donselaar, 1997; Zerbian, 2013; Mennen and De Leeuw, 2014) and following the literature which states that prosodic prominence involves three main acoustic cues: f0, duration and intensity (Gussenhoven, 1983; Pierrehumbert and Hirschberg, 1990; Ladd, 2008). More specifically, this study has adopted Ladd's (2008, p. 4, italics in original) definition of intonation as referring "to the use of *suprasegmental* phonetic features to convey 'postlexical' or *sentence-level* pragmatic meanings in a *linguistically structured* way". Details about acoustic-phonetic cues are provided in sections 2.2.6, 2.2.7 and 2.5.4.

2.2.3 Form and function of intonation

Intonation has its own forms and functions (Gussenhoven, 2004; Wells, 2006; Ladd, 2008). Form refers to various intonation patterns that can be recognised by observing the shape of pitch movement (e.g., fall, rise, fall-rise). It can also refer to the structure of intonation. For example, in the British school (e.g., Crystal, 1969; Cruttenden, 1997; Wells, 2006), each tone group has its own internal structure, composed of pre-head, head, tonic syllable/nucleus and tail. Within the domain of the tone group, the nucleus is the most crucial element in affecting the strongest stressed syllable or nuclear syllable in the utterance. The nucleus is also associated with the choice of tone. In the American school (e.g., Beckman and Pierrhumbert, 1986; Ladd, 2008), each intonational phrase has at least one intermediate phrase with one nuclear pitch accent. Additionally, intonation in AM theory is treated in terms of pitch accents, phrase accents and boundary tones. More detailed descriptions of these two approaches to intonation are given later in the section 2.6 covering the theoretical framework and transcription models of intonation.

Meanwhile, functions refer to the meanings that each intonation pattern carries. There are a variety of intonational functions, just as there are a variety of forms, and these are classified differently from scholar to scholar. Roach (2009) divided the functions of intonation into four categories: grammatical, accentual, discursive, and attitudinal functions. O'Connor and Arnold (1973) stated that one of the most obvious roles of intonation is to express attitudes and emotions conveyed by different tones in relation to sentence types. Ladefoged and Johnson (2011) pointed out that intonation can be used to express different kinds of information,

17

including the individual identity of the speaker (e.g., gender, age, emotional state), apart from other functions like syntactic and discourse functions. According to Wells (2006), the functions of intonation can be elaborated as follows:

- *Attitudinal function, such as to show shock or surprise, pleasure or anger, interest or boredom. This function can be expressed by tone.*
- *Grammatical function, such as to identify grammatical structure such as the beginning and the end of a clause and sentence, as well as to distinguish questions from statements. The former can be expressed by tonality, and the latter by tone.*
- *Focusing function, such as to show what information in the utterance is new and what is already known. This function can be mainly expressed by tonicity and tone choice.*
- *Discourse function, such as to show how clauses and sentences go together in spoken discourse, to show how they contrast or cohere.*
- *Psychological function, such as to organize speech into units that are easy to perceive, memorize and perform. This function can be expressed by tonality.*
- *Indexical function, such as to act as a marker of personal or social identity. (Wells, 2006, pp. 11-12).*

Despite these differences in classification, there often seem to be cases of overlapping intonational functions. This accords with Roach's (2009, p. 147) point of view about the functions of intonation in that:

> … it is difficult to see how they could be treated as separate; for example, the placement of tonic stress is closely linked to the presentation of "new" information, while the question/statement distinction and the indication of contrast seem to be equally important in grammar and discourse. What seems to be common to accentual, grammatical and discourse functions is the indication, by means of intonation, of the relationship between some linguistic element and the context in which it occurs.

To conclude, it is obvious that there are many different points of view about the definitions, forms and functions of intonation. To deal with this, the current study takes a stance by defining the term 'intonation' as part of 'prosody.' Intonation refers to or is realised through variations in pitch patterns over a phrase or an utterance and this applies to the sense of intonation used in this study. With respect to its functions, it seems that the overall concepts relate to the need for communication. In general, the intonational functions discussed above are closely related to communicating the linguistic (e.g., grammatical and discourse functions), paralinguistic (e.g.,

attitudinal function), and non-linguistic (e.g., indexical function) information to accommodate the flow of communication. In this study, one of the above-mentioned communicative functions of intonation is selected for investigation; that is, the focusing function.

2.2.4 Definition of focus

As discussed previously, focus is related to one of the communicative functions of intonation. It can refer to the highlighting of a particular piece of information or a part of a sentence as a response to someone's questions (Bolinger, 1972b; Ladd, 2008). This is known as prosodic focus, narrow focus, or simply focus (Xu and Xu, 2005; Xu, 2017). In research into focus, the term is defined differently based on the specific perspectives and theoretical approaches of individual researchers. For instance, focus can be considered to be one of the basic notions within information structure besides givenness and topic, whereas complementary notions of information structure include background, newness and comment. According to Krifka (2008), focus "indicates the presence of alternatives that are relevant for the interpretation of linguistic expressions" (p. 247), and givenness indicates that "the denotation of an expression is present in the immediate CG [Common Ground] content" (p. 262). Topic is "the entity that a speaker identifies, about which then information, the comment, is given" (Krifka, 2008, p. 265). In addition, Jackendoff (1997, p. 230) defined the term 'focus' as being "the information in the sentence that is assumed by a speaker not to be shared by him and the hearer". According to Büring (2007, 2009, 2012), focus is informative since it is related to new information in the sentence, and that new information is required in order to answer a question. In this sense, it is quite often seen that pairs of questions and answers are commonly used as a tool when investigating focus in a given language.

Focus is usually discussed in terms of two dimensions: size or breadth and type. Focus size includes narrow and broad focus (Selkirk, 1995; Gussenhoven, 1999, 2008; Ladd 1984, 2008), where the distinction between them is indicated by the size of the focused domain and the number of words in that domain. Breen *et al.* (2010) emphasised that narrow focus can involve part of an utterance or one word that is produced with greater prominence, while broad focus can be given to the whole utterance and the entire event conveyed by that utterance. Given below are examples of narrow and broad focus size.

(1)   a.   Q: Who read a novel?

               A: [MARY]~FOCUS~ read a novel

    b.   Q: What did Mary read?

               A: Mary read a [NOVEL]~FOCUS~

    c.   Q: What did Mary do with a novel?

               A: Mary [READ]~FOCUS~ a novel

    d.   Q: What happened?

               A: [Mary read a novel]~FOCUS~

As in the examples above, new information in sentences (1a), (1b) and (1c) tends to become accented by applying pitch accents to the first syllable of the words 'Mary', 'read', and 'novel'. This are cases of narrow focus, with the accent sometimes referred to as a focal accent. More specifically, a question like 'Who read a novel?' is aimed at seeking new information about the person who did read a novel. Thus, in the answer, special emphasis is expected to be placed on the word 'Mary' with the help of a nuclear pitch accent. The segmental domain corresponding to narrow focus is limited to the word 'Mary'. The rest of the post-focal words bear no pitch accents.

On the other hand, there is broad focus in sentence (1d) since there is a tendency for the first and last content words of the utterance in English to be accented. More specifically, the question is not aimed at eliciting any specific new information. It is asked to restate the information given previously, or it broadly focuses on the entire event. In this case, the whole sentence forms one focal domain with two words bearing pitch accents, 'Mary' and 'novel'. In this case, the word 'Mary' is prenuclear accented while the word 'novel' is nuclear accented. Besides this, the domain of focus of sentences can be larger constituents, ranging from a single word (i.e., the subject, or object) to the entire verb phrase or sentence. It is also interesting that the presence of focus and the location of a nuclear accent seem to correlate. However, whatever the focus size (a single word, a phrase), the distribution of nuclear accents is irrelevant.

For the second dimension, focus type includes contrastive and non-contrastive focus. According to Breen *et al.* (2010, p. 1046), contrastive focus refers to "indicating that the element in question is one of a set of explicit alternatives or serves to correct a specific item already present in the discourse." This is similar to Cruttenden's (1997, p. 82) definition of contrastive focus "as involving comparison within a limited set". Furthermore, there is another dimension of focus location which refers to any elements of a sentence (e.g., subject, verb, object) that focus

can be placed on. This dimension is evident in real communicative functions where the locations of focus need to be shifted in order to highlight a particular element of a sentence as most important.

In the literature about focus (e.g., Gussenhoven, 1985, 2008; Ladd, 2008; Büring, 2009, 2012), there are still controversial issues, even with the notion of focus itself. For instance, when it comes to narrow and contrastive focus, the two concepts are not wholly clear even though there may be an attempt to divide narrow focus into two categories of informative and contrastive (Gussenhoven, 2004; Zimmermann and Onea, 2011). Ladd (2008) defines narrow informative focus as mentioning what has already been said or highlighting new information in response to preceding wh-questions, and he defines narrow contrastive focus as either highlighting a concept or correcting a piece of information mentioned previously. These definitions seem to overlap to some extent such that it is not surprising that focus is treated differently between researchers.

For instance, any words can receive focus to convey new or contrastive information to listeners (Bolinger, 1972a, 1972b; Chafe 1974, 1976). Marking focus of this kind merely depends on the speaker's discretion and the context. On the other hand, according to Chomsky and Halle (1968), words would have normal stress unless they carry specific or contrastive meanings. And then those words would receive focus via contrastive stress. In addition, in the Focus-to-Accent (FTA) approach (Gussenhoven, 1983; Gussenhoven and Rietveld, 1985; Ladd, 2008), it is claimed that focus is realised by placing a nuclear pitch accent. When a syllable or word receives a nuclear pitch accent, it is the only syllable or word that has intonational prominence to carry meaning. By default, the location of the nuclear pitch accent is the last fully stressed syllable in an intonational phrase.

2.2.5 Marking focus

It is known that an element of a sentence can be singled out to identify the focus of information in different ways; for example, changing the word order, using particular lexical items, altering syntactic constructions, and using prosody. Since languages differ, the choices and combinations of these focus-marking means and strategies can vary depending on language-specific characteristics such as prosodic structure and segmental phonology, according to Koreman, Andreeva and Barry (2008). For instance, marking focus such as in English and

German can be prosodically achieved by means of pitch accent placement and its acoustic salience. In Thai, marking focus can be achieved by several means; for instance, either using particular word order, lexical focus particles, prosody, or combinations of these. Therefore, there is no one-to-one mapping between focus-marking strategies and focus based on a cross-language perspective and within particular languages (Zimmermann and Onea, 2011).

Given the prosodic marking of focus, a number of linguistic scholars have directed their attention to the phonological and phonetic means used (e.g., Cooper, Eady. and Mueller, 1985; Selkirk, 1995; Gussenhoven, 1999; Frota, 2002; Koreman, Andreeva and Barry, 2008; Adamou and Arvaniti, 2010; Jeon and Nolan, 2017). Some studies (Gussenhoven, 1999; Ladd, 2008; Xu and Xu, 2005) have reported that deaccentuation is one of the crucial phonological strategies used in marking focus in English. Some others (Eady and Cooper, 1986; Xu, 1999; Pell, 2001; Xu and Xu, 2005; Chen, Wang and Xu, 2009; Chen, 2010) report that compression in pitch range, duration and intensity can occur across languages regardless of whether lexical tone is present as in Thai, a pitch accent as in Japanese, or lexical stress as in English. Cross-linguistically, the phonological means employed to mark focus drawn from the relevant literature and studies (e.g., Beckman and Pierrehumbert, 1986; Gussenhoven, 1999; German, Pierrehumbert and Kaufmann, 2006; Féry and Kügler, 2008; Hartmann, 2008; Ladd, 2008) include the following:

- *nuclear pitch accent types and placement on new information*
- *re-phrasing or prosodic phrasing*
- *post focal pitch range compression*
- *lengthening of the focus domain*
- *deaccenting*

The phonetic means for marking focus drawn from the relevant literature and studies (e.g., Eady and Cooper, 1986; Cooper, Eady and Mueller, 1985; Nolan and Farrar, 1999; Xu and Xu, 2005; Baumann, Grice and Steindamm, 2006; Chen, 2006; Breen *et al.,* 2010; Genzel, Ishihara and Surányi, 2015; Nagy, 2015) include:

- *higher f0 peaks on the focused syllables*
- *increased duration of the focused syllables*
- *increased intensity of the focused syllables*

- *greater pitch excursion on accented words*

- *alignment and scaling in f0, such as a delay in the nuclear accent peak in narrow focus*

- *compressing the pitch range (and/or duration, or intensity) of the pre-focal and post-focal materials*

To conclude, it is obvious that the notion of focus and focus marking are treated in different ways from study to study. Therefore, it is necessary for the present study to take a stance on this issue. In this study, focus is generally defined as a communicative intonational function which can be mainly expressed by tonicity (i.e., the location of the nucleus placement) and nuclear tone choice as it is termed in the British approach to intonation. More specifically, this study takes narrow (informative or non-contrastive) focus into consideration in conducting experimentation. This study investigates the phonological characteristics of the marking of focus in terms of nuclear pitch accent types within the American approach to intonational phonology in the first instance. After that, the phonetic realisation of focus is examined. More details of studies of the acoustic-phonetic characteristics of focus are provided in sections 2.2.6, 2.2.7 and section 2.5.4.

2.2.6 Acoustic cues to prominence

Intonation and focus are relevant to the use of prominence at sentence level. Intonation can be considered as the manipulation of accentual prominence, and focus can be considered as the manipulation of focal prominence. When investigating the realisations of focus in intonation, several acoustic-phonetic correlates are taken into consideration, ranging from fundamental frequency (e.g., Eady and Cooper, 1986), duration (e.g., Breen *et al.*, 2010), intensity (e.g., Kochanski *et al.*, 2005), and vowel formant characteristics to more complex ones such as spectral tilt and spectral balance (e.g., Sluijter and Van Heuven, 1996; Heldner, 2003). The first three acoustic correlates of fundamental frequency, duration, and intensity are traditionally considered to be the essential and underlying prosodic elements of speech, and they have been widely studied within and across languages. Their influence varies and is language-specific; however, variations in fundamental frequency (f0) are best known in the literature for having an influential role pertaining to intonation and focus marking. For the present study, fundamental frequency, duration, and intensity are of special interest. Brief descriptions of these three acoustic-phonetic cues to prominence are given below.

In the literature, the acoustic term 'fundamental frequency (f0)' is not synonymous with the term 'intonation'. Even though intonation has the measure of f0 as its main phonetic component, f0, in turn, can be used for other purposes such as a use in lexical contrast. In terms of perception, the term pitch is commonly used, as it is the perceptual correlate of fundamental frequency. Definitions of fundamental frequency have been given by several scholars. For instance, according to Ladefoged and Johnson (2011, p. 24), f0 is "the number of complete repetitions (cycles) of variations in air pressure occurring in a second", which is normally measured in Hertz (Hz). On the other hand, pitch is also an auditory sensation. When we hear a regularly vibrating sound such as a vowel produced by the human voice, we hear a high pitch if the rate of vibration is high and a low pitch if the rate of vibration is low regardless of other acoustic properties. Ladd (2008, p. 5) defines f0 as "a physical property and pitch is its psychophysical correlate". Taking its influence into account, Wells (1986) found that higher f0 or pitch peak is commonly employed to mark the focus of information and is present within an intonational phrase. In addition, f0 is even higher when used to mark narrow and contrastive focus, and it is lower or stepped down following the focused elements (Cooper, Eady and Mueller, 1985; Xu and Xu, 2005). A higher pitch excursion has also been observed to signal prominence (Rietveld and Gussenhoven, 1985; Eady and Cooper, 1986).

Duration is referred to as the physical temporal distance or measurable time interval of any sound or noise. Duration is usually measured in seconds (s) or milliseconds (ms). When studying the time dimension from the point of view of what the listener hears, the term 'length' is common and is preferred for use in considering the listener's subjective impression of how long a sound lasts for. In English, for example, stressed syllables tend to be longer than unstressed ones. Taking its influence into account, Cooper, Eady and Mueller (1985) found that the duration of focused words was greater compared with non-focused words. This accords with the findings of Breen *et al.* (2010).

Intensity is referred to as the physical magnitude of any sound or noise. It is usually measured in decibels (dB). The intensity of a sound is dependent on both the amplitude of the sound wave and its frequency. In the study of speech, the term 'loudness' is also commonly used in describing intensity. It is actually perceived by listeners, representing the listener's sensation of how strong any sound or noise is. In an investigation on seven dialects of British and Irish English, for example, Kochanski *et al.* (2005) found that the interplay of intensity and duration represents prosodic cues as to focus, whereas f0 play a small role. They found loudness is the best acoustic cue to prominence. This accords with the findings of Turk and Sawusch (1996).

To conclude, fundamental frequency, intensity and duration are the acoustic correlates of pitch, loudness and length. In turn, pitch, loudness and length are their percepts, or perceptual correlates. Among these observable and measureable acoustic parameters, fundamental frequency is widely acknowledged to be the most important parameter in prosodic domains such as tone and intonation. Furthermore, in English intonation and focus, duration and intensity interplay with fundamental frequency in a systematic way. Out of the three, it is generally accepted that fundamental frequency (f0) is the most crucial variable (Hirst and Di Cristo, 1998, p. 4). In the present study, the two terms 'fundamental frequency(f0)' and 'pitch' are used interchangeably in the sense that f0 deals with the physical aspect of speech, and pitch involves the way it is perceived by humans; that is, as its auditory correlate. Intensity is used in terms of the acoustic correlate of loudness. Duration is the length of time that for a syllable it takes to be articulated.

2.2.7 Tonal alignment and scaling

The term 'alignment' refers to the "temporal coordination [of tonal targets] with consonants and vowels of the segmental string" (Ladd 2008, p. 169). In other words, it is the phonetic realisation of a tonal target on the horizontal time dimension, where the tonal targets associated with prominent syllables can be mapped onto the segmental strings in an utterance. Recent research within AM theory (Prieto, van Santen and Hirschberg, 1995; Arvaniti, Ladd and Mennen, 2006; Ladd, 2008; Prieto, 2011; Arvaniti, 2012; D'Imperio, 2012) have paid more attention to the phonetic implementation of intonation in terms of tonal alignment and scaling. This might be because tonal alignment offers the possibility of fine-gained phonetic interpretation in terms of examining differences in intonational meanings and categories. It also allows for comparative descriptions within and across languages. However, several issues with tonal alignment are in need of further investigation. For instance, there is an assumption that pitch accents have a primary association with prominent syllables and edge tones have a primary association with the edges of intermediate or intonational phrases. Both pitch accents and edge tones are also assumed to be associated with the tone-bearing units of a given language. In other words, AM theory leaves questions unanswered about the secondary association of the two. As a consequence, most studies of tonal alignment have directed attention to the investigation of the primary phonological association.

Tonal alignment plays a crucial role in examining the distinction between L+H* (the rising peak accent) and L*+H (the scooped accent). Both the L+H* and L*+H accents have a primary tonal association and alignment with the tone-bearing unit; that is, prominent syllables which can be recognised using the (*) diacritic. Nevertheless, the two accents are in different phonological categories. In terms of meanings, L+H* refers to the expression of contrastive focus and L*+H refers to pragmatic uncertainty (Pitrelli, Beckman and Hirschberg, 1994). More specifically, the H* pitch accent in L+H* is associated and aligned with the stressed/accented syllable while having the L tone as a leading tone. A similar explanation can be applied to the L*+H (scooped accent). This means that, in the L*+H accent, it is the L* pitch accent that is associated and aligned with the stressed/accented syllable while having the H tone as a trailing tone. In the literature (Pierrehumbert, 1980; Pierrehumbert and Steele, 1989; Grice, 1995; Arvaniti and Garding, to appear), such alignments of the L+H* and L*+H accents can be referred to as early or late peaks respectively.

More recent studies (Grabe *et al.,* 2000; Atterer and Ladd, 2004; Fletcher, Grabe and Warren, 2005) have revealed that tonal alignment is language- or dialect-specific. That is, within or across languages, the f0 peak (H tones) may be delayed and aligned after the end of the stressed syllables with which it is phonologically associated. On the other hand, the f0 may be aligned earlier in nuclear accents. For example, in L*+H accents, the L tonal target is aligned within accented syllables in English, as the work of Beckman and Pierrehumbert (1986) and Arvaniti and Garding (to appear) shows. In Greek, it has been reported that, in L*+H accents, both the L and H tones are aligned outside the accented syllables (Arvaniti, Ladd and Mennen, 1998). In addition, there are reports of the effects of several prosodic and phonological factors on tonal alignment. These factors include syllable duration or prosodic lengthening affecting differences in peak alignment between prenuclear and nuclear accents (Silverman and Pierrehumbert, 1990), peak alignment occurring beyond the stressed syllable in Greek (Arvaniti, Ladd and Mennen, 1998), syllable structure in Neapolitan Italian (D'Imperio, 2001), and syllable structure and speech rate in Peninsular Spanish (Prieto and Torreira, 2007).

Apart from tonal alignment, another important phonetic realisation of tonal targets is 'scaling'. This term refers to the f0 values of tonal targets in the speaker's pitch range or the f0/pitch level of linguistically distinctive pitch points (Ladd, 2008). There are several factors that affect tonal scaling; for instance, declination, tonal identity, and tonal context. Among these factors, a study by Daly and Warren (2001) on two dimensions of pitch range; pitch level and pitch span, has suggested that pitch range can have a substantial influence on tonal scaling. Paralinguistic

effects come into play for tonal scaling as well. Some studies (Pierrehumbert, 1980; Beckman and Pierrehumbert 1986; Arvaniti and Garding, to appear; Dilley, 2010) regarding declination, tonal identity, and tonal context are briefly outlined next.

Declination is a systematic and gradual lowering of tonal targets over the course of an utterance and it can also be reset throughout that utterance (Ladd, 1984; Connell and Ladd, 1990). This phenomenon has been traditionally analysed as a global phonetic effect triggered by an automatic physiological mechanism (e.g., Liberman, 1975). As for tonal identity, earlier studies (e.g., Gussenhoven and Rietveld, 1985; Arvaniti and Garding, to appear) reported that pitch range and span can affect the scaling of high and low tones. For example, pitch span can cause low tones to become lower and high tones to get higher. A difference in tonal scaling in the American English L*+H and L+H* was found in work by Arvaniti and Garding (to appear). They reported that the scaling of the L was lower and that of the H was higher in L*+H, when compared with those of the low and high targets in L+H*. This evidence strongly supports notions of the effects of tonal identity on tonal scaling. As for tonal context, there are cases of downstepping and upstepping.

Downstepping is a phenomenon where the second H tone of an HLH sequence is stepped down (e.g., Pierrehumbert, 1980; Beckman and Pierrehumbert, 1986). The downstep can be triggered by bi-tonal pitch accents; for example, L*+H !H* (Beckman and Pierrehumbert, 1986). Because of such contextual effects and differences in meanings, the !H* downstep and H* pitch accents are considered to represent independent phonological features so that they are contrastive (Ladd, 2008). Upstepping is another case of tonal context. This case is common in a combination of phrase accents and boundary tones. For example, the L% boundary tone in the H-L% edge tone is not realised as a lowering in pitch, but it is stepped up by the H- boundary tone in the H-L edge tone. A similar explanation can be made with the case of the H-H% edge tone.

To conclude, tonal alignment is the timing of the occurrence of an H or an L tone relative to the segmental string, and scaling is the value levels of fundamental frequency at which a high peak or a low valley occurs. Phonologically, tonal alignment has been used to examine the validity of proposed tonal categories (Pierrehumbert and Steele, 1989; Arvaniti, Ladd and Mennen, 1998; Ladd and Schepman, 2003). This is possible due to the assumption that a tone in a phonological category shows a relatively stable alignment at certain segmental landmarks in a text. Meanwhile, the scaling of tones helps in determining whether pitch movements are to be

considered tones or mere f0 transitions (Ladd and Schepman, 2003). Obviously, a number of factors can affect both tonal alignment and scaling, as mentioned previously. Furthermore, the alignment of tonal targets can vary depending upon a particular language and its varieties. Consequently, tonal targets are not necessarily aligned with segmental strings or anchors due to the effects of those factors. This has led to different interpretations of the segmental anchoring hypothesis, which assumes that the two tonal targets align independently. Having said that, tonal alignment remains a controversial issue among researchers. However, it is of interest for the current study to investigate and compare this phonetic alignment of rising accents in English focus as produced by native speakers and non-native speakers to make a contribution to the SLA research context. This is because of the fact that the two systems of L1 and L2 English may be different from each other either at the phonological level or at the phonetic level, and that the same phonological representations of accent types may be different from each other in terms of phonetic alignment.

2.3 Prosody, Intonation and Focus in English

2.3.1 Relative prominence: lexical and post-lexical

As one characteristic of a stress-accent language (or an intonation language), English lexical stress can be defined as the relative prominence given to a certain syllable in a word. Therefore, a syllable which stands out in this way is a prominent syllable. This phenomenon is referred to as 'word stress'. Its phonetic properties, such as longer duration, higher pitch, and greater intensity, result in the full articulation of that stressed syllable, making it more prominent than other syllables in the same word. These properties of stressed syllables are supported by several studies of acoustic correlates (Lieberman, 1960; Pierrehumbert, 1980; Cooper, Eady and Mueller, 1985; Xu and Xu, 2005; Ladd, 2008; Breen *et al.*, 2010).

Stress in English involves rather unpredictable positions in a word. In other words, it is not fully predictable; for instance, the word 'dinner' is made up of a pattern of strong and weak syllables whereas the word 'about' is made up of a pattern of weak and strong syllables. According to Roach (2004, p. 243, upper case in original), stress in English is "both FREE (in that any syllable is capable in principle of receiving stress) and FIXED (since it only rarely happens in a particular context that more than one stress placement is acceptable).'' Obviously, there is more than one level of stress; for example, primary stress, secondary stress, and a lack of stress.

The primary stress or the word with primary stress is considered to be the potential location of intonational prominence; that is, a pitch accent (Bolinger, 1958, 1989; Pierrehumbert, 1980; Fletcher and Evans, 2002; Ladd, 2008).

In stress placement, there is a phenomenon called 'stress-shift', where the normal location of a stressed syllable is altered due to word compounding or similar phenomena. For example, the words 'fifteenth' and 'afternoon' are stressed on the second and third syllables respectively, whereas the stress is on the first syllable in 'fifteenth place' and 'afternoon tea'. The positions of lexical stress in a word as in these examples may depend on certain rules applicable in English. It is known that the shifting of stress is accomplished in order to avoid two strong stresses occurring close together and to preserve the rhythmical regularity of the speech stream (Wells, 2006; Roach, 2009). Stress-shifting is widely known in the second language acquisition context as one of the difficulties that foreign learners experience (Cheng, 1987; Nava, 2008).

Apart from there being several levels of prominence at the word level (e.g., primary stress, secondary stress) with rhythmic alternations of strong and weak syllables, English is described as having higher levels of prominence signalled by pitch accents. Pitch accents are distinctive pitch patterns or tonal patterns associated with a stressed syllable in a word and they reflect relative prominence patterns in the utterance. Therefore, words with pitch accents are more prominently accented than other words in the same utterance. A pitch accent may be realised as a high (H) pitch target, or a low (L) pitch target relative to the speaker's pitch range for each intermediate or intonational phrase.

As a matter of fact, there can be several pitch accents distributed across an utterance, depending on the number of words with a syllable marked as metrically strong, as shown in Figure 2.2 below. As can be seen in the figure, the words "Nanny", "read" and "novel" can receive pitch prominence. However, there is only one word that receives more pitch prominence than other surrounding words. In the sentence, the word "novel" is considered to bear a nuclear pitch accent. The term 'nuclear pitch accent' can also be referred to as 'nuclear stress', or 'sentence stress'. In the literature (Cruttenden, 1997), nuclear pitch accents are defined as the perceptually most prominent accents and they are assumed to be the last pitch accents in a prosodic phrase. Therefore, a word with a nuclear pitch accent is significantly more prominent relative to other words with the pitch accent only. At this point, it is also worth noting that, if the nuclear pitch accented word comes first in the phrase, all the following words in that phrase must inevitably be deaccented or have no more pitch prominence (Gussenhoven, 1985; Ladd, 2008).

```
                                        x               nuclear pitch accent
          x         (x)                 x               pitch accent
          x          x                  x               primary stress
          x    x     x        x         x  x            syllable

    [     Nanny    read     a        novel    ]ip    ]IP
          |         |                 |         |       |
          H*        (H*)              H*        L-      L%      tone events
```
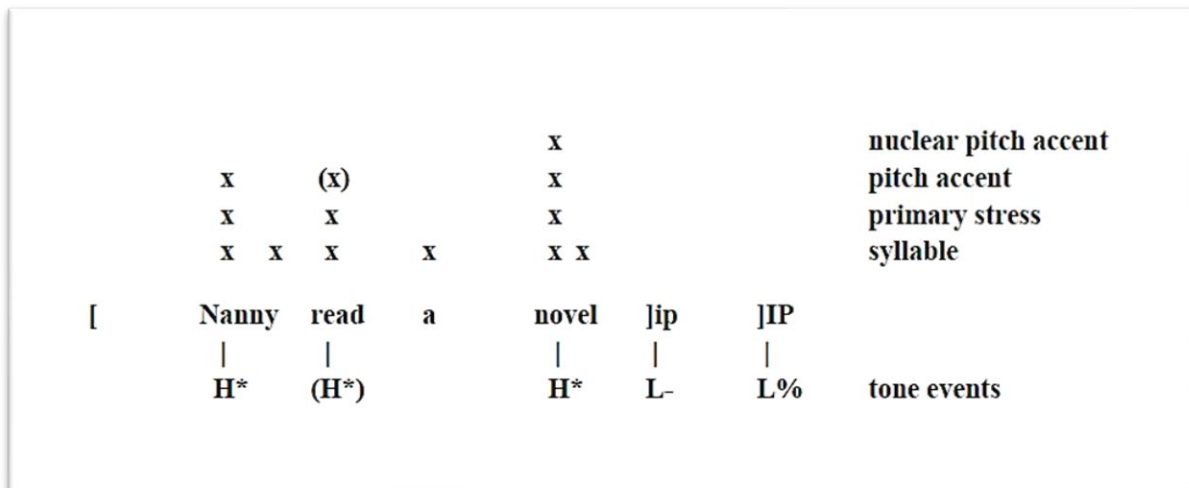
Figure 2. 2 Prominence structure and the intonational contour in the Autosegmental-Metrical (AM) model

In English, it is widely accepted that content words, for example, nouns, verbs, adjectives and adverbs, are categories that typically carry important meanings or information in an utterance. However, there are some exceptions where function words like pronouns and prepositions may receive prominence. In these cases, it apparently depends on the speaker's discretion and the context of the conversation between the interlocutors, such as when marking broad or narrow focus. Given that these words are content words located at or near to the end of the utterance, there is a general tendency for the last pitch accent to fall on that last content word. In turn, the so-called last pitch accent is a nuclear (pitch) accent which is the last accent with the additional prominence associated with the highest level of prosodic phrasing; that is, an intonational phrase.

2.3.2 Pitch accents and focus marking

When a word is given a pitch accent, the listener will hear the word as accented or more prominent compared to other words nearby. In other words, the speaker uses the pitch accent to draw the listener's attention to important information or to highlight the meaningful parts of the utterance. Using pitch accents like this demonstrates the marking of focus. Besides this, locations of pitch accent placement correspond with the locations of the focus. In this sense, pitch accented words are not only perceived more prominently by the listener, but are also used

to show the locations of highlighted information. The locations of information can be utterance-initial (Subject), utterance-medial (Verb), and utterance-final (Object). In cases of the last content word with the most prominence, it usually bears the nuclear pitch accent if located at the right-hand edge of the utterance. Words after this focused content word are deaccented with no pitch accents, while words before it can have a pitch accent but with less prominence. These characteristics are specific to post-focused words and pre-focused words respectively. The latter cases with less prominence can be considered to be prenuclear accents which are less relevant to the marking of focus (Welby, 2003).

Since there are many varieties of English, the marking of focus by means of pitch accent types varies. For instance, according to Ladd (2008), Ladd and Schepman (2003), and Dilley, Ladd and Schepman (2005), focus in English can be marked by placing nuclear pitch accents (H*, L+H*) on any focused elements of a sentence followed by post-focal deaccenting, as mentioned above. Xu and Xu (2005) reported that focus in English can be expressed by the L+H* accent and contrastive focus can be expressed by a wider pitch range. In American English, the marking of focus, including contrastive focus, can be accomplished by such pitch accents as the H* and L+H* accents (Pierrehumbert and Hirschberg, 1990; Selkirk, 1995). A gradual rising movement towards a high target (H* pitch accent) is used for informative focus whereas a steep rising movement in pitch from a low to a high tonal target (L+H* rising peak accent) is used for contrastive focus (Ito, Speer and Beckman, 2004). Hedberg and Sosa (2008) also reported on the use of either H* accents or (L+H*) rising accents for the expression of informative focus and contrastive focus in spontaneous speech in English.

Examples in Figure 2.2 below from Estebas-Vilaplana (2000) demonstrate in detail how focus is phonetically realised. Example (a) can be either broad focus or narrow focus on the object of the sentence. Example (b) demonstrates narrow focus on the verb whereas example (c) demonstrates narrow focus on the subject. These focused elements, which are capitalised, are realised with H* pitch accents. It is noticeable that pre-nuclear accent patterns occur, as in (a) where the nuclear accent falls on the last content word by default, and in (b) where the nuclear accent falls on the focused verb. Any post-focal materials are deaccented as in (b) and (c). The phrase-final nuclear contours of examples (a), (b), and (c) end with a low and flat pitch, resulting in L-L% edge tones.
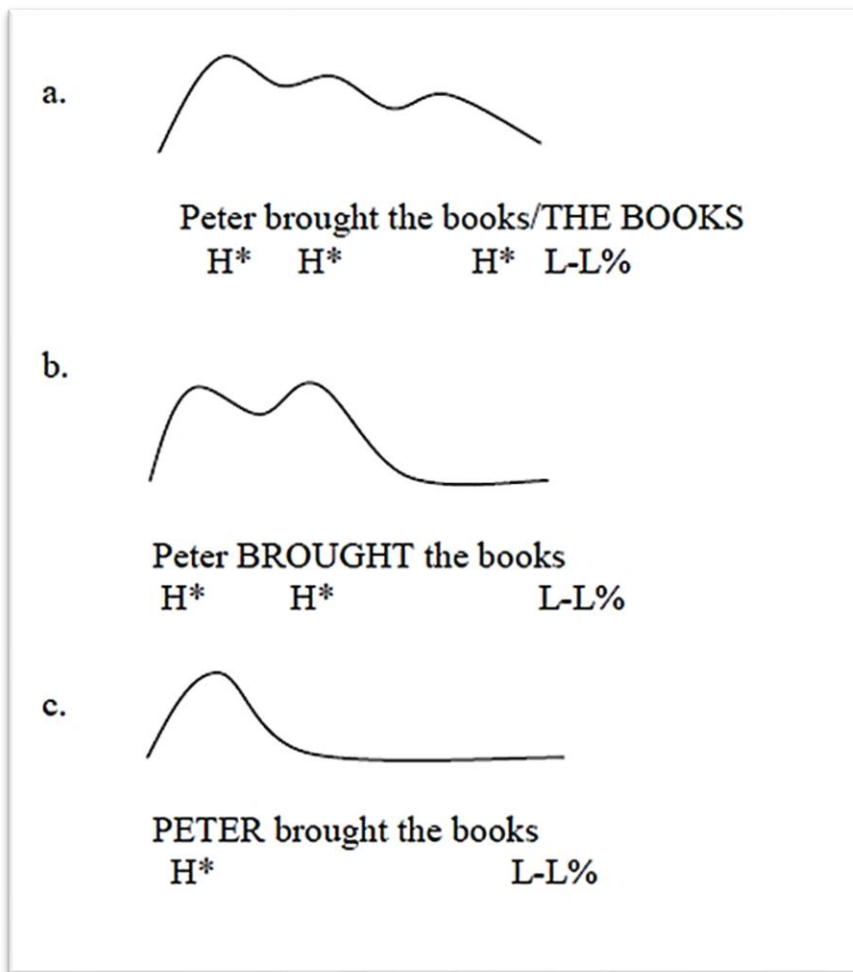
Figure 2. 3 Phonological and phonetic realisations of focus in English (Estebas-Vilaplana, 2000, p. 60)

In acoustic studies (Cooper, Eady and Mueller, 1985; Xu, 1999; Xu and Xu, 2005; Breen *et al.*, 2010), focus of this kind can be marked with a peak in fundamental frequency, greater intensity and longer duration. These acoustic parameters affect the details of the actual phonetic realisation of the focus to a great extent. Brief summaries of some studies of acoustic correlates of focus in English are provided as follows.

Sityaev and House (2003) investigated the phonetic and phonological correlates of broad, narrow and contrastive focus in British English declaratives. They used recorded sentences to examine pitch accent realised on a monosyllabic NP constituent. The accented monosyllable was placed sentence-initially, sentence-medially, and sentence-finally. The data were collected from four male and two female native speakers of Southern British English. The results showed that narrow focused words and contrastive focused words had longer durations than broad

focused words. Downstepping accents in which the second pitch peaks are lower than the first also occurred, particularly in the case of contrastive focus. Besides this, they did not find evidence supporting the relationship between focus type and the distribution of accent types.

Xu and Xu (2005) investigated how focus is phonetically realised in declarative sentences in American English. The participants were native speakers of American English who were asked to utter short declarative sentences under with or without conditions of narrow focus. The results revealed that fundamental frequency (f0) in narrow focus was realised by increasing the pitch range. Such pitch range expansion was observed on the stressed syllables of the focused words. The post-focal pitch range was also compressed or suppressed whereas pre-focal pitch range was left intact. Interestingly, they reported that focus did not play a role in indicating the presence or absence of pitch peaks, and also that a local pitch target was present in every syllable.

Breen *et al.* (2010) investigated the acoustic correlates of information structure in English. Native speakers of American English were recruited, and sentences with the SVO structure and a question-and-answer task were used to investigate words under narrow focus, contrastive focus and broad focus. In addition, the positions of focus varied among subjects, verbs, and objects. The results showed that native English speakers marked focus at these positions by means of longer duration and greater intensity as well as higher mean and maximum f0. Higher degrees of duration and intensity, and increased f0 were found in narrow focus objects. When it came to the marking of contrastive focus, native English speakers in this study could also differentiate between contrastively focused and noncontrastive words with greater degrees of intensity and duration, or even marking the focus with lower pitch. The latter is an alternative way to reduce ambiguity in meaning as studies by Schafer *et al.* (2000) and Snedeker and Trueswell (2003) cited in Breen *et al.* (2010) show.

2.3.3 Levels of prosodic structure

In English, the same utterance can be divided into prosodic phrasing in different ways. As shown in Figure 2.4, there are two levels of prosodic phrasing above the word; that is, an intermediate phrase (ip) and an intonational phrase (IP). With the aid of a change in the rhythm of the speech or a pause, especially at the phrasal boundaries, the right edges of intermediate and intonational phrases can be recognised. In addition, two pitch events are associated with

the two levels of prosodic phrasing. Firstly, phrase accents (T-) are tonal events that mark the right edges of intermediate phrases. Secondly, boundary tones (T%) are tonal events that mark the right edges of intonational phrases. Reflecting a complete utterance, a full intonational phrase can consist of one or more intermediate phrase. The prosodic phrasing of this kind is marked as T-T%, comprising a combination of the phrase accents and boundary tones, which are then often referred to as edge tones.
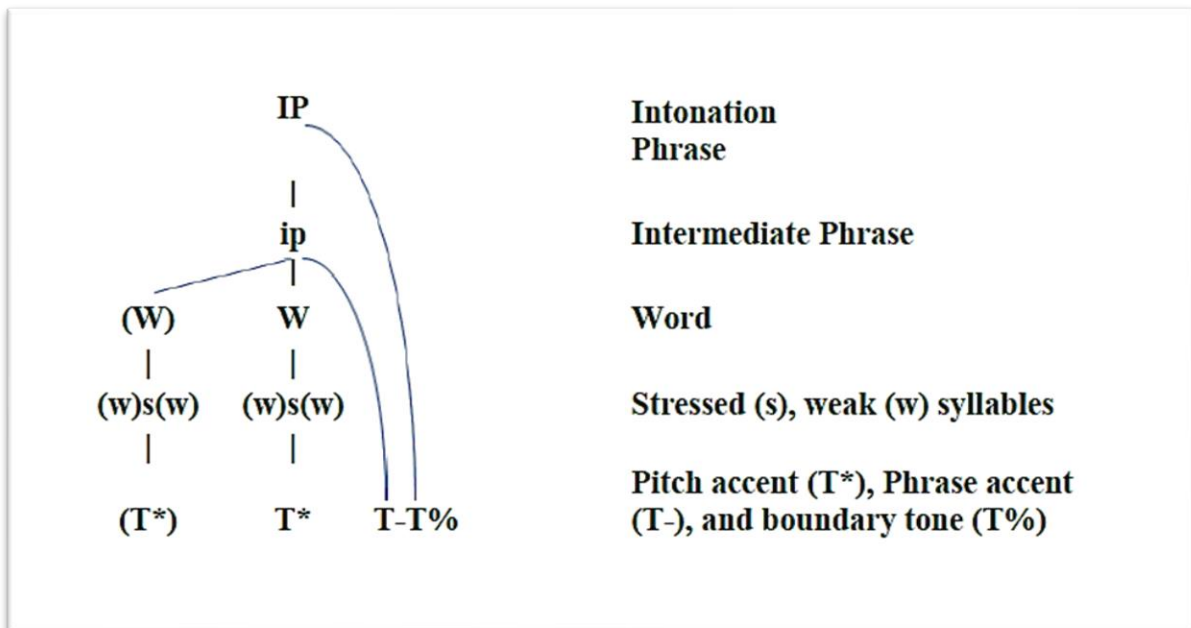
Figure 2. 4 Schematic representation of prosodic hierarchical organisation in English based on intonational phonology (adapted from Jun and Oh, 2000, p. 73)

Figure 2.4 gives a representation of prosodic structure based on intonational phonology (Beckman and Pierrehumbert, 1986; Gussenhoven, 2004**;** Jun and Oh, 2000; Ladd, 2008). The figure shows levels of prosodic constituency starting with the highest level of the intonational phrase and the lower level of the intermediate phrase. The intonational and intermediate phrases have been mostly used in the work of Fletcher and Evans (2002), Ladd (2008), and Arvaniti and Garding (to appear), whereas the intermediate phrase was not included in the work of Grabe *et al.* (2000), Grabe, Kochanski and Coleman (2005), and Gussenhoven (2004). A metrical tree shows the associations of the three tonal events; that is, associations of the pitch accents with the stressed syllables, of the phrase accents with the intermediate phrase, and of the boundary tones with the intonational phrases. In addition, both intonational and intermediate phrases have their own nuclear accents, carried by the last accented words of the respective phrases. The

patterns of tonal events corresponding to the patterns of a nuclear accent are composed of pitch accents and edge tones. In phonological representation, a nuclear accent contour or tune is generally given the notation T* T-T%.

It is obvious that prosodic edges can be identified from phrase accents and boundary tones. Apart from the two different tonal events (T-, T%), the edges of prosodic phrases may be identified according to one or more of the phonetic correlates; for example, pauses, phrase-final lengthening, and reductions in amplitude. These phonetic cues are applied to both levels of prosodic phrasing. A study by Hellmuth (2006, 2007) identified several cues to prosodic phrasing. For example, for an intonational phrase (IP), there may be a long silent pause, a final extra low f0 value at its right edge, more complex tonal changes, and the presence of the glottalization of vowel-onset words at the left edges of intonational phrases, which is not permitted at the onset of intermediate phrases. For an intermediate phrase (ip), the same phonetic cues are still applied but to a lesser extent; for instance, unmarked or short pauses, and less elongation of the final syllable, foot or word at its edge. In addition, pitch range can be reset for a new intermediate phrase, suggesting that there is more information to come.

2.3.4 AM-based descriptions of English intonation

English intonation has been predominantly studied based on the two schools of thought concerning prosody: the American and the British approaches. Each approach offers diverse theoretical perspectives on and practical implications for intonational studies. Furthermore, in the ToBI Annotation Convention in the Guidelines for ToBI Labelling (Beckman and Elam, 1994), Beckman and Hirschberg stated that the tonal categories that they proposed for the analysis of intonation (pitch accent, phrase accent, and boundary tones) can be applied for use with Standard Australian English and RP British English with the addition of tonal categories specific for each variety. Therefore, it is not surprising that there has been further work on the language-specific development of ToBI. Following the AM approach, the descriptions of intonation are applied to the current study regarding L1 and L2 English in the following sections.

The anatomy of intonation patterns in English is composed of three tone events: pitch accents, phrase accents and boundary tones. Pitch accents are associated with prominently accented words in order to communicate different types of intonational meaning (Pierrehumbert and

Hirsberg, 1990). The pitch accents in English used in the ToBI prosodic transcription system consist of H*, L*, L*+H, L+H*, and they can be downstepped to give !H*, L*+!H, L+!H*, and H+!H* (Beckman, Hirschberg, and Shattuck-Hufnagel, 2005). Phrase accents in English are typically low (L-) or high (H-) tones, functioning to mark the boundary of a minor prosodic phrase; that is, an intermediate phrase. They influence the pitch shape after the nuclear-accented syllable and the intermediate phrase edge. Phrase accents can be downstepped to give, for example, !H-, !H-H%, and !H-L%. Boundary tones are either low (L%) or high (H%) tones, and they function to mark the boundaries of a major prosodic phrase; that is, an intonational phrase. They are locally aligned with the intonational phrase edge. Obviously, each tone event functions differently. Combined together, the three tone events constitute well-formed intonational patterns or tunes in English. More importantly, a combination of pitch accents and edge tones at the phrase-final positions of a prosodic phrase, either at an intermediate or intonational phrase, is considered to be a nuclear accent or nuclear contour. According to Wells (2006), the most basic distinctions among English nuclear contours can be grouped into two broad categories of falling and non-falling intonation. Ladd (2008, p. 91) provides an inventory of British-style nuclear tunes equivalent to Pierrehumbert's (1980) work, as in Table 2.1 below.

Table 2. 1 Ladd's inventory of British-style nuclear tunes and Pierrehumbert's model of intonation (Ladd, 2008, p. 91)

Pierrehumbert	British-style
H* L L%	Fall
H* L H%	Fall-rise
H* H L%	Stylised high rise
H* H H%	High rise
L* L L%	Low fall
L* L H%	Low rise (narrow pitch range)
L* H L%	Stylised low rise
L* H H%	Low rise
L+H* L L%	Rise-fall
L+H* L H%	Rise-fall-rise
L+H* H L%	Stylised high rise (without low head)
L+H* H H%	High rise (with low head)
L*+H L L%	Rise-fall (emphatic)
L*+H L H%	Rise-fall-rise (emphatic)
L*+H H L%	Stylised low rise
L*+H H H%	Low rise
H+L* L L%	Low fall (with high head)
H+L* L H%	Low fall-rise (with high head)
H+L* H L%	Stylised high rise (low rise?)
H+L* H H%	Low rise (high range)
H*+L H L%	Stylised fall-rise ('calling contour')
H*+L H H%	Fall-rise (high range)

As shown in Table 2.1, some common intonation patterns or tunes in English include H* L-L%, H* H-H%, and H* L-H%. The intonation pattern H* L-L% is standard or neutral for declarative sentences. It is a falling nuclear tune that signals finality and suggests that the phrase is potentially complete. The same H* L-L% nuclear tune can also be found in wh-questions (Cruttenden, 1997; Wells, 2006). The pattern H* H-H% is very common in yes/no questions, and is known as the high rising terminal (HRT) or high rise of Australian English declaratives. Sometimes, the L* H-H% tune (low-onset high rise tune) can be found. The pattern H* L-H% is known as the continuation rise or fall-rise. It is typical when a speaker suggests that they have something more to say. In addition, this tune H* L-H% is referred to as the 'implicational fall-rise' (Wells, 2006, p. 27). The tune suggests that the speakers imply something without actually saying it.

There are other varieties of nuclear tunes that can be found in an English utterance; for example, in the rise-fall L+H* L-L% when expressing assertion, and the rise-fall-rise L*+H L-H% in expressing emphasis and suggestion. However, this is not to say that tunes and meanings have a one-to-one relationship. It is only that there is a general tendency for that to be the case in line with evidence from earlier studies (Pierrehumbert and Hirschberg, 1990; Pitrelli, Beckman and Hirschberg, 1994; Ladefoged and Johnson, 2011). Ending a prosodic phrase with L% generally indicates completeness, while an ending with H% often indicates seeking information. Similarly, the L-L% combination is acknowledged for standard declarative statements while the H-H% combination is used to ask yes/no questions. H* pitch accents typically emphasise what is expected to be new information added to the discourse. In yes/no questions, it is usually introduced by a monotonal low L* pitch accent. As for the fall onto stress, H+!H* indicates a pragmatic inference and implies familiarity. The H-L% boundary tone is perceived as a level pitch and is common in enumerations or making lists (Pitrelli, Beckman and Hirschberg, 1994).

As can be seen above, there is evidence that the AM approach to intonation has been adopted and used for intonational studies in the British English context. For example, the IViE system for intonation analysis was developed as part of the project of Intonational Variation in English by Grabe, Post and Nolan (2001) and Grabe and Post (2002), and in the intonation system for Glasgow English by Mayo (1996). These systems have been developed under the AM framework, but each has their own criteria in presenting language/dialect-specific phonological representations. The literature on the Autosegmental-Metrical approach and the British approach to intonation analysis is described separately in subsection 2.6.1.

2.4 Prosody, Intonation and Focus in Thai

2.4.1 Introduction to Thai and its phonology

Thai is the official language of the Kingdom of Thailand, formerly known as Siam. Thai or Siamese is a tone language spoken with different accents and dialects from the northern to southern and eastern to western parts of the country. Standard Thai (henceforth referred to simply as Thai) has among its varieties the prestigious regional dialect of Bangkok and surrounding provinces. Thai has been widely used as the principal language for the purposes of formal education and communication throughout the country.

In Thai, there are 21 consonantal phonemes: /p, pʰ, b, t, tʰ, d, kʰ, k, ʔ, m, n, ŋ, f, s, h, tɕ, tɕʰ, l, r, w, j/. They all appear in syllable-initial positions. There are 9 consonantal phonemes occurring in syllable-final positions, four being members of the stops /p, t, k, ʔ/, three of the nasals /m, n, ŋ/ and two of the glides /j, w/ which are permitted after a vowel. The velar nasal /ŋ/ as well as the glides /j, w/ can appear in both the onset and coda of Thai syllables, as in /jáːj/ 'to move', /wâːw/ 'kite', and /ŋûaŋ/ 'sleepy', for instance. Moreover, there are 11 Thai consonantal clusters, /pr, pl, pʰr, pʰl, tr, kr, kl, kʰr, kʰl, kw, kʰw/. Even though /tʰr/ and /fr/ can be found, they are considered as clusters from loanwords. It is noteworthy that Thai clusters can produce well-formed syllable onsets only, and they combine with either /r/, /l/, or /w/ exclusively as the second consonantal phonemes (Kamphikul, 2015).

The inventory of Thai consists of 21 phonemic vowels. There are 9 distinctive vowel sounds, /i, e, ɛ, a, ɔ, o, ɤ, u, ɯ/, which are all short monophthongs, as well as 9 long monophthongs, /iː, eː, ɛː, aː, ɔː, oː, ɤː, uː, ɯː/. Duration plays a crucial role in distinguishing Thai monophthongs into these two categories. Besides monophthongs, there are three diphthongs: /ia, ua, ɯa/. The quality of the unstressed vowel /a/ in each diphthong can be represented as phonemically /ə/ or phonetically [ə] (Luksaneeyanawin, 1983; Tingsabadh and Abramson, 1993; Tumtavitikul, 1997).

2.4.2 Prosodic features in Thai

1) Lexical tones

Lexical tones are language-specific for Thai, and five different lexical tones are used exclusively to contrast lexical meaning (Abramson, 1978; Tingsabadh and Abramson, 1993; Morén and Zsiga, 2006). As shown in Figure 2.5, the first three are level or static tones; that is, 'mid' (ˉ) or no tone mark, 'low' (ˋ), and 'high' (ˊ). Another two are contour or dynamic tones: 'falling' (ˆ), and 'rising' (ˇ). The mid tone is most likely to appear as a level; the high tone as a high rise in contrast with the low rise of the rising tone; the low tone is a low fall in contrast with the high fall of the falling tone. The language itself has also been reported to have an emphatic high tone which is produced with a higher degree in pitch than the normal lexical high tone. This type of tone is used when expressing or intensifying meanings in emphatic situations (Hass, 1946; Abramson, 1962; Iwasaki and Ingkaphirom, 2005).
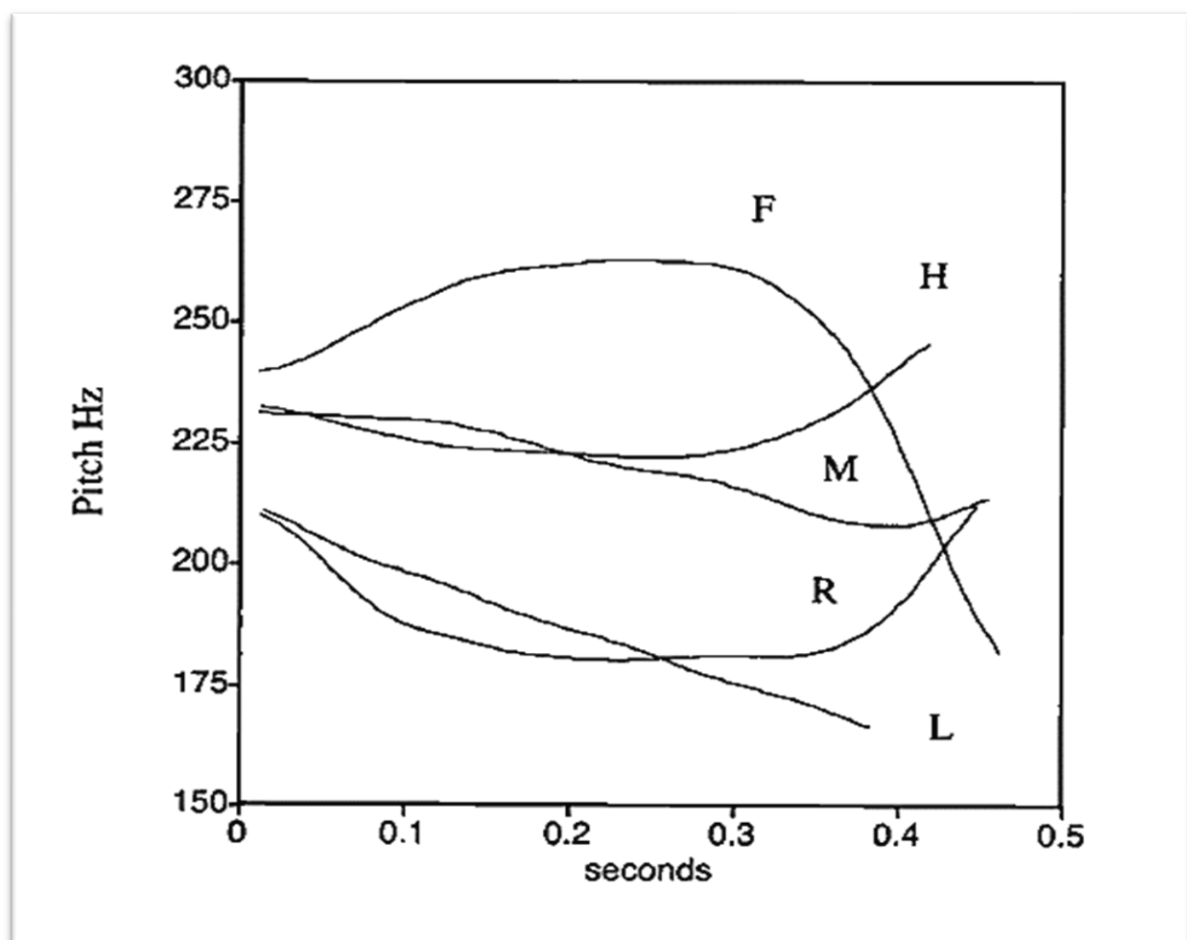
Figure 2. 5 Five lexical tones in Thai phonology (Zsiga and Nitisaroj, 2007, p.347)

These lexical tones of mid, low, high, falling and rising are considered to be compulsory autosegmental elements for every single syllable or word in the Thai language. However, not all lexical tones can occur in every syllable structure, and vice versa. To convey a speaker's intended meaning, a Thai word must be spoken with its correct tone. Examples of words with differences in lexical tones and meanings are shown next (2):

(2)	**Lexical tones**	**Words**	**Meanings**
	mid tone (m)	/pa:/	'to throw'
	low tone (l)	/pà:/	'forest'
	falling tone (f)	/pâ:/	'aunt'
	high tone (h)	/pá:/	'a term for addressing someone's father in a loving way'
	rising tone (r)	/pǎ:/	'father (a title of respect for elder males), or sugar daddy' (when used by younger women)

2) Word stress

In the literature on the prosodic features of Thai (Henderson, 1949; Noss, 1964; Hiranburana, 1971; Luangthongkum, 1977; Luksaneeyanawin, 1983; Warotamasikkhadit, 2000), it is reported that the Thai language uses word stress. Hiranburana (1971) and Luangthongkum (1977) claimed that the language itself uses syllable duration as one of the most salient features to separate stressed from unstressed syllables. This is always true, as shown in evidence from the influence of stress over the realisation of short and long vowel lengths in Thai.

According to Luksaneeyanawin (1983) and Tumtavitikul (1997), Thai has a fixed system of word stress. There are two broad categories of lexemes in the language: content words and function words. Normally, content words include nouns, adjectives, and most verbs and adjectives which are accented in neutral speech. Function words such as pronouns and prepositions are usually not accented in normal speech, but words of this kind can receive stress on special occasions, such as for emphasis and contrastive focus. In Thai, the primary stress is always placed on the last syllable of an accented word (Tingsabadh and Abramson, 1993). This rule of stress assignment also applies to the last syllable of a polysyllabic content word. The secondary stress of the word is governed by rules relating to the syllable structures of the rest of the syllables composing the words. A study of the fundamental frequency correlates of stress in Thai (Potisuk, Gandour, and Harper, 1994) has reported that stressed and unstressed syllables

do not affect the shapes of Thai lexical tones. In other words, all five lexical tones maintain a five-way contrast. Given below (3) are examples of stress, marked with "ˈ", in Thai words.

(3)   **Thai Words**   **Meanings**   **Speech styles**

Thai Words	Meanings	Speech styles
/sà.wàt.ˈdiː/	'greetings'	citation form, careful speech
/sə. wàt.ˈdiː/	'greetings'	normal, casual speech
/má.hăː.wít.tʰá.jaː. ˈlaj/	'university'	citation form, careful speech
/mə.hăː.wít.tʰ ə.jaː.ˈlaj/	'university'	normal, casual speech
/mə.hăː.wít.tʰ ə.jə.ˈlaj/	'university'	normal, casual, fast speech

3) Intonation

Thai employs intonation at the phrase or sentence level, which is sometimes referred to as sentence stress or sentence accent. Considered to involve variations in pitch patterns over an entire utterance, intonation in Thai is viewed as a complex interplay between tone and lexical and post-lexical prominence (Luksaneeyanawin, 1998). In this case, Thais assign tonal targets at the lexical as well as phrase level. Thais not only contrast lexical items and meanings using lexical tones, but also use intonation to distinguish between different meanings and functions for utterances. For instance, a falling intonation at the end of an utterance signals that the interlocutor has finished relaying information and is ending their part in the conversation. According to Abramson (1979) and Luksaneeyanawin (1998), a sentence accent, or prominently accented words, in Thai can be marked by the lengthening of a syllable, a tonal contour that approaches the form of an ideal tone, and an increase in amplitude. These methods can be used independently or in combinations in signalling Thai intonation.

In addition, several findings from earlier work have led to the claim that fundamental frequency (f0) has a potential influence in distinguishing between utterance types and marking prominence, as suggested in studies by, for example, Abramson (1979), Abramson and Svastikula (1983) and Luksaneeyanawin (1983). Brief summaries of these works are given next. Abramson's (1979) work included the first experiment to be conducted on intonation that was based on auditory and acoustic analyses. The participants were two Thai adults who were asked to produce a minute-long conversation and monologue. The findings showed that there were three types of falling, rising, and sustained pitch contours, and that their presence distinguished between sentence types. This study supports his earlier (1962) findings on Thai intonation. In that work, he claimed that the falling pitch can be found in statements, while the rising pitch

can be found in questions and expressions of surprise and doubt, and the sustained pitch can be found in unfinished statements. The following example in Figure 2.6 gives evidence for this claim.
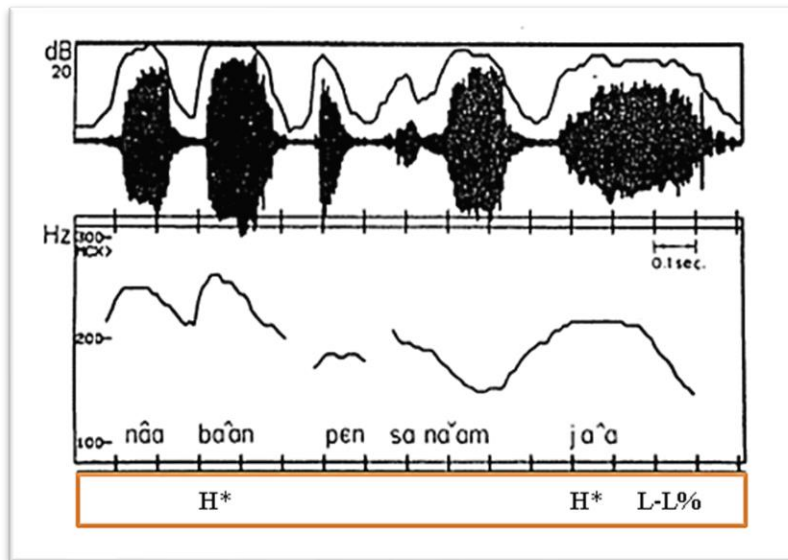
Figure 2. 6 Waveform and the contour of falling intonation for a Thai statement (adapted from Abramson, 1979, p. 156)

A more detailed explanation concerning several aspects of Thai prosody can be derived from sentence (4) below drawn from the statement in Figure 2.6. For the sake of convenience in these examples, lexical tones are annotated with hyphens and small letters: -m for mid, -h for high, -l for low, -f for falling, and -r for rising tone (after Luksaneeyanawin, 1998).

(4)   /naː-f. ˈbaːn-f.   ˈpen-m.   sa-l ˈnaːm-r.   ˈjaː-f/
       [front   house   be   lawn   grass]
       (At the front of the house is the lawn.)

Firstly, the words /ˈbaːn-f/ and /sa-l ˈnaːm-r/ represent good examples of word stress in Thai, which is realised on its own for monosyllabic words and on the last syllable for disyllabic words. Secondly, it is notable that the last syllable /ˈjaː-f/ of the compound word /sa-l ˈnaːm-r. ˈjaː-f/ bears sentence stress, which is realised with a lower f0 at the end of the utterance. The final syllable or word /ˈjaː-f/ is also signalled by syllable lengthening and the shape of tonal contour (Abramson, 1979; Kallayanamit, 2004). This is a typical falling intonation in a Thai statement. Descriptions like this correspond to Abramson's (1979) viewpoint, in that a sentence

accent or prominently accented words in Thai can be marked by a lengthening of the syllable, a tonal contour that approaches the form of ideal tones, and an increase in amplitude.

Abramson and Svastikula (1983) investigated declination in Thai declaratives and its effect on lexical tones. The participants were one male and one female, and they were asked to read a set of declarative sentences. In this study, 125 three-word simple declaratives and 125 longer complex declaratives were used and were then analysed auditorily and acoustically analysed. The findings showed that longer complex declaratives showed a gradual fall in fundamental frequency compared to simple declaratives. They reported that the ideal shapes of five lexical tones were intact or remained the same despite the interaction with declination. They also reported on the effects of neighbouring segments and tones in leading to slight perturbations in tonal contours. However, this work by Abramson and Svastikula (1983) has some limitations; for example, the results are for final words only, and no detailed descriptions of f0 values are given.

To date, the elaborate work on Thai intonation by Luksaneeyanawin (1983, 1998) seems to have influenced intonational studies by other Thai scholars. She investigated intonation in Thai by means of auditory and acoustic analyses. The participants were one male and two females. They were asked to read from cue cards one-word utterances in various grammatical and attitudinal contexts such as statements, yes-no questions, surprise, and anger. The findings suggested that four different tunes can be produced. These are a falling tune, a rising tune, a sustained tune, and a complex tune. After that, she proposed that the four intonation contours or tunes in Thai can be grouped into 3 classes: the fall class, the rise class, and the convolution class. Her classification also corresponds with intonation universals.

It has been acknowledged that there is an interaction between lexical tones and intonation in Thai. This is so because they are both primarily represented by the same acoustic parameter; that is, changes in fundamental frequency (f0). For lexical tones, f0 represents the pitch pattern over a syllable or word, distinguishing one word from another according to meaning. For intonation, f0 represents the pitch pattern over a phrase or an utterance. Thus, it is not surprising that an interaction between lexical tones and intonation has been reported. Previous work by Gandour (1974a, 1974b), Abramson (1979), and Abramson and Svastikula (1983) provides evidence for this. In citation forms, lexical tones have the ideal contours. In running speech, their ideal contours were found to undergo certain changes or were somewhat perturbed in their onset and offset parts. This reflects the influence of coarticulation with neighbouring tones and

of certain consonantal contexts. Additionally, a study by Pittayaporn (2007) showed that some final particles systematically preserve their lexical tones while others are overridden by boundary tones.

However, in spite of the effects of segmental features and tonal coarticulation, Abramson (1979), and Abramson and Svastikula (1983) claimed that the full tonal system of five tones is preserved in Thai. In this respect, each tone still keeps its own shape (i.e., f0 contours) in all contexts, which can be identified by both auditory and graphical inspection. The work on Thai intonation by Luksaneeyanawin (1998) revealed that the tune system in Thai determines the configuration of pitch for an utterance, whereas the lexical tone system does not affect the overall shape of pitch over an utterance at all. The phonetic features of each lexical tone are still distinct from one another and they contribute to word meanings only. However, Luksaneeyanawin seemed to claim that different tunes have certain effects on the characteristics of each lexical tone. For example, the pitch may be raised or lowered, and the pitch range is wider or narrower.

4) *Intonation groups and focus marking*

In the writing system of Thai, words are written or put together to form a piece of information without spaces between them. There are no actual punctuation marks such as the full stop in English. To signal to the reader or listener where a piece of information ends, the writer or the speaker normally leaves blank spaces in writing or pauses at the time of speaking. Using blank spaces or pauses of this kind corresponds to the syntactic boundary of that information. In other words, the beginning and ending of a new word, phrase, clause, or utterance, as well as smaller information units such as a syllable and each segmental, are explicitly marked by spaces or pauses.

Pausing in speech conveys various different meanings for speakers and listeners. The speaker can use pauses to terminate sentences. Using pauses in this way signals that the speaker has no more to say or wants to end the transmission of a piece of information. This function of pauses is quite universal across languages (Luksaneeyanawin, 1983, 1998). Speakers can also use pauses to signal that they themselves need time to think, or otherwise that they want to allow time for the listeners to catch up with what they have said. In addition, pauses are used to group written or spoken materials into chunks of information. Chunking information with pauses like this depends on the speaker's discretion, and the same piece of information can have different

numbers of chunks, which are known to represent or to be equivalent to intonation phrases or intonation groups as in English.

Before further describing intonation groups, it makes sense to explore and understand the prosodic structure in Thai. Thai has the following prosodic constituents: the syllable, the foot, the prosodic word (PW), the intermediate phrase (ip) or phonological phrase, and the intonation phrase (IP). Studies by Luksaneeyanawin (1983, 1998), and Tumtavitikul and Thitikannara (2006) regarding stress and intonation in Thai provide primary evidence for these prosodic constituents. The following is a schematic representation of the Thai prosodic structure based on intonational phonology, adapted from Luksaneeyanawin's (1998, p. 388) study of textual organisation.
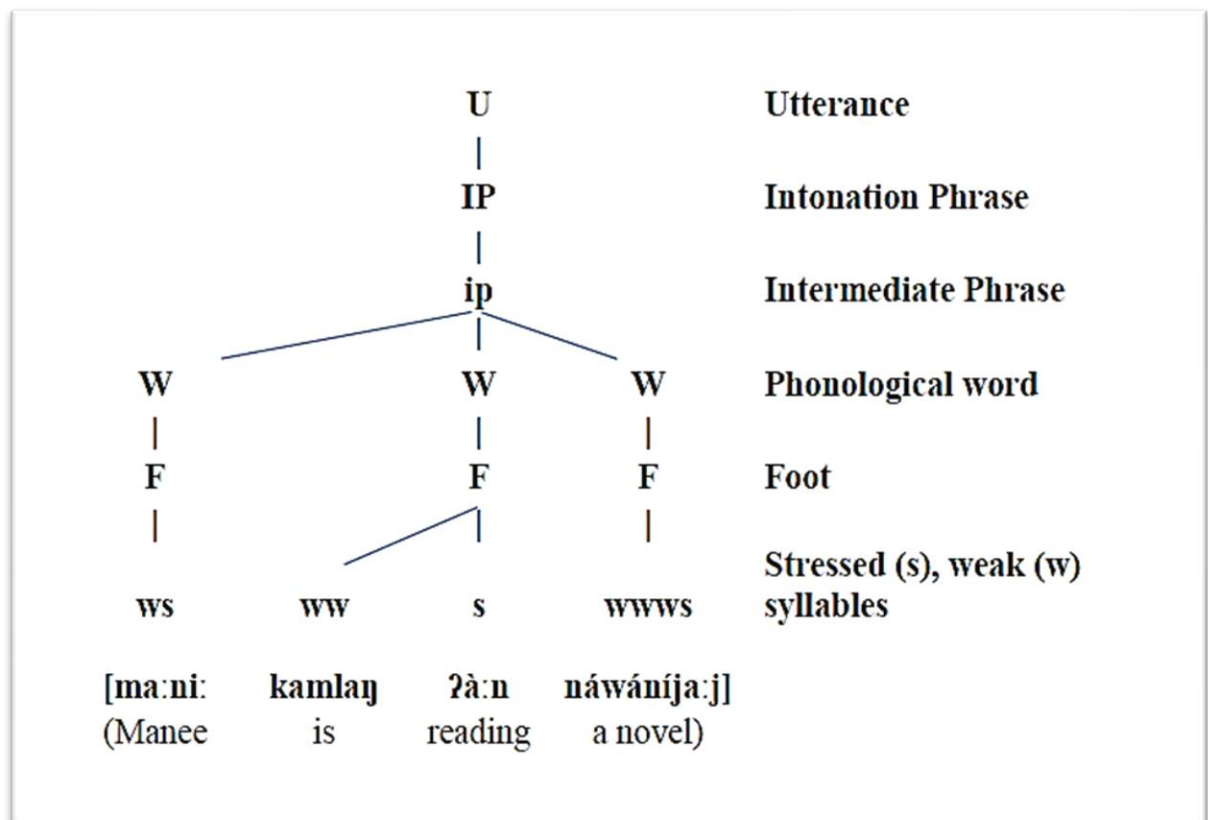
Figure 2. 7 Schematic representation of Thai prosodic structure based on intonational phonology (adapted from Luksaneeyanawin, 1998, p. 388)

As discussed above, intonation groups or intonational phrases in Thai are demarcated by phonological pauses. In addition to this, they can be defined in terms of phonological prominence (Luksaneeyanawin, 1998). When chunking information into an intonation group,

there is only one prominently stressed syllable or accented word in each group. A word with the most prominence is almost always the last content word in the information group or intonation phrase. In Thai, the final syllable of a word has the primary stress, so that the final syllable is always the tonic syllable and the word containing the tonic syllable like this is always the right-hand word of the intonation group (Hiranburana, 1971; Luangthongkum, 1977; Tingsabadh and Abramson, 1993; Luksaneeyanawin, 1998). This firstly functions to signal the end of one piece of information that the writer or speaker wants to convey. Secondly, it functions to indicate that the information is unmarked or marked in the sense that it sounds neutral in normal speech, or alternatively that it shows contrast. The latter case also shows that the focus of information can be marked by the prominence of the final word in an intonational group. However, the location for expressing the focus of information can be shifted to a non-final word. This means that non-final words can also receive prominence, especially in situations where the word needs to be emphasised or highlighted so as to convey specific meanings and to make contrasts. Function words can also be put into focus. In this respect, Luksaneeyanawin (1998, p. 389) called cases of terminating a piece of information by using either final words with prominence or non-final words with prominence as 'end focus'. She termed "expressive focus" cases of highlighting part of the information to show contrast or newness.

The Thai language can thus express a focus on information in various different ways. These include phonological, morphological, and syntactical means. For instance, the marking of expressive focus can be accomplished by using additional linguistic devices such as pausing to divide the syntactic structure, changing word choices, and placing phonological prominence. Luksaneeyanawin also suggested that acoustic cues such as a combination of fundamental frequency (a rise in pitch), duration (vowel lengthening) and intensity (greater loudness) can be used to signal prominent words or focused words in Thai. In addition to this, there may be accompanying phenomena such as the lowering of pitch throughout sentences. These features are in line with the findings of Henderson (1949), Abramson (1962, 1979) and Noss (1964). A more recent study by Silpachai (2013) found that prosodic features in emphatic context in Thai differed between broad and narrow focus. The values of f0 maximum, minimum and range in narrow focus were higher than in broad focus. The values of f0 were lowered and the pitch was reset after focused words. Syllable duration in narrow focus was longer, as well.

The following examples illustrate how focus in Thai is marked differently. Sentences (5) and (6) show the marking of focus in different syntactic structures. Sentences (7), (8) and (9), which are from the work of Luksaneeyanawin (1998, p. 389), show the marking of focus by the placement of phonological prominence or tonic syllables, division of the syntactic structure using a phonological pause, and the choice of lexical items respectively. Again, for the sake of convenience in these examples, lexical tones are annotated with the hyphen and lower case: -m for mid, -h for high, -l for low, -f for falling, and -r for rising tone, after Luksaneeyanawin (1998).

(5)   For neutral/unmarked syntactic structure: SVO
        /ˈphɔm-r.   ruː-h ˈtɕak-l. phuː-f ˈjiŋ-r.   ˈkhɔn-m.   ˈnan-h/
        [I   know   woman   person   that]
        (I know that woman.)

(6)   For marked syntactic structure: OSV
        /phuː-f ˈjiŋ-r.   ˈkhɔn-m.   ˈnan-h.   ˈphɔm-r.   ruː-h ˈtɕak-l/
        [woman   person   that   I   know]
        (I know that woman.)

(7)   /ˈdæŋ-m.   ˈtat-l.   kraʔ-l ˈproːŋ-m.   ˈniː-f/
        [Dang   cut   dress   this]
        (Dang made this dress.)

(8)   /kraʔ-l ˈproːŋ-m.   niː-f.   ˈdæŋ-m.   ˈtat-l/
        [dress   this   Dang   cut]
        (This dress, Dang made it.)

(9)   /ˈdæŋ-m.   ˈtat-l. tɕaːw-f.   kraʔ-l ˈproːŋ-m.   ˈniː-f/
        [Dang   cut   PREFFIX   dress   this]
        (Dang made this dress.)

(10)   /ˈphæŋ-m.   ˈkɤːn-m.   ˈpaj-m/
        [expensive   too   go]
        (It is too expensive.)

Apart from the focus-marking strategies mentioned above, the tune systems in Thai play a crucial role in expressing unmarked and marked information. In the case of marked information, such as it being emphatic, surprised, agreeable or disagreeable, the tunes are always accompanied by a higher peak in fundamental frequency, longer duration and increasing loudness. In sentence (10) /ˈphæŋ-m. ˈkɤːn-m. ˈpaj-m/, the last word /paj-m/ carries the sentence accent marking the focus of information and signalling the end of information being provided at the same time. When it is said with the falling class of intonation, notated as H* L-L%, its meaning is neutral in that, for example, 'I am talking about the price of this product, and I am telling you it is too expensive'. When it is said with the rising class of intonation, notated as H* H-H%, or L* H-H%, its meaning can be emphatic or disagreeable in that, for example, 'I do not think the price of this product is so expensive because of its low quality'. When it is said with a sustained level pitch as H* H-L%, it can be used to express, for example, doubts, or disbelief as well as unfinished information.

To conclude, at this point, it is sufficient to state that the existence of lexical tone does not necessarily prevent the existence of intonation (Hirst and Di Cristo, 1998; Ladd, 2008). This is the case for Thai, a language which relies on lexical tone, stress and intonation. Thai is a tonal language where pitch is fundamentally used in conveying the meaning of each word. Not surprisingly, stress plays less of or even no distinctive role in Thai even though it is reported that the language itself does involve stress. This may be one of several reasons why Thai speakers tend to treat stress in English in the same way as in Thai. Previous studies (Luangthongkum, 1977; Luksaneeyanawin, 1983, 1998; Potisuk, Gandour and Harper, 1994; Tumtavitikul, 1997) have demonstrated that stress assignment in Thai is well-defined and governed by rules. Besides this, for prosody at the utterance level, intonation is considered to have well-known forms in terms of tune systems. Different tunes can produce different types of utterances, such as wh-questions and yes/no questions. They can also be used in the expression of focus, as well as in conveying attitudes and emotions (Yimngam, Premchaisawadi and Kreesuradej, 2011).

However, little is known as yet about several issues with prosodic features such as stress and intonation in the Thai language. In particular, the intonation system of Thai is considered to be underresearched in spite of the groundbreaking work of Luksaneeyanawin (1983, 1998), among others. This is also the case for L2 English intonation as produced by Thai learners. To the best of the knowledge of the present researcher, no comparable studies have been conducted regarding the patterns of L2 English intonation in the terms elaborated on here. This section is,

thus, devoted to a detailed account of Thai prosody, intonation, the marking of focus, and prosodic structure obtained from the existing literature. The present study uses read speech as a point of departure, and it is hoped that future research will continue this line of investigation with the help of other types of data.

2.4.3 Contrastive analysis of L1 English and Thai, and learning models

Contrastive analysis offers a method of comparing the phonological, morphological and syntactic systems of two different languages. The central notions of the contrastive analysis hypothesis (CAH) have been proposed by its two prominent advocates. Firstly, Lado (1975) stated that the linguistic characteristics of an L1 which are the same as those of an L2 will help enhance the learning process of L2 learners, while those linguistic characteristics which are different can cause difficulties in their learning. Secondly, Weinreich (1953) stated that the extent of the differences between two languages results in various degrees of difficulties in learning as well as various degrees of interference from a learner's native language. The CAH aims to describe L2 errors from both production and perception on the basis of influence of first language. With the CAH, researchers can predict all L2 errors by means of examining L2 learners' language, comparing the differences between L1 and L2 patterns based on the transfer of L1 characteristics to L2. The possible results can be either positive or negative as discussed earlier.

To conduct a contrastive analysis, the present study follows the steps outlined by Ellis (2008). A formal description of the two languages (English and Thai) is first accomplished; secondly, certain linguistic characteristics are selected for examination; thirdly, comparisons are made by identifying areas of difference and similarity between the two languages; finally, predictions are made about which areas are likely to cause errors. Possible difficulties may arise from one or more points of difference in the systems of the two different languages. As in Table 2.2 below, differences between L1 English and Thai are summarised. Points of contrasts have been derived from the literature review in sections 2.3 and 2.4 as detailed previously.

Difficulties, as reflected in the English produced by Thai learners, are predicted from points of difference in the systems of the two different languages. Such difficulties as the use of duration in distinguishing stressed and unstressed syllables in the accented words, as well as the use of intonation to signal information structure, are examined and verified in the present study.

Considering those differences, the hypotheses of the current study and the predictions of difficulties were formulated and are described at the end of this chapter.

Table 2. 2 Contrastive analysis of English and Thai

English	Thai	Remarks on Thai
Intonation language	Tone language	Thai also has stress and intonation.
SVO structure	SVO in neutral structure	OSV when marked
Lexical stress; property of a syllable that serves to make it more prominent	Lexical tones; pitch variations that change the lexical meaning of a word	Stress plays no distinctive role, so Thai speakers tend to give stress no distinctive value in English.
Variable stress; primary stress is unpredictable, but with certain exceptions.	Fixed stress; primary stress almost always falls on the same syllable.	Primary stress falls on the last syllable of the word only. The secondary stress of a word is governed by specific rules.
Content words are always accented and grammatical words are unaccented unless marked.	Content words are always accented and grammatical words are unaccented unless marked.	
Intonation is variation in pitch over units larger than a syllable.	Intonation is pitch variation over units larger than a syllable, a complex of the interplay between tone, lexical and post-lexical prominence. Tone and intonation co-exist in the intonation contours.	Intonation can be used to convey linguistic (e.g. statements, questions) and non-linguistic information e.g. attitudes, emotions).
Acoustic cues to prominence (e.g. accentual and focal prominence) are primarily signaled by f0, duration and intensity.	Acoustic cues to prominence are primarily a combination of f0 (a rise in pitch), longer duration and a higher degree of intensity.	Duration is one of the most prominent features used to distinguish stressed and unstressed syllables.
Signalling information structure/focus via default pitch-accent assignment, deaccenting	Signalling information structure/focus via word order, choice of lexical items, syntactic constructions, and prosodic aspects (e.g. lengthening, changing tonal contours)	Thai uses either one of these strategies or a combination of them to signal the marking of the focus.

2.4.4 Models of L2 speech acquisition

Apart from contrastive analysis, there are other models currently used as research frameworks regarding L2 speech acquisition; for example, the Speech Learning Model (Flege, 1995), the Perceptual Assimilation Model (Best, 1995; Best and Tyler, 2007), and the Second-Language Linguistic Perception model (Escudero, 2005). This section first gives a brief summary of these three models, but for more details, see Flege (1995) and Flege and Hillenbrand (1984) for SLM, Best (1995) and Best and Tyler (2007) for PAM and PAM-L2, and Escudero (2005) for L2LP. Then, a more detailed account is given of the Intonational Learning theory (L1Lt) as proposed by Mennen (2015), which is most relevant to the present study and is of special interest.

The Speech Learning Model (SLM) was developed to give accounts of phonological and phonetic features of the L1 and L2 based on contrastive analysis, (Flege and Hillenbrand, 1984; Flege, 1995). This model posits that an L1 sound can be the same as, similar to, or different from an L2 sound (Flege, 1987). It emphasises that the perception of a sound affects its production. Learners' experience in the L2 will also influence the accuracy of the production of new and similar sounds. In other words, producing an L2 sound correctly is a consequence of perceiving that sound correctly. According to Flege, similar sounds which share some acoustic characteristics will be more difficult to acquire than new sounds that are different from any sounds in the L1. Meanwhile, the Perceptual Assimilation Model (PAM) is based on perception more than the production of speech. This model defines phonetic differences and similarities between the L1 and L2 sounds in terms of articulatory gestures rather than acoustic properties (Best, 1995). It gives accounts of perception of L2 segmental patterns for learners whose L2 experience is insufficient. Recently, PAM-L2 (Perceptual Assimilation Model-L2) has been proposed, which is an expanded version of PAM and developed by (Best and Tyler, 2007) to predict more advanced L2 learners' performance. This model aims to describe the perception of non-native sound contrasts by L2 listeners. The Second-Language Linguistic Perception (L2LP) model was developed to give accounts and predictions concerning the perception of L2 sounds over the entire process of development (Escudero, 2005). Predictions from this model are based on comparison of sound contrasts between the L1 and L2. Escudero (2005) proposed several main ingredients for success in L2 learning. These consist of an optimal L1 and target L2, the initial state, the learning task, development and the end state. All in all, the SLM, PAM/PAM-L2 and L2LP accept that the L1 influences L2 speech acquisition. Like the CAH, the SLM emphasises the description and prediction of issues of production and

perception, whereas PAM/PAM-L2 and L2LP emphasise the description and prediction of issues of perception only.

According to Rasier and Hiligsmann (2007) and Mennen (2015), no models have been proposed so far to describe and predict difficulties in producing and perceiving L2 intonation. Flege's (1995) SLM, Best's (1995) PAM, and Best and Tyler's (2007) PAM-2 predict "the relative difficulties or ease of production and perception of non-native speech on comparisons of L1 and the to-be-learned segments" (Mennen, 2015, p. 172). Thus, Mennen proposed the L2 Intonation Learning theory (LILt). This model aims to account for difficulties that L2 learners have when producing L2 intonation. It builds on the basis of cross-language difference in the four dimensions proposed by Ladd (2008) in the AM theory. To account for similarities and dissimilarities between L1 and L2 intonation, these four dimensions are detailed below:

- *The inventory and distribution of categorical phonological elements ('systemic' dimension)*
- *The phonetic implementation of these categorical elements ('realisational' dimension)*
- *The functionality of the categorical elements or tunes ('semantic' dimension)*
- *The frequency of use of the categorical elements ('frequency' dimension)*
  (Mennen, 2015, p. 173)

The systematic dimension concerns typological similarities and differences in the inventory of phonological categories; for example, in pitch accents and edge tones. This dimension also concerns how different phonological categories combine with one another and the relationships between text and tune as discussed in Ladd (2008). The realisational dimension concerns how phonological categories are phonetically realized; for example, in tonal alignment and scaling. The semantic dimension concerns the use of phonological categories or tunes in expressing meaning; for example, the use of a rising intonation to signal yes/no questions. The frequency dimension concerns similarities and differences in the frequency of use of and distribution of phonological categories; for example, the frequency and distribution of the use of pitch accents and boundary tones. In addition to this, Mennen (2015) pointed out that the LILt can allow for a systematic comparison between L2 learners with different levels of proficiency, different ages of arrival, different L1, or any other relevant variables.

The following section 2.5 deals with research into L2 prosody, intonation and focus. It also provides descriptions of factors affecting L2 speech production.

2.5 Research and Studies on L2 Prosody, Intonation and Focus

2.5.1 Overview of studies of L2 prosody, intonation and focus

Studies of the acquisition of prosodic characteristics such as lexical tone, stress, intonation, and the rhythm of a second language have recently started to receive more attention. Similarly, either the British or the AM approaches have been implemented in a number of experimental investigations in this field. Some studies have concentrated on L2 learners' acquisition and awareness of prosodic aspects of speech in terms of the interpretation of forms and meaning (Cruz-Ferreira, 1987; Pierrehumbert and Hirschberg, 1990; Ramiŕez Verdugo, 2005, 2006; Wayland, *et al.*, 2006). Just as L2 learners' errors may be due to the incorrect placement of prominence (McGory, 1997; Mennen, 1999), differences in pitch range have also been discovered (McGory, 1997; Gut and Pillai, 2014). There have been numerous studies that have reported on transfer from the L1 to L2 learner performance in terms of intelligibility and foreign accentedness (Tahta, Wood and Loewenthal, 1981; van Els and de Bot, 1987; Munro and Derwing, 1995; Derwing and Munro, 1997; Ortega-Llebaria and Laura, 2014). Some studies have examined the effects of intra- and inter-speaker factors, such as the age of arrival, length of residence, and L2 experience (Trofimovich and Baker, 2006; Huang and Jun, 2011). In addition to this, some studies have documented L2 learners' production compared with that of native speakers as a benchmark, or otherwise compared non-native speakers' production among those with differing L1 backgrounds (Wennerstrom, 1994; Ueyama and Jun, 1998; Swerts, Krahmer and Avesani, 2002; Mennen, 2006; 2007). Cross-language studies have also found differences in the pragmatic effects of intonation (Cruz-Ferreira, 1987; Chen *et al.,* 2001).

Also, there have been several studies of the functions of the prosodic elements of speech in communication. These studies maintain that, while miscommunication between native speakers and non-native speakers is often caused by stress patterns, the role of intonation should not be underestimated (Wennerstrom, 1994; Clennell, 1997; Pickering, 2001, 2009). The sources of problems vary, but failure in communication is one of the consequences. According to Jenkins (2002, p. 88), mistakes in prominence or "tonic (or nuclear) stress" are one category of phonological error which can cause breakdowns in communication, while mistakes in pitch patterns have smaller effects. Grabe, Kochanski and Coleman (2005) emphasised that it is worth L2 learners learning where to place nuclear accents in the first instance. And then it is also worthwhile for them to recognise why native English speakers place nuclear accents at

particular locations or positions. Pickering (2001, 2009) also pointed out that tone choice should be integrated into L2 English instruction programmes since it contributes to communication failure and success.

From this discussion, it is quite clear that certain aspects of the learners' L2 English, for example, its intonation patterns, exhibit different characteristics from those of native speakers. Also, several factors such as L1 transfer and experience may exert an influence on these patterns. Raising L2 learners' awareness of intonation and the marking of focus will enable them to avoid misunderstandings when talking to native English speakers as well as non-native speakers from different L1 backgrounds. This is so because the appropriate intonation can facilitate the more effective communication of messages.

The next section 2.5.2 deals with factors influencing L2 learners' speech production and its characteristics. In sections 2.5.3 - 2.5.5, research in relation to areas of cross-linguistic differences in L2 intonation and focus marking is then summarised and fully cross-referenced.

2.5.2 Factors affecting L2 speech production

Numerous studies have examined and reported on the factors that have an influence on the production of L2 learners' speech (e.g., Piske, Mackay and Flege, 2001; Rasier and Hiligsmann, 2007; Major, 2008; Ng and Chen, 2011; Kainada and Lengeris, 2015). Notable studies concerning first language transfer, learners' variability such as age, language exposure and proficiency, as well as language input are summarised below:

1) First language influence

First language (L1) background plays a major role in second language learning. This is normal for L2 learners, especially when adult speakers learn to speak a foreign or second language. The production of a new language may share the characteristics of their mother tongue, and, for example, this results in foreign-accented pronunciation (Lado, 1957; Derwing and Munro, 1997; Piske, Mackay and Flege, 2001; Rasier and Hiligsmann, 2007). When a native language affects the linguistic characteristics of the L2, the processes of learning in terms of production and perception, or even the overall development of a second language among L2 learners, this is known as language transfer or influence. This accords with Lado's (1957, p. 2) statement that

... when learning a foreign language, we tend to transfer our entire native language system in the process. We tend to transfer to that [or a foreign] language our transitions, our intonation patterns and their interaction with other phonemes ...

First language influence can be primarily divided into two kinds, namely that with positive effects or with negative effects (Wells, 2006; Lu and Kim, 2016). The former case is termed 'positive transfer', which takes place when L1 characteristics consolidate or enhance L2 acquisition, particularly when the two language systems share similar linguistic elements. On the other hand, the latter case is termed 'negative transfer', which takes place when L1 linguistic characteristics interfere with the L2 learning process, particularly when the systems of the two languages are different. It can be noticed that the degree of differences between the languages has an effect on L2 learning. That is to say, the greater the differences between an L1 and L2, the greater will be the effects of first language influence. Therefore, cross-language differences can result in learning problems or difficulties to a great extent. Several studies (Nguyễn *et al.,* 2008; Lu and Kim, 2016; van Maastricht, Krahmer and Swerts, 2016) have identified this issue in the learning of prosodic as well as segmental L2 features. For example, Nguyễn *et al.* (2008, p. 158) demonstrated that "native speakers perceive and produce words and utterances of L2 through a phonetic or phonological 'filter' of their native language (L1)". Besides this, previous studies (Trofimovich and Baker, 2006; Mennen and De Leeuw, 2014) have reported that L2 learners still transfer prosodic features in their first language into a new language even though they have been exposed to the new language for many years.

2) *Language exposure, input and proficiency*

In Thailand, as noted in the Introduction chapter, instruction in English started in the reign of King Rama III (1824-1851). The teaching and learning of English is considered to be a school subject in foreign languages from primary school to university level (Boonkit, 2002; Wongsothorn *et al.*, 2002). It should be noted that, despite many years in studying, English in Thailand is taught as a foreign language by native Thai teachers and the exposure to the language is limited, as mentioned earlier. Therefore, language exposure and input should be taken into consideration. Firstly, the amount of exposure to an L2 affects how proficient speakers may become in using it. This factor has been well documented in earlier studies (Flege, 1995; Piske, Mackay and Flege, 2001; Trofimovich and Baker, 2006) with the same trend of findings. These studies indicate that there is a tendency for L2 learners to improve their

production and perception of L2 speech as the amount of exposure to the second language increases.

Recently, a number of studies have investigated the issue of input in the second language context. Young-Scholten (1994, 1995) found that the linguistic input of L2 learners' teachers and peers can have negative effects, especially when the input deviates from the norms of native English speakers. Young-Scholten pointed out that, if L2 input is insufficient and limited, the chances of achieving native competence for L2 learners are low regardless of their age at the time of exposure to the input. In turn, this kind of input would result in non-target language that is acquired by L2 learners. This statement is in line with the studies of Flege (1991), Sumdangdej (2007) and Flege (2009). In addition to this, there is the question of whether input should be simplified or natural. The input needs at least to be appropriate for L2 learners in terms of their age, and the more exposure to the L2 from native English speakers in the same age range the better (Young-Scholten, 2013). In this sense, L2 learning can be enhanced substantially. If L2 learners aim to become near-native, they ought to receive input from native English speakers, not just from non-native teachers and other L2 learners.

3) Learners' variability: age

Among other variables, the age of the L2 learner seems to be commonly acknowledged as a potentially important factor in learning a second language. Lenneberg's (1967) notion regarding the critical period for language learning supports this idea. In turn, the critical period has been evidenced in numerous studies in that learning an L2 or a new target language after puberty is more difficult compared to learning it at an earlier age. L2 learners are likely to achieve less mastery over the linguistic elements of the target language; or otherwise, their L2 will still have irregular and incomplete linguistic characteristics. This is particularly the case for phonology, as shown in the work of Herschensohn (2007), Ioup (2008) and Moyer (2009). To achieve success with L2 phonology, it is suggested that L2 learners should start learning it at the age of 6 or at the earliest age possible. For L2 syntactical aspects, it is suggested that L2 learners start learning it before the age of 15 (Long, 1990; Flege, Munro and MacKay, 1995). More specifically, younger learners are found to acquire phonological (and morphosyntactical) features faster than older learners according to Flege *et al.* (2006) and MacKay, Flege and Imai (2006). Besides this, younger learners with more time in using L2 English exhibit a greater level of proficiency (Major, 2014).

To summarise, several factors influence the production of L2 learners' speech. The main factors that the present study take into consideration are L1 transfer, L2 English proficiency levels of the Thai learners. Other factors such as length of residence, motivation, and learning goals are not discussed here and are beyond the scope of this study. The following subsections deal with previous investigations concerning the transfer of L1 prosody, word stress, nuclear accent placement, and focus marking, and tonal alignment and scaling.

2.5.3 Studies of the transfer of L1 prosody

Several studies (e.g., McGory, 1997, Ueyama, 2000, Nguyễn *et al.*, 2008; van Maastricht, Krahmer and Swerts, 2016) have confirmed that the first language has an influence on the L2 intonation of non-native learners. Notable findings from previous studies such as those by McGory (1997), Ueyama (2000) and Nguyễn *et al.* (2008) as to the L1 influence are summarised as follows:

McGory (1997) investigated the acquisition of intonation patterns in L2 English. The participants were native Korean and native Mandarin Chinese speakers. They were grouped according to English proficiency levels. McGory investigated the locations of word stress, intonation in declaratives and yes/no questions, and levels of prominence corresponding to phonetic correlates in terms of fundamental frequency, duration, and intensity on the accented words. The findings revealed that there was an influence of L1 Korean and Mandarin Chinese intonation on L2 English intonation patterns. The findings also showed a strong effect of L2 proficiency level. The evidence showed that native Korean speakers tended to increase the duration of stressed syllables, but they did not reduce that of unstressed syllables. This is the same acoustic cue to accentual prominence as in Korean. Meanwhile native Mandarin Chinese speakers tended to transfer lexical tones from their native language into English intonation, producing both accented syllables and stressed syllables by using higher pitch.

Ueyama (2000) investigated how L1 prosodic characteristics influence the prosodic patterns of L2 English and L2 Japanese. The participants were L2 adults from the two different native languages of Japanese and American English, with two levels of L2 proficiency of beginner and advanced. Four types of stimuli were used: L2 English produced by L1 Japanese speakers, L2 Japanese produced by L1 English speakers, L1 English, and L1 Japanese. Ueyama examined several prosodic phenomena such as differences between lexically accented and unaccented

vowels, those between English tense and lax vowels and between Japanese short and long vowels, and temporal aspects or rhythmic structure across syllables. These prosodic phenomena were subject to phonological and phonetic analyses; therefore, fundamental frequency and duration were also examined. The findings of the study showed that there was evidence of L1 prosodic transfer into L2 prosodic patterns, and the degree of transfer varied from language to language according to specific phonological characteristics. There was also an interplay between prosodic and segmental features during the development of L2 speech.

Nguyễn *et al.* (2008) investigated the influence of prosodic transfer in Vietnamese on contrastive stress patterns in English. The participants were native Australian English speakers and Vietnamese learners of English who were grouped as beginners or advanced learners. Production and perception experiments were conducted to examine English contrastive stress patterns at word and prosodic levels. The three patterns of stress tested were: broad-focus noun phrase, narrow-focus noun phrase, and compound, in terms of the acoustic correlates f0, intensity and duration, as well as the perceptual strategies that the participants used. The findings showed that the first language of each group of the participants affected the stress patterns under investigation. That is, the acoustic features or patterns of stress were used differently according to first language phonology. It was found that native speakers used all of the acoustic features of f0, duration, and intensity in combination to contrast stress patterns. Vietnamese learners of English could use f0 and intensity contrastively on accented syllables, but they had difficulty in using durational contrast to signal prominence in compound words and polysyllabic words or phrases. The authors concluded that this was possibly due to Vietnamese-specific phonology involving the limited use of duration for contrast. Another interesting finding was that the effect of language exposure on the acquisition of prosodic features was identified. That is, advanced speakers exhibited a better ability than beginners to deaccent narrow-focused words and to compress compound words.

2.5.4 Studies of word stress, nuclear accent placement, and focus marking

Studies of word stress and the location of nuclear accents have a long history in the field of second language acquisition. Evidence can be traced back from earlier work to more recent research (e.g**.,** Fokes and Bond, 1989; Wennerstrom, 1994; Low and Grabe, 1999; Chen *et al.*, 2001; Nguyễn *et al.*, 2008; Nagy, 2015). Below are summaries of notable findings from selected studies.

As cited in Mennen (2006, 2007), Jenner (1976) found that L2 learners have problems with the correct placement of prominence at the sentence level as well as within a word, and Backman (1979) investigated the intonation contours of L2 learners and then compared them with those of native speakers of English. One of the findings was that L2 learners placed prominence too far to the left. Fokes and Bond (1989) investigated stress patterns in words and sentences produced by non-native speakers and native American English speakers. They examined the acoustic-phonetic correlates of stress, such as amplitude, fundamental frequency, and duration. They found that L2 learners had problems with word stress placement and they differed most from native speakers in terms of the amplitude and duration used. McGory (1997) explored the acquisition of intonational prominence in English by Seoul Korean and Mandarin Chinese speakers. She also found that both groups of L2 speakers appeared to have difficulties in producing this type of prominence. Although the Mandarin Chinese participants did not have difficulty in reducing the duration of unstressed syllables, they did find it difficult to produce and place nuclear pitch accents in English. Korean participants had difficulties in both reducing the duration of unstressed syllables in word-initial positions as well as in producing and placing nuclear pitch accents.

An interesting finding from a study by Low and Grabe (1999) was that differences between British English and Singapore English stress placement were not the result of a difference in lexical stress placement. Instead, their findings suggested that Singapore English and British English are different in the phonetic realisation of stress. The former is produced with more phrase-final lengthening and without decreasing pitch prominence. Wennerstrom (1994) investigated how speakers use intonation to assign prominence in discourse. The participants were from Spanish, Thai and Japanese language groups. Two tasks using read and spontaneous speech were used in the experiments. The findings showed that native English speakers consistently marked new and contrastive information with a higher peak in fundamental frequency (f0) in both tasks. Thai and Japanese participants did not show the use of high f0 for the target words in read speech. In addition, the Thai and Spanish participants did not show the use of higher f0 in spontaneous speech either. Wennerstrom concluded that the non-native speakers of English made fewer pitch contrasts and failed to produce intensity contrasts on words with prominence. Besides this, native speakers and L2 learners with more exposure to English tended to have shorter durations of pauses between pitch peaks.

2.5.5 Studies of tonal alignment and scaling

As discussed previously in section 2.2.7, several aspects of tonal alignment and scaling are still controversial in AM theory, and further research is needed. For example, the segmental anchoring hypothesis (Arvaniti, Ladd and Mennen, 1998; Ladd, 2006) assumes that, throughout the course of accentual f0 movement, its beginning and end points are associated with specific locations and are independently aligned with specifiable points along the segmental string. There have been a growing number of studies in this field (e.g., Silverman and Pierrehumbert, 1990; D'Imperio, 2000, 2001, 2012; Frota 2002; Prieto and Torreira 2007; Ladd *et al.,* 2009; O'Brien and Gut, 2010; Gut and Pillai, 2014; Lu and Kim, 2016; Graham and Post, 2018; Kuronen and Tergujeff, 2018), and the original notion of the segmental anchoring hypothesis has been revised, although it is still accepted at present in studies regarding tonal alignment and scaling in a less strict way. Summaries of the most important studies of tonal alignment and scaling, including L1 English and not limited to L2 English, are provided next.

Atterer and Ladd (2004) investigated the rising peak accent (L+H*) in the prenuclear position in German. They found that this rising accent was aligned later than the same rising accent in British English. It was also revealed that the later alignment in English may be attributed to L1 German transfer. In addition, they reported that there were dialectal differences within Germans in terms of the phonetic timing of peaks. The H tone varies from dialect to dialect whereas the L tone in Southern German has a consistently later alignment. Atterer and Ladd concluded that the tonal targets of rising accents in Northern and Southern German did not align independently, and they showed a higher degree of peak delay. Lu and Kim (2016) investigated the tonal alignment of the L*+H pitch accent in English produced by native speakers of Mandarin Chinese. One aim of their study was to examine the L1 transfer of tonal alignment into L2 English. The results showed that the L1 alignment pattern remains dominant in Mandarin native speakers' production of the L*+H pitch accent in English. In addition, they found that the effect of speech rate on tonal timing was less consistent in the production of Mandarin.

In addition to this, Avaniti, Ladd and Mennen (1998) reported that the tonal alignment of rising accents in Greek was different from those in English. That is, the L and the H tonal targets align beyond the stressed syllable in L*+H accents in Greek, but not in English. This reflects the fact that the same phonological categories may differ in their phonetic alignment cross-linguistically. Such differences have also been reported in the realisation of the same pitch

accent categories in varieties of English. For instance, Arvaniti and Garding (to appear) investigated the phonetic realisation of the two rising accents L+H* and L*+H in varieties of American English. They found that the L of the L+H* accent was aligned on the stressed syllables while the same L in the L+H* accent can be found before the stressed syllables. They also found a shorter duration of the rise in their study and reported that there may be dialect-specific differences in American English.

Moreover, Arvaniti and her colleague reported on tonal scaling between the L*+H accent and the L+H* accent. They reported that, in the L*+H accent, the scaling of the L tone was lower and the scaling of the H tone was higher when compared to the L and H tonal targets in the L+H* accent. Ladd *et al.* (2009) investigated peak alignment in the prenuclear and nuclear positions in British English. Their investigation showed that prenuclear and nuclear accents had different patterns of alignment across language varieties. One finding was that the high element of prenuclear and nuclear accents in Standard Scottish English was aligned later than that in Southern British English or RP English. In addition, prenuclear accents were aligned in a delayed manner with short vowels. They concluded that the syllable structure may be influential.

To conclude, these studies demonstrate that the L2 characteristics of non-native learners differ from those of native speakers of English in various different respects. For instance, L2 learners may produce different pitch accent types both in terms of placement and phonetic realisation. They may use different strategies in the expression of focus phonologically and phonetically. These can be considered as L2 errors, such as in terms of inappropriate nuclear accent placement and deaccenting errors. These errors can mislead hearers about what information is new and what is given whereas the speakers can sound unnatural owing to misplaced nuclear pitch accents. One cause of these errors is likely to be transfer from their first language to a novel language. The next section deals with current status of L2 English in Thailand.

2.5.6 Current status of L2 English in Thailand

Referring back to the Introduction chapter, it is more common to refer to English as a foreign language (EFL) in Thailand than English as a second language (ESL). Thai learners do not learn English in everyday life, nor do they use it to communicate every day (National Identity Board, 2000). The main purposes of learning English in Thailand are to get a better job with a higher

salary, as well as to receive higher education. Thai learners have the chance to use English with their friends or their teachers only in the classroom. Another reason is that they have learnt English because it is part of the compulsory educational curriculum (Wongsothorn, Hiranburana and Chinnawongs, 2002); therefore, aiming to pass examinations or course criteria may be short-term goals for students learning English. Using English beyond the classroom would involve personal goals or motivations. Even though Thailand has joined the ASEAN, and English is needed as a working language for communication among countries in this community, there seem to be limitations concerning its use only for specific purposes (Kirkpatrick, 2008; Hayes, 2010). Therefore, the situation of teaching and learning English in Thailand is quite different from that in other countries.

When considering the status of the model of pronunciation, RP (Received Pronunciation) English or Standard Southern British English (SSBE) is traditionally accepted and highly valued in every educational institution in Thailand. RP can be called "General British", or "Educated Southern British English", or "BBC Pronunciation or BBC Accent" according to Roach (2004, p. 239). This accords with Deterding's (1994) statement that "Standard Southern British pronunciation (sometimes referred to as RP) is still standard for pronunciation … for teaching purposes and also in the media" (p. 61). Hughes and Trudgill (1987, p. 3) illustrate the prestigious status of RP as a model of pronunciation for L2 learners as follows:

> [A] readily available example of RP [is] the speech of BBC newsreaders. Because of its use on radio and television, within Britain RP has become the most widely understood of all accents. …the learner who succeeds in speaking it, other things being equal, has the best chance of being understood. …[RP] is by far the most thoroughly described of British accents. …descriptions of it were [also] made in response to the needs of foreign learners and their teachers.

Such valuing and acceptance of RP are also influential for Thai scholars, educators and teachers. They are also trained to teach English intonation with materials based on the British school (Estebas-Vilaplana, 2017). When they have taught English pronunciation or intonation, they normally use an RP or British English model-based textbook equipped with CD-ROM and Teacher's Manual, for example, a series of 'Ship and Sheep' and 'Tree or Three' by Ann Baker, 'English Pronunciation in Use' and 'Pronunciation Tasks' by Martin Hewings, and 'Gimson's Pronunciation of English' by Alan Cruttenden. However, when using such textbooks and audio-visual materials, teachers make the most of these resources as a teaching medium via translation into Thai (Boonkit, 2002; Hayes, 2008). A study by Jindapitak (2010) also reveals that British

English is still preferable among Thai learners. Even though they can receive other varieties of English outside, the chances of this are remote since communication and mass media in daily life are in Thai, as discussed previously based on the studies by Wongsothorn, Hiranburana and Chinnawongs (2002), Wiriyachitra (2001), Khamkhien (2010), and Hiranburana (2017). This situation accurately reflects the status of learning and teaching English in Thailand. Consequently, the present study examines Standard Southern British English (SSBE) as the starting point of interest within its scope of investigation, due to its prevalence in instructional contexts in Thailand. The term 'Standard Southern British English (SSBE)' is used to cover the term 'RP' English here as stated in Deterding (1994), Roach (2004), and Ladd *et al.* (2009). The term 'English as a second language (ESL)' and 'English as a foreign language (EFL)' are also used here interchangeably as well.

2.6 Approaches to Intonational Studies and Transcription Models

2.6.1 Approaches to intonational studies

The following sections firstly elaborate on the two prominent approaches to intonational studies; the British nuclear tone approach and the Autosegmental-Metrical (AM) approach. The ToBI and IViE models of intonation transcription are then explained. Justifications for implementing the AM approach and ToBI are subsequently discussed in detail.

1) The British nuclear tone approach

The British nuclear tone approach (Crystal, 1969; O'Connor and Arnold, 1973; Cruttenden, 1997; Wells, 2006) is a tone-unit or contour-based system that has been well established for the investigation of British English intonation. This approach is known as the British school or tradition. Some scholars or researchers (e.g., Crystal, 1969; O'Connor and Arnold, 1973; Roach, 2009) prefer to refer to the structure of British English intonation using the term 'tone unit' or 'intonation group' and the like. Within the tone unit, there is a clearly defined internal structure which is still used in intonational studies within the British nuclear tone tradition. A complete structure of a tone unit is provided in example (11) as follows:

(11)   (prehead) (head) nucleus (tail)

<div align="center">(Crystal, 1969, p. 208)</div>

According to Crystal (1969), O'Connor and Arnold (1973), Roach (1991), and Wells (2006), the prehead is the first part of the intonational phrase, which usually consists of any unstressed material. The head is the part of the intonational phrase to which the first accent (onset accent) belongs and it extends up to the last syllable before the nucleus. The prehead and the head in combination can be referred to as the 'pre-nuclear accent' pattern, which is an optional term but it can convey meaning when used with a particular nuclear tone in the intonational phrase. Cruttenden (1997) states that the tone choices for the head have more variation than the tone choices of the nucleus, and this is a reason why there has been more research into nuclear accent patterns than pre-nuclear and head patterns. Another reason is that the nucleus is considered to be an obligatory component of the tone unit or intonational phrase. It is the location of the nuclear accent, as mentioned earlier. The tail is the last part of the intonational phrase, carrying no stressed material. The nucleus and the tail can be referred to as the 'nuclear accent' and 'post-nuclear accent' patterns respectively. Therefore, example (11) can be modified to:

(12)   (pre-nuclear accent) nuclear accent (post-nuclear accent)

The principal notions of the British tradition are based on the three components of tonality, tonicity, and tone (or the 3Ts). These are still in widespread use among many researchers who have the ultimate goal of investigating intonation in British-style descriptions to date. This approach to intonation analysis has also been taken as a model for the teaching of English intonation to foreign language learners dating back to the time of Halliday's (1967, 1970) work. With explicit instructions about which nuclear tones to use and where to place tonic syllables in teaching materials, the British approach has played a central role, particularly in second language classrooms (Deterding, 1994).

The British tradition accounts for intonation patterns as a series of pitch movements. There are three main elements commonly used in descriptions of intonation patterns in the British approach. These are tonality, tonicity, and tone. Brief summaries of these elements are provided as follows, mainly based on work by Cruttenden (1997), Wells (2006), and Roach (2009). Firstly, the term 'tonality' refers to the division of spoken material into chunks of information or intonational phrases (IPs). The size and number of intonational phrases vary depending on the speaker's discretion as well as different styles of speech (Wells, 2006). For example, an

intonational phrase tends to be longer and has more accents in reading-aloud tasks, whereas it tends to be shorter with fewer accents in ordinary conversation. Intonational phrases have intonational breaks or boundaries which often correspond to the syntactic boundaries of words, phrases, clauses, and sentences. Sometimes, it is called a 'thought group' or 'intonation group' or 'tone unit' or 'tone group'. It is typically recognised by the presence of a pause, final syllable lengthening, or a slow speech rate at the end of the intonational phrase, followed by a faster speech rate at the beginning of the next phrase, and a pitch reset. Example (13) illustrates how an utterance can be broken up into intonational phrases using the symbol '|'.

(13)   We don't know who she is.
        We don't know | who she is.
        We | don't know who she is.
        We don't | know who she is.
        We | don't know | who she is.
                    (Wells, 2006, p. 7)

In this example, there are several possibilities of dividing up or presenting the sentence 'We don't know who she is'. It completely depends on the speakers' discretion as to whether to say the sentence as a whole sentence, or as two pieces of information (Wells, 2006) If it is the former case that the speaker decides to utter, the sentence will have one intonation phrase. If it is the latter case, the sentence will have two intonation phrases, when an intermediate phrase in the AM-styled description is not taken into consideration.

The term 'tonicity' refers to the location of the nucleus (nucleus placement), where the syllable bears the nuclear tone. Sometimes the term 'accentuation' can be used instead for locating where an accented syllable is. The accented syllable here is often known as the 'tonic syllable' or 'nuclear syllable'. It bears the focus of the information that the speaker wants to convey in terms of what is meant for listeners, so that the decision is made to highlight the most important words by giving pitch prominence. The location of a nuclear accent is usually on the last content word which will be most prominently accented in the intonational phrase; however, it can be function words in cases where the speaker is making a contrastive focus, or when he/she wants to emphasise particular information. Example (14) illustrates how words in an utterance can be accented by placing pitch prominence using the symbol '_'.

(14)   She was ˈtrying to lose ˈ<u>weight</u>.

   She was ˈtrying to <u>lose</u> weight.

   She was ˈ<u>trying</u> to lose weight.

   She ˈ<u>was</u> trying to lose weight.

   ˈ<u>She</u> was trying to lose weight.

<div align="center">(Wells, 2006, p. 139)</div>

In this example, Wells (2006) gave explanations as to the effects of changing the location of nucleus placement in the sentence to change the status given to information, making new information stand out, and expressing different kinds of contrastive focus. Besides their own discretion on chunking information, speakers very often vary nucleus placement to draw the listener's attention to whichever piece of information they want to highlight. To do this, they can bring every piece of information into focus. For example, 'She was ˈtrying to lose ˈ<u>weight</u>' could be considered to have broad focus, where the last content word is commonly put into focus or there is neutral tonicity. On the other hand, the same sentence would be considered to have narrow focus, or contrastive focus, if it was chosen to put the focus on a particular word; for example, 'weight', 'lose', 'trying', 'was' and 'She' as in example (14).

The term 'tone' in the British tradition refers to the pitch movements or the directions of pitch movements found on the nucleus or tonic syllable in an intonation unit. After the speakers have made decisions on tonicity, they will choose a specific nuclear tone, such as a fall or rise associated with the nucleus for that intonational phrase. In English, there are several possible tones for a speaker to choose. Basically, the system of English nuclear tones consists of four main tones: falls, rises, fall-rises, and rise-falls, whereas level tones can be considered in some systems (Wichmann, 2000). These nuclear tones can be categorised with the most basic distinction such as falling and non-falling, or with finer distinctions as high-falls, low-falls, rise-falls, low-rises, high-rises, fall-rises, and mid-levels (Cruttenden, 1997; Wells, 2006).

Since the British nuclear tone model and the AM model have long exerted an influence on studies on English intonation as an L1/L2, it is not an easy task to decide which model of intonation analysis should be used. In the second language acquisition (SLA) context, no language teacher can avoid the influence of the British nuclear tone model for teaching L2 English intonation, so that this approach is "widely used both descriptively and pedagogically" (Deterding, 1994, p. 62). The systems of transcription of intonation within this model vary from scholar to scholar (e.g., Halliday, 1967, 1970; Brazil, Coulthard and Johns, 1980; Cauldwell

and Hewings, 1996; Wells, 2006). For instance, orthographic transcription with iconic notation (tonetic-stress marking) supplemented in the written text itself offers detailed descriptions of tonality, tonicity, tone and even finer tonal distinctions. Example (15) illustrates how utterances are orthographically transcribed with iconic symbols (ˈ for high level; ˌ for low rise; ´ for high rise; > for mid-level, and ˌ for low fall):

(15)   ˈFirst we have ˌone thing, | ˈthen we have aˌnother.
        ˈFirst we have ´one thing, | ˈthen we have aˌnother.
        ˈFirst we have >one thing, | ˈthen we have aˌnother.

                                            (Wells, 2006, p. 224)

Interlinear tonetic or tadpole transcription offers graphic representations of the intonation patterns which accommodate the learning and teaching of the pitch movements of the British system par excellence. Example (16) below illustrates how utterances are transcribed with the tadpole diagrams used by Crystal (1969), O'Connor and Arnold (1973), and Cruttenden (1986). In these examples, it should be noted that the British system of intonation transcription falls into two categories in itself. That is, the interlinear tadpole transcription in example (16) is similar to narrow transcription (Cruttenden, 1997), and transcription with iconic notation inserted in ordinary written text as in examples (13, 14 and 15) is broad transcription (Wichmann, 2000).

(16)
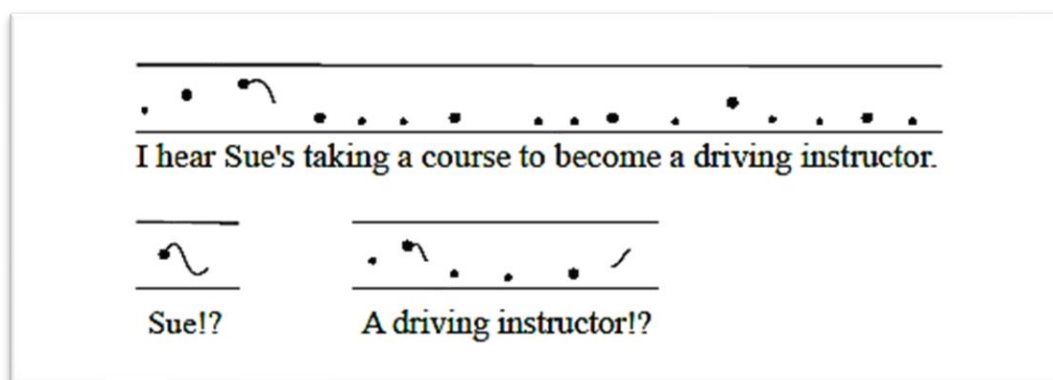
                                            (Ladd, 2008, p. 48)

To conclude, the notions of intonation systems in British English with the 3Ts of tonality, tonicity, and tone are very often used to account for intonation in the literature (Halliday, 1967, 1970; Wells, 2006; Roach, 2009). By definition, tonality is used to refer to the division of

speech materials into intonation groups or tone units. Tonicity is used to refer to the placement of the tonic syllable or nucleus. Tone is used to refer to tone or pitch movement. With some modifications and adjustments of notation systems concerning the transcription of intonation such as the interlinear tadpole transcription and the transcription with tonetic-stress marking, language teachers and researchers aiming to develop teaching methodologies for intonation can make the most use of this model. Therefore, it is not surprising that the British model is still used across the board in studies of the SLA of intonation (e.g., Halliday, 1970; O'Connor and Arnold, 1973; Cruttenden, 1997; Wells, 2006; Roach, 2009).

2) The Autosegmental-Metrical approach

As proposed by Ladd (2008), the Autosegmental-Metrical (AM) approach is referred to as a phonological theory of intonational and prosodic structure. This approach has long been developed from work on intonation systems (Liberman, 1975; Pierrehumbert, 1980; Gussenhoven, 1984; Beckman and Pierrehumbert, 1986; Pierrehumbert and Hirschberg, 1990; Ladd, 2008, Beckman and Venditti, 2011). In essence, tone-bearing units under the AM approach are considered to be primary components which a string of tones can be associated with. Thus, the two level tones of high (H) and low (L) constitute categorical elements or primitives in phonological representations of intonation. A complex intonation contour can be transcribed as combinations of these two tonal targets (e.g., L*+H LL%, or L*+H L-L%).

The AM approach treats intonation contours in terms of three distinctive tonal events: pitch accents, phrase accents and boundary tones (Pierrehumbert, 1980; Pierrehumbert and Hirschberg, 1990; Beckman and Ayers-Elam, 1997; Cruttenden, 1997: Ladd, 2008). A very basic distinction can be made between two different kinds of tones: head-marking tones, and edge-marking tones. Head-marking tones are pitch accents which are associated with metrically stressed syllables in the utterance, while edge-marking tones are phrase accents and boundary tones which are associated with the edges of intonational phrases (the intermediate phrase and the full intonational phrase respectively). AM theory was developed from two approaches: the Autosegmental approach and the Metrical approach (e.g., Liberman, 1975; Goldsmith, 1976; Pierrehumbert, 1980). It is autosegmental in the sense that it describes a discrete phonological element of intonation as consisting of a string of independent tonal autosegments which may be high or low. These tonal targets are represented on a separate autonomous tier. It is metrical in the sense that tone is associated with lexically stressed syllables that are metrically stronger than nearby unstressed syllables.

In addition, AM theory aims to provide descriptions mapping from phonological elements to continuous acoustic parameters of intonation. In theory, according to Pierrehumbert (1980) Grice (1995), and Ladd and Schepman (2003), the high and low tonal targets are not static points, and their positions are not fixed levels in the speaker's overall pitch range. The positions of high or low targets are relative levels of high or low tonal targets in the same utterance. The tonal targets are aligned with parts of the segmental strings so that all intervening pitch is derived simply by phonetic interpolation between these targets. This mapping process is known as text-tune association (Ladd, 2008). In this case, phonological representations of tones are realised as tonal targets that can be defined along two dimensions of alignment and scaling. Details of tonal alignment and scaling were previously provided in section 2.2.7 and more details regarding research in this area are provided in section 2.5.5.

Finally, the Autosegmental-Metrical (AM) theory has exerted a strong influence on intonational studies of languages other than English, such as Japanese (Beckman and Pierrehumbert, 1986; Venditti, 2005), Neapolitan Italian (D'Imperio, 2001), German (Grice, Baumann and Benzmuüller, 2005), Greek (Arvaniti and Baltazani, 2005), Korean (Jun, 2005), and Mandarin (Peng *et al.,* 2005). AM descriptions and analyses of intonation also offer substantial potential in comparing intonational aspects across languages. Its theoretical framework of intonation underlies the Tones and Break Indices (ToBI) system, which is a widely used notational convention for prosodic transcriptions in American English (Silverman *et al.,* 1992, Beckman and Ayers-Elam, 1997; Beckman, Hirschberg and Shattuck-Hufnagel, 2005). Recently, AM theory and ToBI have influenced a growing body of research into intonational phonology within and across languages.

Since its development, the Autosegmental-Metrical approach has been well-known for analysing pitch contours as distinctive pitch levels, whereas the British approach analyses pitch contour as pitch configurations such as falling and rising patterns (O'Connor and Arnold, 1973; Deterding, 1994; Cruttenden, 1997; Wichmann, 2000; Wells, 2006). In the AM theoretical framework, there are four basic assumption concerning sequential tonal structure, distinctions between pitch accent and stress, the analysis of pitch accents in terms of level tones, and local sources for global trends. These four assumptions are important in the present study and are presented according to Ladd's (2008) description as follows:

(1) Sequential tonal structure: tonal structure consists of a string of local events associated with certain points in the segment string. Between such events the pitch contour is phonologically unspecified and can be described in terms of transitions from one event to the next. In languages like English, the most events of the tonal string are pitch accents, which are associated with prominent syllables in the segmental string, and edge tones, which are associated with the edges of intonational tunes at major prosodic boundaries.

(2) Distinction between pitch accent and stress: pitch accents, in languages that have them, may serve as concrete perceptual cues to stress or prominence. However, they are in the first instance intonational features, which are associated with certain syllables in accordance with various principles of prosodic organization. The perceived prominence of accented syllable is, at least in some languages, a matter of metrical strength and/or dynamic stress, which can be distinguished from pitch accent.

(3) Analysis of pitch accents in terms of level tones: pitch accents and edge tones in intonational languages can be analysed as consisting of primitive level tones or pitch targets, High (H) and Low (L).

(4) Local sources for global trends: the phonetic realization or scaling of any given H and L tone depends on a variety of factors (degree of emphasis, position in utterances, etc.) that are essentially orthogonal to its identity as H or L. Overall trends in pitch contours (e.g., gradual lowering of overall range) mostly reflect the operation of localised but iterated changes in scaling factors. (Ladd, 2008, p. 44, italics in original)

These AM assumptions have become a basis for the development of the ToBI system (Silverman *et al.,* 1992; Pitrelli, Beckman and Hirschberg, 1994**;** Beckman and Ayers-Elam, 1997). Since their development, the AM theory and the ToBI system have begun to influence work on intonational aspects in different languages, including but not limited to varieties of English, German, Greek, Spanish, Korean, and Mandarin Chinese. Besides its four well-defined assumptions, AM theory proposes a taxonomy of cross-language differences in intonation in four dimensions: "semantic, systematic, realisational, and phonotactic" (Ladd, 2008, p. 115). Ladd clarified that these cross-linguistic variations involve differences in the meaning or use of the same tune, the inventory of phonologically distinct tune types irrespective of semantic differences, of detail in the phonetic realisation of the same tune, and tune-text associations and the permitted structure of tunes.

2.6.2 Models of the transcription of intonation

      1) *The ToBI*

The term 'ToBI' stands for the Tones and Break Indices system for prosodic transcription. The ToBI system can be used to transcribe not only intonation patterns, but also other prosodic aspects like stress, rhythm, and tempo. As its name suggests, ToBI refers to two core components of prosodic analysis. These are *To* for tonal patterns, and *BI* for boundaries in intonational phrasing. This system was first devised by multi-disciplinary researchers (Silverman *et al.,* 1992; Pitrelli, Beckman and Hirschberg, 1994; Beckman and Ayers-Elam, 1997**)** in the USA since 1992. It has been used as a prosodic notation system for the analysis of English utterances, especially Mainstream American English. It was also developed on the basis of Autosegmental-Metrical (AM) theory (Ladd, 2008), and from Pierrehumbert's (1980) work analysing American English intonation which proposed that any pitch movements of an utterance can be described as sequences of the two level tones, high (H) and low (L). Since its development, the ToBI system has been revised and is still operationalized in a large number of recent studies of speech analysis and synthesis.

It is worth noting that the ToBI is not an International Phonetic Alphabets (IPA) transcription system for prosody. It is a tool for a researcher to annotate prosodic features of a language under analysis only. This is because a language has its own intonational and prosodic systems which differ from language to language and among its own varieties. Apart from its notational conventions, the intonational descriptions of the ToBI system can be adapted to apply to other work, such as speech recognition and synthesis, and speech technology (Silverman *et al.,* 1992; Grabe, Kochanski and Coleman, 2007). Therefore, it is not surprising that several language-specific ToBI systems have been developed on the basis of the same concepts as in the original ToBI, by using the same inventory of tone labels (L and H) with some modifications in the use of diacritics such as + and * to suit a particular language. For instance, there are several ToBI systems for English and its varieties: MAE_ToBI, (Beckman, Hirschberg, and Shattuck-Hufnagel, 2005), AusE_ToBI (Fletcher *et al.*, 2002), and Glasgow English_ToBI (Mayo, 1996). There are also instances of ToBI for other languages, such as K-ToBI for Seoul Korean (Jun, 2005), ToDI for standard Dutch (Gussenhoven, 2005), GRToBI for Athens Greek (Arvaniti and Baltazani, 2005), IToBI for Neapolitan, Bari, Palermo and Florentine Italian (Grice *et al.*, 2005), Pan-Mandarin ToBI for Mandarin (Peng *et al.,* 2005), and C-ToBI for

Cantonese (Wong, Chan and Beckman, 2005), as well as a developing IPrA (an International Prosodic Alphabet) system (Hualde and Prieto, 2016).

The current ToBI system of intonation transcription consists of four tiers: 1) tone tier, 2) orthographic tier, 3) break index tier, and 4) miscellaneous tier (Beckman and Ayers-Elam, 1997; Cruttenden, 1997; Ladefoged, and Johnson, 2011). According to Beckman and Ayers-Elam (1997), the tone tier displays an analysis of the phonological categories of tones, as pitch accents, phrase accents and boundary tones. On this tier, the diacritics *, +, - , % are applied to the tonal H and L targets. The diacritic '*' is used to indicate the simple H*, L* pitch accents, the diacritic '+' for the bitonal pitch accents L+H*, L*+H, the diacritic '–' for the phrase accents H-, L-, and the diacritic '%' for the phrase-initial or phrase-final boundary tones %H, %L, and H%, L% respectively. Identifying these tonal targets can be accomplished in a visual inspection of fundamental frequency (f0) traces. In essence, the pitch accent is preferably labelled within the vowel interval of the stressed syllable, and the labels of H* or L* tones are aligned with the realisations of the f0 maximum or minimum respectively. This is the stage where phonetic implementations of the actual f0 turning points or curves are mapped onto phonological representations of underlying tones.

The orthographic tier displays word-by-word or segmental transcriptions of all the words in the utterance using ordinary spelling. The orthographic labels are time-aligned with the end of each word. The break index tier displays the perceived degree of disjuncture in terms of the presence or absence of boundaries among strings of words using 5 digits: '0' for no word boundary or clitic groups (e.g., gotta), '1' for a break between words, '2' for uncertainty as to the disjuncture between words, '3' for breaks at the end of an intermediate phrase, and '4' for breaks at the end of an intonational phrase. The miscellaneous tier displays additional comments pertaining to other prosodic information, such as disfluencies, audible breaths, false starts, and voice quality (e.g., a creaky or breathy voice). These phenomena can perturb the smooth movements of pitch contours. Example (17) illustrates an alternative way of analysing the intonation patterns using the ToBI:

(17)

Two clause.	When you are winning,	I will run away.
BREAK INDEX	[1 1 1 4	1 1 1 4]
TONE TIER	[H* L-L%	H* H* L-L%]
SEGMENTAL TIER	[wən ju: ɑ: wɪnɪŋ	aɪ wɪl rʌn əweɪ]

(Ladefoged and Johnson, 2011, p. 130)

The ToBI analyses the pitch movement over an utterance as strings of the two level tones (H and L), and not the pitch contours of rises and falls as used in the British approach. Either a single H or L or some combinations of them create several pitch patterns such as H* L-L%, where H* marks phrasal prominence and L-L% marks phrasal boundaries, which are considered as phonological targets appearing on the tone tier. Based on the AM approach, sequences of H and L are associated with three types of tonal events; that is, pitch accents, phrase accents, and boundary tones. At this point, it can be said that the pitch movements over an utterance are represented as a string of pitch accents, phrase accents, and boundary tones.

The tone inventory in the ToBI (Ladefoged and Johnson, 2011, p. 128) is provided below. As in the table, pitch accents are associated with the primary stressed syllable of the most prominent word in an utterance. They are represented as a single high (H) or low (L) tone, or a combination of the H and L tones, with a * diacritic. These single pitch accents H*, and L* can be taken as examples, where the H* and L* tones are relative to the pitch range for each intermediate phrase or intonational phrase. When there are two tones associated with the same primary stress of an accented word, they will form complex or bi-tonal pitch accents. The L+H* bitonal accent or a rising peak accent can be taken as an example, where the H* tone has a leading L tone starting and rising from a low part of the speaker's pitch range. If there are two or more pitch accents in an utterance, the last pitch accent constitutes the nuclear pitch accent by default. In addition, the inventory of pitch accents is the same for pre-nuclear accent and nuclear accent patterns in an intonational phrase (IP).

Table 2. 3 English tone inventory available in the ToBI

Optional Pre-nuclear Pitch Accents on Stressed Syllables	Nuclear Pitch accent	Phrase Accent	Boundary Tone
H*	H*		
L*	L*		
L+H*	L+H*		
L*+H	L*+H	H-	H%
H+!H*	H+!H*		
(!H*)	(!H*)	L-	L%

Phrase accents are associated with pitch movements from the last pitch accent to the right edge of an intermediate phrase (ip). Sometimes, the term 'phrase tone' can be used to refer to the 'phrase accent'. They are represented as a single high (H) or low (L) tone with a phrase accent (-) diacritic. If we take the phrase accents H-, and L- as examples, the H- phrase accent accounts for a high pitch level sustained after the nuclear pitch accent to the edge of the intermediate phrase. The L- phrase accent accounts for a long low stretch of pitch after the nuclear pitch accent to the edge of intermediate phrase. A downstepped high phrase accent (!H-) is an additional category of phrase accents in the ToBI standard. It is generally used for the pitch movement that remains high or does not reach as low a level of the speaker's pitch range as the pitch movement of the L- phrase accent. H- and !H- (downstepped) phrase accents are known as triggers for the raising of the level of the following boundary tone.

Boundary tones are associated with the end of an intonational phrase. They are represented as a single high (H) or low (L) tone with a boundary tone (%) diacritic. The boundary tones H%, and L% can serve as examples. Typically, the H% and L% boundary tone accounts for pitch movement ending in the high and low level of the speaker's pitch range respectively. When phrase accents combine with boundary tones, they will define the edge of an intonation phrase. In this case, there are several combinations of phrase accents and boundary tones. For instance, a combination of L-L% is perceived as a fall in pitch at the boundary, which is normally found in declaratives. A combination of H-H% is perceived as a rise in pitch at the boundary, which is normally found in yes/no questions. Two tone events in combination like this are sometimes referred to as 'edge tones'.

2) The IViE

Modelled on the ToBI system for American English (e.g., Silverman *et al.,* 1992; Beckman and Ayers-Elam, 1997), the IViE was developed by Grabe, Post and Nolan (2001). The term 'IViE' stands for Intonational Variation in English. It is a labelling system for prosodic analysis developed as part of the corpus project with the same name which investigated the patterns of intonation in urban dialects of English spoken in the British Isles. The complete IViE corpus project consists of spoken material from nine dialects representing Belfast, Dublin, Newcastle, Bradford, Leeds, Cambridge, London, Liverpool, and Cardiff. The IViE speech data comprises five speaking styles: read sentences, read text, retold text, map task, and free conversation. The speech data in the corpus were annotated and transcribed by means of the Autosegmental-Metrical analysis of intonation. The ultimate aims of the IViE are to provide comparable speech data across varieties of English from different speaking styles, as well as to provide the public with systematic, linguistic annotations and analyses of prosodic data.

According to Grabe, Post and Nolan (2001), the IViE transcription method consists of 5 tiers: 2 orthographic tiers and 3 prosodic tiers. One of the two orthographic tiers is used for the transcription of spoken material in ordinary written English and for locating word boundaries. The other is the comments tier for transcribers to add further notes and alternative transcriptions of the tonal patterns. The three prosodic tiers consist of the phonological tier, the target tier, and the prominence tier. The phonological tier is used by transcribers to make a generalisation of the phonological classification of tones. The target tier is for transcribers to examine the phonetic realisation of intonational events. On this tier, the phonetic transcriptions are syllable-based with the alignment of finer pitch patterns surrounding prominent syllables. The prominence tier is for the transcribers to locate prominently stressed and accented syllables as well as rhythmic boundaries.

As can be noticed, the IViE adds two tiers to the original ToBI; that is, the phonetic tier and the prominence tier**.** Grabe (2004) claimed that these tiers help increase the transparency and replicability of the tonal labels on the phonological tier. Thus, before embarking on phonological transcriptions, two steps need to be dealt with: first identifying prominently stressed syllables, and secondly identifying the shape of the pitch movements surrounding those prominent syllables. Grabe further maintains that IViE transcriptions can be used to make explicit comparisons across the varieties of English because they are based on the same set of tonal labels compiled from several dialects, not just one dialect. In addition, the IViE system

works well for the capturing of rhythmic differences across dialects as well as differences in phonetic realisation.

One more advantage of the IViE system is that it generates a pool of tonal label options for transcribers to choose from. The options for tonal labels in the IViE include a wide range from intonation contours from several dialects; therefore, transcribers can draw up intonation patterns that are specific for a language/dialect in question. As mentioned earlier, even though the concept of the IViE was developed from the ToBI, the notational conventions of the IViE have been developed from work on the intonational phonology of English by Gussenhoven (1984) and Grabe (1998).

In terms of the tonal inventory, the IViE has a set of tonal labels for pitch accent types and boundary tones only. Since it excludes an intermediate phrase from the prosodic structure in English, there are no phrase accents. The tonal labels for pitch accents commonly include L*, H*, H*L, L*H, L*HL, H*LH, and so on. There are three boundary tones: H%, L%, and 0%. Phrase-initial and phrase-final boundary tones use the same pool of these boundary tone labels. A zero boundary tone refers to the pitch on the last syllable in the intonational phrase, when it is not different from that of the immediately preceding tone. In other words, the pitch level of the last tone preceding the boundary is sustained to the end of the intonational phrase. According to Grabe (1998), Grabe, *et al.* (2000), Grabe (2004) and Grabe, Kochanski and Coleman (2005), general Southern British English has two options for boundary tones of rising (H%) or level (0%), while Belfast English has three options of rising (H%), level (0%) and falling (L%). Grabe and her colleagues claim that the IViE can capture this dialect-specific difference transparently. This is so because the IViE provides transcribers with such advantages as tonal label options and one level of intonational phrasing; that is, the intonational phrase, as mentioned earlier.

To sum up, it can be said that the IViE constitutes another language-specific transcription system like KToBI, MAE_ToBI, and ToDI. Although the IViE system is not in widespread use internationally compared to the ToBI, previous studies (Grabe, *et al.*, 2000; Grabe and Post, 2002; Grabe, 2004; Grabe, Kochanski and Coleman, 2005) have demonstrated that this transcription system is effective and straightforward. It allows for cross-dialect transcriptions and readily gives accounts of differences concerning phonological structure, the phonetic implementation of phonological categories and the placement of stressed and accented syllables. Furthermore, the IViE is an example that shows the ultimate exploitation of the

British tradition and the AM approach in combination for investigations of the intonational variation in English. In other words, it makes use of British-style descriptions of intonation modelled on the ToBI within the AM approach. Accordingly, it goes without saying that the IViE is one of the most useful hybrid annotation systems for the transcription of intonation.

The above paragraphs have pointed out that either ToBI or IViE have their own outstanding points in serving as methods for prosodic annotation. The present study employs the ToBI system within the AM approach to intonational phonology for several reasons as discussed below.

2.6.3 The approach and model chosen as appropriate for this study

There have been different theoretical approaches to intonational phonology as well as different models of intonation transcription, as described in the previous sections. It is thus reasonable to address the potential of the existing approaches and models for intonation analysis as well as their possible pitfalls before implementing one of them in the current study. Based on the research purposes: 1) investigating the use of intonation (choices of accent types) for marking narrow focus produced by Thai learners; and 2) investigating the intonational means, in particular the use of rising accents, for marking narrow focus in terms of its acoustic-phonetic realisations (tonal alignment and scaling) as well as the review of relevant literature, the present study applies the Autosegmental-Metrical (AM) theory and the ToBI transcription standard in investigating the characteristics of L2 English intonation and focus produced by Thai learners, compared to L1 English intonation and focus produced by native speakers. The implementation of the AM approach and ToBI are justified in detail below.

1) Justifications for implementing the AM approach

There seems to be no consensus regarding which approach should be used to describe intonation, and which model generally prevails. Part of the reason for this might be the fact that different languages vary in terms of having their own intonational systems. Consequently, language- or dialect-specific systems for intonation analysis have been developed, as discussed earlier.

The AM and the British approaches still have similar points. Ladd (2008) provides evidence of twenty-two instances of the perfect phonological mapping of nuclear tones in the British approach with the pitch accents and edge tones in the AM approach; for example, with high-falls as H* L-L%; low-falls as L* L-L%; high-rises as H* H-H%; low rises as L* H-H%; and fall-rises as H* L-H%. Additionally, rise-falls can be either L+H* L-L% or L*+H L-L% in cases of emphasis. As can be noted, intonation patterns in terms of pitch contours in the British model are decomposed into underlying pitch targets as sequences of H and L, and various combinations of them.

However, the British nuclear tone approach and the Autosegmental-Metrical (AM) approach differ. Firstly, the British and the AM approaches illustrate how English intonation works based on differing theoretical and experimental perspectives concerning intonational phonology. Secondly, the British model analyses the patterns of pitch variation acoustically as configurations of falling and rising contours for downward and upward pitch movements. By contrast, the AM model analyses pitch patterns as a sequence of the two level tones: high (H) and low (L) tones. Hence, a falling contour can be described as having a H* pitch accent followed by a L-L% edge tone whereas a rising contour can be described as having a H* (or L*) pitch accent followed by a L-H% (or H-H%) edge tone, for example. Due to the H and L labels, the AM model can offer finer distinctions with reference to pitch range or to onset accent type, whereas the British model requires various other information in order to describe these distinctions (Ladd, 2008).

Taking this into account, it has been decided in the current study to follow the theoretical framework of the Autosegmental-Metrical (AM) approach in investigating L2 English intonation and focus as used by Thai learners and English native speakers. Justifications for this decision include the reasons given above and the additional reasons below.

First of all, it has already been acknowledged that the AM approach provides researchers in the field with both a theoretical framework and practical implications for analysing intonation. In theory, this approach offers descriptions of intonation in terms of two level tonal targets, with a tune resulting from interpolations between these tonal targets. In practice, this approach allows for the incorporations of both acoustic and auditory analysis, while the British tradition (Crystal, 1969; Halliday, 1970; O'Connor and Arnold, 1973; Cruttenden, 1997) primarily investigates intonational features in an impressionistic way, and requires "here and there to describe certain distinctions" such as when referring to pitch range and onset accents (Ladd,

2008, p. 92). As a result, the influence of the AM approach is reflected in many studies in terms of the descriptions and analysis of intonation and focus; for example, in work on the acquisition of intonation patterns in English (McGory, 1997), the phonetic realisation of pitch accents in British English (Grabe *et al.,* 2000), and the phonetic realisation of focus in English declaratives (Xu and Xu, 2005).

Besides this, the AM approach underlies many cross-language studies on intonation and focus. The approach can also be used to investigate intonation and focus in under-described languages. It is flexible in the sense that it allows for the comparison of intonation patterns across languages at differing phonetic and phonological levels. This is because it is widely accepted that two languages may be different from each other at one level of representation while being similar at another level. Ladd (2008, p. 115) proposed that cross-linguistic differences in intonation can be examined in terms of "semantic, systematic, realisational, and phonotactic" differences. Semantic differences involve differences in the meanings or uses of the same tune. Systematic differences involve differences in the phonological inventory of distinct tune types irrespective of semantic differences. Realisational differences involve differences in the phonetic realisation of the same tune, and phonotactic differences involve differences in terms of tune-text association and the permitted structure of tunes.

2) *Justifications for implementing the ToBI*

There are several reasons why the ToBI is used for the transcription of intonation and focus for Thai learners and native speakers of English in this study. Evidence from previous studies and the relevant literature was drawn on in making this decision as follows.

Firstly, unlike the IViE which is intended for inter-dialectal comparison, "the ToBI labelling system was originally developed to cover the three most widely used varieties of spoken English - namely, General American, standard Australian, and southern British English" (Beckman and Ayers-Elam, 1997, p. 8). This claim is not at all overexaggerated when considering more recent research into intonation with the incorporation of the ToBI. Instances of more recent studies include those of English and its varieties, such as the MAE_ToBI (Beckman, Hirschberg, and Shattuck-Hufnagel, 2005), AusE_ToBI (Fletcher *et al.*, 2002), and Glasgow English_ToBI (Mayo, 1996). There are also a number of instances of language research that has already applied the ToBI to investigate specific intonational systems; to name but a few, K-ToBI for Seoul Korean (Jun, 2005), ToDI for standard Dutch (Gussenhoven, 2005), GRToBI for Athens

Greek (Arvaniti and Baltazani, 2005), IToBI for Neapolitan, Bari, Palermo and Florentine Italian (Grice *et al.*, 2005).

Secondly, the ToBI labelling system is transparent and learnable. It has thus been widely used by and accepted among researchers who work on a language, or across languages, or even within dialects of a given language. In addition, researchers who are working on differing types of data can make the most use of the ToBI since it allows for the use of a wide range of spoken and read materials. There have recently been a growing number of studies on intonation and more work on the development of language-specific ToBI systems with the use of the same tone labels from a single inventory of tonal categories in the original ToBI. Various sources of evidence support this case, as in the paragraph above, with some additional examples such as Pan-Mandarin ToBI for Mandarin (Peng *et al.,* 2005), C-ToBI for Cantonese (Wong, Chan and Beckman, 2005), and GToBI for German (Grice, Baumann and Benzmüüller, 2005).

Thirdly, the ToBI is flexible in terms of its tonal categories that can be modified to suit specific languages. Each language and its dialectal varieties require different ToBI transcription systems. In this sense, it is suggested that researchers modify it and add categories specific to the language they are working on. Decomposing pitch contours into high and low tones, and into various combinations of these allows the ToBI to capture language-specific intonational differences. Evidence from the studies cited above also supports this. Even though there are controversial issues among researchers investigating intonation in the same language, such controversies or disagreements are normal. One solution for a researcher as a ToBI labeller is to be consistent in using tonal labels for the annotation of the intonation under analysis. In addition to this, researchers should bear in mind that intonation in some dialectal varieties of a language can be analysed and transcribed using the same annotation criteria, such as in the transcription of intonation in General American English and Australian English.

Fourthly, the ToBI is not only useful for first language, but also has practical implications for studies of the second language acquisition (SLA) of intonation. By using the ToBI, it can be beneficial for language teachers to analyse and examine the intonation patterns produced and perceived by L2 learners, and to follow up their development of intonation acquisition. In this case, comparative and contrastive analyses can be conducted by comparing the L2 learner's intonation patterns with those of other learners with the same or different first languages. This claim does not overestimate the potential of the ToBI transcription method. At least, it is

affirmed by Hirschberg's (2002) statement that the ToBI analysis of intonation can account for cross-linguistic differences as well as comparisons from large-scale analyses of speech corpora.

Finally, good values of inter-labeller reliability for the ToBI have been reported. The percentage agreement between labellers reaches 88% for the presence or absence of tonal categories, 81% for tonal labels, and 91% for break indices (Silverman *et al.,* 1992; Pitrelli, Beckman and Hirschberg, 1994). Thus, the ToBI system can be considered to be reliable, coherent, and learnable. Furthermore, the ToBI has well-defined notational conventions. Combining the phonological representations of the tonal targets H and L with symbols or diacritics like *, -, and % can reflect what functions a tonal target serves (Hirst, 2005).

3) *Summarising the application of the AM and the ToBI*

The present study definitely requires a specific notation system that works best in fulfilling the research objectives. It makes sense if a system of interest allows for further possibilities of modification and adaptation such as in tone and tune categories and the structure of tiers in order to capture the intonational characteristics of a language as well as to allow for cross-language comparison. As per the justifications discussed above, the Autosegmental-Metrical (AM) theory and the ToBI transcription system are assumed to be a good starting-point for the investigation of under-described intonation systems in L2 English as used by Thai learners in comparison with the intonation of native speakers.

Taking the AM approach as the main theoretical framework for the research, the present study describes and analyses pitch contours as two level pitch or tone targets which are high (H) and low (L), depending on the relative height of pitch along each contour. Intonation patterns or tunes are derived from the interpolation of these H and L tones. Furthermore, the current study uses a standard and practical implication of the Autosegmental-Metrical (AM) approach employing both acoustic and auditory analyses (Beckman and Ayers-Elam, 1997; Ladd, 2008), together with the notational convention of the ToBI (Beckman and Ayers-Elam, 1997), instead of having pitch contours analysed only by means of an impressionistic auditory method.

As an analytic tool, the ToBI system is used in this study to transcribe and annotate English intonation as used by native speakers and Thai learners. The phonological analysis of intonation has been conducted with the help of auditory analysis and the visual inspection of time-aligned fundamental frequency (f0) contours and time-aligned words using speech analysis software

such as Praat (Boersma and Weenink, 2019). This method has long been employed in the field of research into intonational phonology. Furthermore, according to Beckman and Ayers-Elam (1997, p. 8), "The ToBI labelling system was originally developed to cover the three most widely used varieties of spoken English, namely, general American, standard Australian, and southern British English". A more detailed account of the analysis and annotation of the types of intonation under investigation is provided in the next chapter on research methodology.

Finally, to date, there have been a number of language-specific ToBI developments, for example, K-ToBI for Seoul Korean (Jun, 2005), Pan-Mandarin ToBI for Mandarin (Peng *et al.,* 2005), C-ToBI for Cantonese (Wong, Chan and Beckman, 2005), and J-ToBI for Tokyo Japanese (Venditti, 2005). This demonstrates that the ToBI is widely accepted among scholars who are working on prosodic annotation and transcription. It allows for further modifications and adaptations such as for tonal labels and inventories in investigating the prosodic features of different languages and their varieties. Having said that, the ToBI can be considered as a transcription standard for making comparisons within and across languages as well. This claim accords with Cruttenden's (1997, p. 64) statement that the ToBI can be assumed to be one of the best annotation systems available because of the complete correspondence of graphic "representation to phonetic reality and to semantics and pragmatics."

2.7 Summary of the Chapter

As an attempt to develop this literature review, several areas related to the literature and research on prosody, intonation and focus in L1 and L2 English were covered. This helps in setting a good understanding of the groundwork of the research topic. Therefore, this literature review is divided into several sections and subsections. Brief summaries of each are given as follows:

Section 2.2 elaborated on the definitions of prosody, intonation, focus, acoustic cues to prominence, and tonal alignment and scaling used in the current study. Prosody is defined in the same way as suprasegmentals; that is, the study of those elements of speech beyond segmental elements. Prosody includes several domains ranging from lexical tones, intonation, stress, and rhythm. Research into each of these prosodic domains involves investigating their relevant prosodic features which can be either perceptual correlates (pitch, loudness, length), acoustic correlates (fundamental frequency, intensity, duration), or both in combination. Intonation is one part of the prosodic domains, and it is known to involve variations in pitch

over units larger than a syllable. Intonation in this study is defined as "the use of s*uprasegmental* *phonetic* features to convey 'postlexical' or *sentence-level* pragmatic meanings in a *linguistically structured* way" (Ladd, 2008, p. 4, italics in original). As a communicative function of intonation, focus is defined as narrow informative focus. Finally, tonal alignment is referred to as temporal relationships between tonal targets and segmental strings and scaling is referred to as fundamental frequency values of H and L tonal targets. Taking the above-mentioned concepts into consideration, it can be stated that the present study is concerned with an investigation of a cross-language comparison as to the phonological and phonetic realisations of focus using intonation or nuclear pitch accents in its expression.

Section 2.3 provided detailed descriptions of prosody, intonation and focus in English. It presented details of relative prominence in English (primary stress, secondary stress, unstressed). Primary stress is considered to be the potential location of an intonational prominence or pitch accent. There can be several pitch accents in an intonational phrase (IP), but there is only one nuclear pitch accent which can be normally located as the last accented content word in that IP. Furthermore, rising pitch accents are used to mark focus intonationally, together with other focus-marking strategies such as deaccenting the materials coming after the focused components. Two levels of prosodic structure, intermediate and intonational phrases, were discussed and both concepts were adopted in this present study, as well as intonational descriptions within the AM approach.

Section 2.4 introduced prosody, intonation and focus in Thai and gave a brief overview of Thai phonology. The language itself has lexical tones, but it also has stress and intonation as well. Stress in Thai is fixed, always falling on the same last syllable of content words. According to the literature concerning Thai discussed earlier, syllable duration is considered to be one of the most salient features in separating stressed from unstressed syllables. Apart from pitch variation affecting the lexical meaning of a word, it has a potential influence over units larger than a syllable in distinguishing between utterance types and marking prominence. As for the marking of focus, native Thai speakers can make the most of their first language in several ways. These include changing word order, changing word choices, using an emphatic high tone, lengthening the last syllable of a word, and even using pauses. The section ended with a contrastive analysis of L1 English and Thai, and current models of L2 speech acquisition, particularly the L2 Intonation Learning theory (LILt).

Section 2.5 discussed previous research and studies into L2 prosody, intonation, and focus as well as tonal alignment and scaling. It discussed issues of intonational aspects from the point of view of second language acquisition research. This includes studies regarding L1 transfer, word stress, accent placement and focus marking, as well as tonal alignment and the scaling of tonal targets. A link was also made with studies of L1 English, where applicable. An attempt was made to point out that the research topic in question are subject to a variety of factors, ranging from first language interference, learner variability, and exposure to the target language as well as language teaching in classrooms. Evidence of the effects of these factors, for example, on nuclear pitch accent types and placement, as well as tonal alignment and scaling, was drawn from previous studies and the literature related to prosody, intonation and focus across languages. The section ended with a discussion of the status of English as a foreign language in Thailand.

Section 2.6 presents the main theoretical approaches to prosody and the models used for prosodic transcription. The two main approaches are the British nuclear tone approach and the American Autosegmental-Metrical approach and both were elaborated in detail. The British approach concentrates on nuclear tones described as pitch contours or tunes (e.g., falling, rising), and it has the following elements in common: tonality, tonicity and tone, otherwise known as the 3Ts. The American approach concentrates on pitch levels. Pitch contours or intonation contours in the British approach are decomposed into two level H and L tones. H and L tones are associated with three pitch events identified by adding diacritics, for example, H* for the pitch accent, L- for the phrase tone, and L% for the boundary tone. Regarding the model chosen for transcribing prosody, the ToBI system is used in this study. This is not just because the ToBI is an intonation transcription system developed within the Autosegmental-Metrical approach to intonational phonology, but also since it is in widespread use for phonetic analysis in terms of tonal alignment and scaling. The section ended with a discussion of the AM and the ToBI selected for the investigation of the intonational features and marking of focus in the present study.

In the light of the literature review, the present study has been established to examine how native speakers of Thai, a tonal language, use prosodic features to mark focus in English. The hypotheses below have been formulated depending on the findings of previous studies and the literature concerning nuclear pitch accents, the location of nucleus placement, focus-marking strategies, and tonal alignment and scaling. This study intends to shed light on and increase the understanding of cross-linguistic differences in L2 intonation and to make contributions to the

existing relevant literature. The specific questions addressed by this study are restated here with the following hypotheses:

Table 2. 4 Research questions (RQ) for qualitative analysis

Research questions		Chapter
RQ1	What are the accent types that native speakers of English use to mark narrow informative focus?	Chapter 4
RQ2	What are the accent types that Thai learners of English use to mark narrow informative focus?	Chapter 4
RQ3	To what extent do Thai learners differ from native speakers in terms of the use of accent types to mark narrow informative focus?	Chapter 4
RQ4	To what extent do the L2 Intonation Learning theory (LILt) and the Contrastive Analysis Hypothesis (CAH) account for the L2 English intonation for focus in this study?	Chapter 7

Table 2. 5 Research questions (RQ) and hypotheses (H) for quantitative analysis

Research questions		Chapter
RQ5	To what extent do the factors such as levels of English proficiency, gender and focus positions affect the acoustic characteristics of focus used by Thai learners for the marking of narrow informative focus in terms of the use of f0, intensity and duration?	Chapter 5
H5.1 - 5.3	Thai learners use f0, duration and intensity as the acoustic parameters to mark focus. However, the use of f0, duration and intensity may differ from native speakers due to language groups, levels of English proficiency, gender and focus positions. Duration could be greater since in Thai duration is one of the most prominent features used to distinguish stressed and unstressed syllables (Luksaneeyanawin, 1998).	
RQ6	To what extent do the factors such as levels of English proficiency, gender and focus position affect the temporal relations between the tonal movement in rising pitch accents and segmental strings in narrow-focused words produced by Thai learners of English?	Chapter 6
H6.1 - 6.5	Thai learners may choose the right nuclear accent for a particular accented syllable or word. However, they may place that right nuclear tone differently from native English speakers somewhere early or late due to language groups, levels of English proficiency, gender and focus positions. This is known as a misalignment in intonation (Mennen, 2006, 2007).	

Chapter 3. Methodology

3.1 Introduction

This chapter presents information about the research methodology used in the current study. It begins with section 3.2 which gives a detailed account of the design of speech material used in investigating focus marking. Section 3.3 includes information about the recruitment of participants, section 3.4 describes the procedures used for data collection and recording, and section 3.5 explains the method applied for segmentation and annotation. Sections 3.6 and 3.7 give more detailed accounts of the qualitative analysis of the prosodic marking of focus and the acoustic-phonetic measurements of focal prominence, while section 3.8 covers tonal alignment and scaling. Section 3.9 then presents the statistical analyses, followed by a summary of the chapter in section 3.10.

3.2 Design of Speech Material

The corpus used in this study consists of L1 and L2 English. A series of questions and answers in English was designed for use in investigating how Thai learners of English and native speakers of English mark focus in information prosodically. The speech materials aimed to elicit mainly narrow focus in the question-and-answer task. The marking of broad focus was elicited to use for cross-reference where relevant, and the questions and answers for broad focus work as the filler sentences as well. This method has been used in previous research (Cooper, Eady and Mueller, 1985; Breen *et al.*, 2010; Jun and Fletcher, 2014).

In the construction of target words, sonorants; for example, /m, n, l, r/, were preferred and used as much as possible to create smooth fundamental frequency (f0) contours. Voiced segments were used more often than voiceless segments. This is because, according to Lehiste (1970), van Santen and Hirschberg (1994) and Frota (2002), voiceless segments and non-sonorant consonants have a tendency to break down the continuity of fundamental frequency contours.

In this study, the target words included both monosyllabic and disyllabic words. The disyllabic words were at the sentence-initial position (Subject) or sentence-final position (Object). They had a stressed syllable structure with a simple onset with a short vowel and no coda consonant: CV. Clearly, although the number of syllables differed, the location of lexical stress in the target disyllabic words was the same. The target monosyllabic words were at the sentence-

medial position (Verb). They had a stressed syllable structure with a simple onset with a short vowel and with a coda consonant: CVC. The one exception was the word LENT which has the CVCC structure. Thus, the focus position varied, as in the studies of Sityaev and House (2003), and Breen *et al.* (2010). Table 3.1 below gives more details of all target words, where each target's stressed syllables is shown as CV in upper case. In addition to this, each of the target words and its focus position were made clear to the participants by means of capitalised and bold-typed letters (e.g., Sityaev and House, 2003; Oliver and Andreeva, 2004). All of the target words were embedded in carrier sentences which all had the simple sentence structure of SVO.

Table 3. 1 Target words, their stressed syllable structure and focus positions in the sentences

Target words	Stressed syllable structure in target words	Focus position in target sentences
NANNY	CVcv	Subject
RONNEY	CVcv	Subject
JIMMY	CVcv	Subject
MANNY	CVcv	Subject
READ	CVC	Verb
WON	CVC	Verb
HAD	CVC	Verb
LENT	CVCC	Verb
NOVEL	CVcvc	Object
MEDAL	CVcvc	Object
MELON	CVcvc	Object
MONEY	CVcv	Object

Overall, the total number of test sentences was 1920 (4 test sentences x 4 positions x 3 repetitions x 20 subjects x 2 language groups). Among these, there were 1440 test sentences for narrow focus (4 test sentences x 3 positions x 3 repetitions x 20 subjects x 2 language groups), and 480 for broad focus (4 test sentences x 3 repetitions x 20 subjects x 2 language groups). Sentences with broad focus was included here as an aim to work as the filler sentences. Some examples of test sentences with the prompt questions used in the present study are shown in Table 3.2 below. A full set of prompt questions and answers, as well as a list of the filler sentences, is given in Appendix D.

Table 3. 2 Examples of prompt questions and test sentences

Prompt questions	Test sentences with target words	Focus size/type
Who won a medal?	**RONNEY** won a medal.	Narrow focused subject
What did Ronney win?	Ronney won a **MEDAL**.	Narrow focused object
What did Ronney do with a medal?	Ronney **WON** a medal.	Narrow focused verb
What did you say?	Ronney won a medal.	Broad focus/whole sentence

3.3 Participants and Language Groups

The participants in this study included two groups: 1) native Thai speakers; and 2) native English speakers. There were 10 female participants and 10 male participants in each group who were university students in Newcastle. The number of participants was 40 in total. The names of all participants in this study were coded in order to maintain anonymity. Examples of the consent form (Appendix A) and the questionnaires (Appendix B for native English speakers and Appendix C for native Thai speakers) are provided. Details of the sampling are described below. A summary of all information about the participants are also provided in Table 3.5.

1) Native Thai speakers

The researcher recruited the native Thai speakers through personal contacts. The Thai learners varied in age with a group mean of 33.10 years, language proficiency levels or IELTS (International English Language Testing System) scores of speaking skills with a group mean of 6.4, and length of residence in the UK with a group mean of 31 months. The Thai female learners' ages ranged from 21 to 42, with a mean of 34 years. Their IELTS scores for speaking skills varied between 5.5 to 7.5, with a mean of 6.3. Their length of residence ranged from 10 months to 60 months with a mean of 33.9 months. The Thai male learners' ages were from 21 to 38, with a mean of 32.2 years. Their IELTS scores for speaking skills ranged from 5.5 to 7.5, with a mean of 6.5. Their length of residence ranged from 12 months to 51 months with a mean of 28.1 months. In the researcher's judgement, they were all classified as educated Thai participants. These subjects came from various different provinces or cities in Thailand. Table 3.3 below shows more information about the native Thai speakers.

Table 3. 3 Demographic data for Thai learners of English

Participant	Gender	Age	Length of Residence (month)	IELTS (speaking)	Home town
en02_f01	female	40	48	6.5	Yala
en02_f02	female	37	44	7.5	Nakhon Si Thammarat
en02_f03	female	38	60	7.0	Songkhla
en02_f04	female	34	48	5.5	Nakhon Si Thammarat
en02_f07	female	38	14	6.0	Chaiyaphum
en02_f08	female	39	48	6.0	Bangkok
en02_f09	female	26	17	6.0	Nonthaburi
en02_f10	female	42	10	6.0	Phuket
en02_f11	female	21	36	6.5	Phuket
en02_f13	female	25	14	6.0	Ratchaburi
en02_m01	male	38	42	7.5	Bangkok
en02_m02	male	35	12	5.5	Nonthaburi
en02_m03	male	36	51	5.5	Nakhon Si Thammarat
en02_m04	male	33	25	6.5	Krabi
en02_m05	male	31	37	6.5	Chumphon
en02_m06	male	31	15	6.5	Bangkok
en02_m07	male	38	12	6.5	Surat Thani
en02_m08	male	38	36	6.0	Buriram
en02_m09	male	21	24	7.5	Bangkok
en02_m11	male	21	27	6.5	Nakhon Ratchasima

2) Native English speakers

The native speakers of English were recruited initially via a university email call-out, and then according to the results of self-evaluation questionnaires (see, e.g., Breen *et al.*, 2010; Dilley, 2010). The questionnaires were used to screen for English native participants who spoke Southern Standard British English (SSBE). They varied in terms of age, with a group mean of 23.05 years and in their home towns. The female native English speakers' ages ranged from 20 to 26, with a mean of 22.6 years. The male native English speakers' ages were between 18 and 32, with a mean of 23.5 years. The native English speakers came from different areas of the UK, but they self-evaluated as native speakers of Standard British English (SSBE). Table 3.4 below shows more information about the native English speakers.

Table 3. 4 Demographic data for native speakers of English

Participant	Gender	Age	Home town
en01_f01	female	21	Wiltshire
en01_f03	female	22	Tamworth
en01_f06	female	20	Stamford
en01_f07	female	24	Southampton
en01_f11	female	23	Surrey
en01_f12	female	21	Essex
en01_f13	female	25	Milton Keynes
en01_f15	female	22	Gloucestershire
en01_f17	female	22	Gloucestershire
en01_f21	female	26	Hampshire
en01_m02	male	21	London
en01_m04	male	23	Harrogate
en01_m06	male	31	Cambridge
en01_m07	male	20	Oldham
en01_m08	male	18	Surrey
en01_m09	male	21	London
en01_m10	male	32	London
en01_m11	male	22	Harrogate
en01_m13	male	20	London
en01_m20	male	27	Northampton

Table 3. 5 A summary of all information about the participants, including the mean values of IELTS scores for the speaking skill, age and length of residence in the UK.

Group	Gender (n)	Age (mean, years)	IELTS speaking score (mean)	Length of residence (mean, months)
Native Thai speakers (L2 English)	20 (10f, 10m)	33.10	6.4	31
Native English speakers (L1 English)	20 (10f, 10m)	23.05	N/A	N/A

3.4 Data Collection and Recording Procedure

3.4.1 Questionnaire and preparation session

All participants first completed questionnaires (Appendix B and C) which aimed to obtain some of their demographic information such as their age, their home town, and language learning history. Then, they received information giving an overview of the present research and its goals. Before recording, the researcher gave an explanation about the prominence of the target words and the participants were asked to relate the target word to the prompt questions. In this way, they could express their communicative intentions more effectively in response to the questions. No further details about broad focus were given to the participants. Next, each participant was allowed to practise their reading with the test materials for about 5 minutes in order to familiarise themselves with the sentences under investigation. If they had questions, the researcher did not hesitate to assist them. After that, they were asked to sit inside the sound-recording booth. When they were ready, the researcher asked them to read the answers on the computer screen after they heard the questions from the researcher (e.g., Ladd *et al.*, 2009). Each answer that they read was presented as a series of PowerPoint slides. The participants were also asked to read aloud the answers using their normal natural speech.

3.4.2 Recording session

In this experiment, the information structure was controlled by having the participants utter each carrier sentence as an answer to the researcher's prompt question (Appendix D). The researcher's questions required that the participants should change the focus positions in response to the question from sentence-initial (Subject), to sentence-medial (Verb) and sentence-final (Object) positions to entire sentences. In other words, the questions were exploited to trigger how the participants would highlight the target focused words for the marking of narrow and broad focus. In addition, the participants were asked the same questions three times by the researcher, but in a randomised order, so that this gave the participants the chance to read the test sentences three times in randomised order as well. This process would be like three separate internally-randomised blocks, as the following example in Table 3.6. The filler sentences (Appendix D) were inserted randomly among the test sentences (see, e.g., Ladd *et al.*, 2009). There was a short pause after the reading of each sentence. The Thai learners of English and native speakers of English were engaged in the same tasks.

Table 3. 6 Examples of the so-called three separate internally-randomised blocks

First block	
The researcher asked:	The participant answered:
Q1 Who read a novel?	**NANNY** read a novel.
Q3 What did Nanny do with a novel?	Nanny **READ** a novel.
	Nanny read a **NOVEL**.
Q2 What did Nanny read?	Nanny read a novel.
Q4 What did you say?	**RONNEY** won a medal.
---	Ronney won a medal.
Q5 Who won a medal?	Ronney won a **MEDAL**.
Q8 What did you say?	Ronney **WON** a medal.
Q6 What did Ronney win?	
Q7 What did Ronney do with a medal?	
---and so on.	
Second block	
Q12 What did you say?	Jimmy had a melon.
Q9 Who had a melon?	**JIMMY** had a melon.
Q11 What did Jimmy do with a melon?	Jimmy **HAD** a melon.
Q10 What did Jimmy have?	Jimmy had a **MELON**.

Q1 Who read a novel?	**NANNY** read a novel.
Q4 What did you say?	Nanny read a novel.
Q3 What did Nanny do with a novel?	Nanny **READ** a novel.
Q2 What did Nanny read?	Nanny read a **NOVEL**.
---and so on.	
Third block	
Q7 What did Ronney do with a medal?	Ronney **WON** a medal.
Q5 Who won a medal?	**RONNEY** won a medal.
Q8 What did you say?	Ronney won a medal.
Q6 What did Ronney win?	Ronney won a **MEDAL**.

Q10 What did Jimmy have?	Jimmy had a **MELON**.
Q12 What did you say?	Jimmy had a melon.
Q11 What did Jimmy do with a melon?	Jimmy **HAD** a melon.
Q9 Who had a melon?	**JIMMY** had a melon.
---and so on.	

The recording sessions were conducted in the soundproof recording room of the Phonetics Laboratory at Newcastle University. The speech materials were recorded using an Edirol CS-50 microphone and Edirol R-44 recorder, and the data were stored in separate files for each subject. The data recordings were made with a sampling frequency of 44100 Hz in mono wave

format (16 bits quantisation). The researcher used headphones connected to the computer that controlled another (or the participants') computer and the recorder in the recording booth in order to listen to the participants' speech. In cases of any obvious errors or mispronunciations produced by the participants, they were allowed to repeat the same sentences. This is a requirement for investigating temporal relationships among tonal alignment and segmental elements (e.g., Ladd *et al.*, 2009). After the recording session with each participant, their sound files were coded for each sentence with the use of Praat software (Boersma and Weenink, 2019). Therefore, the data from the two different groups were ready for intonational annotation and further analysis:

(1) L1 English from native speakers of English as reference data,
(2) L2 English from Thai learners of English

3.5 Methods of Segmentation and Annotation

3.5.1 Pitch tracking and smoothing

The procedures of segmenting and annotating test sentences into target syllables or words were conducted with the help of Praat (Boersma and Weenink, 2019). Firstly, pitch tracking was performed using the Praat standard algorithm based on the autocorrelation method. Adjustment were made for male speakers with 75-350 Hz and for female speakers with 100-600 Hz (cf. Fuchs and Maxwell, 2015). Secondly, pitch smoothing was performed using the Praat script 'Manual and Automatic Smoothing' of f0 tracks (or mausmooth) developed by (Cangemi, 2015). Following Cangemi's guideline for using mausmooth, this tool first extracts f0 candidates, and then pauses and gives time for researchers to manually correct pitch octave jumping or errors such as pitch halving and pitch doubling. The researcher finally interpolates and smooths the f0 contours according to the default setting; that is, using SMOOTH1 for smoothing the manually corrected f0 points and using SMOOTH2 for smoothing the contour again after interpolation. Cangemi also suggested that the smoothed contours can be used as reliable input for further analysis in other environments such as R or Matlab. According to this, the researcher used the smoothed f0 contours in Hertz for plotting in Praat with the aims of displaying stylistic intonation contours and further analysis.

3.5.2 Segmentation

Segmentation of the words in each test sentence was accomplished manually, following the methods recommended in previous studies (Peterson and Lehiste, 1960; Ladd *et al.*, 2009; Prieto, 2009). As shown in Figure 3.1 in section 3.5.3 below, the word tier was built by hand on the basis of displays of f0 on the Praat screen (Boersma and Weenink, 2019). In addition to auditory impressions and visual inspections of f0 traces, other cues such as duration and intensity, or all of the above in combination, were used in the procedures of segmentation and annotation (Peterson and Lehiste 1960; Grabe, Kochanski and Coleman, 2005; Ladd, 2008; Ladd *et al.*, 2009; Gut and Pillai, 2014). Relevant guidelines for segmentation are provided below:

> … vowel-consonant boundaries were located at breaks in the formant structure (generally with a corresponding drop in waveform amplitude), while consonant-vowel boundaries were located at regular vocalic formant structure, using the amplitude and shape of successive pitch periods as a subsidiary guide.
> (Ladd *et al.*, 2009, p. 148)

More specifically in the case of sonorants (m, n, l), additional guidelines were employed as in Peterson and Lehiste (1960, cited in Prieto, 2009) as follows:

> The beginning or end of a sonorant consonant was identified at the start of the abrupt change from the steady-state period in the spectrogram to the onglide transition movement to the vowel. When the formant transitions were not abrupt enough, the criterion used was the expected change in amplitude displayed in the waveform. (Prieto, 2009, p. 869)

However, segmenting approximants (j, w, r) was problematic to some degree because the approximant do not involve complete closure of the vocal tract. Turk *et al.* (2006) advised that approximants should be avoided. In this case, the researcher's segmentation would rely on visual inspections of the formant structure breaks, a decrease of intensity in a waveform as well as impressive auditory inspections of the words in question from the soundfiles.

3.5.3 ToBI-style annotation

Prosodic annotation was also completed manually, following the guidelines for ToBI labelling (Silverman *et al.*, 1992; Beckman and Ayers-Elam, 1997). In addition to this, auditory impressions and the visual inspection of f0 traces, and other cues such as duration and intensity,

or all the above in combination were used in the procedures of the annotation of prosodic features. This method is commonly used as in previous studies (see, e.g., Grabe, Kochanski and Coleman, 2005; Turk, Nakai and Sugahara, 2006; Ladd, 2008; Gut and Pillai, 2014). The focused words were also marked with capital letters (see, e.g., Oliver and Andreeva, 2004).

Figure 3.1 below illustrates this method of segmentation and prosodic annotation, together with examples of a waveform, a wideband spectrogram and an f0 contour. With the help of Praat (Boersma and Weenink, 2019), textgrids were made for every single sentence of each participant's sound files. As can be seen in Figure 3.1, each textgrid consists of five tiers. The word, phonological and break index tiers are adopted from the ToBI system. The two additional tiers are the CV tier and LH tonal target tier, for which more details are given in section 3.8. As shown in Figure 3.1, the top tier is the time-aligned word tier. The second tier is the CV tier which presents the landmarks segmenting the syllables into consonants (Cs) and vowels (Vs). The third tier is the tonal target landmark tier. For this tier, the f0 contours are decomposed into sequences of low (L) and high (H) tonal events. The CV and the tonal target landmark tiers were used in investigating tonal alignment and scaling and are further discussed in the section 3.8. The fourth tier is the phonological tone tier which presents intonational features; for example, in terms of pitch accents such as bitonal rising accents (L+H*). In this study, the pitch accents on the focused words in the initial, medial, and final sentence positions under investigation were the main aim of investigation and analysis, but edge tones such as LL% or L-L% are also provided in cases where the whole intonation contour is described (e.g., L+H* LL%). The break index tier is also used as the fifth tier to observe the continuity of the sentences produced by the participants. In this tier, the ToBI criteria for displaying the degree of disjuncture in the speech stream were adopted as follows:

0 = no word boundary

1 = break between words

2 = disjuncture with pauses and no clear tonal target marking the phrasal edge

3 = intermediate phrase boundary marked by phrase tone

4 = full intonational phrase boundary, or final boundary tone, or complete break at the end of an utterance
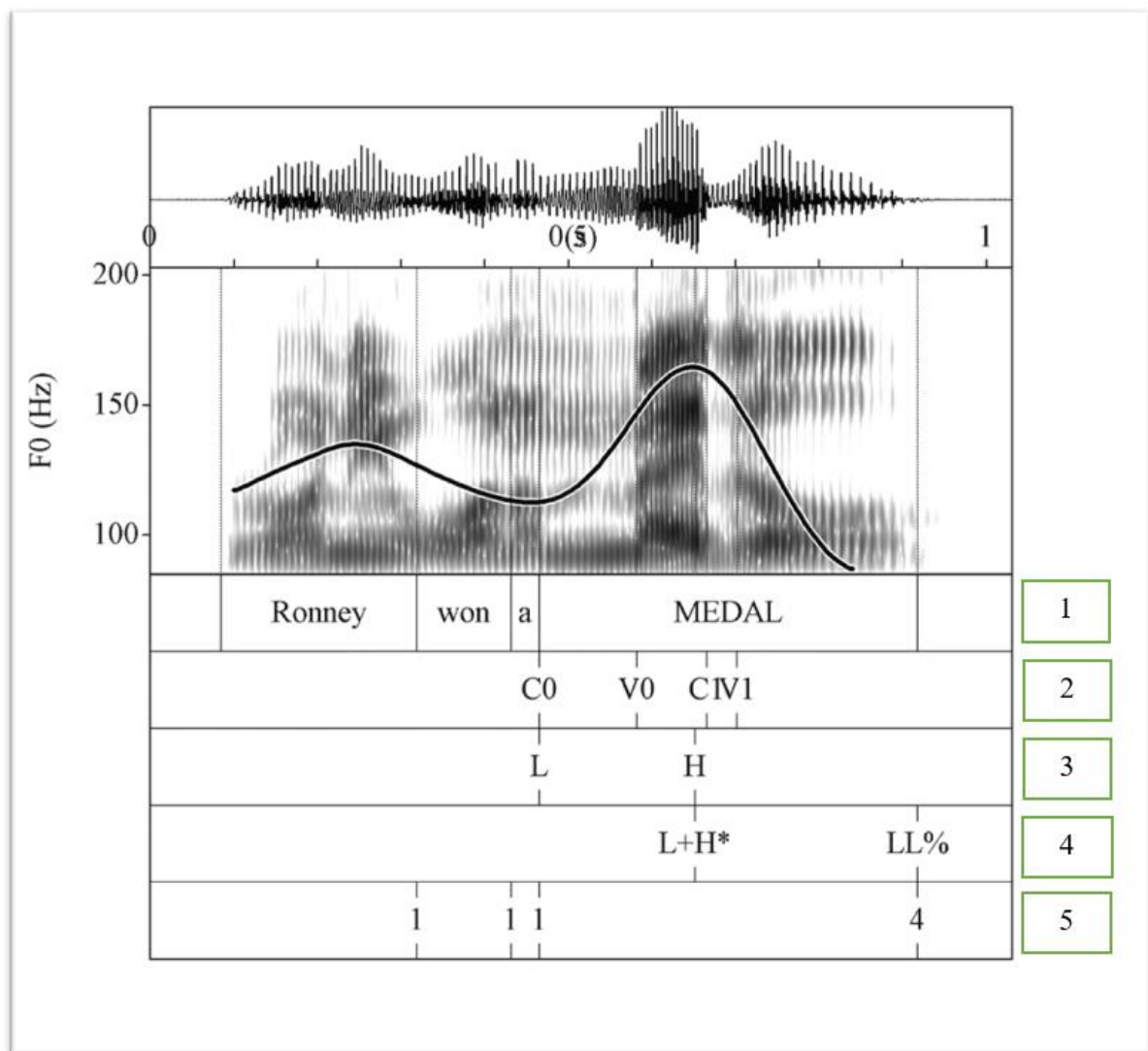
Figure 3. 1 Examples of segmentation and prosodic annotation, illustrating three tiers: word tier (no.1), phonological tone tier (no.4), break index tier (no.5) adopted from ToBI labelling, and two additional tiers: CV tier (no.2) and tonal target tier (no.3) as produced by a native English speaker en1_m09_q22022

3.5.4 Reliability of labelling

All measurements were carried out by the researcher and rechecked. Then two research colleagues whom the researcher had briefed on the ToBI method of prosodic annotation were asked for help and to check the tonal labels that the researcher had marked on the focused words. They verified the labelling by analysing random sentences (i.e., 5 tokens of each position --initial, medial, and final -- from the two groups) and then comparing their labels with the researcher's labelling as in Table 3.7 below. A similar method was used in Oliver and Andreeva

(2004). No disagreements were found between the researcher and the two colleagues in the labelling of accent choices produced by native English speakers. Only a few cases of disagreement on accent choices (i.e., 2 L*+H accents changed to 2 L+<H* accents for final positions) produced by Thai learners were discussed, and the solutions were made based on the agreement among two of the three labellers on the same tonal labels.

Table 3. 7 The number of tokens used for verifying TOBI labelling

Labellers	Tokens for narrow focus					
	L1 English			L2 English		
	Initial	Medial	Final	Initial	Medial	Final
The researcher	5	5	5	5	5	5
1st Research colleague	5	5	5	5	5	5(2)*
2nd Research colleague	5	5	5	5	5	5

*5 = the number of tokens verified,
(2) = the number of the cases of disagreement on accent choices and then the researchers made an agreement on them

3.6 Phonological Analysis

3.6.1 Tokens for phonological analysis

For phonological analysis, the ways in which Thai learners and native speakers of English marked the focus in the target words embedded in the carrier sentences were observed. For narrow focus, the locations of focused words were kept constant at the sentence-initial (Subject), sentence-medial (Verb), and sentence-final (Object) positions. The target focused words correspond to these three positions as given in Table 3.1 in section 3.2. There were 1440 tokens for narrow focus (4 test sentences x 3 positions x 3 repetitions x 20 subjects x 2 language groups) to be analysed as provided in chapter 4. With the aim to serve for cross-reference, details of the strategies for marking broad focus are provided in Appendix E and F.

3.6.2 Analysis of the data

The following prosodic aspects of marking focus were investigated and described:

- *The choice of accent type (e.g., L+H*) which is aligned with the stressed syllable of the focused word. This information can be obtained from the phonological tone tier (the fourth tier). Such an accent type comes from observing the shape of the pitch contour,*

listening to sound files, and drawing a picture of the pitch contour. The distribution of pitch accents was mainly observed here. The LL% edge tone was expected to occur at the end of an utterance. This is so because the carrier sentences in this study were declarative, and so a falling intonation would usually happen. Where relevant, other types of edge tone would be given.

- *The presence of deaccentuation such as pre-focal and post-focal deaccenting as well as pitch range compression. This information can be obtained from the phonological tier (the fourth tier) by observing the drop in pitch after a nuclear accent, or the sustained low pitch before a nuclear accent. For the former, the post-focal contour is deaccented since there are no more accents following the focus elements, whereas for the latter the pre-focal contour is deaccented since there are no more accents preceding the focus elements (Hartmann, 2008).*

- *Phrasing or grouping, if any, by observing the presence of pauses. This information can be obtained from the phonological tier (the fourth tier) and the break index tier (the fifth tier). Normally, the focused elements are located at the right edge of intermediate phrases and/or intonational phrases producing the whole pitch contour. Such phenomenon as prosodic phrasing or re-phrasing are identified from observation of the intonational phrases if they are produced separately followed by a pause. Pauses "ranging from 50 ms up to several hundred ms ..." were also indicated by "ear perception" (Chen, 2015, p. 756). Besides this, the pitch at the end of prosodic phrasing can be realised as a rise, fall, or sustained level, as well as a lengthening of final syllables or words (Hellmuth, 2007; Hartmann, 2008).*

3.7 Acoustic Measurements

3.7.1 Tokens for acoustic study

In the acoustic analysis, narrow focused words were examined using the same number of tokens as in the phonological analysis. The locations of focused words were kept constant at the sentence-initial (Subject), sentence-medial (Verb), and sentence-final (Object) positions. The target focused words correspond to these three positions given as in Table 3.1 in section 3.2. These were 4800 tokens in normalised data for narrow focus (from 4 test sentences x 3 positions x 3 repetitions x 20 subjects x 2 language groups) which are analysed in chapter 5.

3.7.2 Measurement and analysis

In the analysis of the acoustic characteristics of narrow focus marking, the measurements and analysis were mainly conducted based on three phonetic correlates of post-lexical prominence in English. These are f0, duration, and intensity. In the literature it is suggested that vowel quality is one of the factors which is related to the phonetic realisation of prominence in English; however, this variable was not considered in the present study. In this case, this study followed the same strategy as in previous studies (Cutler, Dahan and van Donselaar, 1997; Zerbian, 2013; Mennen and De Leeuw, 2014; Muntendam and Torreira, 2016) and according to the literature (Gussenhoven, 1983; Pierrehumbert and Hirschberg, 1990; Ladd, 2008) which states that prosodic prominence involves the three main acoustic cues of f0, duration and intensity.

The procedures of acoustic measurement and analysis were performed using the Praat script, ProsodyPro (Xu, 2013). This script can be used to obtain several measurements, such as maximum f0, minimum f0, and mean f0 in either Hertz or semitone units with time normalization, as well as syllable or word duration. To investigate the acoustic properties of focus, the mean f0 (Hz) of each repetition of each stressed syllable of the focused words in the test sentences was calculated and converted into semitones (ST) relative to 1 Hertz, so that a mean f0 (ST) of each three repetitions were obtained. This is because the f0 values in semitones can allow for "cross-participant comparison" (Gut and Pillai, 2014, p. 291). In the same way as calculating f0, the mean intensity (dB) of each repetition of each stressed syllable of the focused words in the test sentences was calculated and then the mean intensity of the three repetitions was obtained. The mean duration (ms) of each repetition of each stressed syllable of the focused words in the test sentences was also calculated and then the mean duration of the three repetitions was obtained. All measurement and analysis followed the guideline of ProsodyPro (Xu, 2013).

3.8 Analysis of Tonal Alignment and Scaling

3.8.1 Tokens of rising accents

To analyse the alignment and scaling of rising tonal targets in marking narrow focus in Chapter 6, tokens of rising accents were chosen and obtained from the results regarding the distribution of accents in Chapter 4, where the use of intonation (accent choices) to mark focus was

qualitatively analysed. A total of 424 narrow focused words with rising pitch movement were used in the analyses. There were 207 of rising tokens from Thai learners and 217 rising tokens were from native speakers of English. The rising accents under broad focus marking were beyond the scope of the current study. This is due to variations in the word choices and the locations of the placing of rising accents among the participants and sentences.

3.8.2 Annotation criteria

Figure 3.1 in section 3.5 shows the annotation tiers corresponding to the acoustic landmarks which are used in measuring and analysing the alignment and scaling of the tonal targets of the rising pitch movement. For the measurement and analysis of tonal alignments, the CV tier (2nd tier) and the tonal target landmark tier (3rd tier) were taken into consideration. In the CV tier, consonants and vowels are labelled "C" and "V" respectively. On the basis of previous research on tonal alignment and scaling (e.g., Atterer and Ladd, 2004; Arvaniti, Ladd and Mennen, 2006; Ladd *et al.*, 2009), more detailed descriptions of the labels on the CV tier are provided in Table 3.8 below.

Table 3. 8 Criteria and description for CV annotation

CV landmark	Descriptions
C0	the beginning of the initial consonant in the stressed syllable of the focused word
V0	the beginning of the vowel in the stressed syllable of the focused word, or the beginning of the rhyme.
C1	the beginning of the initial consonant in the syllable after the stressed syllable of the focused word. C1 can be considered as the offset of the stressed vowel of the focused words, or as the beginning of the coda consonants in a closed syllable.
V1	the beginning of the vowel in the syllable after the stressed syllable of the focused word.
C2	the beginning of the second coda consonant in a closed syllable as in the target word "Lent".

For example, the target word "Nanny" was labelled as the series C0V0C1V1; the target word "Read" was labelled C0V0C1, and the target word "Lent" C0V0C1C2. The actual points used in extracting the lengths of time intervals in the CV tier were only C0, V0, C1, since the first syllable in all of the target words in question was accented. In the tonal target landmark tier, the f0 contours were decomposed into sequences of Low (L) and High (H) tones. The method used in detecting tonal targets or f0 turning points was on the basis of visual inspection (Ladd *et.al*, 2009; Prieto, 2009). This method was used in combination with the Praat functions for pitch analysis which involved 'move cursor to minimum pitch' and 'move cursor to maximum pitch' (Peters, Hanssen and Gussenhoven, 2014). This means that the low (L) tonal target was placed in the position of the lowest f0 turning point that was in the vicinity of the target stressed syllable of the focused word, and the high (H) tonal target was placed in the position of the highest f0 turning point that was in the vicinity of the target stressed syllable of the focused word. After the L and H tonal targets were identified as presenting a rising movement, they were all selected in order to examine the alignment of tonal targets with segmental strings.

3.8.3 Measurement and analysis

Consistent with previous research on f0 alignment (Ladd and Schepman, 2003; Prieto and Torreira, 2007; Ladd *et al.*, 2009), the durational measurements shown in the following table 3.9 were taken, and then the values obtained were converted into relative proportional measurements. For example, the values of lowest f0 (L) turning points were measured relative to the syllable onset (C0), and then the alignment of the f0 valley was calculated as a proportion of the duration of the accented syllable of the focused word; that is, C0toL/C0toC1. Temporal distances and time intervals in this study were measured in milliseconds (ms). Values of f0 were taken in Hertz, and were not converted into semitones relative to 1 Hertz. All of the values were extracted from the relevant measurement points. The results obtained are reported in the section on tonal alignment in Chapter 6 as the relative alignment of the f0 valley, the relative alignment of the f0 peak, and the f0 excursion. Table 3.9 below shows the measurements relevant to tonal alignment and scaling.

Table 3. 9 Measures of tonal alignment and scaling, together with corresponding measurement points

Measures	Corresponding measurement points	Descriptions
Syllable onset duration	C0toV0	The distance from C0 to V0, equivalent to the duration of the accented consonant
Vowel duration	V0toC1	The distance from V0 to C1, equivalent to the duration of the accented vowel
Syllable duration	C0toC1	The distance from C0 to C1, equivalent to the duration of the accented syllable
Rise size	Maxf0 – Minf0	The f0 difference between the peak and the preceding minimum f0, equivalent to f0 excursion size
Maximum f0 (H)	Absolute f0 value at H turning point	The value of the f0 peak on the accented syllable of the focused word
Minimum f0 (L)	Absolute f0 value at L turning point	The value of the f0 valley preceding the f0 peak (H)
F0 valley alignment	C0toL / C0toC1	The alignment of the f0 valley relative to the accented syllable onset as a proportion of the duration of the accented syllable
F0 peak alignment	C0toH / C0toC1	The alignment of f0 peak relative to the accented syllable onset as a proportion of the duration of the accented syllable

3.9 Statistical Analysis

The current study used R, which is a language and environment for statistical computing and graphics (R Development Core team, 2019), and R studio (RStudio Team, 2017) to manipulate the data set. In addition to this, other R packages such as 'dplyr' (Wickham *et al.*, 2018) and 'reshape2' (Wickham, 2017) are used to manipulate data along with 'ggplot2' (Wickham, 2016; Wickham *et al.*, 2019) for make graphical visualisations. The statistical methods of mixed-effects models, using R studio/software were applied to analyse a range of variables in terms of main and random effects.

In the statistical analyses of the accent types, a series of GLMM analyses was performed to determine whether or not native English speakers and Thai learners differ in terms of their choices of accent types for marking narrow focused words. This includes:

- Accent types as dependent variables
- English proficiency levels and sentence positions as independent variables

In the statistical analyses of the acoustic correlates of focus, narrow-focused words were taken into account. A series of linear mixed-effect models was constructed separately for each case of acoustic measurements or for each case of dependent acoustic variables. These included measurements of:

- *Fundamental frequency (f0)*
- *Intensity*
- *Duration*

In the statistical analyses of tonal alignments in rising accents, narrow-focused words were again taken into account. Linear mixed-effect models were constructed separately for each case of tonal alignment and scaling measurements as follows:

- Alignment and scaling of the f0 valley
- Alignment and scaling of the f0 peak or relative peak delay
- Alignment and scaling of the rising slope

For each of the above cases, the main question of interest was how the first language of the participants, English proficiency, gender and focus position affected their levels of f0, intensity, and duration in marking English focus, as well as f0 valley alignment, f0 peak alignment, and the scaling slope of L and H tonal targets.. To examine this, the "Participant" and "Word" variables were treated as random effects, while the 'Group' (English native speakers or Thai native speakers), 'Position' (sentence-initial, sentence-medial or sentence-final), 'English proficiency', and 'Gender' variables were included as fixed effects. The results were considered statistically significant when the p-value associated with the relevant statistics was smaller than 0.05. When significant effects were found, especially for either 2-way or 3-way interaction effects, post-hoc pairwise comparisons using the post-hoc Tukey test were performed to examine which pairs of factor levels were significantly different from each other.

3.10 Summary of the Chapter

This chapter has outlined the design of the experiment on the prosodic marking of focus in English. It has covered the design of the speech materials, the recruitment of participants, the procedures of data collection and recording, and the adoption of the ToBI system for the annotation and analysis of prosodic data. The methods of statistical analysis used in the current study have been explained as well. Furthermore, the laboratory experiment has been briefly described. Attempts were made to control as far as possible several variables, as listed below, in the design of the speech materials. In this study, the focus positions of target words and the locations of lexical stress in the target words were exploited in an investigation into the prosodic marking of focus. In addition to this, the experiment was designed to investigate both phonological and acoustic-phonetic means for marking focus. Factors affecting the realisations of focus were also explored. The procedure of the research is illustrated as in Figure 3.2 below in terms of examining:

1. *the accent types used by Thai learners and native speakers of English*
2. *the acoustic properties of accented syllables in focused words*
3. *the tonal alignment and scaling of rising accents on accented syllables in focused words. Of special interest, the rising accents were selected.*

Examination of Accent Choices	Examination of Acoustic Realisation of Focus	Examination of Phonetic Realisation of Rising Accents
• L+H* • L*+H • L+<H* • H*, L*	• f0 • duration • intemsity	• Alignment • Scaling
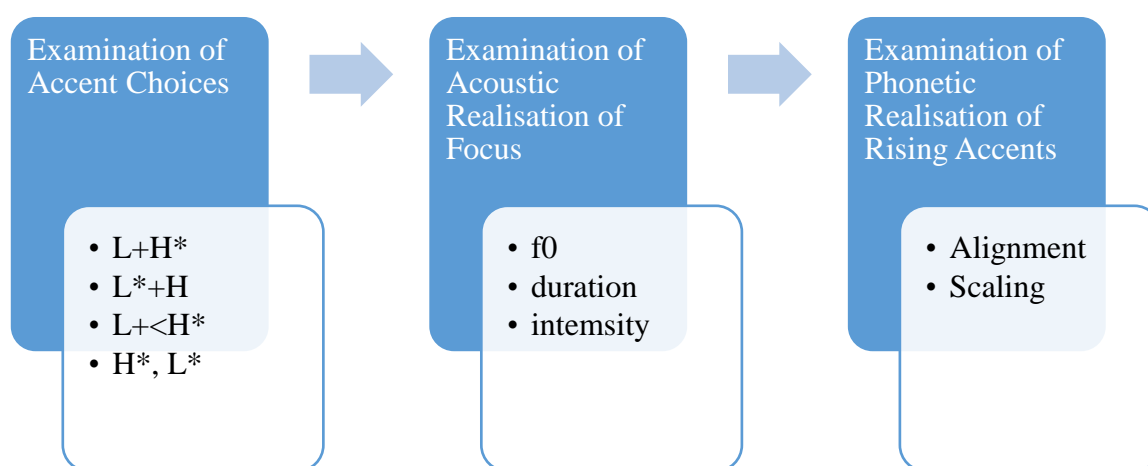
Figure 3. 2 Procedure of the research on the intonational marking of narrow focus

The results of the acoustic-phonetic and phonological analyses are reported in separate chapters. Chapter 4 presents the phonological results for focus marking, and Chapter 5 presents the acoustic results, while those for alignment and scaling results can be found in Chapter 6.

Chapter 4. Accent Choices for Marking Narrow Focus in English

4.1 Introduction

The chapter begins by considering the marking of narrow focus in the L1 by native English speakers in section 4.2. This section covers the accent types used by native English speakers as well as other relevant strategies from the qualitative analysis. Section 4.3 presents the results of the qualitative analysis of the marking of narrow focus by Thai learners of English in terms of the use of accent types with the addition of other relevant strategies. Section 4.4 then statistically reports on similarities and differences in the use of accent types to mark the focus. A summary of the chapter is provided in section 4.5.

4.2 Marking Narrow Focus in L1 English by Native Speakers

This section presents the results of the phonological analysis of the use of intonation to mark focus. The research question that this section addresses is shown below, and the answers are given in section 4.2.1 concerning the use of accent types and other accompanying strategies.

Research question 1:   What are the accent types that native speakers of English use to mark narrow informative focus?

4.2.1 Use of accent types

Table 4.1 shows the descriptive statistics of the overall number of English native speakers or NSE (n = 720). The data in the table is made up of the distribution of accents that were produced by all the subjects. The number of response accents were counted and displayed in the columns related to the positions in the sentences that the accents occurred, for example, sentence-initial, sentence-medial, and sentence-final. The number and accent types varied a great deal across sentence positions. In total, the number of accents were 720 items used by native speakers of English. The accents included: H*, L*+H, L*+H, and L+H*. The native English speakers in this study did not use L* accent to mark narrow focus.

Table 4. 1 Descriptive statistics for the frequency of pitch accents by the native English subjects by positions

Subjects	English Proficiency Levels	Pitch Accents	Focus Positions			Total
			Initial	Medial	Final	
Native English speakers (NSE, n = 720)	Native (n = 720)	H*	58 (0.24)	31 (0.13)	8 (0.03)	97
		L*	0 (0.00)	0 (0.00)	0 (0.00)	0
		L*+H	8 (0.03)	0 (0.00)	0 (0.00)	8
		L+<H*	52 (0.22)	45 (0.19)	22 (0.09)	119
		L+H*	122 (0.51)	164 (0.68)	210 (0.88)	496
		Total	240	240	240	720

*The digits in () represent a proportion of pitch accents occurring at the same position.
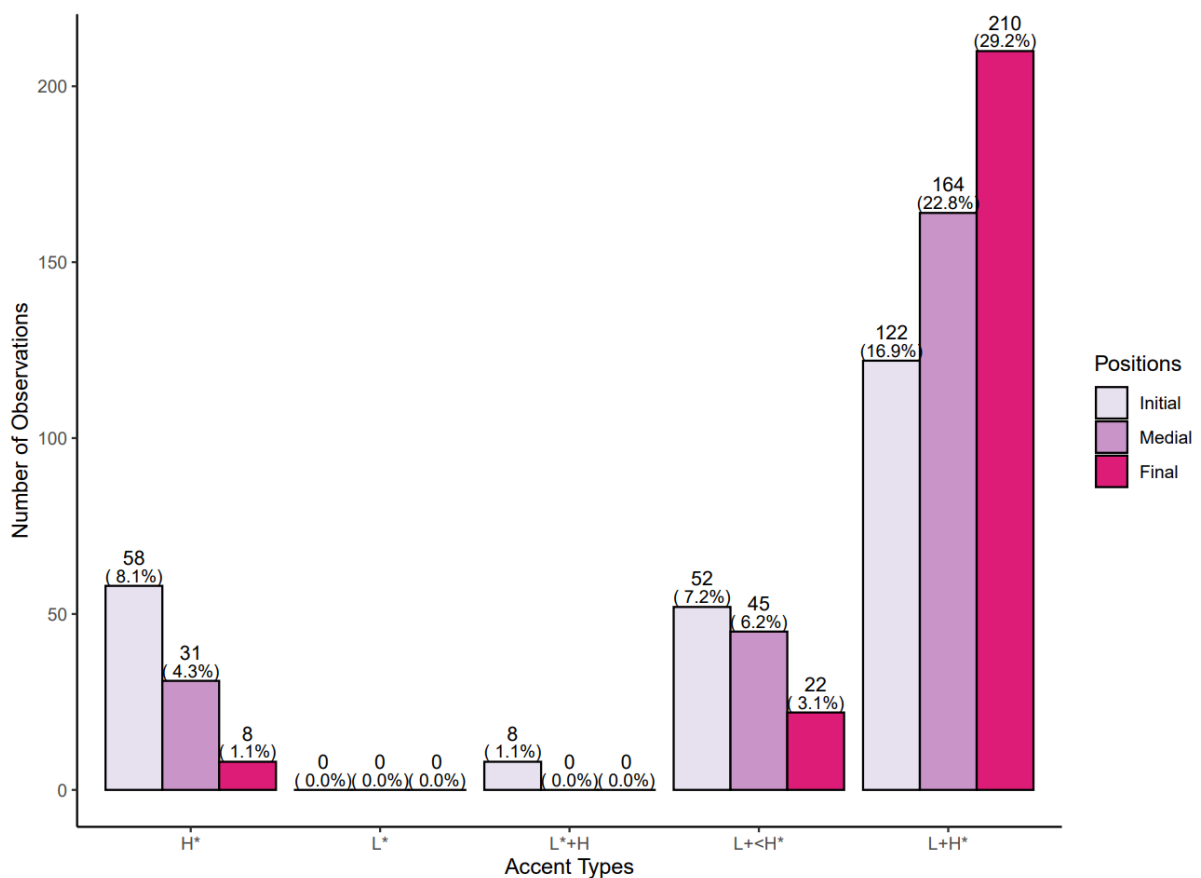
Figure 4. 1 Proportional distribution of accents as produced by native English speakers across narrow focus positions

Figure 4.1 above shows the proportional distribution of accents as produced by native English speakers. The horizontal axis represents the accent types occurring on the focused words at the three different positions (sentence-initial, -medial, and -final). The vertical axis represents the number of observations (or the proportion of values) in each category of accents across positions.

After the data were analysed, the results reveal that native English speakers in this study predominantly used the H* accent, the (L+H*) rising accent, and the (L+<H*) rising accent with a delayed peak in the expression of focus. When considering a particular location of accent placement, the following accent types were found. The H* accent was used 8.1% (58 observations), or a proportion of 0.24 at the beginning, 4.3% (31 observations), or a proportion of 0.13 at the middle, and 1.1% (8 observations), or a proportion of 0.03 at the end of the sentence. The L+H* accent was mostly found in 16.9% (122 observations), or a proportion of 0.51 at the beginning, 22.8% (164 observations), or a proportion of 0.68 at the middle, and 29.2% (210 observations), or a proportion of 0.88 at the end of the sentence. The L+<H* accents was used in 7.2% (52 observations) of cases, or a proportion of 0.22 at the beginning, 6.2% (45 observations), or a proportion of 0.19 at the middle, and 3.1% (22 observation), or a proportion of 0.09 at the end of the sentence. However, individual speakers used these accents differently. That is, while one speaker preferred to use the H* accents, others preferred the L+H* accents to mark the same focused words. In addition, using or placing these pitch accents varied from one position to another. In other words, both accent types can be placed at any position in the sentence. Apart from the H* and L+H* accents, there were other accent types that native speakers of English in this study employed. These included the (L*+H) scooped accents in 1.1% (8 observations) of cases, or a proportion of 0.03 at the beginning. No L* accents were found to be used for the marking of focus.

Given the overall use of accent types for marking focus at the three positions in the sentence, native English speakers used the rising accents (L+H*) predominantly, while H* accents and L+<H* accents with peak delay were used to some degree. They used the rising accents (L+H*) sentence-finally, sentence-medially and sentence-initially in proportions of 0.88 (29%, 210 from a total of 720 observations), 0.68 (22.8%, 164 observations), and 0.51 (16.9%, 122 observations) respectively. They employed H* accents sentence-initially with 8.1% (58 observations), or in a proportion of 0.24. Finally, they also employed the L+<H* accents sentence-initially with 7.2% (52 observations) of cases, or in proportion of 0.22.
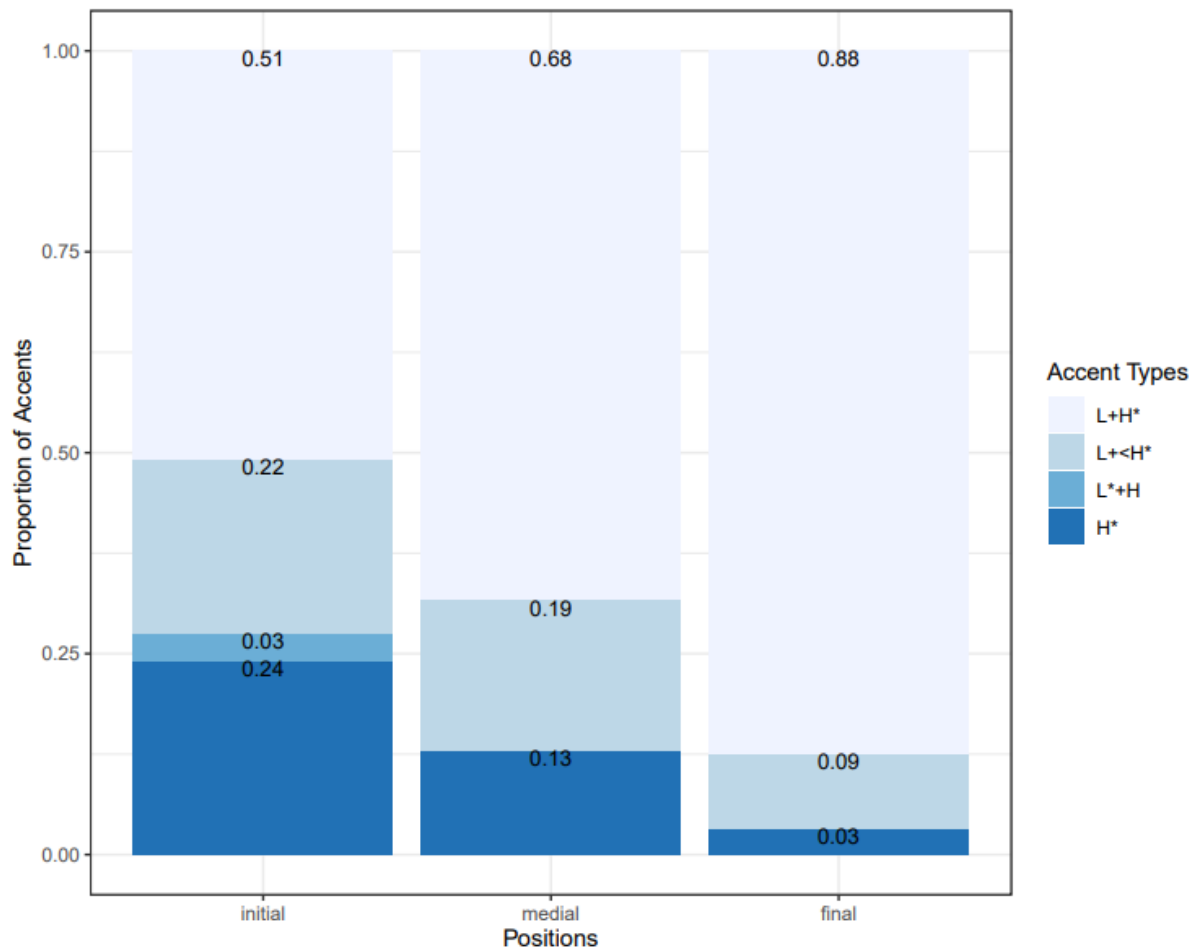
Figure 4. 2 The proportion of accents to positions as produced by native English speakers

4.2.2 Use of other prosodic strategies

Besides the use of accent types as reported above, there were several accompanying strategies used by native speakers of English to work as cues for the marking of narrow focus in the sentences; for example, deaccenting, using a compressed pitch range, and prosodic phrasing. These strategies were found in different positions in the sentence; across sentences and participants. The use of deaccenting was more frequent compared to other strategies. These strategies can be accompanied by any accent (H*, L*+H, L+H*, L+<H). The results regarding this are exemplified and reported separately in relation to positions in the sentence as follows.

1) Sentence-initial position

Native English participants in this study tended to mark focus sentence-initially to a great extent. They produced focused words with a nuclear focal accent type, followed by post-focal deaccenting. For example, a rising (L+H*) accent could be followed by post-focal deaccenting, as shown in Figure 4.3, and a rising (L+<H*) accent with a delayed peak followed by post-focal deaccenting as in Figure 4.4.
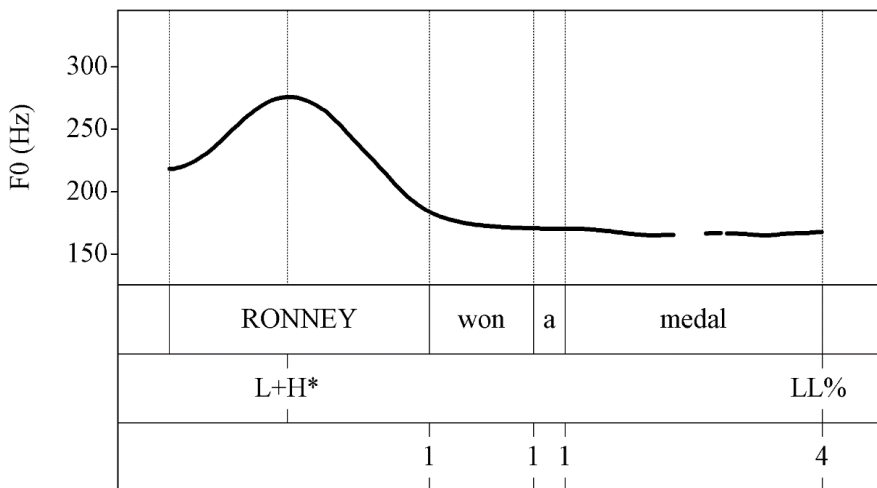
Figure 4. 3 Rising accent (L+H*) followed by post-focal deaccenting as produced by the female native English speaker en1_f13_q11021
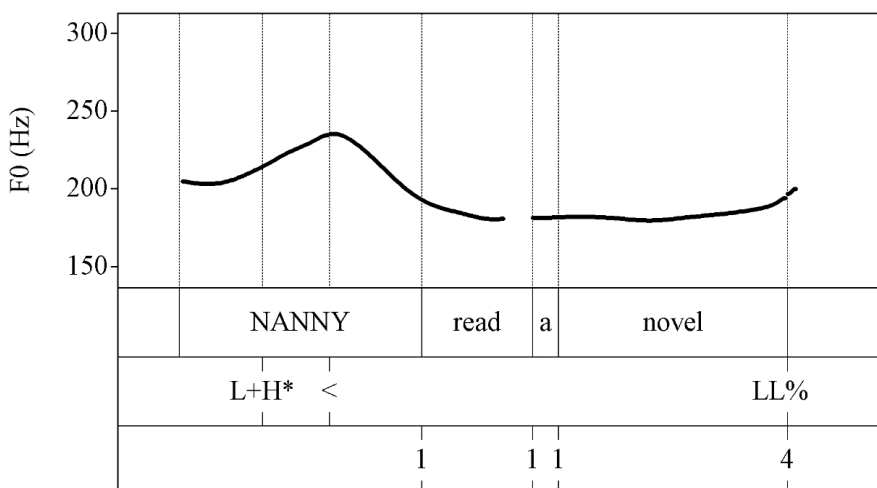
Figure 4. 4 Rising accent (L+<H*) with delayed peak followed by post-focal deaccenting as produced by the female native English speaker en1_f17_q11011

2) Sentence-medial position

English native participants in this study marked focus sentence-medially by producing focused words with a nuclear focal accent type, preceded by pre-focal deaccenting and followed by post-focal deaccenting. For example, there was pre-focal deaccenting before a rising (L+H*) focal accent and then post-focal deaccenting after the same rising (L+H*) focal accent, as shown in Figure 4.5. Figure 4.6 illustrates a rising accent with delayed peak preceded by pre-focal deaccenting and then followed by post-focal deaccenting. However, other strategies were also used by some native speakers of English. These included a rising (L+H*) accent preceded by a pre-nuclear accent, as in Figure 4.7, and a rising (L+<H*) accent with a delay peak as in Figure 4.8.
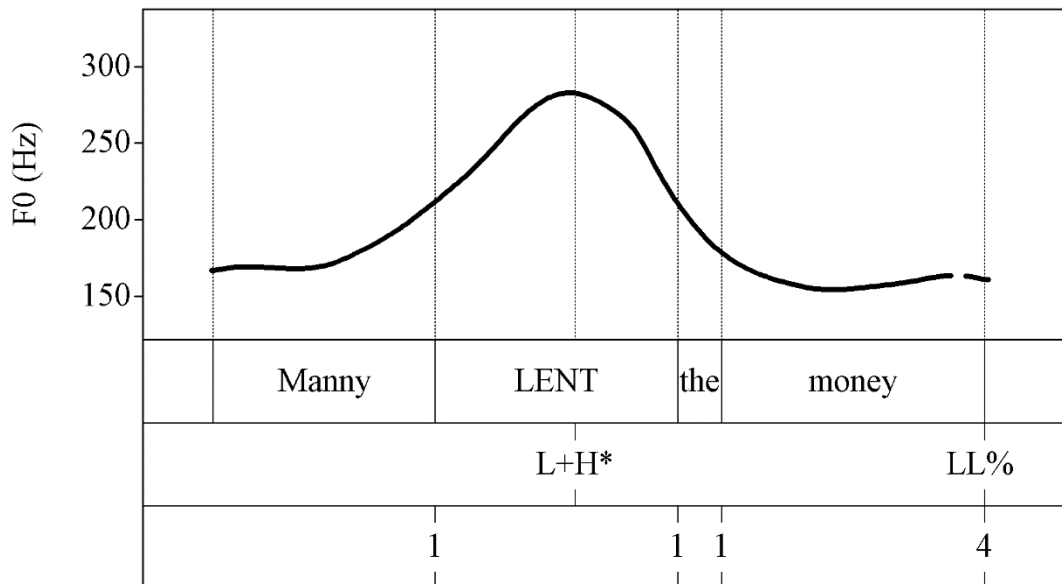
Figure 4. 5 Rising accent (L+H*) with pre-focal and post-local deaccenting as produced by the female native English speaker en1_f01_q33062
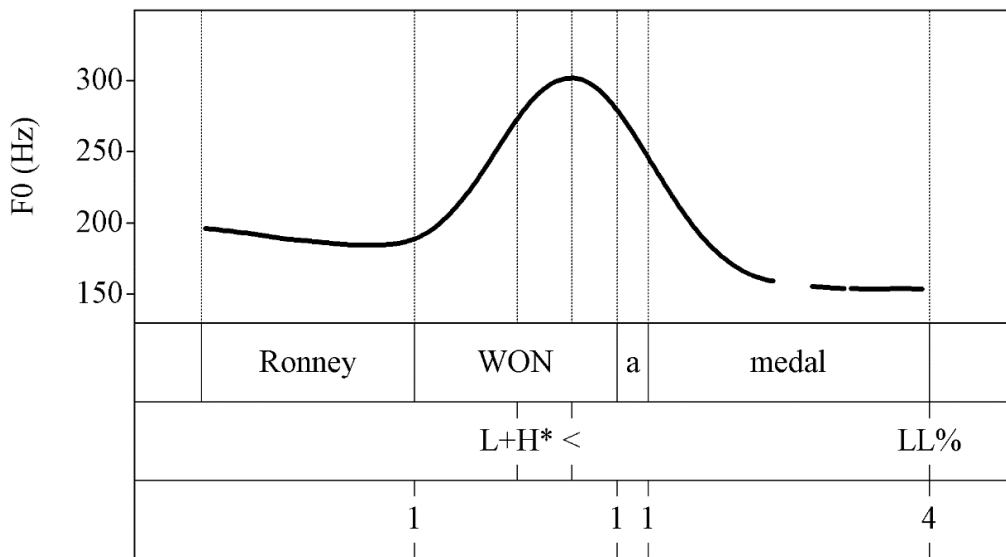
Figure 4. 6 Rising accent (L+<H*) with delayed peak alongside pre-focal and post-local deaccenting as produced by the female native English speaker en1_f13_q33023
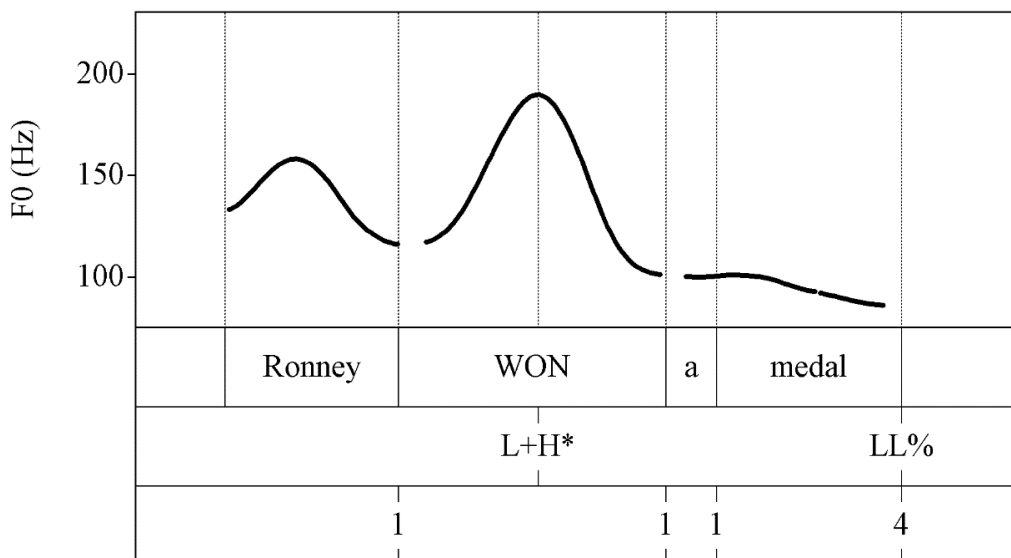
Figure 4. 7 Rising accent (L+H*) with pre-nuclear accent as produced by the male native English speaker en1_m08_q33021
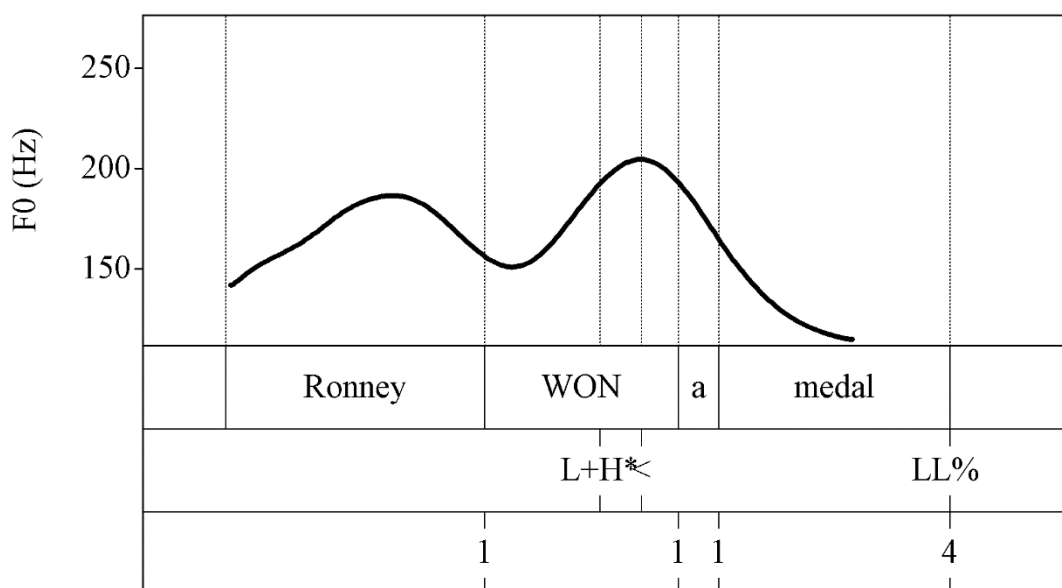
Figure 4. 8 Rising accent (L+<H*) with a delayed peak preceded by pre-nuclear accenting as produced by the male native English speaker en1_m07_q33021

3) Sentence-final position

In the final position of the sentence, English native speakers in this study employed several means to mark focus. First, there was the production of focused words with a nuclear focal accent preceded by pre-focal deaccenting as in Figure 4.9. Second, there was a pre-nuclear accent before the production of focused words with a nuclear focal accent as shown in Figure 4.10. In addition, there were other strategies used by native speakers of English. For example, there was a plateau-like pattern which can be transcribed as either H* H* LL% or L+H* L+H* LL% as in Figure 4.11. Producing one sentence as two separate minor or major intonational phrases, which is normally referred to as prosodic phrasing, was also employed as shown in Figures 4.12, and 4.13.
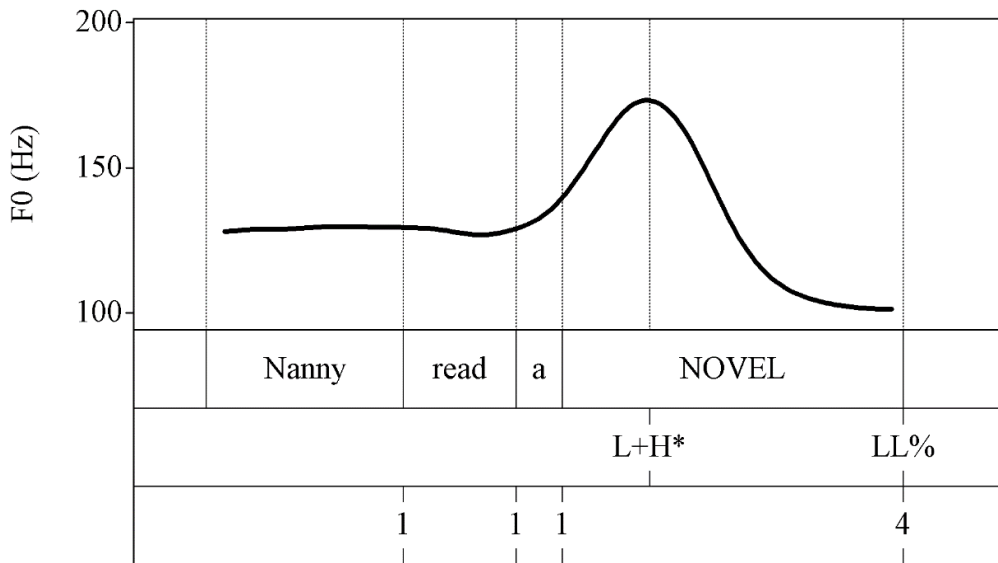
Figure 4. 9 Rising accent (L+H*) with pre-nuclear accent or pre-focal deaccenting as produced by the male native English speaker en1_m10_q22012
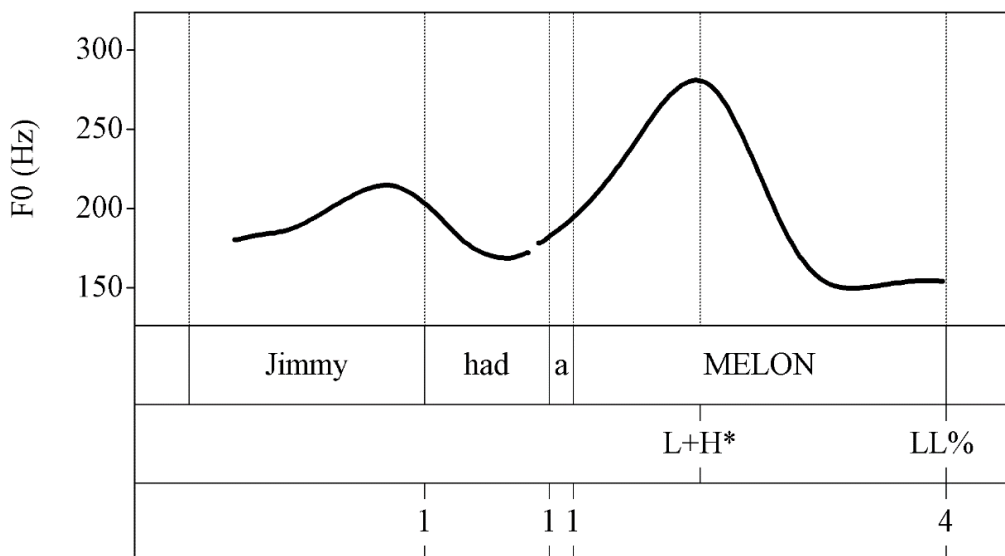
Figure 4. 10 Rising accent (L+H*) with the pre-nuclear accent as produced by the female native English speaker en1_f01_q22031
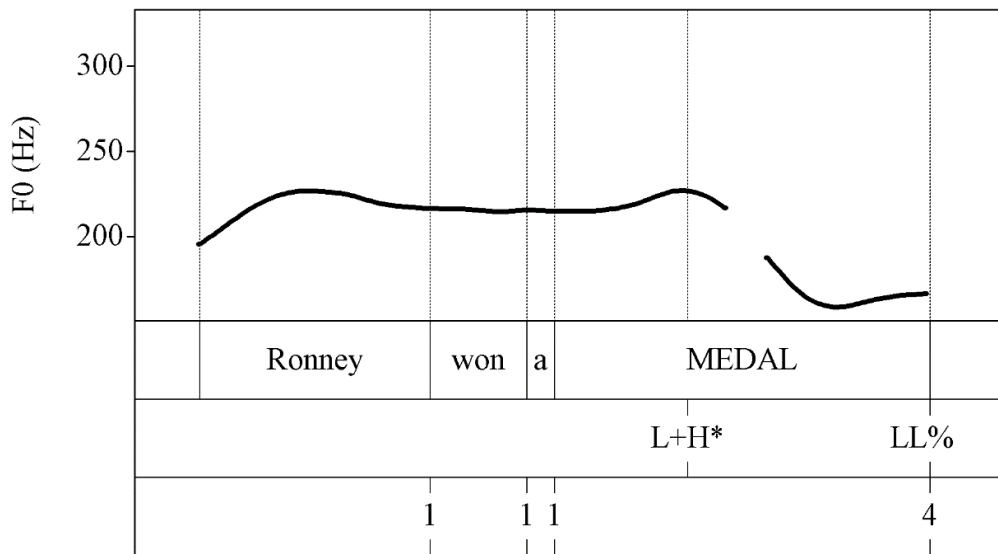
Figure 4. 11 Characteristics of high plateau-like pattern as produced by the female native English speaker en1_f07_q22021
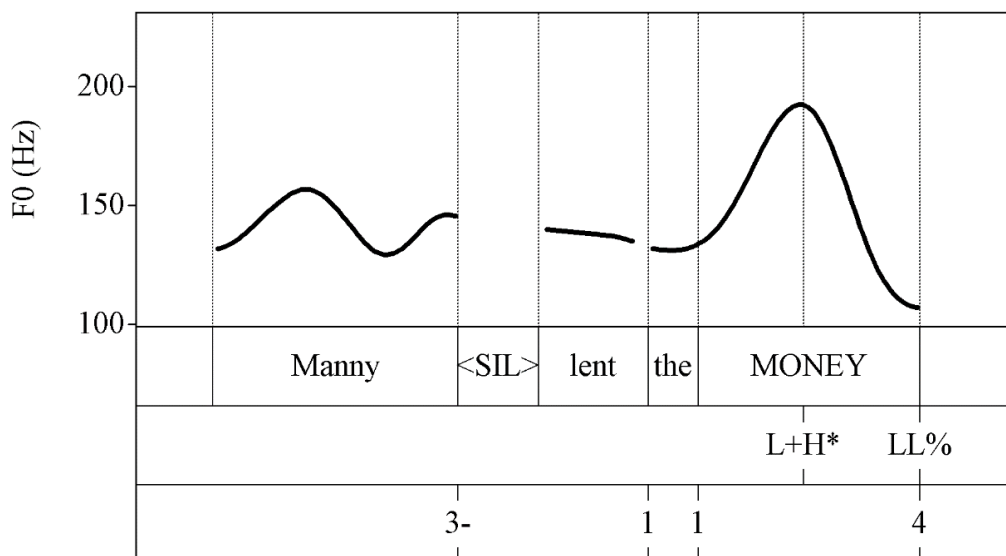
Figure 4. 12 Characteristics of prosodic phrasing or re-phrasing with a rising accent in the first phrase as produced by the male native English speaker en1_m07_q22061
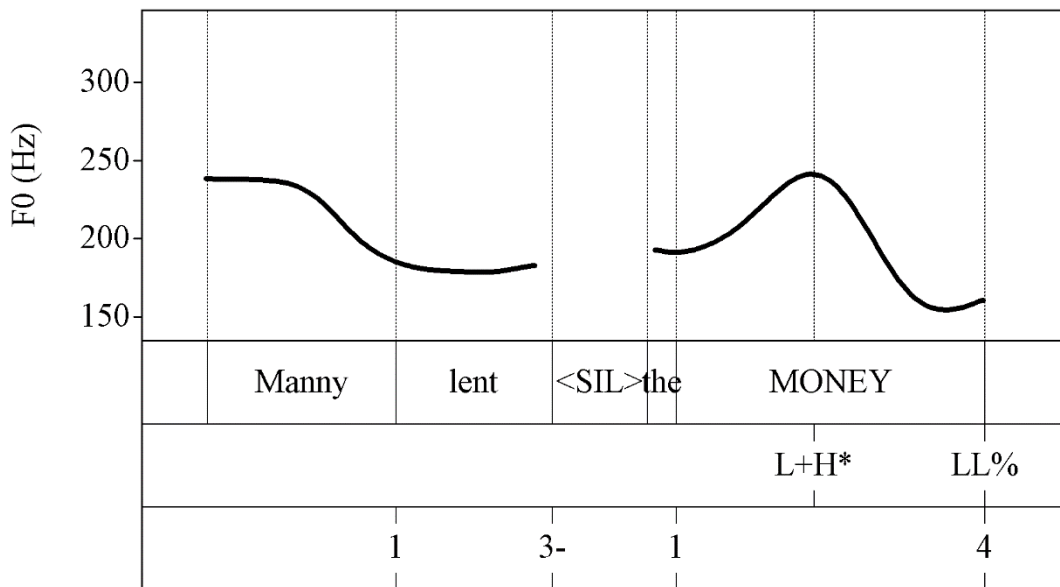
Figure 4. 13 Characteristics of prosodic phrasing or re-phrasing with H* accent in the first phrase as produced by the female native English speaker en1_f13_q226

4.2.3 Summary of the marking of narrow focus in L1 English by native speakers

From section 4.2.1, the native speakers of English predominantly employed rising accents (L+H*) for marking narrow informative focus. They also employed L+<H* accents with delayed peak and H* accents. They were found to employ the L*+H accent, but they employed it least often. Figure 4.14 below shows the 3 rising accents as mentioned here: L+H*, L+<H* and L*+H. In addition, all of these accents, even H*, were used in combination with other prosodic strategies as shown in section 4.2.2 according to different focus positions. The native speakers in this study employed predominantly post-focal deaccenting while other strategies were found to be used less; for instance, pre-nuclear accenting and prosodic phrasing.
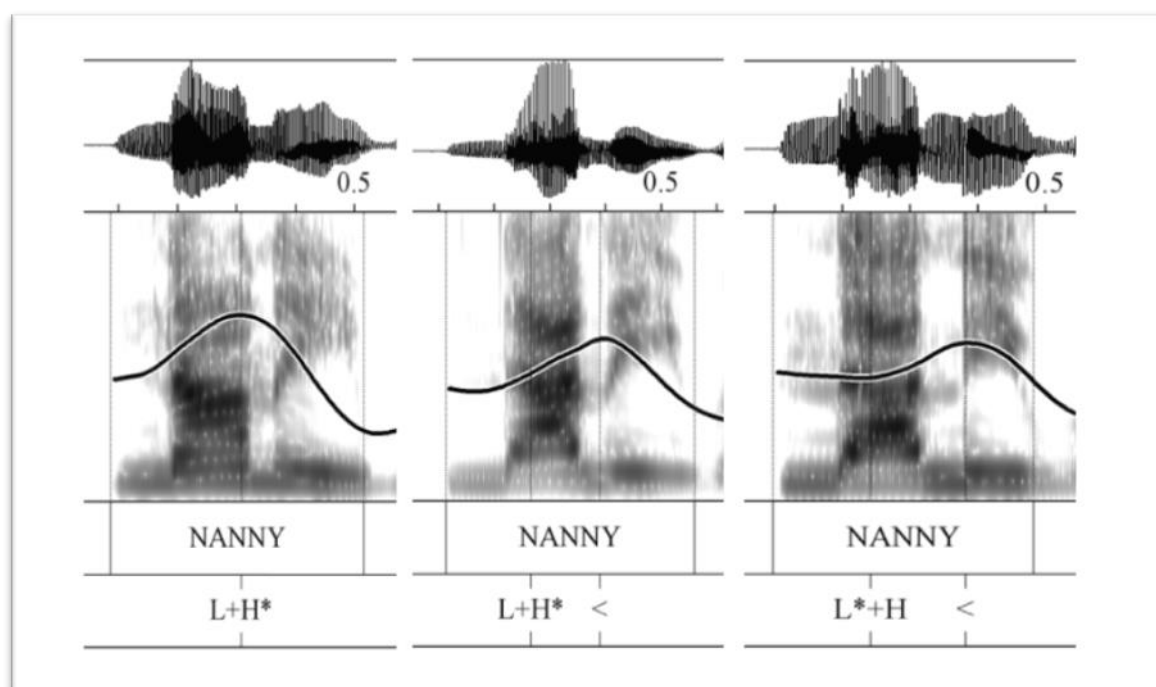
Figure 4. 14 Schematic representations of 3 rising accents: rising accent with aligned peak (L+H*), rising accent with delayed peak (L+H* < or L+<H*), and scooped accent (L*+H) on focus accented words produced by the native speakers

4.3 Marking Narrow Focus in L2 English by Thai Learners

The research question that this section addresses is as follows, and the answers to it are given concerning the use of accent types and other accompanying strategies in section 4.3.1.

Research question 2:   What are the accent types that Thai learners of English use to mark narrow informative focus?

4.3.1 Use of accent types

Table 4.2 shows the descriptive statistics of the overall number of Thai learners or English L2 Learners or ELL (n = 720). They were separated into two groups according to their levels of English proficiency: Low for Thai learners with low English proficiency (n = 360), and High for Thai learners with high English proficiency (n = 360). The data in the table is made up of the distribution of accents that were produced by all the Thai subjects. The number of response

accents were counted and displayed in the columns related to the positions in the sentences that the accents occurred, for example, sentence-initial, sentence-medial, and sentence-final. In total, the number of accents were 720 items used by Thai learners of English. The accents included: H*, L*, L*+H, L*+H, and L+H*. Some of Thai learners of English in this study used L* accent to mark narrow focus.

Table 4. 2 Descriptive statistics for the frequency of pitch accents by Thai learners of English by positions

Subjects	English Proficiency Levels	Pitch Accents	Focus Positions			Total
			Initial	Medial	Final	
English L2 learner (ELL) or Thai learners (n = 720)	Low & High (n = 360*2)	H*	52 (0.22)	79 (0.33)	39 (0.16)	170
		L*	0 (0.00)	12 (0.05)	0 (0.00)	12
		L*+H	5 (0.02)	1 (0.004)	0 (0.00)	6
		L+<H*	127 (0.53)	46 (0.19)	51 (0.21)	224
		L+H*	56 (0.23)	102 (0.43)	150 (0.63)	308
		Total	240	240	240	720

*The digits in () represent a proportion of pitch accents occurring at the same position.

Figure 4.15 below shows the distribution of accents as produced by native Thai speakers. The horizontal axis represents the accent types occurring on the focused words at the three different positions (sentence-initial, -medial, -final). The vertical axis represents the percentage of values in each category of accents across positions. As shown in the figure, there were a wide range of accent types that the Thai learners of English employed in this study. To mark narrow focus, the L* accent, the H* accent, the scooped accent (L*+H), the rising accent (L+H*), and the rising accent (L+<H*) with a delayed peak were used.
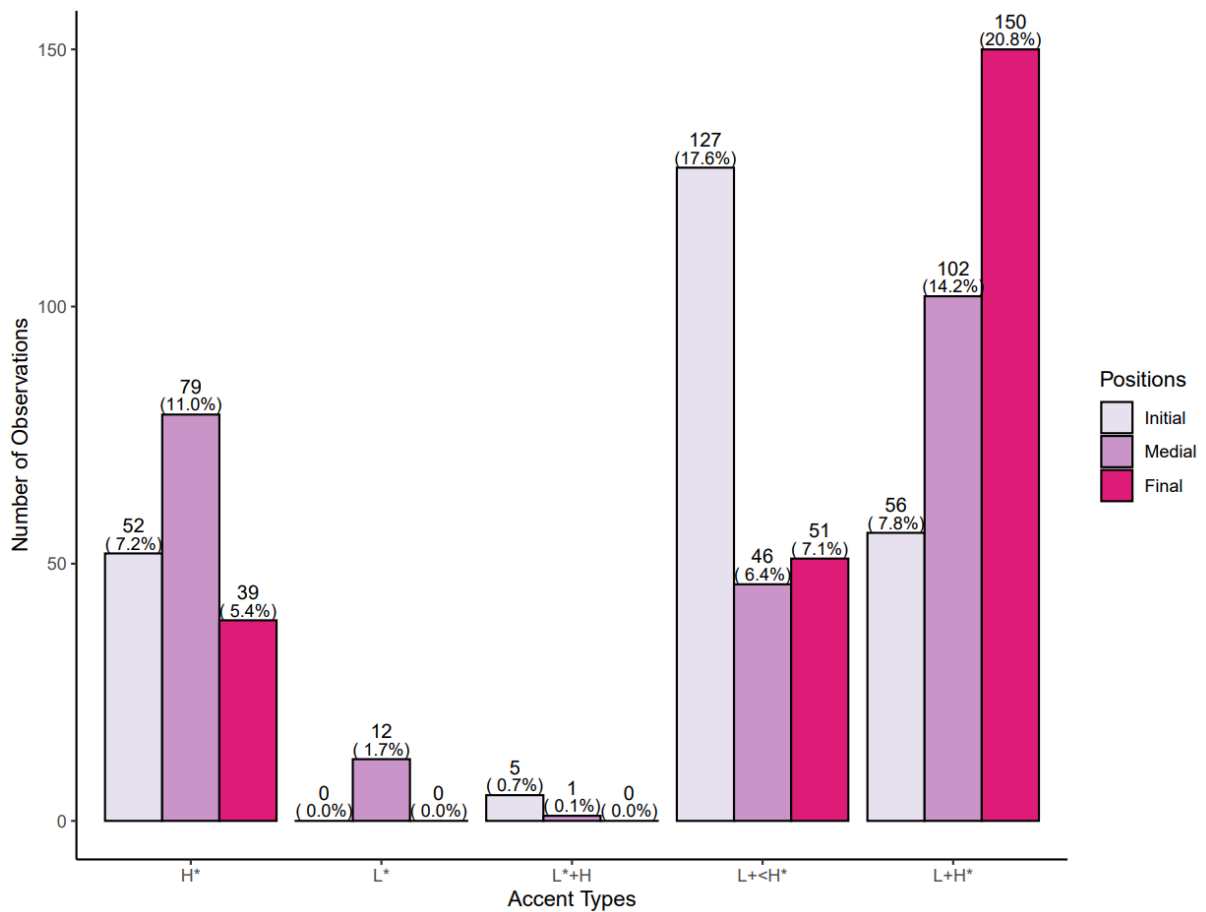
Figure 4. 15 Distribution of accents as produced by native Thai speakers across narrow focus positions

When looking at the use of accent types to mark narrow focus according to sentence position, Thai learners used the L* accent in 1.7% (12 observations) of cases, or a proportion of 0.05 sentence-medially. The H* accent was used in 7.2% (52 observations) of cases, or a proportion of 0.22 at the beginning, 11.0% (79 observations), or a proportion of 0.33 at the middle, and 5.4% (39 observations), or a proportion of 0.16 at the end of the sentence. The L+H* accent was found to be used in 7.8% (56 observations) of cases, or a proportion of 0.23 at the beginning, 14.2% (102 observations), or a proportion of 0.43 at the middle, and 20.8% (150 observations), or a proportion of 0.63 at the end of the sentence. In addition, Thai learners in this study employed other accent types. These included 0.7% (5 observations) of the scooped accents (L*+H), or a proportion of 0.02 at the beginning, and 0.1% (1 observation), or a proportion of 0.004 at the middle of the sentence. The L+<H* (rising accent with a delayed peak) was used in 17.6% (127 observations) of cases, or a proportion of 0.53 at the beginning,

6.4% (46 observations), or a proportion of 0.19 at the middle, and 7.1% (51 observations), or a proportion of 0.21 at the end of the sentence.

Given the overall use of accent types when marking focus at the three different positions in the sentence as in Figure 4.16, Thai learners used L+H*, L+<H*, and H* predominantly. First, they marked narrow focus using mostly rising accents (L+H*) sentence-finally in 20.8% (150 observations), or a proportion of 0.63. Second, they used L+<H* accents sentence-initially in 17.6% (127 observations), or a proportion of 0.53. Finally, they used rising accents (L+H*) sentence-medially 14.2% (102 observations), or a proportion of 0.43. They also used the H* accents to express this kind of focus sentence-medially in 11.0% (79 observations), or a proportion of 0.33.
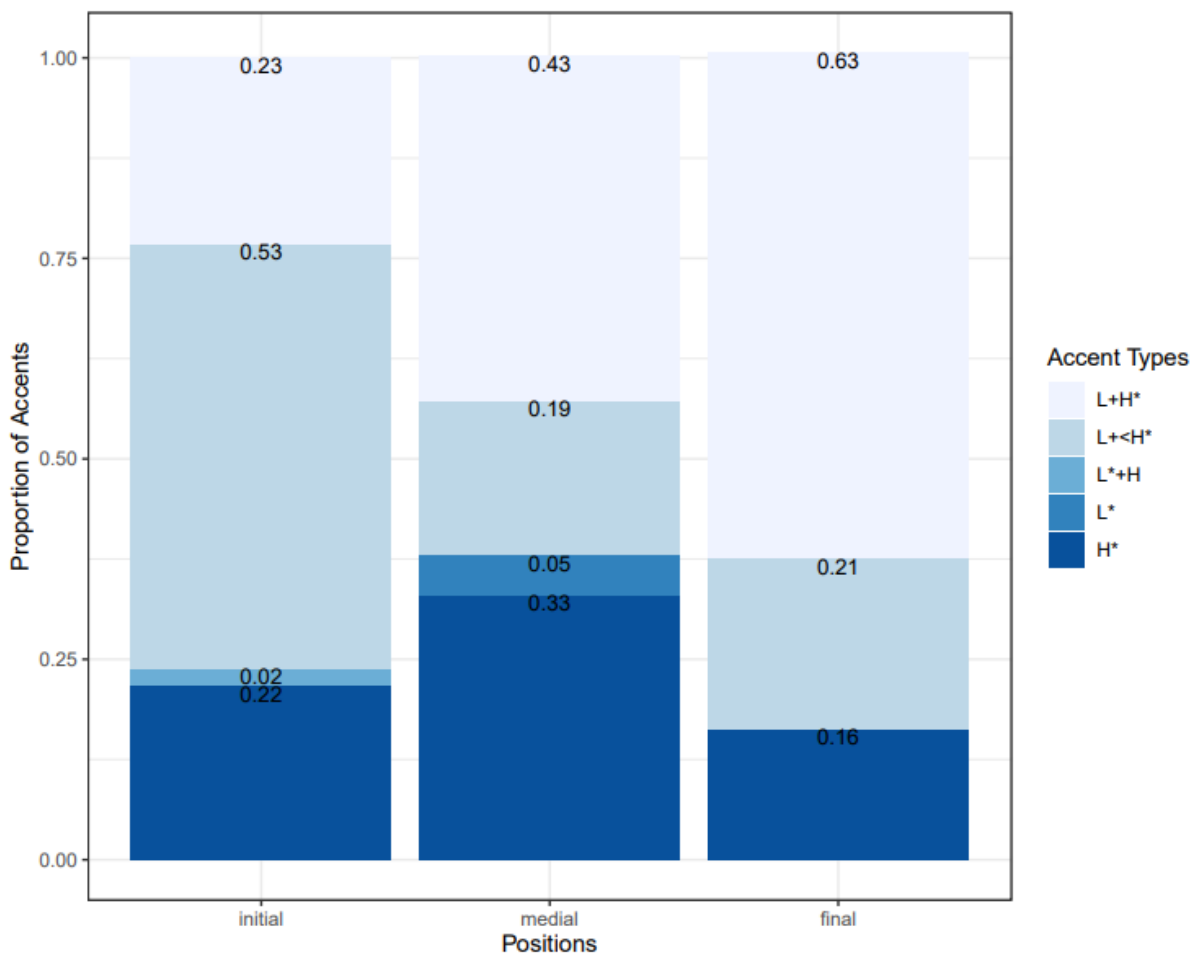

Figure 4. 16 The proportion of accents to positions as produced by native Thai speakers

4.3.2 Use of other prosodic strategies

There were several accompanying strategies used by Thai learners of English when marking narrow focus in sentences. These included deaccenting, using a compressed pitch range, and prosodic phrasing. Similar to native English speakers, these strategies were found in different positions of the sentence across sentences and participants. The use of prosodic phrasing was relatively more common compared to other strategies. The results regarding this aspect are exemplified and reported separately in relation to positions in the sentence as follows.

1) Sentence-initial position

Thai learners of English in this study tended to mark focus sentence-initially by means of several strategies. They might produce focused words with a nuclear focal accent type, followed by deaccenting. For example, there was a rising (L+H*) accent followed by a compressed pitch range as shown in Figure 4.17. In addition to this, they employed post-focal declination and re-phrasing as shown in Figure 4.18 and 4.19. Some of the Thai participants produced a declination pattern where the peak showed a lowering throughout the utterance.
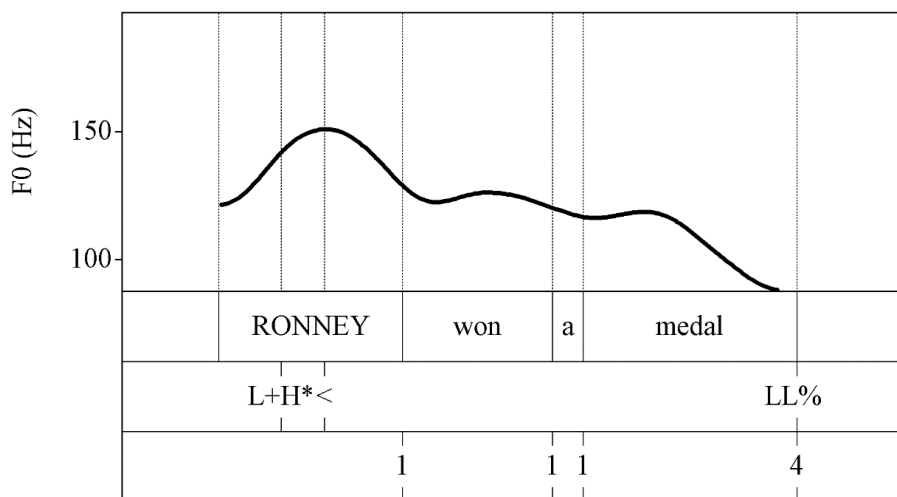
Figure 4. 17 Characteristics of pitch range compression as produced by the male native Thai speaker en2_m03_q11021
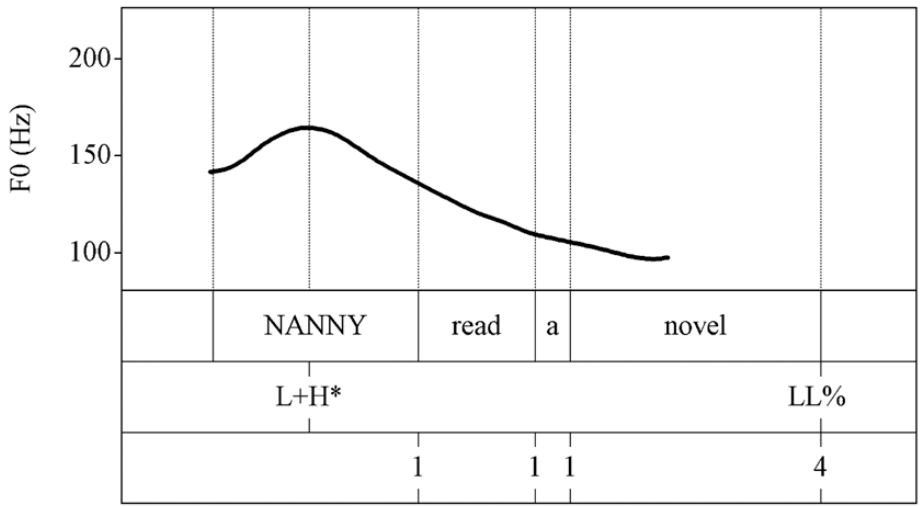
Figure 4. 18 Characteristics of post-focal declination as produced by the male native Thai speaker en2_m09_q11013
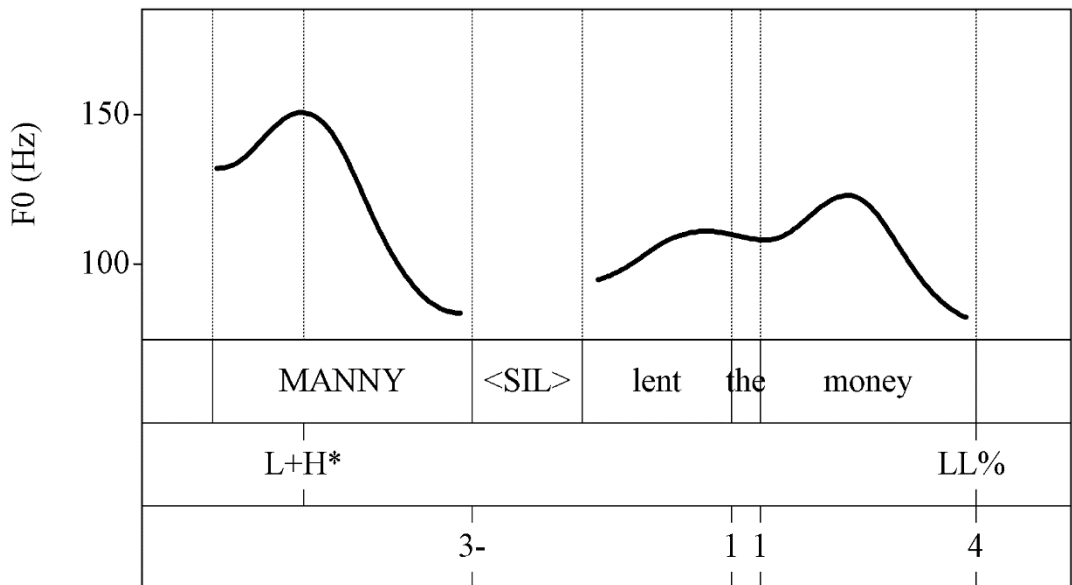
Figure 4. 19 Characteristics of prosodic phrasing or re-phrasing as produced by the male native Thai speaker en2_m03_q11061

As shown in Figure 4.19, the sentence "MANNY lent the money." was produced as two separate phrases. The first is the narrow-focused noun "MANNY", which was produced as a separate phrase (ip) and was realised to have an intermediate contour (L+H* L-), followed by a pause. The second is the phrase "lent the money" which was also produced as a separate phrase (IP) and was realised to have an intonational contour (H* H* LL% or H* L+H* LL%). A full intonation contour of this sentence would be L+H* L- H* H* LL% or L+H* L- H* L+H* LL%. As for re-phrasing, the f0 height on the narrow-focused word was often realised in a lower pitch range compared to the f0 on other accented words of that utterance. Thus, the speakers did not always produce higher peaks when the focused word was produced in a separate phrase.

2) *Sentence-medial position*

To mark focus sentence-medially, Thai learners of English in this study used different strategies. They could produce the focused words with a nuclear focal accent, preceded by deaccenting or a pre-nuclear accent. For example, there was a pre-nuclear H* accent before a rising (L+H*) nuclear focal accent which was probably followed by a lowered pitch range (or in some cases, post-focal deaccenting). In the instances when post-focal material was not deaccented and the utterance was produced as one phrase, the words after the narrow focused syllable were realised in a compressed pitch range. This is the case shown in Figure 4.20.
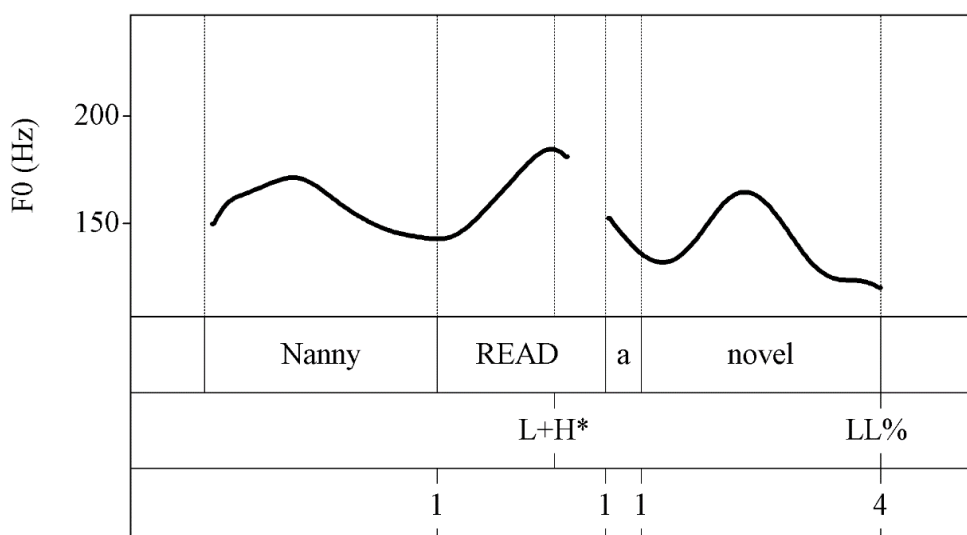
Figure 4. 20 Rising accent (L+H*) with a pre-nuclear accent alongside a post-focal compressed pitch range as produced by the male native Thai speaker en2_m08_q33011

In addition to this, they could employ a compressed pitch range or the downstep-like patterns (L+H* L+!H*) after the focused word as shown in Figure 4.21. Another strategy can be observed in Figure 4.22, which is re-phrasing. In the sentence "Manny LENT the money," the sentence was produced as two separate phrases. The first was "Manny LENT" with a preceding pre-nuclear accent. This intermediate phrase has an intonation contour of L+H* L+H* L- or L+H* L+!H* L-. The second was "the money" produced with an intonation contour of L+H* LL% as another phrase. The full intonation contour of this sentence would be L+H* L+H* L- L+H* LL%, or L+H* L+!H* L- L+H* LL% when taking the downstepping into account.
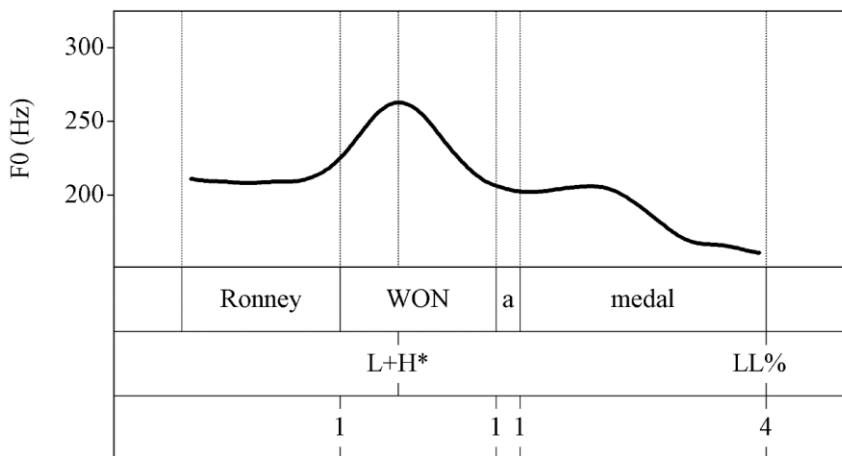
Figure 4. 21 Rising accent (L+H*) alongside a pre-focal deaccenting and post-focal compressed pitch range as produced by the female native Thai speaker en2_f13_q33022
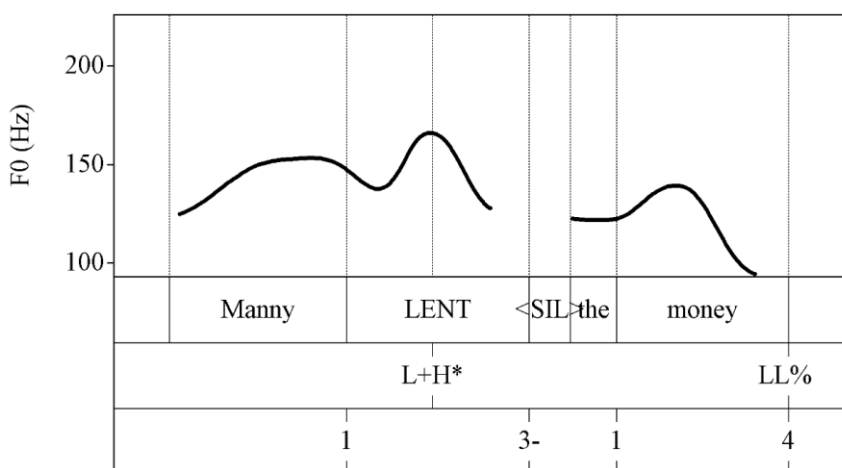
Figure 4. 22 Rising accent (L+H*) with prosodic phrasing or re-phrasing as produced by the male native Thai speaker en2_m05_q33061

124

3) Sentence-final position

In the final position of the sentence, Thai learners of English in this study employed several means to mark focus. First, there was the production of focused words with a nuclear focal accent preceded by pre-focal deaccenting as in Figure 4.23. Second, there might be a pre-nuclear accent before the production of focused words with a nuclear focal accent as in Figure 4.24. In addition, there were other strategies used by Thai learners, such as the prosodic phrasing as shown in Figure 4.25.
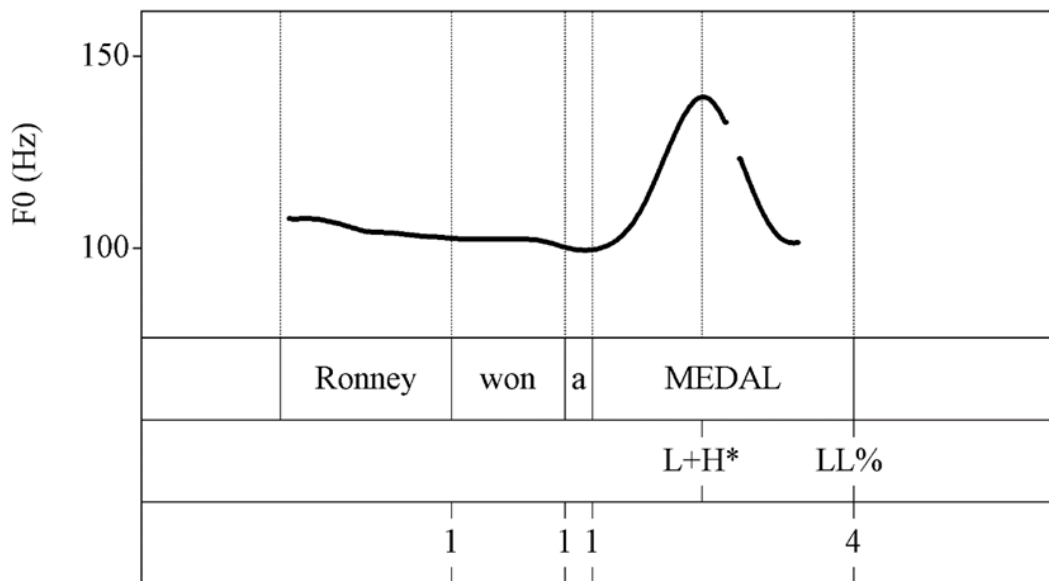
Figure 4. 23 Rising accent (L+H*) with pre-focal deaccenting as produced by the male native Thai speaker en2_m09_q22022
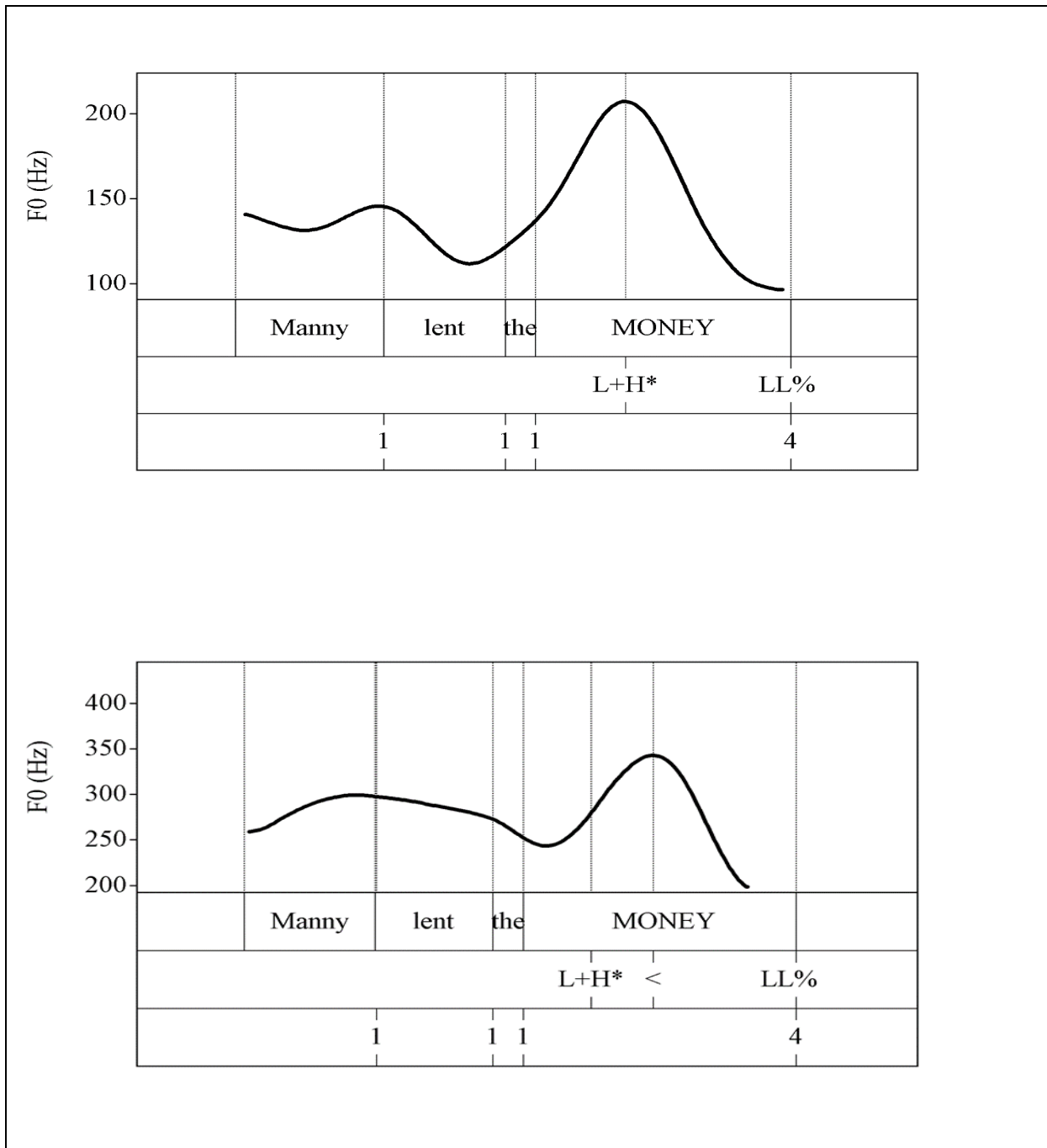
Figure 4. 24 Rising accent (L+H*) and pre-nuclear accenting as produced by the male native Thai speaker en2_m01_q22061 (top panel), and rising accent (L+<H*) with delayed peak alongside pre-nuclear accenting as produced by the female native Thai speaker en2_f13_q22032 (bottom panel)
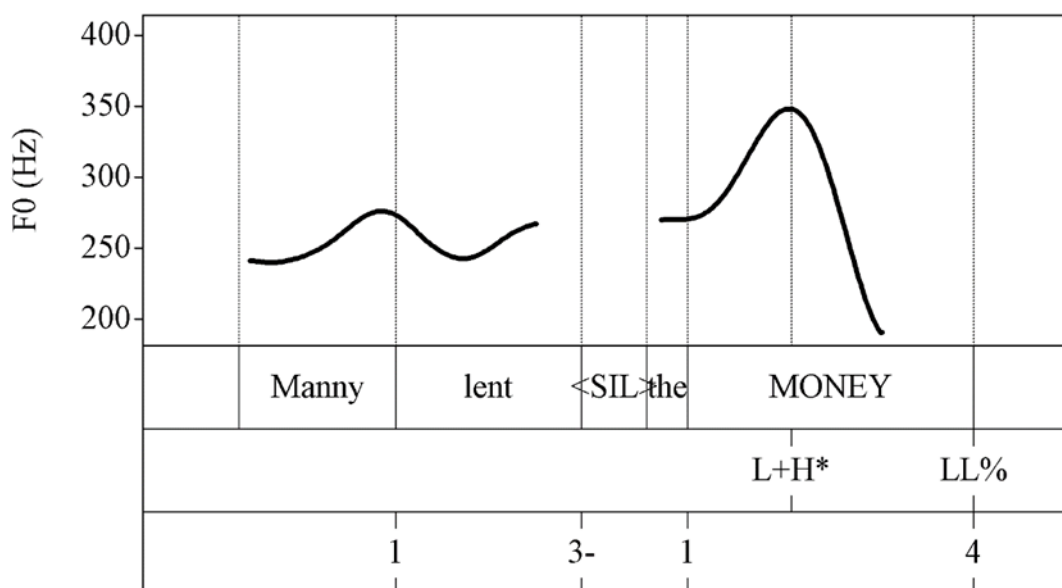
Figure 4. 25 Characteristics of prosodic phrasing or re-phrasing as produced by the female native Thai speaker en2_f09_q22062

4.3.3 Summary of the marking of narrow focus in L2 English by Thai learners

From section 4.3.1, the Thai learners of English predominantly employed rising accents (L+H*) for marking narrow informative focus. They also employed L+<H* accents with delayed peak and H* accents. They were found to employ less of the L*, and L*+H accents. Figure 4.24 below shows the 3-way rising accents as mentioned here: L+H*, L+<H* and L*+H. In addition, all three of these accents, and even H*, were used in combination with other prosodic strategies, as shown in section 4.3.2 according to the difference in focusing location. Thai learners in this study employed predominantly prosodic phrasing or separating a piece of an utterance into two or more separate phrases. Other strategies were found to be used less, for instance, compressed pitch range and pre-nuclear accenting.
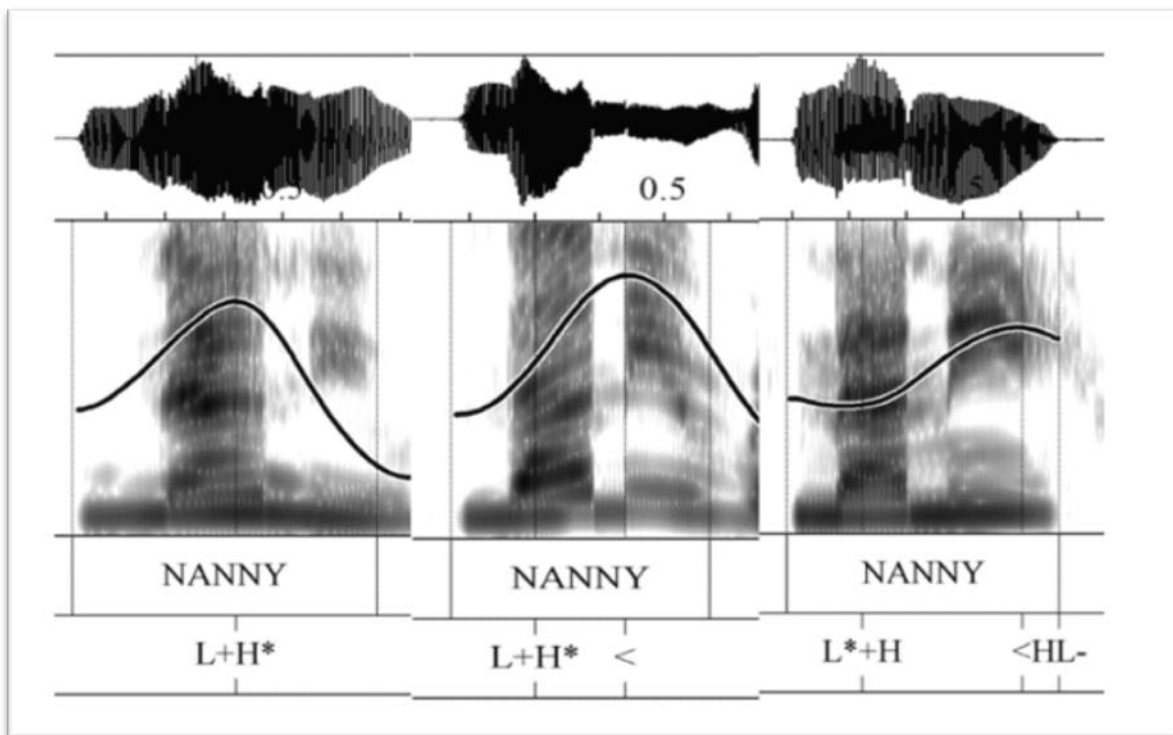
Figure 4. 26 Schematic representations of 3-way rising accents: rising accent with aligned peak (L+H*), rising accent with delayed peak (L+H* < or L+<H*), and scooped accent (L*+H) on focus accented words produced by the Thai learners

4.4 Similarities and Differences between L1 and L2 English

This section presents the findings from the results of the statistical analysis to answer the research question as follows.

Research question 3:   To what extent do Thai learners with different levels of English proficiency differ from native speakers in terms of the use of accent types to mark narrow informative focus?

4.4.1 Use of accent types and effects of English proficiency and focus positions

Table 4.3 shows the descriptive statistics of the overall number of English native speakers or NSE (n = 720) and Thai learners or English L2 Learners or ELL (n = 720) and were separated into three groups according to their levels of English proficiency: Native for native English

speakers (n = 720), and Low for Thai learners with low English proficiency (n = 360) and High for Thai learners with high English proficiency (n = 360). In addition, the data in the table is made up of the distribution of accents that were produced by all the subjects. The accents included 5 types of accents: H*, L*, L*+H, L*+H, and L+H* for each group of different English proficiency levels. The number of response accents were counted and displayed in the columns related to the positions in the sentences that the accents occurred, for example, sentence-initial, sentence-medial, and sentence-final. The number and accent types varied a great deal across subject groups, levels of English proficiency, and sentence positions. In total, the number of accents were 720 items used by native speakers of English, 360 items used by Thai learners with low English proficiency and 360 items used by Thai learners with high English proficiency. Interestingly, the data in the table also shows that the native English speakers in this study did not use L* accent to mark narrow focus, whereas Thai learners with low English proficiency never used scooped accents or L*+H at all. The proportion of accent types across the English proficiency levels and the three different positions are also visualised in Figure 4.27.

Table 4. 3 Descriptive statistics for the frequency of pitch accents by subjects by English proficiency levels and by positions

Subjects	English Proficiency Levels	Pitch Accents	Focus Positions			Total
			Initial	Medial	Final	
Native English speakers (NSE, n = 720)	Native (n = 720)	H*	58 (0.12)	31 (0.06)	8 (0.02)	97
		L*	0 (0.00)	0 (0.00)	0 (0.00)	0
		L*+H	8 (0.02)	0 (0.00)	0 (0.00)	8
		L+<H*	52 (0.11)	45 (0.09)	22 (0.05)	119
		L+H*	122 (0.25)	164 (0.34)	210 (0.44)	496
English L2 learner (ELL) or Thai learners (n = 720)	Low (n = 360)	H*	22 (0.05)	47 (0.10)	9 (0.02)	78
		L*	0 (0.00)	1 (0.002)	0 (0.00)	1
		L*+H	0 (0.00)	0 (0.00)	0 (0.00)	0
		L+<H*	66 (0.14)	15 (0.03)	25 (0.05)	106
		L+H*	32 (0.07)	57 (0.12)	86 (0.18)	175

Table 4.3 continued

Subjects	English Proficiency Levels	Pitch Accents	Focus Positions			Total
			Initial	**Medial**	**Final**	
English L2 learner (ELL) or Thai learners (n = 720)	High (n = 360)	H*	30 (0.06)	32 (0.07)	30 (0.06)	92
		L*	0 (0.00)	11 (0.02)	0 (0.00)	11
		L*+H	5 (0.01)	1 (0.002)	0 (0.00)	6
		L+<H*	61 (0.13)	31 (0.06)	26 (0.05)	118
		L+H*	24 (0.05)	45 (0.09)	64 (0.13)	133
		Total	480	480	480	1440

*The digits in () represent a proportion of pitch accents occurring at the same position.
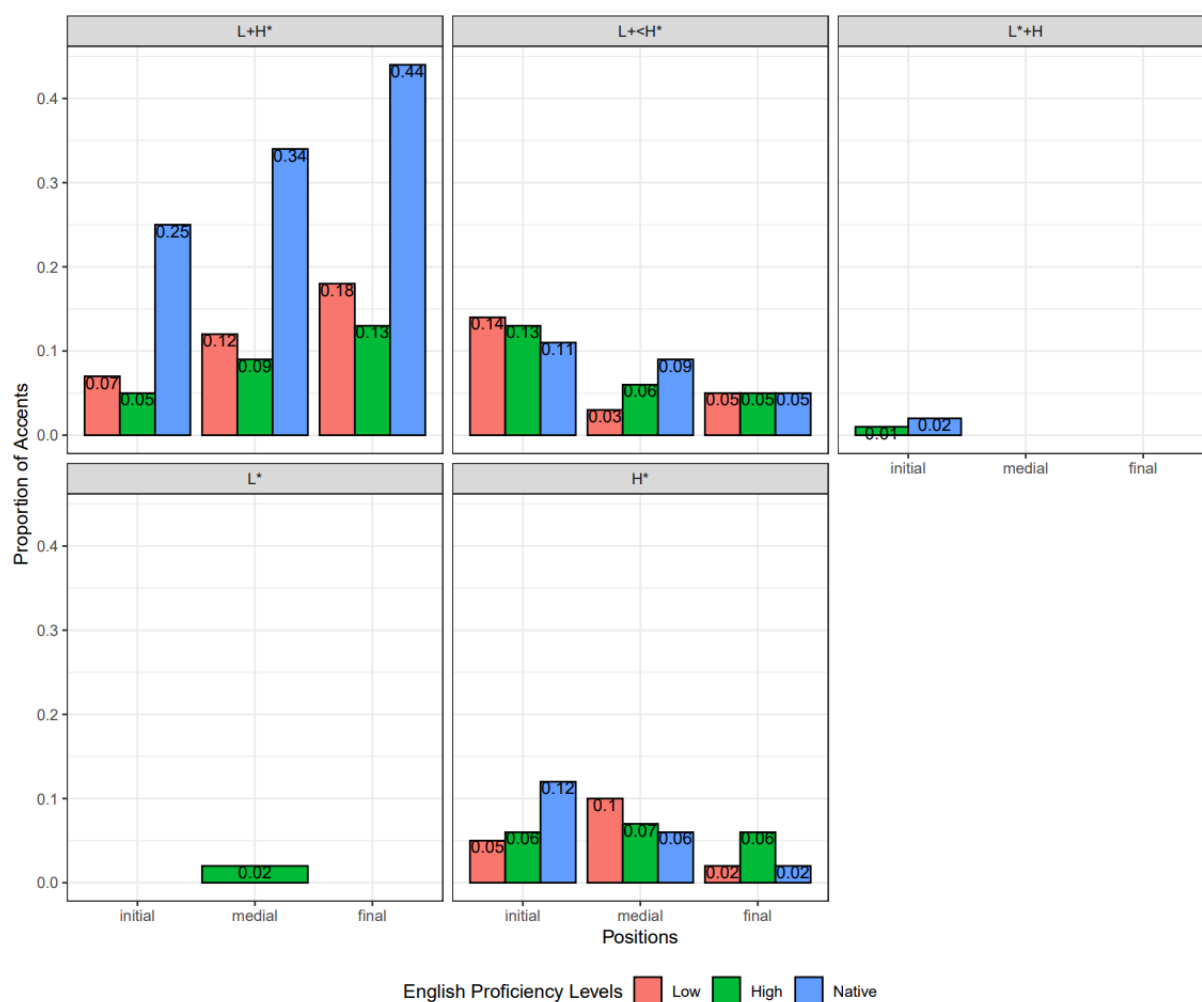
Figure 4. 27 The bar plots represent the proportion of accent types across the English proficiency levels and the three different positions.

The main statistical manipulations were performed by a series of GLMM analyses with the help of the 'lme4' package in R. In GLMM, Speaker or Participant and Word were treated as a random factor, while English Proficiency Level (EPL: low, high and native), and Position (sentence-initial, sentence-medial and sentence-final) were included as fixed factors. When the interaction effect is statistically significant, all group means from the interaction are compared by conducting pairwise comparisons. Focusing on the groups in the interaction better describes the results of the analysis. For the pitch accents in accented and focal syllables as in Table 4.4, the final optimal model with the smaller AIC value was chosen and displayed:

model_PAint <-   glmer(PitchAccent ~ EPL * Position + (1|Subject) +
                      (1|Word), family = binomial, data = mydata1,
                      control = glmerControl(optimizer = "bobyqa",
                      optCtrl = list(maxfun=2e5)))

model_PA   <-   glmer(PitchAccent ~ EPL + (1|Subject) + (1|Word),
                      family = binomial, data = mydata1)

nullmodel_PA <-   glmer(PitchAccent ~ 1 + (1|Subject) + (1|Word),
                      family = binomial, data = mydata1)

Table 4. 4 GLMM results for the pitch accents in target accented syllables produced by Thai learners and native speakers of English

Random Effects	
s^2	3.29
$\tau_{00 \ Subject}$	1.24
$\tau_{00 \ Word}$	1.85
ICC	0.48
$n_{\ Subject}$	40
$n_{\ Word}$	12
Observations	1440
Marginal R^2 / Conditional R^2	0.130 / 0.551
AIC	953.968
log-Likelihood	-465.984

Table 4.2 continued.

	Pitch Accent					
Predictors	*Odds Ratios*	*std. Error*	*Log-Odds*	*std. Error*	*Statistic*	*p*
(Intercept)	16.12	13.71	2.78	0.85	3.27	**<0.001**
EPL[High]	0.45	0.31	-0.80	0.68	-1.17	0.240
EPL[Native]	0.47	0.28	-0.74	0.59	-1.25	0.210
Position[medial]	0.11	0.12	-2.21	1.05	-2.10	**<0.036**
Position[final]	1.25	1.36	0.22	1.09	0.20	0.840
EPL[High]*Position[medial]	5.22	2.96	1.65	0.57	2.91	**<0.004**
EPL[Native]*Position[medial]	16.72	8.73	2.82	0.52	5.39	**<0.001**
EPL[High]*Position[final]	0.45	0.28	-0.80	0.63	-1.26	0.207
EPL[Native]*Position[final]	5.69	3.73	1.74	0.66	2.65	**<0.008**
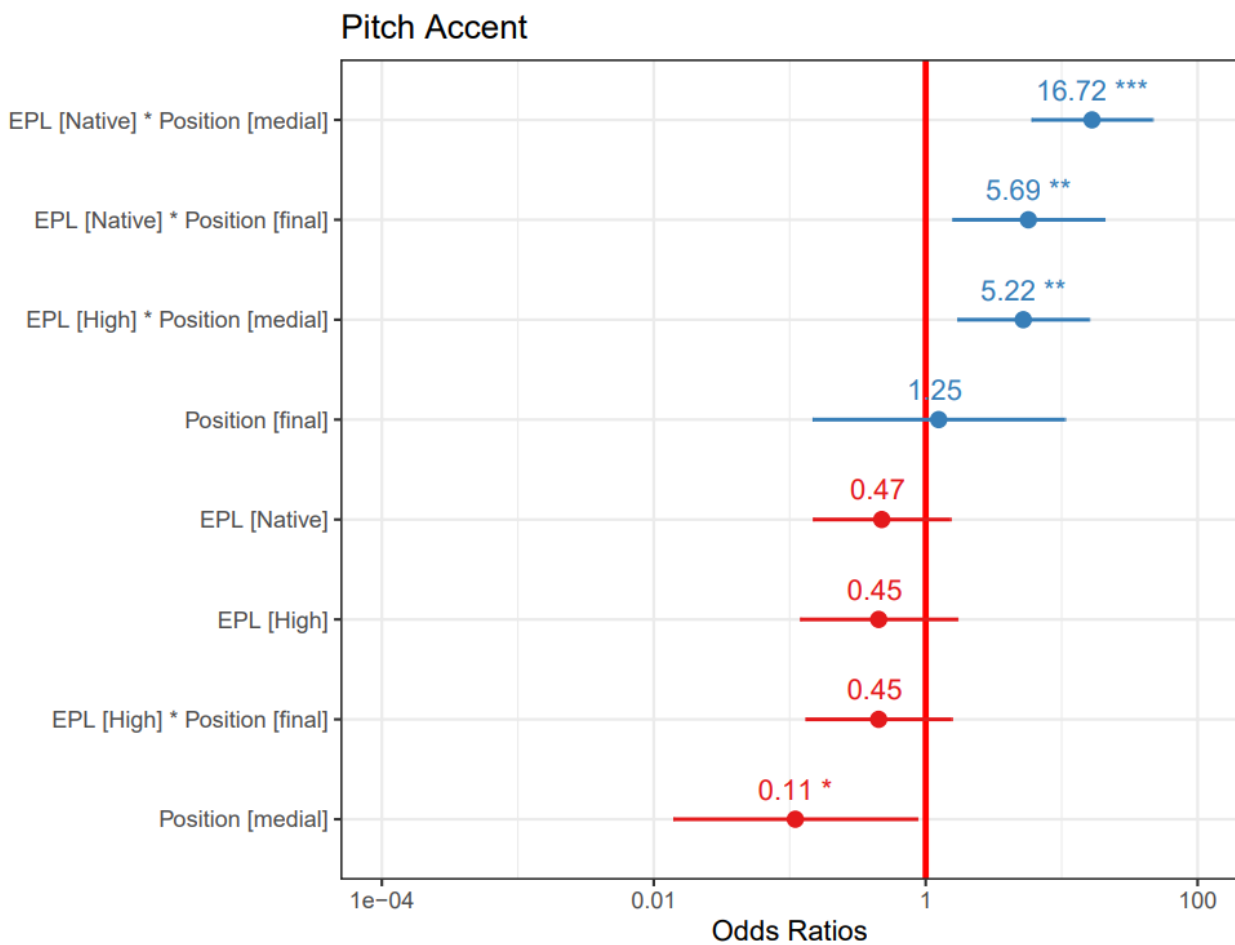
Figure 4. 28 Plot of the estimates of model_PAint from highest to lowest values and the red vertical line indicating no effect

Table 4.4 shows the statistical results of analysing the model_Paint without gender since the result of comparing models showed that there was no significant improvement when the variable "Gender" was added. From Table 4.4, it reveals that there were significant main effects of positions on the distribution of pitch accents used by the participants. Differences were among Position[medial] versus Position[initial] [t = -2.10, p < 0.05]. There were also interaction effects of English proficiency and focus positions.

A significant difference was reported for an interaction of EPL[High] and Position[medial] [t = 2.91, p < 0.05], an interaction of EPL[Native] and Position[medial] [t = 5.39, p < 0.05], and an interaction of EPL[Native] and Position[final] [t = 2.65, p < 0.05]. From these results, there was a significant positive relationship or effect between focusing positions and English proficiency levels, reflecting variations in the use of accent types across the subjects. Rather, it seemed the interaction was significant, and it significantly improved the model. Note that the model can account for English proficiency levels and positions interacting: EPL may not be significant on its own, but may have a significant interaction with Position. Predicted probabilities of pitch accents can be tracked from Figure 4.28 and 4.29.

As mentioned earlier, if main effects, or 2-way interaction effects or 3-way interaction effects are statistically significant, the mean-separation tests for either a 2-way interaction or a 3-way interaction will be conducted only and the mean-separation tests for the main effects will be of less interest. To examine differences among groups, all pairwise comparisons were done only on the optimal model. The results containing the means and SEs of English proficiency or EPL (low, high and native) and positions (initial, medial and final) for the use of accent types are tabulated in Table 4.5. The formula for 'emmeans' is displayed below:

```
model_PA_emmean   <-   emmeans(model_PA, specs = pairwise ~ EPL * Position,
                              adjust = "tukey")
```

Table 4. 5 Means and SEs of EPL for the pitch accents from pairwise comparison based on GLMM analysis and emmeans

Subject groups	English Proficiency	Focus position	Mean	SE	asymp.LCL	asymp.UCL
Native English speakers (NSE)	Native		2.89	0.50	1.91	3.87
		initial	2.04	0.76	0.54	3.53
		medial	2.64	0.76	1.14	4.14
		final	3.99	0.82	2.39	5.60
English L2 learner (ELL) or Thai learners	Low		2.12	0.56	1.01	3.22
		initial	2.78	0.85	1.11	4.45
		medial	0.57	0.80	-0.99	2.14
		final	3.00	0.85	1.33	4.67
	High		1.61	0.55	0.52	2.69
		initial	1.98	0.83	0.36	3.61
		medial	1.43	0.81	-0.15	3.01
		final	1.41	0.80	-0.17	2.98

Figure 4.29 below shows differences of the use of accent types due to levels of English proficiency by position. Keep in mind that even though the main effects of position in the final optimal model were statistically significant, they might be significant because of their interaction with other independent variables. In this case, position had an interaction with levels of English proficiency (EPL) influencing differences in using accent types. The results can be better visualised and more informative in Figure 4.29, which shows the effects of independent variables (EPL and Position) and their interactions. The plot presents their interactions, wherein the y-axis represents the probabilities of distribution of pitch accents. The x-axis represents different prosodic positions (initial, medial and final). The plot is also separated by levels of English proficiency (Low, High and Native). The dot inside a line corresponds to the mean values obtained from pairwise comparison based on GLMM analysis and emmeans as in Table 4.5.
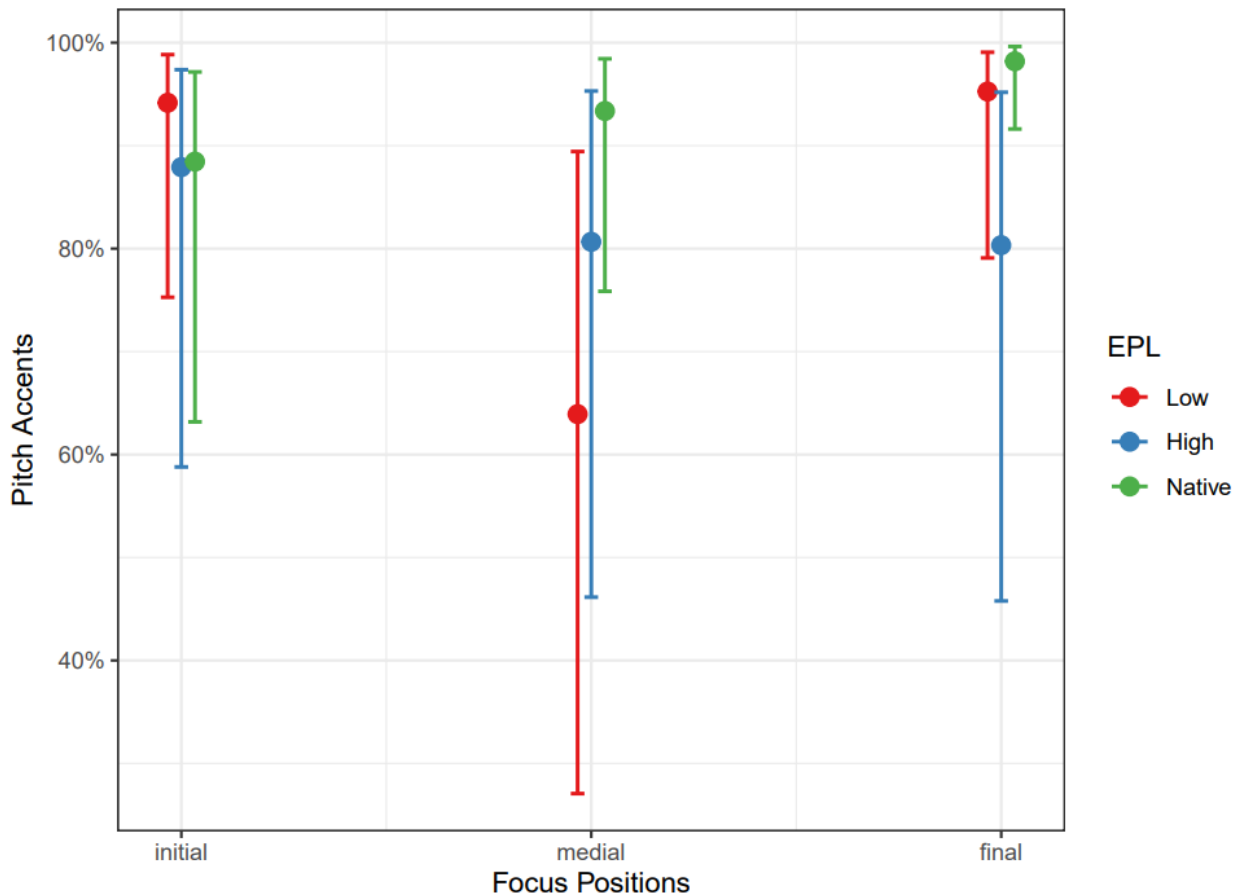
Figure 4. 29 Plot for interaction effects of positions (initial, medial, final) and English proficiency levels (Low, High and Native) on the use of pitch accents

According to Tukey's HSD post-hoc test, the statistical differences were only found in a few pairs and were reported as follows. Thai learners with low English proficiency significantly differed from native English speakers in using accent types at medial position [t = -3.817, df = Inf, p < 0.05]. Further, Thai learners with high English proficiency significantly differed from native English speakers in using accent types at final position [t = -4.168, df = Inf, p < 0.05] as shown in Figure 4.29. Statistical testing of mean separation had found no significant differences in other cases.

4.5 Summary of the Results from Statistical Analyses

The following sections present the results of the qualitative analysis for marking narrow focus or focal prominence by using accent types. The statistical analysis, including English Proficiency levels (Low, High and Native) and three focus positions of the sentence (initial,

medial and final), was also conducted when investigating similarities and differences between two L1 groups.

Native speakers of English in this study marked narrow focus at different positions in a sentence with different accent types. So did Thai learners of English. Regardless of first language group and different positions in the sentence, the common set of accents for narrow focus as produced by both groups consisted of the H* accent, the scooped accent (L*+H), the rising accent (L+H*), and the rising accent (L+<H*) with a delayed peak. The accent choices that native speakers of English and the Thai learners employed to mark narrow focus are summarised in the table below.

Table 4. 6 The summary of accent types used by native speakers of English and the Thai learners

Positions	English Native Speakers (NSE)	Thai learner of English (ELL)		Remark
		Low Proficiency	High Proficiency	
Initial	L+H* L+<H* H* L*+H	L+<H* L+H* H*	L+<H* H* L+H* L*+H	Accent types were arranged according to how often they occurred in the data.
Medial	L+H* L+<H* H*	L+H* H* L+<H* L*	L+H* H* L+<H* L* L*+H	
Final	L+H* L+<H* H*	L+H* L+<H* H*	L+H* H* L+<H*	

There are, however, differences between the two groups to some extent in terms of accent types and placement used. First, at the sentence-initial position, native speakers of English and Thai learners of English employed the same set of accent types to mark focus. These accents were L+H*, H*, L+<H* and L*+H, but they differed in how often they occurred. At the sentence-medial position, there were only three accent types (L+H*, L+<H* and H*) that were used predominantly by native speakers of English. As for Thai leaners, they still used L+H*, H*, L+<H*, L*+H and with one additional accent; that is, the L* accent. At the sentence-final position, native speakers of English still employed L+H*, L+<H* and H*. Similarly, these accents were produced in the Thai learner data, but they differed in how frequent they occurred.

Overall, when taking L2 proficiency levels of Thai learners into consideration, it revealed that English proficiency levels reflected variations in the choice of accent types to use for marking narrow focus more or less. The inventory of accent types used by Thai learners with high English proficiency might be more varied than Thai learners with low English proficiency, especially in the beginning and middle of the sentences as in Table 4.6. This variation also depended on positions in the sentences where the focused accented words were. Interestingly, most of accent types used by Thai learners with high English proficiency and Thai learners with low English proficiency were the ones used by native speakers of English.

4.6 Summary of the Chapter

This chapter has presented an investigation into the prosodic marking of focus in English. In particular, the expression of narrow informative focus in terms of using accent types produced by native speakers of English and Thai learners of English was closely examined. In this study, native speakers of English used L+H* accents predominantly, while H* accents and L+<H* accents were used to some degree. Thai learners of English used L+H*, L+<H*, and H* accents predominantly. There were variations in the choices and use of accents according to the positions in the sentences and how often they occurred.

The overall results show that, in terms of the production of accent types, rising accents were preferred and used for the marking of focus on information by both native English speakers and Thai learners. In other words, the most frequent accent associated with focused and accented words was a rising accent such as L+H*, L*+H, or L+<H*, in which the pitch shape has the characteristics of a rising fundamental frequency. Prosodic strategies for marking focus such as deaccenting and prosodic phrasing were also found to be used to accompany accent placement. From the results in this chapter, the rising accents, which were used most often by the participants were examined in terms of whether or not the acoustic-phonetic characteristics of the three rising accents differed between native speakers and Thai learners due to English proficiency, gender and focus position. Thus, another level of analysis is applied to examine the acoustic realisation of focus as reported in Chapter 5, and tonal alignment and scaling in Chapter 6.

Chapter 5. Acoustic Realisation of Focus

5.1 Introduction

This chapter describes the investigations into the acoustic cues in relation to marking the focus of information, considering the effects of factors such as English proficiency, gender and focus position. It starts with the results for fundamental frequency in section 5.2. The results for intensity are then reported in section 5.3, followed by in section 5.4 the results for duration. A summary of the statistical analyses and of the chapter are provided in section 5.5 and 5.6.

Since Thai learners use such accents as H*, L*+H, L+H*, and L+<H* in the same way as native English speakers do to mark focus (Chapter 4), it is of particular interest to determine the extent to which factors such as English proficiency, gender, as well as focus position affect the use of aspects of the acoustic realisation of focus. The research question that this section addresses is:

Research question 5: To what extent do factors such as levels of English proficiency, gender and focus position affect the acoustic characteristics of focus used by Thai learners and native speakers of English for the marking of narrow informative focus in terms of f0, intensity and duration?

The hypotheses to be tested are given in each of the relevant sections below.

5.2 Fundamental Frequency

To examine the effects of such factors as level of English proficiency, gender and focus position, the hypothesis is that:

Hypothesis 5.1: The values of fundamental frequency produced by Thai learners differ from those of native speakers due to the different levels of English proficiency, gender and focus position.

5.2.1 Fundamental frequency and effects of English proficiency, gender and focus position

Table 5.1 shows the descriptive statistics of fundamental frequency (f0) for English native speakers or NSE (n = 2400) and Thai learners or English L2 Learners or ELL (n = 2400). The mean values of f0 for the formers was 180.13 Hz, SD = 47.16 and those of the latter was 198.70 Hz, SD = 59.61. Thai learners with low proficiency (n = 1200) have the mean f0 values of 209.68 Hz, SD = 62.53, and those with high proficiency (n = 1200) have the mean f0 values of 187.73 Hz, SD = 54.40.

Table 5. 1 Descriptive statistics for mean normalised fundamental frequency in Hertz

Subjects		n	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
NSE		2,400	180.13	47.16	83.48	138.08	217.68	305.11
ELL		2,400	198.70	59.61	86.93	146.50	247.39	373.82
	Low	1,200	209.68	62.53	86.93	156.54	257.50	373.82
	High	1,200	187.73	54.40	97.56	140.10	226.86	341.44

The main statistical manipulations were performed by a series of LMM analyses with the help of the 'lme4' package in R. In LMM, Speaker or Participant and Word were treated as a random factor, while English Proficiency Level (EPL: low, high and native), Position (sentence-initial, sentence-medial and sentence-final), and Gender (female and male) were included as fixed factors. When the interaction effect is statistically significant, all group means from the interaction are compared by conducting pairwise comparisons, especially for 3-way interaction effect only. That is because focusing on the groups in the interaction better describes the results of the analysis. For the mean normalised f0 in accented and focal syllables as in Table 5.2, the final full model with the smaller AIC value was chosen and displayed here after conducting several different model analyses and comparison. In this study, many maximal random effects models (e.g., both random intercept and slope) failed to converge or cannot find a solution. For example:

model4.1   <-   lmer(MeanNormF0 ~ EPL*Gender+Position+(1+EPL+Position

|Subject)+ (1 + EPL + Position|Word), data = mydata1, REML=FALSE)

model5.1   <-   lmer(MeanNormF0~EPL+Gender+Position+

EPL*Gender+(1+EPL+Position|Subject)+

(1+EPL+Position|Word),

data = mydata1, REML=FALSE)

model9   <-   lmer(MeanNormF0~ EPL+Gender+Position+

EPL*Gender*Position+EPL*Position+Gender*Position+

(1 + EPL + Position |Subject)+ (1 + EPL + Position|Word),

data = mydata1, REML=FALSE)

In this case, the researcher had to use a simpler model (e.g., intercepts-only). One common means to test the model's fit is to rerun the analysis but include only the intercept terms which is often called the null model.

model_f0   <-   lmer(MeanNormF0 ~ EPL + Gender + Position +

EPL * Gender * Position + EPL * Gender +

EPL * Position + Gender * Position +

(1|Subject) + (1|Word), data = mydata1, REML = FALSE)

nullmodel_f0 <-   lmer(MeanNormF0 ~ 1+(1|Subject)+ (1|Word), data = mydata1,

REML=FALSE)

Table 5. 2 LMM results for the mean normalised f0 in target accented syllables

Random Effects	
s^2	344.59
$\tau_{00 \text{ Subject}}$	406.07
$\tau_{00 \text{ Word}}$	51.36
ICC	0.57
n_{Subject}	40
n_{Word}	12
Observations	4800
Marginal R^2 / Conditional R^2	0.731 / 0.884
AIC	41950.989
BIC	42087.0
log-Likelihood	-20954.495

Table 5.2 continued

Predictors	Estimates	std. Error	Statistic	p	df
(Intercept)	281.72	9.05	31.12	**<0.001**	4779.00
EPL[High]	-23.19	13.14	-1.76	0.078	4779.00
EPL[Native]	-51.58	10.52	-4.91	**<0.001**	4779.00
Gender[male]	-121.07	13.14	-9.21	**<0.001**	4779.00
Position[medial]	-38.24	5.34	-7.16	**<0.001**	4779.00
Position[final]	-49.44	5.34	-9.25	**<0.001**	4779.00
EPL[High]*Gender[male]	30.53	18.59	1.64	0.101	4779.00
EPL[Native]*Gender[male]	41.00	15.99	2.56	**<0.010**	4779.00
EPL[High]*Position[medial]	11.91	2.68	4.45	**<0.001**	4779.00
EPL[Native]*Position[medial]	24.22	2.14	11.30	**<0.001**	4779.00
EPL[High]*Position[final]	14.83	2.68	5.53	**<0.001**	4779.00
EPL[Native]*Position[final]	31.55	2.14	14.72	**<0.001**	4779.00
Gender[male]*Position[medial]	12.38	2.68	4.62	**<0.001**	4779.00
Gender[male]*Position[final]	29.73	2.68	11.10	**<0.001**	4779.00
(EPL[High]*Gender[male])* Position[medial]	-2.18	3.79	-0.57	0.565	4779.00
(EPL[Native]*Gender[male])* Position[medial]	-13.86	3.26	-4.25	**<0.001**	4779.00
(EPL[High]*Gender[male])* Position[final]	-20.76	3.79	-5.48	**<0.001**	4779.00
(EPL[Native]*Gender[male])* Position[final]	-24.25	3.26	-7.44	**<0.001**	4779.00

Mean Normalised F0

141

Table 5.2 reveals that there were significant effects of English proficiency, gender and position on the values of f0. Differences were among native English speakers EPL[Native] versus Thai learners with low English proficiency [t = -4.91, df = 4779.00, p < 0.05], reflecting greater f0 values for Thai learners with low English proficiency and lower f0 values in the group of native English speakers. There were also effects of gender and focus positions. Gender[female] differs from Gender[male] statistically [t = -9.21, df = 4779.00, p < 0.05] in terms of higher f0. A significant difference was reported for the f0 height in the medial position versus initial position [t = -7.16, df = 4779.00, p < 0.05] and the final position versus initial position [t = -9.25, df = 4779.00, p < 0.05], indicating the lower f0 for the medial and final positions of the sentences.

There were two-way interactions between EPL[Native] and Gender[male] [t = 2.56, df = 4779.00, p < 0.05], an interaction between EPL[High] and Position[medial] [t = 4.45, df = 4779.00, p < 0.05], an interaction between EPL[Native] and Position[medial] [t = 11.30, df = 4779.00, p < 0.05], an interaction between EPL[High] and Position[final] [t = 5.53, df = 4779.00, p < 0.05], an interaction between EPL[Native] and Position[final] [t = 14.72, df = 4779.00, p < 0.05], an interaction between Gender[male] and Position[medial] [t = 4.62, df = 4779.00, p < 0.05], and an interaction between Gender[male] and Position[final] [t = 11.10, df = 4779.00, p < 0.05].

In addition, there were three-way interactions of EPL[Native], Gender[male] and Position[medial] [t = -4.25, df = 4779.00, p < 0.05]. There was an interaction effect of EPL[High], Gender[male] and Position[final] [t = -5.48, df = 4779.00, p < 0.05]. Significant difference was found due to an interaction of EPL[Native], Gender[male]) and Position[final] [t = -7.44, df = 4779.00, p < 0.05], suggesting that main independent variables (English proficiency levels, gender and positions) might not have an effect on dependent variable fundamental frequency (f0) by itself but their effect might depend on another independent variable. As in Figure 5.1, it shows the estimates of model_f0 from highest to lowest values and the red vertical line indicating no effect.
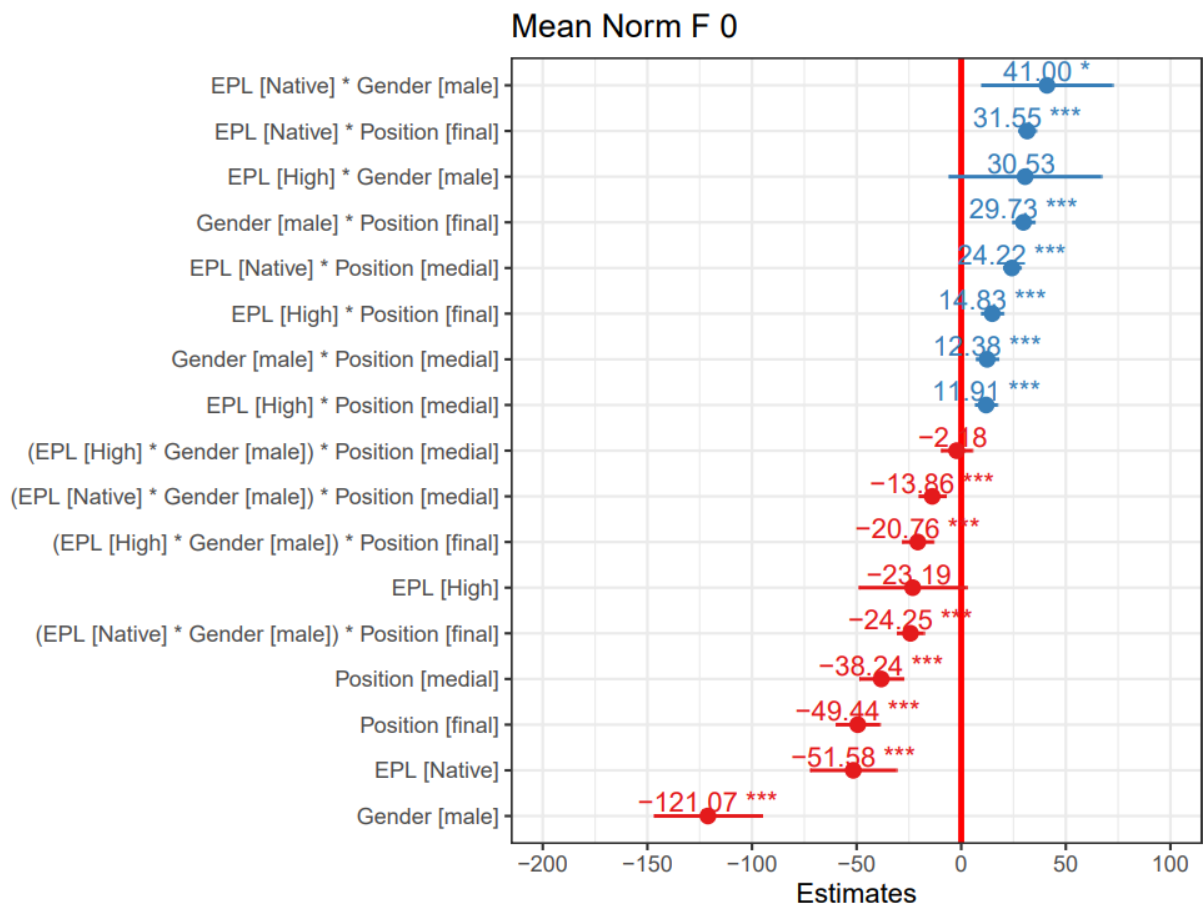
Figure 5. 1 Plot of the estimates of model_f0 from highest to lowest values and the red vertical line indicating no effect

As mentioned earlier, if main effects, or 2-way interaction effects or 3-way interaction effects are statistically significant, the mean-separation tests for a 3-way interaction will be conducted only and the mean-separation tests for the main effects or 2-way interaction effects will be of less interest. This is because focusing on the groups in the interaction better describe the results of the analysis. To examine differences among groups, all pairwise comparisons were done only on the optimal model. The results containing the means and SEs of English proficiency or EPL (low, high and native), gender (female and male) and positions (initial, medial and final) for the f0 are tabulated in Table 5.3. The formula for 'emmeans' is displayed below:

emmf0   <-   emmeans(model_f0, specs = pairwise ~ EPL : Gender :
                   Position, adjust = "tukey")

143

Table 5. 3 Means and SEs of EPL x Gender x Position for the f0 from pairwise comparison based on LMM

Subject groups	English Proficiency	Gender	Focus position	Mean	SE	df
Native English speakers (NSE)	Native			180.13	5.41	59.69
		female		219.50	7.29	55.32
			initial	230.14	8.03	61.29
			medial	216.11	8.03	61.29
			final	212.25	8.03	61.29
		male		140.76	7.29	55.32
			initial	150.06	8.03	61.29
			medial	134.55	8.03	61.29
			final	137.66	8.03	61.29
English L2 learner (ELL) or Thai learners	Low			198.98	7.43	55.04
		female		252.49	9.22	52.23
			initial	281.72	9.84	60.13
			medial	243.48	9.84	60.13
			final	232.28	9.84	60.13
		male		145.46	11.17	50.42
			initial	160.65	11.70	57.46
			medial	134.78	11.70	57.46
			final	140.94	11.70	57.46
	High			196.14	7.43	55.04
		female		238.22	11.17	50.42
			initial	258.53	11.70	57.46
			medial	232.20	11.70	57.46
			final	223.92	11.70	57.46
		male		154.06	9.22	52.23
			initial	167.99	9.84	60.13
			medial	151.85	9.84	60.13
			final	142.35	9.84	60.13

Figure 5.2 below shows differences of f0 due to levels of English proficiency by gender. Keep in mind that even though the main effects of English proficiency, gender, and position are statistically significant, they might be significant because of their interaction with other independent variables. In this case, the results can be better visualized and more informative in Figure 5.3, which shows the effects of independent variables and their interactions on the fundamental frequency (f0) for the target accented syllables. The plot presents their interactions, in which the y-axis represents the distribution of the f0 values measured in Hertz (Hz). The x-axis represents different prosodic positions (initial, medial and final). The plot is also separated by gender and the levels of English proficiency (Low, High and Native). The dots inside the boxes correspond to the means, f0.
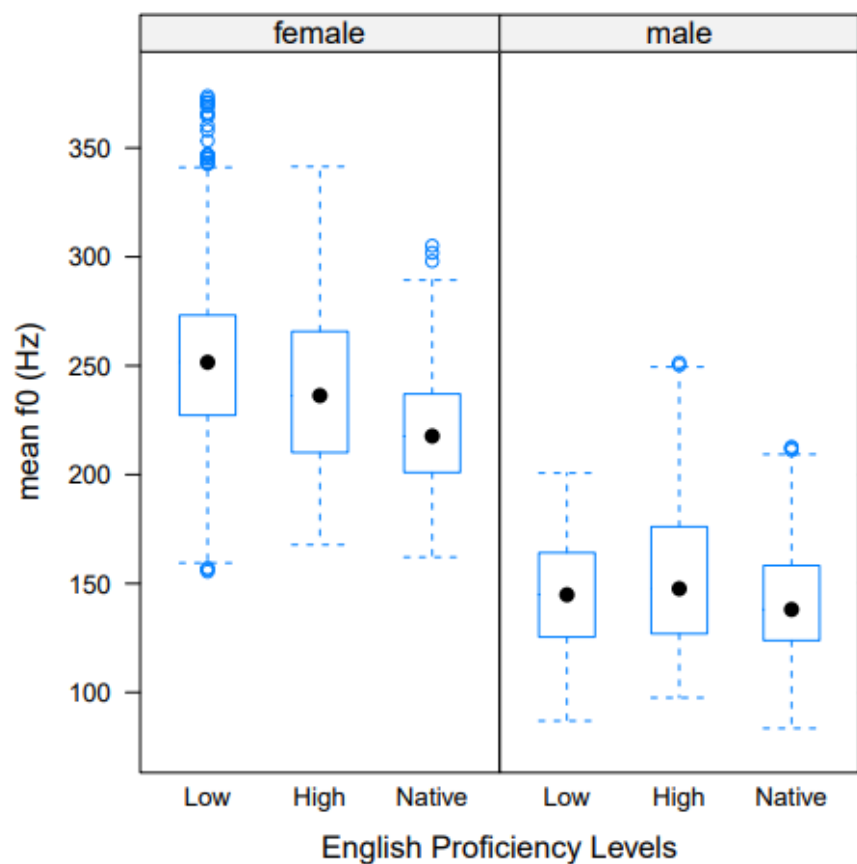
Figure 5. 2 Box plot for mean normalised f0 due to levels of English proficiency (Low and High for Thai learners of English with low and high English proficiency, and Native for native English speakers) presented by gender
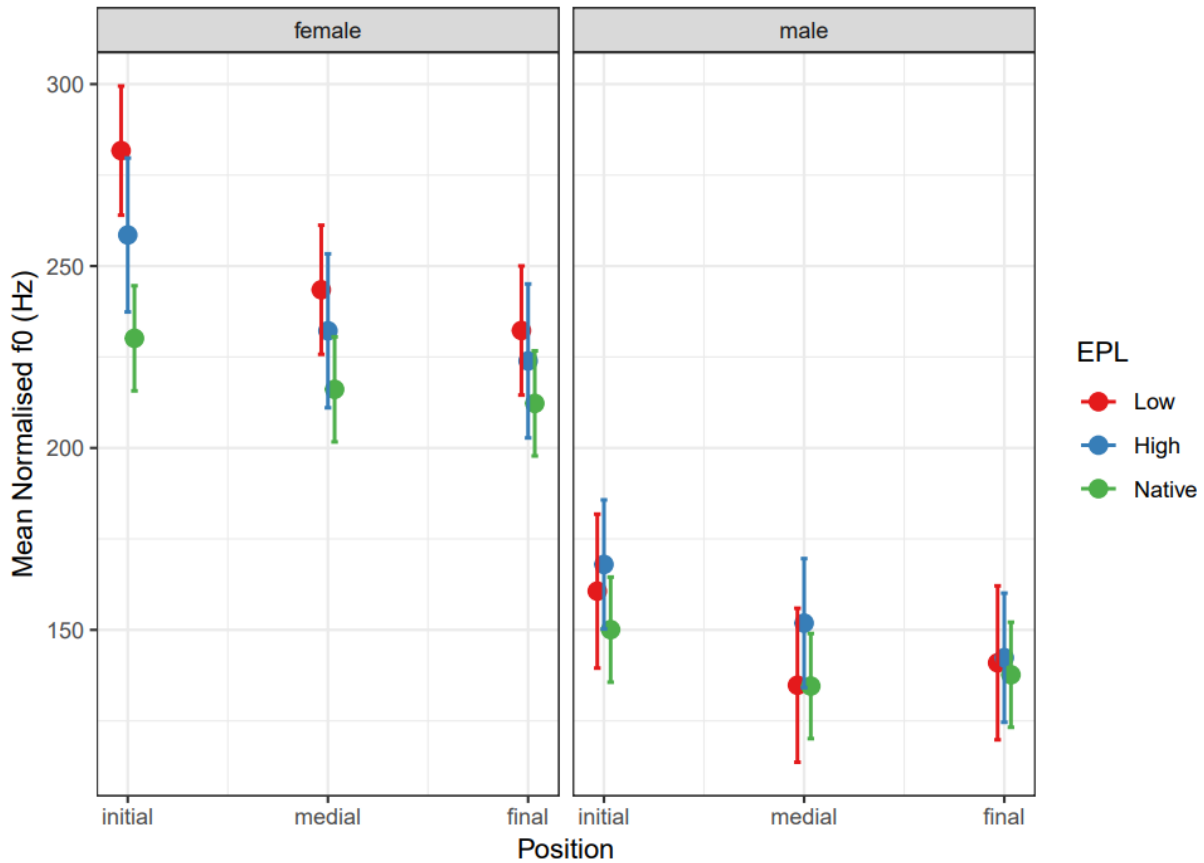
Figure 5. 3 Plot for interaction effects of positions (initial, medial, final) and English proficiency levels (Low, High and Native) grouped by gender for the f0

According to Tukey's HSD post-hoc test, the statistical differences are reported and grouped into sentence-initial, sentence-medial, and sentence-final as follows. The report here takes three-way interactions (Position, Gender, and English proficiency levels) into consideration.

1. Sentence-initial
    1.1 Sentence-initial and Thai females with low English proficiency
        1.1.1 Thai females with low English proficiency had different f0 values at the initial position compared to f0 values at the medial position [t = 6.444, df = 17.0, p < 0.05] and at final position [t = 8.330, df = 17.0, p < 0.05].
        1.1.2 Thai females with low English proficiency had different f0 values at the initial position compared to f0 values at the final position produced by Thai females with high English proficiency [t = 3.781, df = 58.4, p < 0.05].

146

1.1.3 Thai females with low English proficiency had different f0 values at the initial position compared to f0 values at the initial position [t = 4.544, df = 47.7, p < 0.05], medial position [t = 5.167, df = 61.0, p < 0.05], and final position [t = 5.471, df = 61.0, p < 0.05] produced by native English females.

1.1.4 Thai females with low English proficiency had different f0 values at the initial position compared to f0 values at the initial position [t = 8.532, df = 47.7, p < 0.05], medial position [t = 9.611, df = 58.4, p < 0.05], and final position [t = 9.208, df = 58.4, p < 0.05] produced by Thai males with low English proficiency.

1.1.5 Thai females with low English proficiency had different f0 values at the initial position compared to f0 values at the initial position [t = 8.961, df = 47.7, p < 0.05], medial position [t = 9.337, df = 59.9, p < 0.05], and final position [t = 10.020, df = 59.9, p < 0.05] produced by Thai males with high English proficiency.

1.1.6 Thai females with low English proficiency had different f0 values at the initial position compared to f0 values at the initial position [t = 11.597, df = 47.7, p < 0.05], medial position [t = 11.590, df = 61.0, p < 0.05], and final position [t = 11.345, df = 61.0, p < 0.05] produced by native English males.

1.2 Sentence-initial and Thai females with high English proficiency

1.2.1 Thai females with high English proficiency had different f0 values at the initial position compared to f0 values at the medial position [t = 4.349, df = 18.7, p < 0.05], and final position [t = 5.716, df = 18.7, p < 0.05].

1.2.2 Thai females with high English proficiency had different f0 values at the initial position compared to f0 values at the initial position [t = 6.297, df = 47.7, p < 0.05], medial position [t = 7.476, df = 57.2, p < 0.05], and final position [t = 7.104, df = 57.2, p < 0.05] produced by Thai males with low English proficiency.

1.2.3 Thai females with high English proficiency had different f0 values at the initial position compared to f0 values at the initial position [t = 6.381, df = 47.7, p < 0.05], medial position [[t = 6.978, df = 58.4, p < 0.05], and final position [t = 7.600, df = 58.4 , p < 0.05] produced by Thai males with high English proficiency.

1.2.4   Thai females with high English proficiency had different f0 values at the initial position compared to f0 values at the initial position [t = 8.340, df = 47.7, p < 0.05], medial position [[t = 8.734, df = 59.6, p < 0.05], and final position [t = 8.515, df = 59.6, p < 0.05] produced by native English males.

1.3 Sentence-initial and Native English females

1.3.1   Native English females had different f0 values at the initial position compared to f0 values at the initial position [t = 5.343, df = 47.7, p < 0.05], medial position [t = 6.717, df = 59.6, p < 0.05], and final position [t = 6.283, df = 59.6, p < 0.05] produced by Thai males with low English proficiency.

1.3.2   Native English females had different f0 values at the initial position compared to f0 values at the initial position [t = 5.475, df = 47.7, p < 0.05], medial position [t = 6.165, df = 61.0, p < 0.05], and final position [t = 6.914, df = 61.0, p < 0.05] produced by Thai males with high English proficiency.

1.3.3   Native English females had different f0 values at the initial position compared to f0 values at the initial position [t = 8.145, df = 47.7, p < 0.05], medial position [t = 8.415, df = 61.3, p < 0.05], and final position [t = 8.142, df = 61.3, p < 0.05] produced by native English males.

1.4 Sentence-initial and Thai males with low English proficiency

1.4.1   Thai males with low English proficiency had different f0 values at the initial position compared to f0 values at the medial position [t = 4.272, df = 18.7, p < 0.05].

1.5 Sentence-initial and Thai males with high English proficiency

1.5.1   Thai males with high English proficiency had different f0 values at the initial position compared to f0 values at the final position [t = 4.319, df = 17.0, p < 0.05].

2. Sentence-medial

   2.1 Sentence-medial and Thai females with low English proficiency

       2.1.1   Thai females with low English proficiency had different f0 values at the medial position compared to f0 values at the initial position [t = 5.418, df = 58.4, p < 0.05], medial position [t = 7.659, df = 47.7, p < 0.05], and final position [t = 6.707, df = 58.4, p < 0.05] produced by Thai males with low English proficiency.

       2.1.2   Thai females with low English proficiency had different f0 values at the medial position compared to f0 values at the initial position [t = 5.427, df = 59.9, p < 0.05], medial position [t = 7.218, df = 47.7, p < 0.05], and final position [t = 7.271, df = 59.9, p < 0.05] produced by Thai males with high English proficiency.

       2.1.3   Thai females with low English proficiency had different f0 values at the medial position compared to f0 values at the initial position [t = 7.357, df = 61.0, p < 0.05], medial position [t = 9.594, df = 47.7, p < 0.05], and final position [t = 8.333, df = 61.0, p < 0.05] produced by native English males.

   2.2 Sentence-medial and Thai females with high English proficiency

       2.2.1   Thai females with high English proficiency had different f0 values at the medial position compared to f0 values at the initial position [t = 4.323, df = 57.2, p < 0.05], medial position [t = 6.267, df = 47.7, p < 0.05], and final position [t = 5.513, df = 57.2, p < 0.05] produced by Thai males with low English proficiency.

       2.2.2   Thai females with high English proficiency had different f0 values at the medial position compared to f0 values at the initial position [t = 4.200, df = 58.4, p < 0.05], medial position [t = 5.662, df = 47.7, p < 0.05], and final position [t = 5.877, df = 58.4, p < 0.05] produced by Thai males with high English proficiency.

       2.2.3   Thai females with high English proficiency had different f0 values at the medial position compared to f0 values at the initial position [t = 5.786, df = 59.6, p < 0.05], medial position [t = 7.508, df = 47.7, p < 0.05], and final position [t = 6.660, df = 59.6, p < 0.05] produced by native English males.

2.3 Sentence-medial and Native English females

    2.3.1   Native English females had different f0 values at the medial position compared to f0 values at the initial position [t = 3.907, df = 59.6, p < 0.05], medial position [t = 6.253, df = 47.7, p < 0.05], and final position [t = 5.295, df = 59.6, p < 0.05] produced by Thai males with low English proficiency.

    2.3.2   Native English females had different f0 values at the medial position compared to f0 values at the initial position [t = 3.790, df = 61.0, p < 0.05], medial position [t = 5.660, df = 47.7, p < 0.05], and final position [t = 5.809, df = 61.0, p < 0.05] produced by Thai males with high English proficiency.

    2.3.3   Native English females had different f0 values at the medial position compared to f0 values at the initial position [t = 5.815, df = 61.3, p < 0.05], medial position [t = 8.295, df = 47.7, p < 0.05], and final position [t = 6.907, df = 61.3, p < 0.05] produced by native English males.

3.   Sentence-final

3.1 Sentence-final and Thai females with low English proficiency

    3.1.1   Thai females with low English proficiency had different f0 values at the final position compared to f0 values at the initial position [t = 4.685, df = 58.4, p < 0.05], medial position [t = 6.377, df = 58.4, p < 0.05], and final position [t = 6.436, df = 47.7, p < 0.05] produced by Thai males with low English proficiency.

    3.1.2   Thai females with low English proficiency had different f0 values at the final position compared to f0 values at the initial position [t = 4.623, df = 59.9, p < 0.05], medial position [t = 5.782, df = 59.9, p < 0.05], and final position [t = 7.085, df = 47.7, p < 0.05] produced by Thai males with high English proficiency.

    3.1.3   Thai females with low English proficiency had different f0 values at the final position compared to f0 values at the initial position [t = 6.475, df = 61.0, p < 0.05], medial position [t = 7.696, df = 61.0, p < 0.05], and final position [t = 8.335, df = 47.7, p < 0.05] produced by native English males.

3.2 Sentence-final and Thai females with high English proficiency

3.2.1 Thai females with high English proficiency had different f0 values at the final position compared to f0 values at the initial position [t = 3.822, df = 57.2, p < 0.05], medial position [t = 5.385, df = 57.2, p < 0.05], and final position [t = 5.338, df = 47.7, p < 0.05] produced by Thai males with low English proficiency.

3.2.2 Thai females with high English proficiency had different f0 values at the final position compared to f0 values at the initial position [t = 3.659, df = 58.4, p < 0.05], medial position [t = 4.714, df = 58.4, p < 0.05], and final position [t = 5.748, df = 47.7, p < 0.05] produced by Thai males with high English proficiency.

3.2.3 Thai females with high English proficiency had different f0 values at the final position compared to f0 values at the initial position [t = 5.203, df = 59.6, p < 0.05], medial position [t = 6.296, df = 59.6, p < 0.05], and final position [t = 6.632, df = 47.7, p < 0.05] produced by Native English males.

3.3 Sentence-final and Native English females

3.3.1 Native English females had different f0 values at the final position compared to f0 values at the initial position [t = 3.635, df = 59.6, p < 0.05], medial position [t = 5.457, df = 59.6, p < 0.05], and final position [t = 5.482, df = 47.7, p < 0.05] produced by Thai males with low English proficiency.

3.3.2 Native English females had different f0 values at the final position compared to f0 values at the medial position [t = 4.756, df = 61.0, p < 0.05], and final position [t = 6.157, df = 47.7, p < 0.05] produced by Thai males with high English proficiency.

3.3.3 Native English females had different f0 values at the final position compared to f0 values at the initial position [t = 5.475, df = 61.3, p < 0.05], medial position [t = 6.840, df = 61.3, p < 0.05], and final position [t = 7.586, df = 47.7, p < 0.05] produced by native English males.

5.3 Intensity

To examine the effects on intensity of such factors as levels of English proficiency, gender and focus position, the hypothesis is that:

Hypothesis 5.2:   The values of intensity produced by Thai learners differ from those of native speakers due to different levels of English proficiency, gender and focus positions.

5.3.1 Intensity and effects of English proficiency, gender and focus position

Tables 5.4 below shows the descriptive statistics of intensity for English native speakers or NSE (n = 2400) and Thai learners or English L2 Learners or ELL (n = 2400). The mean values of intensity for the former was 63.10 dB, SD = 8.48 and those of the latter was 67.81 dB, SD = 8.88. Thai learners with low proficiency (n = 1200) have the mean intensity values of 67.95 dB, SD = 8.45, and those with high proficiency (n = 1200) have the mean intensity values of 67.67 dB, SD = 9.29.

Table 5. 4 Descriptive statistics for mean normalised intensity in decibels

Subjects		n	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
NSE		2,400	63.10	8.48	22.25	58.57	68.20	87.94
ELL		2,400	67.81	8.88	18.84	64.04	73.26	89.10
	Low	1,200	67.95	8.45	19.53	64.56	73.03	83.96
	High	1,200	67.67	9.29	18.84	63.46	73.47	89.10

The main statistical manipulations of intensity were performed by a series of LMM analyses with the help of the 'lme4' package in R. In LMM, Speaker or Participant and Word were treated as a random factor, while English Proficiency Level (EPL: low, high and native), Position (sentence-initial, sentence-medial and sentence-final), and Gender (female and male) were included as fixed factors.

When their interaction is significant, all group means from the interaction are compared by conducting pairwise comparisons, especially for a 3-way interaction effect. That is, focusing on

the groups in the interaction better describes the results of the analysis. For the mean normalised intensity in accented and focal syllables as in Table 5.5, the final full model with the smaller AIC value was chosen and displayed after running a series of model analyses and comparison. In this study, many maximal random effects models (e.g., both random intercept and slope) failed to converge or cannot find a solution. For example:

model4.1 <-   lmer(NormIntensity ~ EPL*Gender+Position+
(1 + EPL + Position |Subject)+ (1 + EPL + Position|Word),
data = mydata1, REML=FALSE)

model5.1 <-   lmer(NormIntensity ~ EPL+Gender+Position+
EPL*Gender+(1 + EPL + Position |Subject)+
(1 + EPL + Position|Word), data = mydata1, REML=FALSE)

model9 <-   lmer(NormIntensity ~ EPL+Gender+Position+
EPL*Gender*Position+EPL*Position+Gender*Position+
(1 + EPL + Position |Subject)+ (1 + EPL + Position|Word),
data = mydata1, REML=FALSE)

In this case, the researcher had to use a simpler model (e.g., intercepts-only). One common means to test the model's fit is to rerun the analysis but include only the intercept terms which is often called the null model.

model_inten   <-   lmer(NormIntensity ~ EPL + Gender + Position +
EPL * Gender * Position + EPL * Gender +
EPL * Position + Gender * Position +
(1|Subject) + (1|Word), data = mydata1, REML = FALSE)

nullmodel_inten   <-   lmer(NormIntensity ~ 1+(1|Subject)+ (1|Word), data =
mydata1, REML=FALSE)

Table 5. 5 LMM results or the mean normalised intensity in target accented syllables

Random Effects					
s^2	48.80				
$\tau_{00\ Subject}$	20.25				
$\tau_{00\ Word}$	1.49				
ICC	0.31				
$n_{Subject}$	40				
n_{Word}	12				
Observations	4800				
Marginal R^2 / Conditional R^2	0.130 / 0.398				
AIC	32510.129				
BIC	32646.1				
log-Likelihood	-16234.065				
Normalised Intensity (dB)					
Predictors	*Estimates*	*std. Error*	*Statistic*	*P*	*df*
(Intercept)	68.02	1.99	34.22	**<0.001**	4779.00
EPL[High]	-1.58	2.99	-0.53	0.598	4779.00
EPL[Native]	-5.57	2.39	-2.33	**<0.020**	4779.00
Gender[male]	0.89	2.99	0.30	0.766	4779.00
Position[medial]	-3.95	1.07	-3.68	**<0.001**	4779.00
Position[final]	1.93	1.07	1.80	0.072	4779.00
EPL[High]*Gender[male]	2.25	4.23	0.53	0.595	4779.00
EPL[Native]*Gender[male]	1.55	3.64	0.43	0.669	4779.00
EPL[High]*Position[medial]	0.53	1.01	0.53	0.598	4779.00
EPL[Native]*Position[medial]	1.23	0.81	1.53	0.126	4779.00
EPL[High]*Position[final]	0.41	1.01	0.41	0.683	4779.00
EPL[Native]*Position[final]	-0.64	0.81	-0.79	0.428	4779.00

Table 5.5 continued.

Normalised Intensity (dB)					
Predictors	*Estimates*	*std. Error*	*Statistic*	*P*	*df*
Gender[male]*Position[medial]	2.47	1.01	2.45	**<0.014**	4779.00
Gender[male]*Position[final]	-0.62	1.01	-0.62	0.538	4779.00
(EPL[High]*Gender[male])* Position[medial]	-1.90	1.43	-1.33	0.183	4779.00
(EPL[Native]*Gender[male])* Position[medial]	-3.39	1.23	-2.76	**<0.006**	4779.00
(EPL[High]*Gender[male])* Position[final]	-1.44	1.43	-1.01	0.311	4779.00
(EPL[Native]*Gender[male])* Position[final]	0.99	1.23	0.81	0.421	4779.00

The results in Table 5.5 reveal that the main effects were statistically significant. That is, there were statistically significant effects of English proficiency levels and positions on the values of intensity. EPL[Native] versus Thai learners with lower English proficiency [t = -2.33, df = 4779.00, p < 0.05], reflecting the greater values of intensity for Thai learners and the lower values of intensity in the group of native English speakers.

There were also effects of focus positions. A significant difference was reported for the intensity height in the medial position versus initial position [t = -3.68, df = 4779.00, p < 0.05], indicating the lower intensity sentence-medially. Two-way interaction effects of Gender[male] and Position[medial] showed significant differences [t = 2.45, df = 4779.00, p < 0.05].

Furthermore, there was a three-way interaction of English proficiency levels, gender and positions. Differences were statistically significant between EPL[Native], Gender[male] and Position[medial] [t = -2.76, df = 4779.00, p < 0.05] as in Figure 5.4.
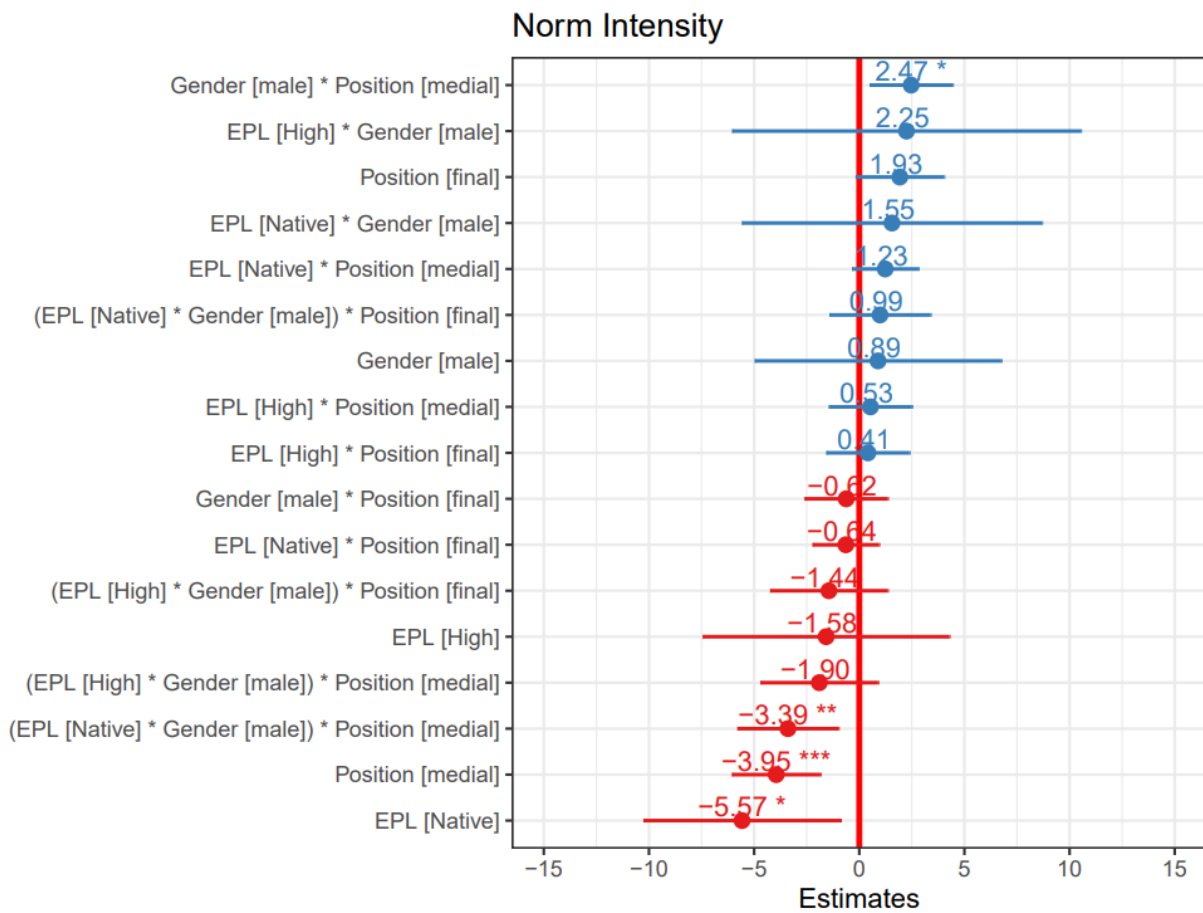
Figure 5. 4 Plot of the estimates of model_inten from highest to lowest values and the red vertical line indicating no effect

To examine differences among groups, all pairwise comparisons were done only on the optimal model. The results containing the means and SEs of English proficiency or EPL (low, high and native), gender (female and male) and positions (initial, medial and final) for the intensity are tabulated in Table 5.6. The formula for 'emmeans' is displayed below:

```
emmint       <-      emmeans(model_inten,
                     specs = pairwise ~ EPL : Gender :
                     Position, adjust = "tukey")
```

Table 5. 6 Means and SEs of EPL x Gender x Position for the intensity from pairwise comparison based on LMM

Subject groups	English Proficiency	Gender	Focus position	Mean	SE	df
Native English speakers (NSE	Native			63.10	1.17	56.35
		female		61.97	1.60	52.27
			initial	62.44	1.72	61.88
			medial	59.73	1.72	61.88
			final	63.73	1.72	61.88
		male		64.23	1.60	52.27
			initial	64.89	1.72	61.88
			medial	61.25	1.72	61.88
			final	66.55	1.72	61.88
English L2 learner (ELL) or Thai learners	Low			68.10	1.64	52.07
		female		67.35	2.05	50.17
			initial	68.02	2.15	59.02
			medial	64.07	2.15	59.02
			final	69.95	2.15	59.02
		male		68.85	2.49	49.03
			initial	68.91	2.59	56.68
			medial	67.43	2.59	56.68
			final	70.22	2.59	56.68
	High			67.40	1.64	52.07
		female		66.08	2.49	49.03
			initial	66.44	2.59	56.68
			medial	63.03	2.59	56.68
			final	68.78	2.59	56.68
		male		68.72	2.05	50.17
			initial	69.58	2.15	59.02
			medial	66.73	2.15	59.02
			final	69.86	2.15	59.02

Figure 5.5 below shows the values of intensity due to levels of English proficiency by gender. Even though the main effects of English proficiency and positions are significant, they might be significant because of their interaction with other independent variables. In this case, the results can be better visualised in Figure 5.6, which shows the effects of independent variables and their interactions on the intensity for the target accented syllables. The plot presents their interactions, in which the y-axis represents the distribution of the intensity values measured in decibels (dB). The x-axis represents different prosodic position (initial, medial and final). The plot is also grouped by gender and the levels of English proficiency (Low, High and Native). The dots inside the boxes correspond to the mean intensity.
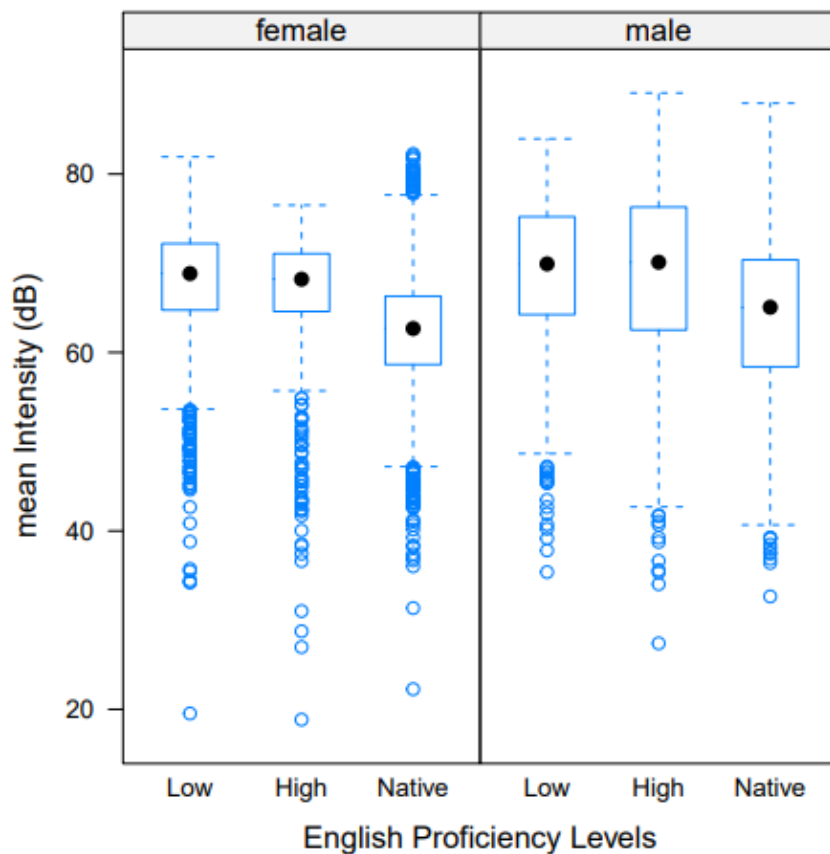
Figure 5. 5 Box plot for mean normalised intensity due to levels of English proficiency (Low and High for Thai learners of English with low and high English proficiency, and Native for native English speakers) presented by gender
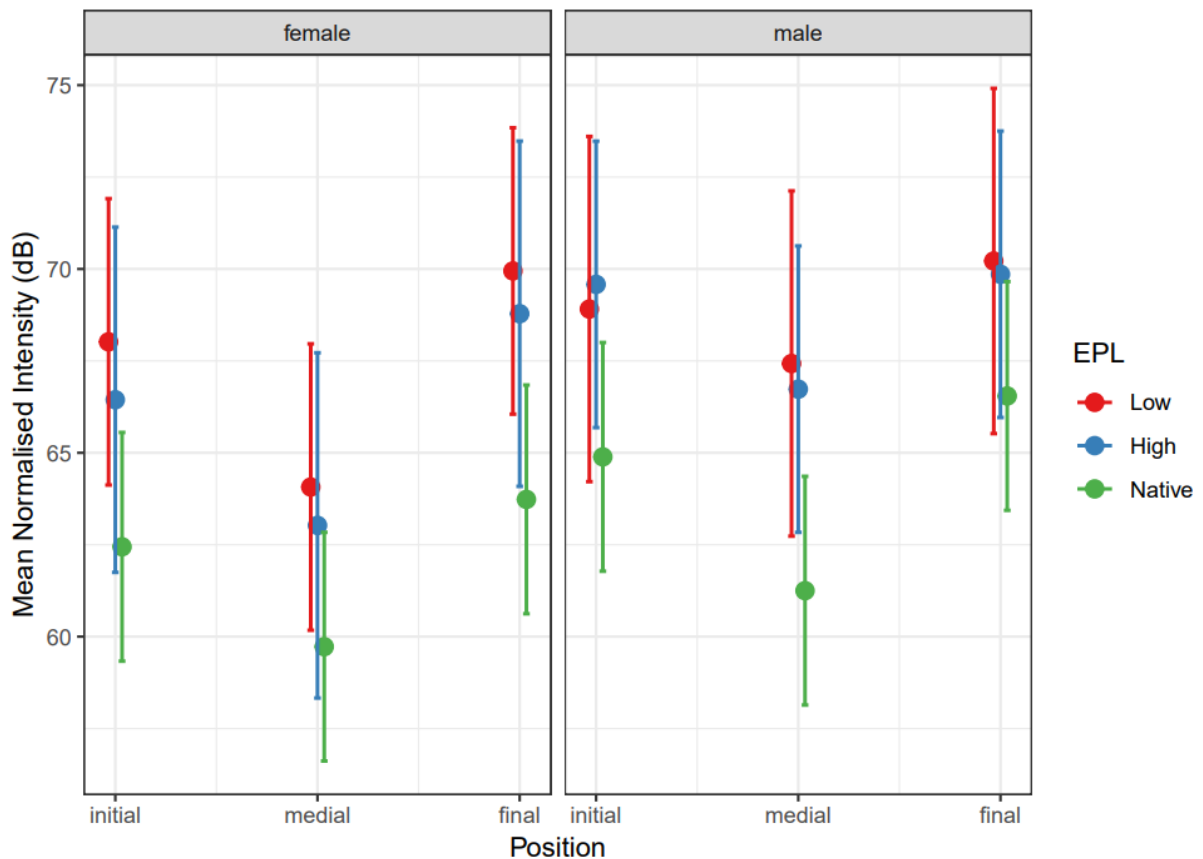
Figure 5. 6 Plot for interaction effects of positions (initial, medial, final) and English proficiency levels (Low, High and Native) grouped by each gender for the intensity

According to Tukey's HSD post-hoc test, the statistical differences are reported and grouped into sentence-initial, sentence-medial and sentence-final as follows. The report here takes three-way interactions (Position, Gender, and English proficiency levels) into consideration.

1. Sentence-initial
    1.1 Sentence-initial and Thai females with low English proficiency
        1.1.1   Thai females with low English proficiency had different values of intensity at the initial position compared to values at the medial position [$t = 3.401$, $df = 28.5$, $p < 0.05$].

2. Sentence-medial

    2.1 Sentence-medial and Thai females with low English proficiency

        2.1.1   Thai females with low English proficiency had different values of intensity at the medial position compared to values at the final position [t = -5.062, df = 28.5, p < 0.05].

    2.2 Sentence-medial and Thai females with high English proficiency

        2.2.1   Thai females with high English proficiency had different values of intensity at the medial position compared to values at the final position [t = -4.622, df = 39.5, p < 0.05].

    2.3 Sentence-medial and Native English females

        2.3.1   Native English females had different values of intensity at the medial position compared to values at the final position [t = -3.680, df = 21.0, p < 0.05].

        2.3.2   Native English females had different values of intensity at the medial position compared to values at the final position [t = -3.673, df = 60.0, p < 0.05] produced by Thai males with high English proficiency.

    2.4 Sentence-medial and Native English males

        2.4.1   Native English males had different values of intensity at the medial position compared to values at the final position [t = -4.864, df = 21.0, p < 0.05].

3. Sentence-final

    3.1 Sentence-final and Thai females with low English proficiency

        3.1.1   Thai females with low English proficiency had different values of intensity at the final position compared to values at the medial position [t = 3.705, df = 60.0, p < 0.05] produced by native English females.

5.4 Duration

To examine the effects on duration of such factors as levels of English proficiency, gender and focus positions, the hypothesis is that:

Hypothesis 5.3:   The values of duration produced by Thai learners differ from those of native speakers due to different levels of English proficiency, gender and focus positions.

5.4.1 Duration and effects of English proficiency, gender and focus position

Tables 5.7 below shows the descriptive statistics of duration for English native speakers or NSE (n = 2400) and Thai learners or English L2 Learners or ELL (n = 2400). The mean values of duration for the former was 118.73 ms, SD = 93.30 and those of the latter was 143.37 ms, SD = 112.63. Thai learners with low proficiency (n = 1200) have the mean duration values of 145.03 ms, SD = 113.83, and those with high proficiency (n = 1200) have the mean duration values of 141.71 ms, SD = 111.44.

Table 5. 7 Descriptive statistics for mean normalised duration in milliseconds

Subjects		n	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
NSE		2,400	118.73	93.30	0	46.2	169.6	568
ELL		2,400	143.37	112.63	0	53.6	207.3	635
	Low	1,200	145.03	113.83	0	53.7	209.5	564
	High	1,200	141.71	111.44	0	53.5	204.5	635

The main statistical manipulations were performed by a series of LMM analyses with the help of the 'lme4' package in R. In LMM, Speaker or Participant and Word were treated as a random factor, while English Proficiency Level (EPL: low, high and native), Position (sentence-initial, sentence-medial and sentence-final), and Gender (female and male) were included as fixed factors. When their interaction is significant, all group means from the interaction are compared by conducting pairwise comparisons, especially for 3-way interaction effect. That is, focusing on the groups in the interaction better describes the results of the analysis. For the mean normalised duration in accented and focal syllables, as in Table 5.8, the final full model with

the smaller AIC value was chosen and displayed after conducting several different model analyses and comparison. In this study, many maximal random effects models (e.g., both random intercept and slope) failed to converge or cannot find a solution. For instance:

```
model4        <-      lmer(NormDurMs ~ EPL*Position+Gender+
                      (1 + EPL + Position |Subject)+ (1 + EPL + Position|Word),
                      data = mydata1, REML=FALSE)


model6        <-       lmer(NormDurMs~ EPL+Gender+Position+
                      EPL*Gender*Position+EPL*Gender+EPL*Position+
                      Gender*Position+(1 + EPL + Position |Subject)+
                      (1 + EPL + Position|Word), data = mydata1,  REML=FALSE)


model9        <-      lmer(NormDurMs ~ EPL+Gender+Position+
                      EPL*Gender*Position+EPL*Position+Gender*Position+
                      (1 + EPL + Position |Subject)+ (1 + EPL + Position|Word),
                      data = mydata1,  REML=FALSE)
```

In this case, the researcher had to use a simpler model (e.g., intercepts-only). One common means to test the model's fit is to rerun the analysis but include only the intercept terms which is often called the null model.

```
model_dur      <-      lmer(NormDurMs ~ EPL + Gender + Position +
                      EPL * Gender * Position + EPL * Gender +
                      EPL * Position + Gender * Position +
                      (1|Subject) + (1|Word), data = mydata1,  REML = FALSE)


nullmodel_dur   <-      lmer(NormDurMs ~ 1+(1|Subject)+ (1|Word), data = mydata1,
                      REML=FALSE)
```

Table 5. 8 LMM results for the mean normalised duration in target accented syllables

Random Effects					
s^2	8560.21				
$\tau_{00\ Subject}$	204.18				
$\tau_{00\ Word}$	103.36				
ICC	0.03				
$n_{Subject}$	40				
n_{Word}	12				
Observations	4800				
Marginal R^2 / Conditional R^2	0.183 / 0.211				
AIC	57201.490				
BIC	57337.5				
log-Likelihood	-28579.745				
Normalised Duration (ms)					
Predictors	*Estimates*	*std. Error*	*Statistic*	*p*	*df*
(Intercept)	99.52	9.77	10.18	**<0.001**	4779.00
EPL[High]	6.59	13.20	0.50	0.617	4779.00
EPL[Native]	-9.75	10.56	-0.92	0.356	4779.00
Gender[male]	-12.40	13.20	-0.94	0.347	4779.00
Position[medial]	115.45	11.09	10.41	**<0.001**	4779.00
Position[final]	31.47	11.09	2.84	**<0.005**	4779.00
EPL[High]*Gender[male]	1.77	18.67	0.09	0.925	4779.00
EPL[Native]*Gender[male]	12.37	16.06	0.77	0.441	4779.00
EPL[High]*Position[medial]	8.53	13.35	0.64	0.523	4779.00
EPL[Native]*Position[medial]	-49.49	10.68	-4.63	**<0.001**	4779.00
EPL[High]*Position[final]	6.81	13.35	0.51	0.610	4779.00
EPL[Native]*Position[final]	-21.87	10.68	-2.05	**<0.041**	4779.00
Gender[male]*Position[medial]	6.84	13.35	0.51	0.608	4779.00
Gender[male]*Position[final]	4.38	13.35	0.33	0.743	4779.00
(EPL[High]*Gender[male])*Position[medial]	-44.48	18.89	-2.36	**<0.019**	4779.00
(EPL[Native]*Gender[male])*Position[medial]	19.63	16.25	1.21	0.227	4779.00
(EPL[High]*Gender[male])*Position[final]	-27.28	18.89	-1.44	0.149	4779.00
(EPL[Native]*Gender[male])*Position[final]	-8.10	16.25	-0.50	0.618	4779.00

The results in Table 5.8 reveal that the main effect of focus positions was statistically significant. There were the effects of Position[medial] versus Position[initial] [t = 10.41, df = 4779.00, p < 0.05], indicating that duration of the accented target syllables were longer. There were the effects of Position[final] versus Position[initial] [t = 2.84, df = 4779.00, p < 0.05], suggesting that duration at this final position was made longer. The group of native English speakers (EPL[Native]) interacting with the focus positions (Position[medial]) [t = -4.63, df = 4779.00, p < 0.05] and Position[final]) [t = -2.05, df = 4779.00, p < 0.05] affected the different values of duration of the target accented syllables.

There was also a 3-way interaction of English proficiency levels, gender and positions for Thai learners. The interaction of EPL[High], Gender[male]) and Position[medial] had significant effects on durational intervals [t = -2.36, df = 4779.00, p < 0.05], as in Figure 5.7.
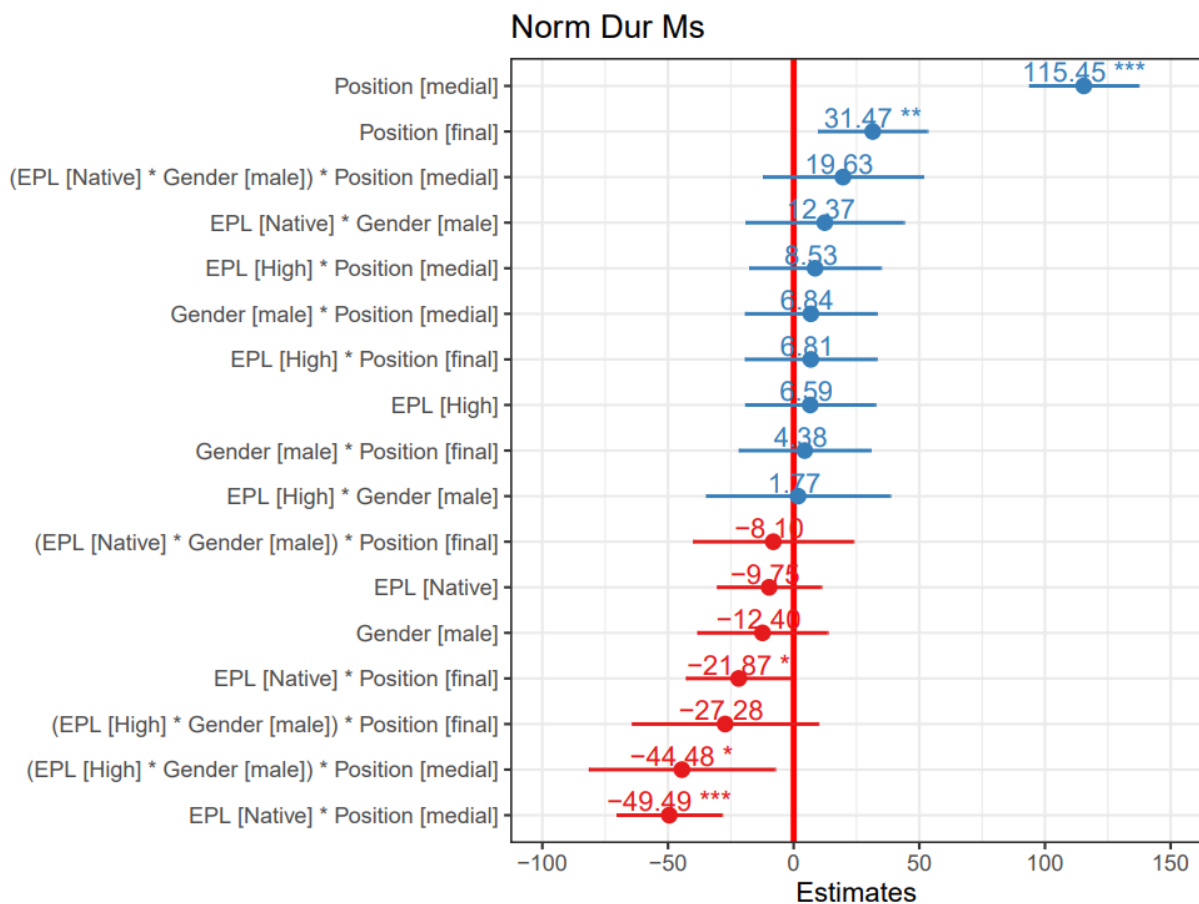
Figure 5. 7 Plot of the estimates of model_dur from highest to lowest values and the red vertical line indicating no effect

To examine further differences among groups, all pairwise comparisons were done only on the optimal model. The results containing the means and SEs of English proficiency or EPL (low, high and native), gender (female and male) and positions (initial, medial and final) for the duration are tabulated in Table 5.8. The formula for 'emmeans' is displayed below:

emmdur   <-   emmeans(model_dur, specs = pairwise ~ EPL : Gender : Position, adjust = "tukey")

Table 5. 9 Means and SEs of EPL x Gender x Position for the duration from pairwise comparison based on LMM

Subject groups	English Proficiency	Gender	Focus position	Mean	SE	df
	Native			118.73	5.26	44.54
		female		114.96	6.61	56.84
Native English speakers (NSE)			Initial	89.77	9.02	56.24
			Medial	155.74	9.02	56.24
			Final	99.36	9.02	56.24
		male		122.51	6.61	56.84
			Initial	89.74	9.02	56.24
			Medial	182.18	9.02	56.24
			Final	95.61	9.02	56.24
	Low			144.16	6.71	57.14
		female		148.49	8.06	57.65
English L2 learner (ELL) or Thai learners			Initial	99.52	10.59	82.96
			Medial	214.98	10.59	82.96
			Final	130.99	10.59	82.96
		male		139.83	9.57	55.58
			Initial	87.12	12.28	102.04
			Medial	209.41	12.28	102.04
			Final	122.96	12.28	102.04
	High			144.79	6.71	57.14
		female		160.20	9.57	55.58
			Initial	106.12	12.28	102.04
			Medial	230.10	12.28	102.04
			Final	144.39	12.28	102.04
		male		129.38	8.06	57.65
			Initial	95.48	10.59	82.96
			Medial	181.82	10.59	82.96
			Final	110.85	10.59	82.96

Figure 5.8 below shows the values of duration due to levels of English proficiency by gender, reflecting no main effects of these two variables except for the main effect of positions as the model result in Table 5.8. In this case, the results can be better visualized and more informative in Figure 5.9, which shows the effects of independent variables and their interactions on the duration for the target accented syllables. The plot presents their interactions, in which the y-axis represents the distribution of the duration values measured in milliseconds (ms). The x-axis represents different prosodic positions (initial, medial and final). The plot is also grouped by gender and the levels of English proficiency (Low, High and Native). The dots inside the boxes correspond to the mean duration.
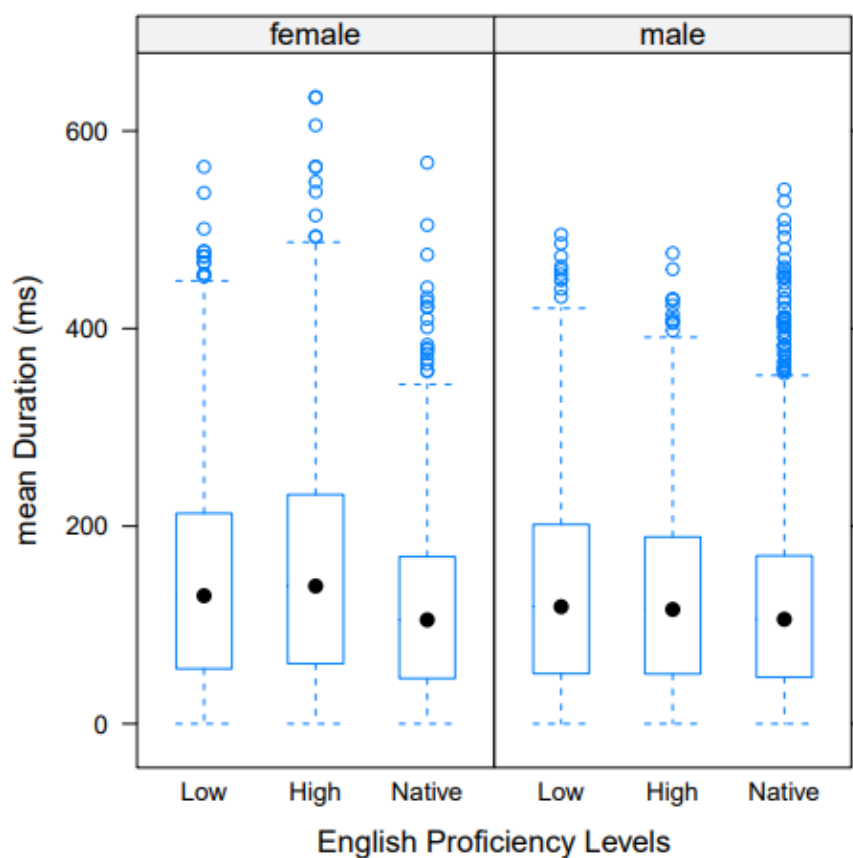
Figure 5. 8 Box plot for mean normalised duration due to levels of English proficiency (Low and High for Thai learners of English with low and high English proficiency, and Native for native English speakers) presented by gender
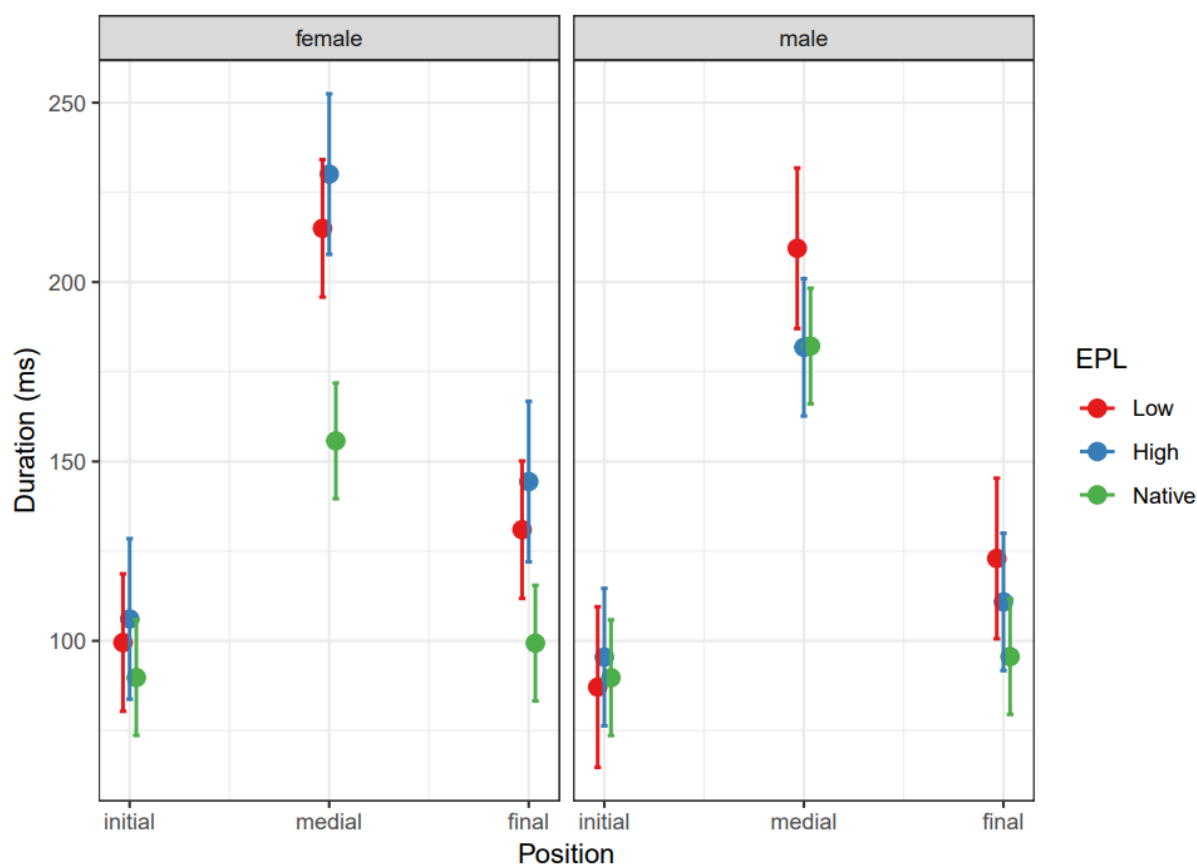
Figure 5. 9 Plot for interaction effects of positions (initial, medial, final) and English proficiency levels (Low, High and Native) grouped by each gender for the mean duration

According to Tukey's HSD post-hoc test, the statistical differences are reported and grouped into sentence-initial, sentence-medial and sentence-final as follows. The report here takes three-way interactions (Position, Gender, and English proficiency levels) into consideration.

1. Sentence-initial
   1.1 Sentence-initial and Thai females with low English proficiency
       1.1.1 Thai females with low English proficiency had different values of duration at the initial position compared to values at the medial position [$t = -9.701$, $df = 58.7$, $p < 0.05$].
       1.1.2 Thai females with low English proficiency had different values of duration at the initial position compared to values at the medial position produced by Thai females with high English proficiency [$t = -8.053$, $df = 93.5$, $p < 0.05$].

1.1.3 Thai females with low English proficiency had different values of duration at the initial position compared to values at the medial position [t = -6.777, df = 71.0, p < 0.05] produced by native English females.

1.1.4 Thai females with low English proficiency had different values of duration at the initial position compared to values at the medial position [t = 8.532, df = 93.5, p < 0.05] produced by Thai males with low English proficiency.

1.1.5 Thai females with low English proficiency had different values of duration at the initial position compared to values at the medial position [t = -5.494, df = 82.7, p < 0.05] produced by Thai males with high English proficiency.

1.1.6 Thai females with low English proficiency had different values of duration at the initial position compared to values at the medial position [t = -5.942, df = 71.0, p < 0.05] produced by native English males.

1.2 Sentence-initial and Thai females with high English proficiency

1.2.1 Thai females with high English proficiency had different values of duration at the initial position compared to values at the medial position [t = -9.309, df = 97.2, p < 0.05].

1.2.2 Thai females with high English proficiency had different values of duration at the initial position compared to values at the medial position [t = -5.949, df = 101.1, p < 0.05] produced by Thai males with low English proficiency.

1.2.3 Thai females with high English proficiency had different values of duration at the initial position compared to values at the medial position [t = -4.669, df = 93.5, p < 0.05] and final position [t = -0.292, df = 93.5, p < 0.05] produced by Thai males with high English proficiency.

1.2.4 Thai females with high English proficiency had different values of duration at the initial position compared to values at the medial position [t = -4.993, df = 85.1, p < 0.05] produced by native English males.

1.3 Sentence-initial and Native English females

1.3.1 Native English females had different values of duration at the initial position compared to values at the medial position [t = -6.204, df = 34.7, p < 0.05].

1.3.2 Native English females had different values of duration at the initial position compared to values at the medial position [t = -7.854, df = 85.1, p < 0.05] produced by Thai males with low English proficiency.

168

1.3.3　Native English females had different values of duration at the initial position compared to values at the medial position [t = -6.617, df = 71.0, p < 0.05] produced by Thai males with high English proficiency.

1.3.4　Native English females had different values of duration at the initial position compared to values at the medial position [t = -7.245, df = 56.6, p < 0.05] produced by native English males.

1.4 Sentence-initial and Thai males with low English proficiency

1.4.1　Thai males with low English proficiency had different values of duration at the initial position compared to values at the medial position [t = -9.182, df = 97.2, p < 0.05].

1.4.2　Thai males with low English proficiency had different values of duration at the initial position compared to values at the medial position [t = -5.841, df = 93.5, p < 0.05] produced by Thai males with high English proficiency.

1.4.3　Thai males with low English proficiency had different values of duration at the initial position compared to values at the medial position [t = -6.240, df = 85.1, p < 0.05] produced by native English males.

1.5 Sentence-initial and Thai males with high English proficiency

1.5.1　Thai males with high English proficiency had different values of duration at the initial position compared to values at the medial position [t = -7.255, df = 58.7, p < 0.05].

1.5.2　Thai males with high English proficiency had different values of duration at the initial position compared to values at the medial position [t = -6.232, df = 71.0, p < 0.05] produced by native English males.

1.6 Sentence-initial and native English males

1.6.1　Native English males had different values of duration at the initial position compared to values at the medial position [t = -8.694, df = 34.7, p < 0.05].

2. Sentence-medial

   2.1 Sentence-medial and Thai females with low English proficiency

      2.1.1   Thai females with low English proficiency had different values of duration at the medial position compared to values at the final position [t = 7.057, df = 58.7, p < 0.05].

      2.1.2   Thai females with low English proficiency had different values of duration at the medial position compared to values at the initial position [t = 6.714, df = 93.5, p < 0.05], and final position [t = 4.353, df = 93.5, p < 0.05] produced by Thai females with high English proficiency.

      2.1.3   Thai females with low English proficiency had different values of duration at the medial position compared to values at the initial position [t = 9.000, df = 71.0, p < 0.05], medial position [t = 5.334, df = 105.2, p < 0.05], and final position [t = 8.311, df = 71.0, p < 0.05] produced by native English females.

      2.1.4   Thai females with low English proficiency had different values of duration at the medial position compared to values at the initial position [t = 7.885, df = 93.5, p < 0.05], and final position [t = 5.675, df = 93.5, p < 0.05] produced by Thai males with low English proficiency.

      2.1.5   Thai females with low English proficiency had different values of duration at the medial position compared to values at the initial position [t = 7.978, df = 82.7, p < 0.05], and final position [t = 6.951, df = 82.7, p < 0.05] produced by Thai males with high English proficiency.

      2.1.6   Thai females with low English proficiency had different values of duration at the medial position compared to values at the initial position [t = 9.003, df = 71.0, p < 0.05], and final position [t = 8.581, df = 71.0, p < 0.05] produced by native English males.

   2.2 Sentence-medial and Thai females with high English proficiency

      2.2.1   Thai females with high English proficiency had different values of duration at the medial position compared to values at the final position [t = 6.435, df = 97.2, p < 0.05].

      2.2.2   Thai females with high English proficiency had different values of duration at the medial position compared to values at the initial position [t = 9.212, df = 85.1, p < 0.05], medial position [t = 5.844, df = 105.2, p < 0.05], and final position [t = 8.582, df = 85.1, p < 0.05] produced by native English females.

2.2.3 Thai females with high English proficiency had different values of duration at the medial position compared to values at the initial position [t = 8.235, df = 101.1, p < 0.05], and final position [t = 6.170, df = 101.1, p < 0.05] produced by Thai males with low English proficiency.

2.2.4 Thai females with high English proficiency had different values of duration at the medial position compared to values at the initial position [t = 8.302, df = 93.5, p < 0.05], and final position [t = 7.354, df = 93.5, p < 0.05] produced by Thai males with high English proficiency.

2.2.5 Thai females with high English proficiency had different values of duration at the medial position compared to values at the initial position [t = 9.213, df = 85.1, p < 0.05], medial position [t = 3.766, df = 105.2, p < 0.05], and final position [t = 8.828, df = 85.1, p < 0.05] produced by native English males.

2.3 Sentence-medial and Native English females

2.3.1 Native English females had different values of duration at the medial position compared to values at the final position [t = 5.302, df = 34.7, p < 0.05].

2.3.2 Native English females had different values of duration at the medial position compared to values at the initial position [t = 4.504, df = 85.1, p < 0.05], and medial position [t = -4.218, df = 105.2, p < 0.05] produced by Thai males with low English proficiency.

2.3.3 Native English females had different values of duration at the medial position compared to values at the initial position [t = 4.332, df = 71.0, p < 0.05] produced by Thai males with high English proficiency.

2.3.4 Native English females had different values of duration at the medial position compared to values at the initial position [t = 5.174, df = 56.6, p < 0.05], and final position [t = 4.714, df = 56.6, p < 0.05] produced by native English males.

2.4 Sentence-medial and Thai males with low English proficiency

2.4.1 Thai males with low English proficiency had different values of duration at the medial position compared to values at the final position [t = 6.491, df = 97.2, p < 0.05].

171

2.4.2   Thai males with low English proficiency had different values of duration at the medial position compared to values at the initial position [t = 7.027, df = 93.5, p < 0.05], and final position [t = 6.079, df = 93.5, p < 0.05] produced by Thai males with high English proficiency.

2.4.3   Thai males with low English proficiency had different values of duration at the medial position compared to values at the initial position [t = 7.856, df = 85.1, p < 0.05], and final position [t = 7.471, df = 85.1, p < 0.05] produced by native English males.

2.5 Sentence-medial and Thai males with high English proficiency

2.5.1   Thai males with high English proficiency had different values of duration at the medial position compared to values at the final position [t = 5.963, df = 58.7, p < 0.05].

2.5.2   Thai males with high English proficiency had different values of duration at the medial position compared to values at the initial position [t = 6.619, df = 71.0, p < 0.05], and final position [t = 6.197, df = 71.0, p < 0.05] produced by native English males.

2.6 Sentence-medial and native English males

2.6.1   Native English males had different values of duration at the medial position compared to values at the final position [t = 8.142, df = 34.7, p < 0.05].

3. Sentence-final

3.1 Sentence-final and Thai females with low English proficiency

3.1.1   Thai females with low English proficiency had different values of duration at the final position compared to values at the medial position [t = -6.112, df = 93.5, p < 0.05] produced by Thai females with high English proficiency.

3.1.2   Thai females with low English proficiency had different values of duration at the final position compared to values at the medial position [t = -4.837, df = 93.5, p < 0.05] produced by Thai males with low English proficiency.

3.1.3   Thai females with low English proficiency had different values of duration at the final position compared to values at the medial position [t = -3.680, df = 71.0, p < 0.05] produced by native English males.

172

3.2 Sentence-final and Thai females with high English proficiency

    3.2.1   Thai females with high English proficiency had different values of duration at the final position compared to values at the initial position [t = 3.586, df = 85.1, p < 0.05] produced by native English females.

    3.2.2   Thai females with high English proficiency had different values of duration at the final position compared to values at the medial position [[t = -3.745, df = 101.1, p < 0.05] produced by Thai males with low English proficiency.

    3.2.3   Thai females with high English proficiency had different values of duration at the final position compared to values at the initial position [t = 3.588, df = 85.1, p < 0.05], and final position [t = 3.834, df = 105.2, p < 0.05] produced by native English males.

3.3 Sentence-final and Native English females

    3.3.1   Native English females had different values of duration at the final position compared to values at the medial position [t = -7.224, df = 85.1, p < 0.05] produced by Thai males with low English proficiency.

    3.3.2   Native English females had different values of duration at the final position compared to values at the medial position [t = -5.927, df = 71.0, p < 0.05] produced by Thai males with high English proficiency.

    3.3.3   Native English females had different values of duration at the final position compared to values at the medial position [t = -6.493, df = 56.6, p < 0.05] produced by native English males.

3.4 Sentence-final and Thai males with low English proficiency

    3.4.1   Thai males with low English proficiency had different values of duration at the final position compared to values at the medial position [t = -3.630, df = 93.5, p < 0.05] produced by Thai males with high English proficiency.

    3.4.2   Thai males with low English proficiency had different values of duration at the final position compared to values at the medial position [t = -3.887, df = 85.1, p < 0.05] produced by native English males.

3.5 Sentence-final and Thai males with high English proficiency

    3.5.1   Thai males with high English proficiency had different values of duration at the final position compared to values at the medial position [t = -5.127, df = 71.0, p < 0.05] produced by native English males.

5.5 Summary of the Results from Statistical Analyses

The following sections present the results of the acoustic correlates of focus or focal prominence: f0, intensity and duration across three focus positions of the sentence, English proficiency levels and gender.

- *F0*

The results of the acoustic correlates of focus showed that the f0 height was a reliable cue to marking narrow focus for all the participants in this study. Clearly from the results females and males differ a great deal in terms of f0 regardless of L1. There are many differences within the same gender. Here, Thai females had a higher f0 than English females regardless of L1. Similarly, Thai males had a higher f0 than English males in most case. In terms of English proficiency levels, Thai females with low English proficiency levels had a significantly greater f0 than native English females as in the table below. In terms of focus positions, the results shows a consistent pattern of f0 from the beginning to the end of the sentence. That is, at the initial position of the sentence, both groups regardless of English proficiency levels and gender produced focused or accented syllables with quite higher values of f0, lower values of f0 at the medial position, and lowest at the final position of the sentence. There were some cases of males who had a re-increase in f0 at the final position of the sentence. In addition, the f0 results revealed differences in the degree of f0 manipulation between the two genders, three levels of English proficiency and three focus positions. When considering f0 differences at the same position, e.g., f0 of the initial position made by females but different English proficiency levels, or f0 of the initial position made by Thai females and Thai males with different English proficiency levels, or those made by native English speakers, the most significant contrast or directions of differences explained above can be summarised in the table below.

Table 5. 10 The significant directions of differences in the duration in accented syllables

Position	Directions of Differences	Explanations
Initial	Thai f. w/low EPL > Native English f. Thai f. w/low EPL > Thai m. w/low EPL Thai f. w/low EPL > Thai m. w/high EPL Thai f. w/low EPL > Native English m. Thai f. w/high EPL > Thai m. w/low EPL Thai f. w/high EPL > Thai m. w/high EPL Thai f. w/high EPL > Thai m. w/high EPL Native English f. > Thai m. w/low EPL Native English f. >Thai m. w/high EPL Native English f. > Native English m.	EPL is English proficiency. f. is for female. m. is for male. > is for directions of differences showing that the lefthand-side group of > produces the accented syllables in the focused words with a higher degree of f0.

174

Table 5.10 continued.

Position	Directions of Differences	Explanations
Medial	Thai f. w/low EPL > Thai m. w/low EPL Thai f. w/low EPL > Thai m. w/high EPL Thai f. w/low EPL > Native English m. Thai f. w/high EPL > Thai m. w/low EPL Thai f. w/high EPL > Thai m. w/high EPL Thai f. w/high EPL > Native English m. Native English f. > Thai m. w/low EPL Native English f. > Thai m. w/high EPL Native English f. > Native English m.	EPL is English proficiency. f. is for female. m. is for male. > is for directions of differences showing that the lefthand-side group of > produces the accented syllables in the focused words with a higher degree of f0.
Final	Thai f. w/low EPL > Thai m. w/low EPL Thai f. w/low EPL > Thai m. w/high EPL Thai f. w/low EPL > Native English m. Thai f. w/high EPL > Thai m. w/low EPL Thai f. w/high EPL > Thai m. w/high EPL Thai f. w/high EPL > Native English m. Native English f. > Thai m. w/low EPL Native English f. > Thai m. w/high EPL Native English f. > Native English m.	

- *Intensity*

Intensity was also used to mark focus for all participants in this study. The results revealed that Thai learners regardless of gender, English proficiency levels , or even focus positions exploited a large degree of intensity to mark focus on the accented words. Thai learners of English produced intensity with higher values than native English speakers (both males and females) in all cases. Interestingly, lowest intensity was manipulated by all participants to mark medial focused positions, and higher intensity was manipulated by all participants to mark focus at the beginning and end of the sentences. According to the results, it shows that the positions of the focused words in the sentences had an effect on intensity, as well as Thai learners differing from native English speakers in terms of using intensity. When considering statistical differences of intensity at the same position, e.g., intensity of the initial position made by females but different English proficiency levels, or intensity of the initial position made by Thai females and Thai males with different English proficiency levels, or those made by native English speakers, Tukey's HSD post-hoc test cannot find the most significant contrast except for the contrast across different three positions as explained above.

- *Duration*

Duration was one of the reliable cues to mark focus for all participants. The results showed that there was variation in duration of each focus position in the sentence. The effect of medial position was found on syllable length. That is, accented syllables sentence-medially were produced longer in duration than accented syllables sentence-initially and sentence-finally regardless of gender, and levels of English proficiency. When considering durational differences at the same position, e.g., duration of the initial position made by females but different English proficiency levels, or duration of the initial position made by Thai females and Thai males with different English proficiency levels, or those made by native English speakers, the most significant contrast or directions of differences can be summarised in the table below.

Table 5. 11 The significant directions of differences in the duration in accented syllables

Position	Directions of Differences	Explanations
Initial	-	EPL is English proficiency. f. is for female.
Medial	Thai f. w/low EPL > Native English f. Thai f. w/high EPL > Native English f. Thai f. w/high EPL > Native English m. Native English f. > Thai m. w/low EPL	m. is for male. > is for directions of differences showing that the lefthand-side group of > has a greater duration of the accented syllables in the focused words.
Final	Thai f. w/high EPL > Native English m.	

Table 5.12 below summarises the results of post-hoc mean-separation tests for significant interaction effects between English proficiency levels, gender and positions on the f0 from statistical analyses, as well as Table 5.13 for intensity and 5.14 for duration. English proficiency levels include low (l), high (h) and native (n). Gender includes female (f) and male (m). Focus positions include sentence-initial, sentence-medial, and sentence-final. The asterisk * indicates statistically significant differences at $p < 0.05$.

Table 5. 12 Summary of the results of interaction effects between English proficiency levels, gender and positions on the f0 from statistical analyses

			Initial						Medial						Final					
			F			m			f			m			f			M		
			l	h	n	l	h	n	l	h	n	l	h	n	l	h	n	l	h	n
Initial	f	l			*	*	*	*	*		*	*	*	*	*	*	*	*	*	*
		h				*	*	*		*		*	*	*		*		*	*	*
		n				*	*	*				*	*	*				*	*	*
	m	l										*								
		h																	*	
		n																		
Medial	f	l				*	*	*				*	*	*				*	*	*
		h				*	*	*				*	*	*				*	*	*
		n				*	*	*				*	*	*				*	*	*
	m	l																		
		h																		
		n																		
Final	f	l				*	*	*				*	*	*				*	*	*
		h				*	*	*				*	*	*				*	*	*
		n				*		*				*	*	*				*	*	*
	m	l																		
		h																		
		n																		

Table 5. 13 Summary of the interaction effects between English proficiency levels, gender and positions on the intensity from statistical analyses

			Initial						Medial						Final					
			F			m			f			m			f			m		
			l	h	n	l	h	n	l	h	n	l	h	n	l	h	n	l	h	n
Initial	f	l							*											
		h																		
		n																		
	m	l																		
		h																		
		n																		
Medial	f	l													*					
		h														*				
		n															*		*	
	m	l																		
		h																		
		n																		*
Final	f	l								*										
		h																		
		n																		
	m	l																		
		h																		
		n																		

Table 5. 14 Summary of the interaction effects between English proficiency levels, gender and positions on the duration from statistical analyses

			Initial						Medial						Final					
			F			m			f			m			f			m		
			l	h	n	l	h	n	l	h	n	l	h	n	l	h	n	l	h	n
Initial	f	l							*	*	*	*	*	*						
		h								*		*	*	*						
		n									*	*	*	*						
	m	l										*	*	*						
		h										*	*							
		n												*						
Medial	f	l		*	*	*	*	*		*					*	*	*	*	*	*
		h			*	*	*	*		*				*		*	*	*	*	*
		n				*	*	*				*					*			*
	m	l					*	*										*	*	*
		h						*											*	*
		n																		*
Final	f	l								*		*		*						
		h		*				*				*								*
		n										*	*	*						
	m	l											*	*						
		h												*						
		n																		

5.6 Summary of the Chapter

This chapter has reported on the results for the acoustic cues to post-lexical prominence reflected in the realisation of focus. Given the acoustic correlates, fundamental frequency, duration and intensity were of special interest. In answer to the research questions, it was found that Thai learners used three acoustic parameters (f0, duration and intensity) in the same ways as native speakers of English; however, they differed in degree as to how each parameter varied. That is, Thai learners produced accented syllables of the focused words with higher f0, higher intensity and longer duration. Factors such as differences in L2 proficiency levels, and gender as well as positions of focus in the sentence also influenced the manipulation of these acoustic cues more or less. In chapter 6, tonal alignment and scaling in rising accents are examined to determine whether or not Thai learners differ from native speakers in these respects.

Chapter 6. Alignment and Scaling of Rising Accents for Marking Focus

6.1 Introduction

This chapter describes the investigations into the alignment and scaling of rising accents in relation to marking the focus of information. The effects of factors such as English proficiency, gender and focus position are considered. The chapter starts with the results for the alignment of the low valley in section 6.2, followed by the results for the alignment of the high peak in section 6.3. Section 6.4 reports on the rising slope of low and high tonal targets in rising accents. A summary of the statistical analyses and of the chapter is provided in section 6.5 and 6.6.

Since Thai learners use rising accents (L*+H, L+H*, L+<H*) in the same way as native English speakers do to mark focus (Chapter 4), it is of special interest to consider the extent to which factors such as English proficiency, gender as well as focus position affect the use of tonal alignment and scaling. The research question that this section addresses is:

Research question 6:   To what extent do factors such as level of English proficiency, gender and focus position affect the temporal relations between the tonal movement in rising pitch accents and segmental strings in narrow-focused words produced by Thai learners and native speakers of English?

The hypotheses tested are given in each of the relevant sections below.

6.2 Alignment of the Low Valley

To investigate the alignment of low (L) tonal target, the temporal intervals from the syllable onset to L (C0toL) were measured and the data was collected from Thai learners and native speakers of English. The C0toL distances were calculated as a relative proportion of the accented syllable duration (C0toC1). To examine the effects of such factors as levels of English proficiency, gender and focus positions, the hypothesis tested is:

Hypothesis 6.1: The alignment of the low valley produced by Thai learners differs from those of native English speakers due to different levels of English proficiency, gender and focus positions

6.2.1 Alignment of the low valley and effects of English proficiency, gender and focus positions

Table 6.1 shows the descriptive statistics of the alignment of the low valley for English native speakers or NSE (n = 217) and Thai learners or English L2 Learners or ELL (n = 207). The mean value of f0 for the former was 7.68, SD = 9.12 and that of the latter was 12.67, SD = 12.24. Thai learners with low proficiency (n = 105) have a mean low valley value of 12.84, SD = 10.02, and those with high proficiency (n = 102) have a mean value of 12.50, SD = 14.22.

Table 6. 1 Descriptive statistics for the f0 valley alignment

Subjects		n	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
NSE		217	7.68	9.12	0.00	1.32	10.29	45.91
ELL		207	12.67	12.24	0.00	3.20	18.62	87.22
	Low	105	12.84	10.02	0.24	4.58	18.84	41.02
	High	102	12.50	14.22	0.00	1.90	18.25	87.22

The main statistical manipulations were performed by a series of LMM analyses with the help of the 'lme4' package in R. In LMM, Speaker or Participant and Word were treated as a random factor, while English Proficiency Level (EPL: low, high and native), Position (sentence-initial, sentence-medial and sentence-final), and Gender (female and male) were included as fixed factors. When their interaction is significant, all group means from the interaction are compared by conducting pairwise comparisons, especially for 3-way interaction effect. That is, focusing on the groups in the interaction better describes the results of the analysis. For the f0 valley alignment in accented and focal syllables as in Table 6.2, the final full model with the smaller AIC value was chosen and displayed after running a series of model comparison. In this study, many maximal random effects models (e.g., both random intercept and slope) failed to converge or cannot find a solution. In this case, the researcher had to use a simpler model. For example:

```
model4.1        <-      lmer(f0ValleyAlignR*100 ~ EPL*Gender+Position+
                        (1 + EPL + Position |Subject)+
                        (1 + EPL + Position|Word),
                        data = mydata1, REML=FALSE)


model5.1        <-      lmer(f0ValleyAlignR*100 ~ EPL+Gender+Position+
                        EPL*Gender+(1 + EPL + Position |Subject)+
                        (1 + EPL + Position|Word),
                        data = mydata1,  REML=FALSE)


model9          <-      lmer(f0ValleyAlignR*100~ EPL+Gender+Position+
                        EPL*Gender*Position+EPL*Position+
                        Gender*Position+(1 + EPL + Position |Subject)+
                        (1 + EPL + Position|Word),
                        data = mydata1,  REML=FALSE)
```

In this case, the researcher had to use a simpler model (e.g., intercepts-only). One common means to test the model's fit is to rerun the analysis but include only the intercept terms which is often called the null model.

```
model_valleyAR      <-      lmer(f0ValleyAlignR*100 ~ EPL + Gender + Position +
                            EPL * Gender * Position + EPL * Gender +
                            EPL * Position + Gender * Position +
                            (1|Subject) + (1|Word), data = mydata1,   REML =
                            FALSE)


nullmodel_valleyAR      <-      lmer(f0ValleyAlignR*100 ~ 1+(1|Subject)+ (1|Word),
                                data = mydata1,  REML=FALSE)
```

Table 6. 2 LMM results for the f0 valley alignment in target accented syllables

Random Effects					
s^2	57.04				
$\tau_{00\ Subject}$	18.25				
$\tau_{00\ Word}$	32.79				
ICC	0.47				
$n_{\ Subject}$	40				
$n_{\ Word}$	12				
Observations	424				
Marginal R^2 / Conditional R^2	0.267 / 0.613				
AIC	3053.772				
BIC	3138.8				
log-Likelihood	-1505.886				
f 0 Valley Align R*100					
Predictors	*Estimates*	*std. Error*	*Statistic*	*p*	*df*
(Intercept)	16.68	3.72	4.48	**<0.001**	403.00
EPL[High]	-0.47	3.75	-0.13	0.899	403.00
EPL[Native]	5.19	3.06	1.70	0.090	403.00
Gender[male]	0.34	3.78	0.09	0.927	403.00
Position[medial]	-2.91	4.78	-0.61	0.543	403.00
Position[final]	-4.36	4.63	-0.94	0.346	403.00
EPL[High]*Gender[male]	4.79	5.37	0.89	0.372	403.00
EPL[Native]*Gender[male]	-8.95	4.64	-1.93	0.054	403.00
EPL[High]*Position[medial]	5.27	3.95	1.33	0.183	403.00
EPL[Native]*Position[medial]	-13.52	3.13	-4.32	**<0.001**	403.00
EPL[High]*Position[final]	6.42	3.55	1.81	0.070	403.00
EPL[Native]*Position[final]	-14.26	2.88	-4.95	**<0.001**	403.00
Gender[male]*Position[medial]	-3.54	3.83	-0.92	0.356	403.00
Gender[male]*Position[final]	-5.44	3.56	-1.53	0.126	403.00
(EPL[High]*Gender[male])*Position[medial]	-14.22	5.47	-2.60	**<0.009**	403.00
(EPL[Native]*Gender[male])*Position[medial]	15.42	4.64	3.32	**<0.001**	403.00
(EPL[High]*Gender[male])*Position[final]	-11.82	5.09	-2.32	**<0.020**	403.00
(EPL[Native]*Gender[male])*Position[final]	13.75	4.37	3.15	**<0.002**	403.00

Table 6.2 reveals that main effects were statistically significant. There were significant interaction effects of English proficiency and positions on the values of f0 valley alignment. There were interactions between EPL[Native] and Position[medial] [t = -4.32, df = 403.00, p < 0.05] and between EPL[Native] and Position[final] [t = -4.95, df = 403.00, p < 0.05].

There were also 3-way interaction effects of English proficiency, gender and focus positions. EPL[High] has a significant interaction with Gender[male] and Position[medial] [t = -2.60, df = 403.00, p < 0.05], as well as Position[final] [t = -2.32, df = 403.00, p < 0.05]. EPL[Native] has a significant interaction with Gender[male] and Position[medial] [t = 3.32, df = 403.00, p < 0.05] as well as Position[final] [t = 3.15, df = 403.00, p < 0.05] as shown in Figure 6.1 below.
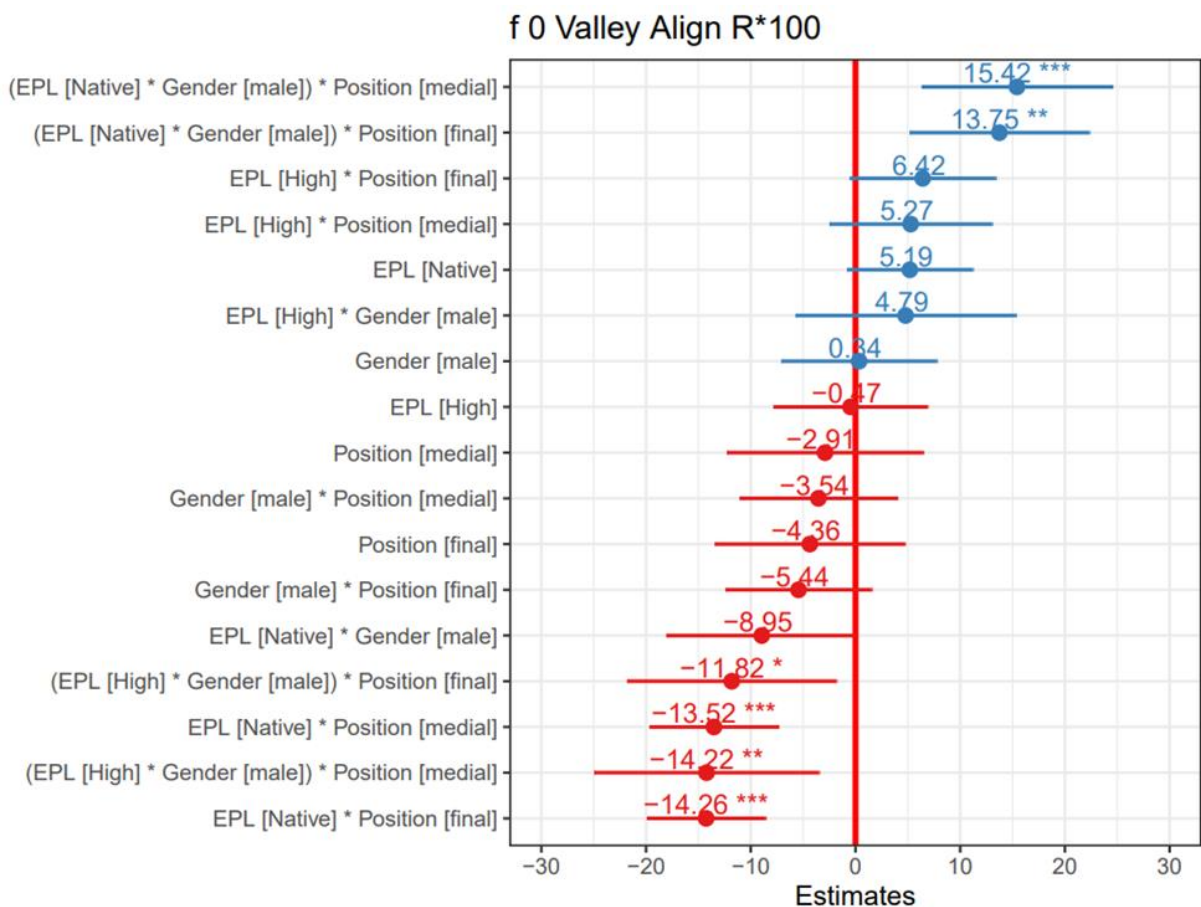
Figure 6. 1 Plot of estimates of model_valleyAR from highest to lowest values and the red vertical line indicating no effect

183

To examine differences among groups, all pairwise comparisons were done only on the optimal model. The results containing the means and SEs of English proficiency or EPL (low, high and native), gender (female and male) and positions (initial, medial and final) for the f0 valley alignment are tabulated in Table 6.3. The formula for 'emmeans' is displayed below:

emmvalley   <-   emmeans(model_valleyAR, specs = pairwise ~ EPL : Gender : Position, adjust =   "tukey")

Table 6. 3 Means and SEs of EPL x Gender x Position for the f0 valley alignment from pairwise comparison based on LMM

Subject groups	English Proficiency	Gender	Focus position	Mean	SE	df
Native English speakers (NSE)	Native			9.25	2.26	27.17
		female		10.19	2.54	39.94
			initial	21.87	3.92	29.78
			medial	5.45	3.86	27.68
			final	3.25	3.85	27.48
		male		8.31	2.54	39.96
			initial	13.27	3.90	29.06
			medial	8.72	3.88	28.39
			final	2.95	3.85	27.48
English L2 learner (ELL) or Thai learners	Low			12.93	2.57	41.21
		female		14.26	2.89	52.75
			initial	16.68	4.18	38.12
			medial	13.77	4.32	44.61
			final	12.32	4.15	37.21
		male		11.61	3.25	58.57
			initial	17.02	4.57	52.20
			medial	10.58	4.61	54.47
			final	7.22	4.50	49.32
	High			14.41	2.58	41.69
		female		17.68	3.28	60.27
			initial	16.20	4.53	50.35
			medial	18.57	4.75	61.65
			final	18.27	4.53	50.79
		male		11.13	2.89	52.57
			initial	21.33	4.24	40.78
			medial	5.93	4.21	39.71
			final	6.14	4.18	38.56

The results of Means and SEs of EPL x Gender x Position for the f0 valley alignment from pairwise comparison can be visualised in Figure 6.2, which shows the effects of independent variables and their interactions on the f0 valley alignment for the target accented syllables. The plot presents their interactions, wherein the y-axis represents the distribution of the values of f0 valley alignment by the levels of English proficiency (Low, High and Native). The x-axis represents different prosodic position (initial, medial and final). The dot inside a line corresponds to the mean f0 valley alignment, and the plot is also grouped by gender. Basically, the values of f0 valley alignment > 0 indicate that alignment of Lis within the accented syllables. For both groups of subjects, their L target was still timed within the focused words. At the beginning of the sentence, focused words took more time to align with accented syllables, especially for Thai subjects. Sentence-final alignment of L target was significantly different from native English females with [t = 2.908, df = 84.0, p < 0.05] to Thai females with low English proficiency, and with [t = 4.159, df = 87.3, p < 0.05] to those with a higher level of English proficiency.
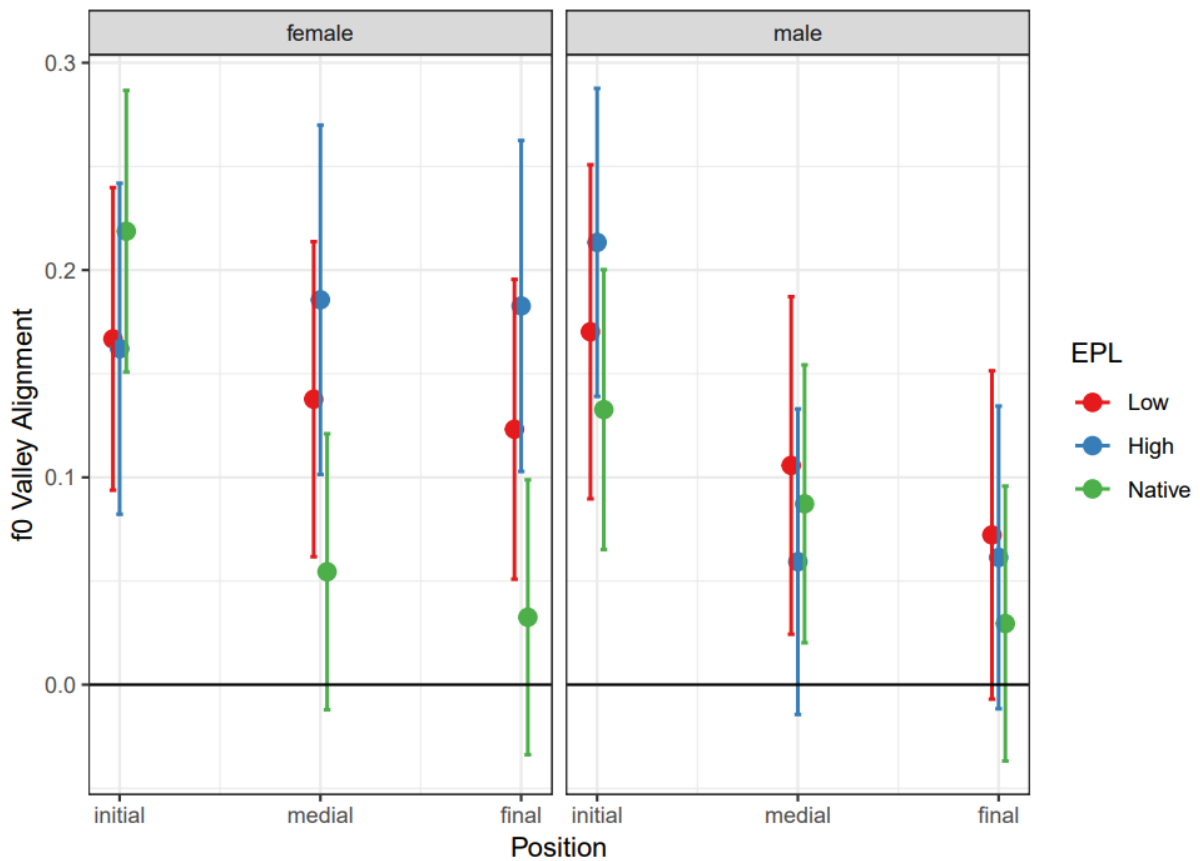
Figure 6. 2 Plot for interaction effects of positions (initial, medial, final) by English proficiency levels (Low, High and Native) for each gender

According to Tukey's HSD post-hoc test, the statistical differences are reported and grouped into sentence-initial, sentence-medial, and sentence-final as follows. The report here takes three-way interactions (Position, Gender, and English proficiency levels) into consideration.

1. Sentence-initial

    1.1 Sentence-initial and native English females

    1.1.1 Native English females had different values of f0 valley alignment at the initial position compared to values at the initial position [t = 3.032, df = 102.8, p < 0.05], and final position [t = 3.445, df = 29.1, p < 0.05], produced by native English males.

    1.1.2 Native English females had different values of f0 valley alignment at the initial position compared to values at the medial position [t = 3.226, df = 21.4, p < 0.05], and final position [t = 3.665, df = 21.3, p < 0.05], produced by native English females.

2. Sentence-medial

    2.1 Sentence-medial and Thai females with high English proficiency

    2.1.1 Thai females with high English proficiency had different values of f0 valley alignment at the medial position compared to values at the medial position [t = 2.978, df = 112.2, p < 0.05] produced by Thai males with high English proficiency.

    2.1.2 Thai females with high English proficiency had different values of f0 valley alignment at the medial position compared to values at the medial position [t = 3.375, df = 111.6, p < 0.05] produced by native English females.

3. Sentence-final

    3.1 Sentence-final and Thai females with low English proficiency

    3.1.1 Thai females with low English proficiency had different values of f0 valley alignment at the final position compared to values at the final position [t = 2.908, df = 84.0, p < 0.05] produced by native English females.

    3.1.2 Thai females with low English proficiency had different values of f0 valley alignment at the final position compared to values at the final position [t = 3.005, df = 84.1, p < 0.05] produced by native English males.

3.2 Sentence-final and Thai females with high English proficiency

    3.2.1   Thai females with high English proficiency had different values of f0 valley alignment at the final position compared to values at the final position [t = 3.058, df = 89.3, p < 0.05] produced by Thai males with high English proficiency.

    3.2.2   Thai females with high English proficiency had different values of f0 valley alignment at the final position compared to values at the final position [t = 4.243, df = 87.3, p < 0.05] produced by native English males.

    3.2.3   Thai females with high English proficiency had different values of f0 valley alignment at the final position compared to values at the final position [t = 4.159, df = 87.3, p < 0.05] produced by native English females.

6.3 Alignment of the H peak

To investigate the alignment of high (H) tonal target, temporal intervals from C0 to H were measured for the data collected from Thai learners and native English speakers. The distances of C0toH or peak delay was calculated as a relative proportion of the accented syllable duration (C0toC1). To examine the effects of such factors as levels of English proficiency, gender and focus positions, the hypothesis is:

Hypothesis 6.2:   The alignment of the high peak produced by Thai learners differs from those of native English speakers due to different levels of English proficiency, gender and focus positions.

6.3.1 Alignment of the high peak and effects of English proficiency, gender and focus positions

Table 6.4 below shows the descriptive statistics of f0 peak alignment for English native speakers or NSE (n = 217) and Thai learners or English L2 Learners or ELL (n = 207). The mean value of f0 peak alignment for the former was 87.90, SD = 16.10 and that of the latter was 102.78, SD = 39.08. Thai learners with low proficiency (n = 105) have a mean f0 peak alignment value of 101.86, SD = 42.22, and those with high proficiency (n = 102) have a mean value of 103.72, SD = 35.74.

187

Table 6. 4 Descriptive statistics for the f0 peak alignment

Subjects		n	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
NSE		217	87.90	16.10	46.44	77.18	97.15	137.02
ELL		207	102.78	39.08	36.67	80.59	112.70	343.27
	Low	105	101.86	42.22	36.67	80.47	111.05	343.27
	High	102	103.72	35.74	40.85	81.14	117.06	268.04

The main statistical manipulations were performed by a series of LMM analyses with the help of the 'lme4' package in R. In LMM, Speaker or Participant and Word were treated as a random factor, while English Proficiency Level (EPL: low, high and native), Position (sentence-initial, sentence-medial and sentence-final), and Gender (female and male) were included as fixed factors.

When their interaction is significant, all group means from the interaction are compared by conducting pairwise comparisons, especially for 3-way interaction effect. That is, focusing on the groups in the interaction better describes the results of the analysis. For the f0 peak alignment in accented and focal syllables as in Table 6.5, the final full model with the smaller AIC value was chosen and displayed here after conducting several different model analyses and comparison. In this study, many maximal random effects models (e.g., both random intercept and slope) failed to converge or cannot find a solution. For example:

```
model4.1    <-    lmer(f0PeakAlignR*100 ~ EPL*Gender+Position+
                   (1 + EPL + Position |Subject)+ (1 + EPL + Position|Word),
                   data = mydata1, REML=FALSE)


model5.1    <-    lmer(f0PeakAlignR*100 ~ EPL+Gender+Position+
                   EPL*Gender+(1 + EPL + Position |Subject)+
                   (1 + EPL + Position|Word), data = mydata1,  REML=FALSE)


model9      <-    lmer(f0PeakAlignR*100 ~ EPL+Gender+Position+
                   EPL*Gender*Position+EPL*Position+Gender*Position+
                   (1 + EPL + Position |Subject)+ (1 + EPL + Position|Word),
                   data = mydata1,  REML=FALSE)
```

In this case, the researcher had to use a simpler model (e.g., intercepts-only). One common means to test the model's fit is to rerun the analysis but include only the intercept terms which is often called the null model.

model_peakAR   <-   lmer(f0PeakAlignR*100 ~ EPL + Gender + Position + EPL * Gender * Position + EPL * Gender + EPL * Position + Gender * Position + (1|Subject) + (1|Word), data = mydata1, REML = FALSE)

nullmodel_peakAR   <-   lmer(f0PeakAlignR*100 ~ 1+(1|Subject)+ (1|Word), data = mydata1, REML=FALSE)

Table 6. 5 LMM results for the f0 peak alignment in target accented syllables

Random Effects					
s^2	330.80				
$\tau_{00 \text{ Subject}}$	135.09				
$\tau_{00 \text{ Word}}$	154.41				
ICC	0.47				
$n_{\text{ Subject}}$	40				
$n_{\text{ Word}}$	12				
Observations	424				
Marginal R^2 / Conditional R^2	0.407 / 0.684				
AIC	3804.294				
BIC	3889.3				
log-Likelihood	-1881.147				
f0 Peak Alignment *100					
Predictors	*Estimates*	*std. Error*	*Statistic*	*p*	*df*
(Intercept)	153.24	8.74	17.54	**<0.001**	403.00
EPL[High]	-20.13	9.67	-2.08	**<0.037**	403.00
EPL[Native]	-46.40	7.89	-5.88	**<0.001**	403.00
Gender[male]	-36.67	9.76	-3.76	**<0.001**	403.00
Position[medial]	-77.87	10.69	-7.28	**<0.001**	403.00

Table 6.5 continued.

f0 Peak Alignment *100					
Predictors	*Estimates*	*std. Error*	*Statistic*	*p*	*df*
Position[final]	-60.92	10.31	-5.91	**<0.001**	403.00
EPL[High]*Gender[male]	39.65	13.83	2.87	**<0.004**	403.00
EPL[Native]*Gender[male]	30.80	11.96	2.57	**<0.010**	403.00
EPL[High]*Position[medial]	37.99	9.53	3.99	**<0.001**	403.00
EPL[Native]*Position[medial]	57.72	7.54	7.65	**<0.001**	403.00
EPL[High]*Position[final]	26.48	8.55	3.10	**<0.002**	403.00
EPL[Native]*Position[final]	38.85	6.94	5.59	**<0.001**	403.00
Gender[male]*Position[medial]	32.35	9.23	3.50	**<0.001**	403.00
Gender[male]*Position[final]	30.42	8.57	3.55	**<0.001**	403.00
(EPL[High]*Gender[male])*Position[medial]	-40.56	13.19	-3.08	**<0.002**	403.00
(EPL[Native]*Gender[male])*Position[medial]	-36.19	11.17	-3.24	**<0.001**	403.00
(EPL[High]*Gender[male])*Position[final]	-45.22	12.25	-3.69	**<0.001**	403.00
(EPL[Native]*Gender[male])*Position[final]	-25.04	10.53	-2.38	**<0.017**	403.00

As in Table 6.5, it reveals that the main effects and interaction effects were statistically significant. English proficiency, gender and position have significant effects on the values of the f0 peak alignment in target accented syllables. First, there were main effects of EPL[High] versus EPL[Low] with [t = -2.08, df = 403.00, p < 0.05], EPL[Native] versus EPL[Low] with [t = -5.88, df = 403.00, p < 0.05], indicating that the participants with low English proficiency had later f0 peak alignment than the participants with high and native English proficiency. Differences between Gender[male] versus Gender[female] with [t = -3.76, df = 403.00, p < 0.05] suggested an earlier f0 peak alignment. There were main effects of Position[medial] versus Position[initial] with [t = -7.28, df = 403.00, p < 0.05],and Position[final] versus Position[initial] with [t = -5.91, df = 403.00, p < 0.05], indicating a greater distance of H to C0toL sentence-initially, sentence-finally and sentence-medially, respectively.

There were 2-way interaction effects of English proficiency, gender and positions. Significant interaction effects were between EPL[High] and Gender[male] with [t = 2.87, df = 403.00, p < 0.05], between EPL[Native] and Gender[male] with [t = 2.57, df = 403.00, p < 0.05], between

EPL[High] and Position[medial] with [t = 3.99, df = 403.00, p < 0.05], between EPL[Native] and Position[medial] with [t = 7.65, df = 403.00, p < 0.05], between EPL[High] and Position[final] with [t = 3.10, df = 403.00, p < 0.05], between EPL[Native] and Position[final] with [t = 5.59, df = 403.00, p < 0.05], between Gender[male] and Position[medial] with [t = 3.50, df = 403.00, p < 0.05], and between Gender[male] and Position[final] with [t = 3.55, df = 403.00, p < 0.05].

In addition, there were 3-way interactions of English proficiency, gender and positions. Interaction effects of (EPL[High] and Gender[male]) and Position[medial] were significantly different with [t = -3.08, df = 403.00, p < 0.05], (EPL[Native] and Gender[male]) and Position[medial] were significantly different with [t = -3.24, df = 403.00, p < 0.05], (EPL[High] and Gender[male]) and Position[final] were significantly different with [t = -3.69, df = 403.00, p < 0.05], as well as (EPL[Native] and Gender[male]) and Position[final] were significantly different with [t = -2.38, df = 403.00, p < 0.05]. All of these interaction effects can be clearly seen in Figure 6.3 below.
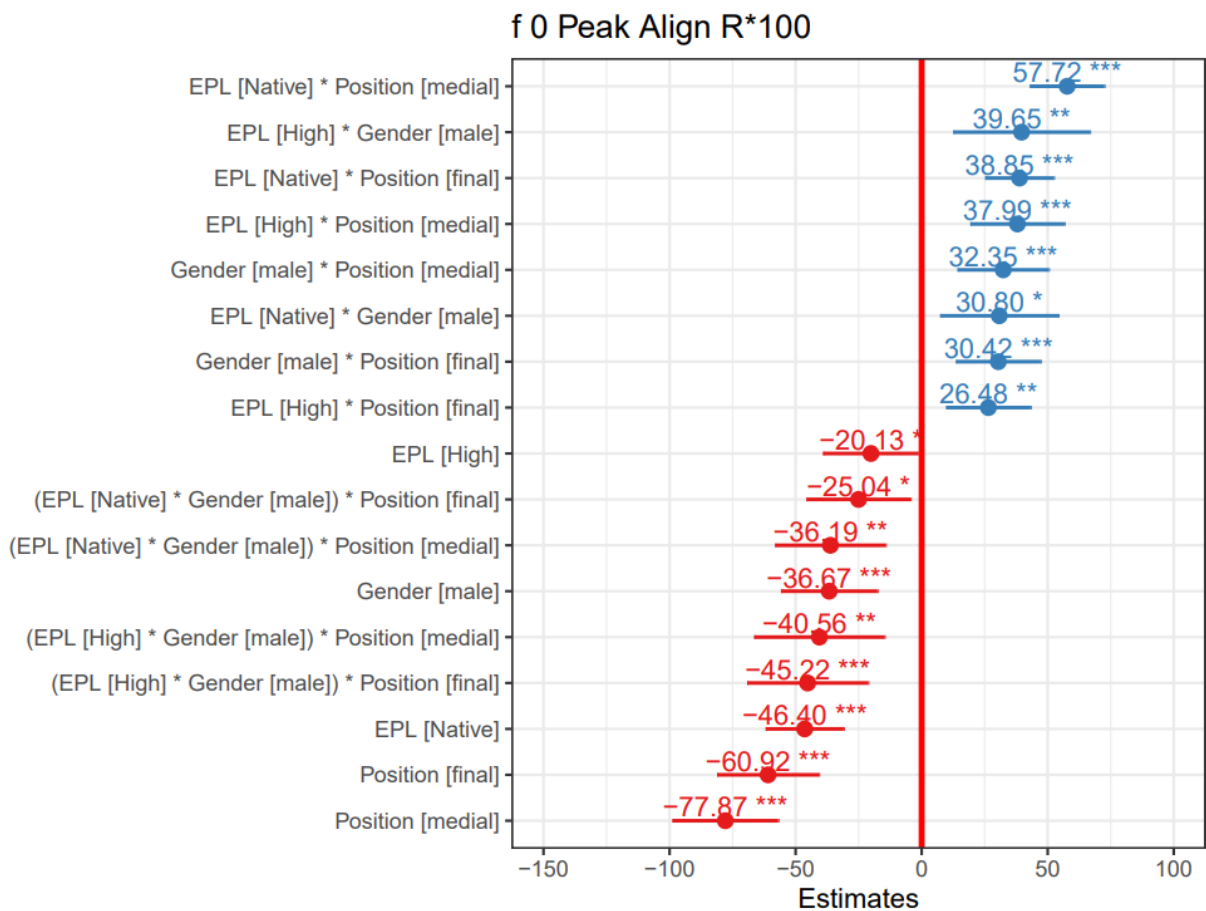
Figure 6. 3 Plot of estimates of model_peakAR from highest to lowest values and the red vertical line indicating no effect

To examine differences among groups, all pairwise comparisons were done only on the optimal model. The results containing the means and SEs of English proficiency or EPL (low, high and native), gender (female and male) and positions (initial, medial and final) for the f0 peak alignment are tabulated in Table 6.6. The formula for 'emmeans' is displayed below:

emmpeak   <-   emmeans(model_peakAR, specs = pairwise ~ EPL : Gender : Position, adjust = "tukey")

Table 6. 6 Means and SEs of EPL x Gender x Position for the f0 peak alignment from pairwise comparison based on LMM

Subject groups	English Proficiency	Gender	Focus position	Mean	SE	df
Native English speakers (NSE)	Native			90.08	5.20	31.97
		female		92.76	6.05	47.46
			initial	106.84	8.95	34.89
			medial	86.68	8.79	32.25
			final	84.76	8.76	31.99
		male		87.40	6.05	47.49
			initial	100.96	8.90	33.99
			medial	76.97	8.84	33.15
			final	84.27	8.76	31.99
English L2 learner (ELL) or Thai learners	Low			99.10	6.13	48.80
		female		106.97	7.06	58.20
			initial	153.24	9.72	46.21
			medial	75.37	10.07	54.39
			final	92.31	9.64	45.04
		male		91.22	8.09	59.86
			initial	116.56	10.81	63.05
			medial	71.04	10.91	65.79
			final	86.06	10.65	59.52
	High			105.99	6.15	49.36
		female		108.33	8.16	61.51
			initial	133.11	10.72	60.77
			medial	93.23	11.26	74.61
			final	98.67	10.72	61.28
		male		103.64	7.06	58.01
			initial	136.08	9.88	49.56
			medial	87.99	9.80	48.23
			final	86.84	9.72	46.77

The results can be visualised in Figure 6.4, which shows the effects of independent variables and their interactions on the f0 peak alignment for the target accented syllables. The plot presents their interactions, wherein the y-axis represents the distribution of the values of f0 peak alignment by the levels of English proficiency (Low, High and Native). The x-axis represents different prosodic position (initial, medial and final). The dot inside a line corresponds to the mean f0 peak alignment, and the plot is also grouped by gender. The values of f0 peak alignment < 1 indicate that alignment of H is within the accented syllables. For both groups of subjects, their H target was still anchored well into the focused words. Except for the beginning of the sentence, Thai subjects produced significantly delayed peaks. Differences were sentence-initial alignment of H target produced by Thai females and by Thai males.
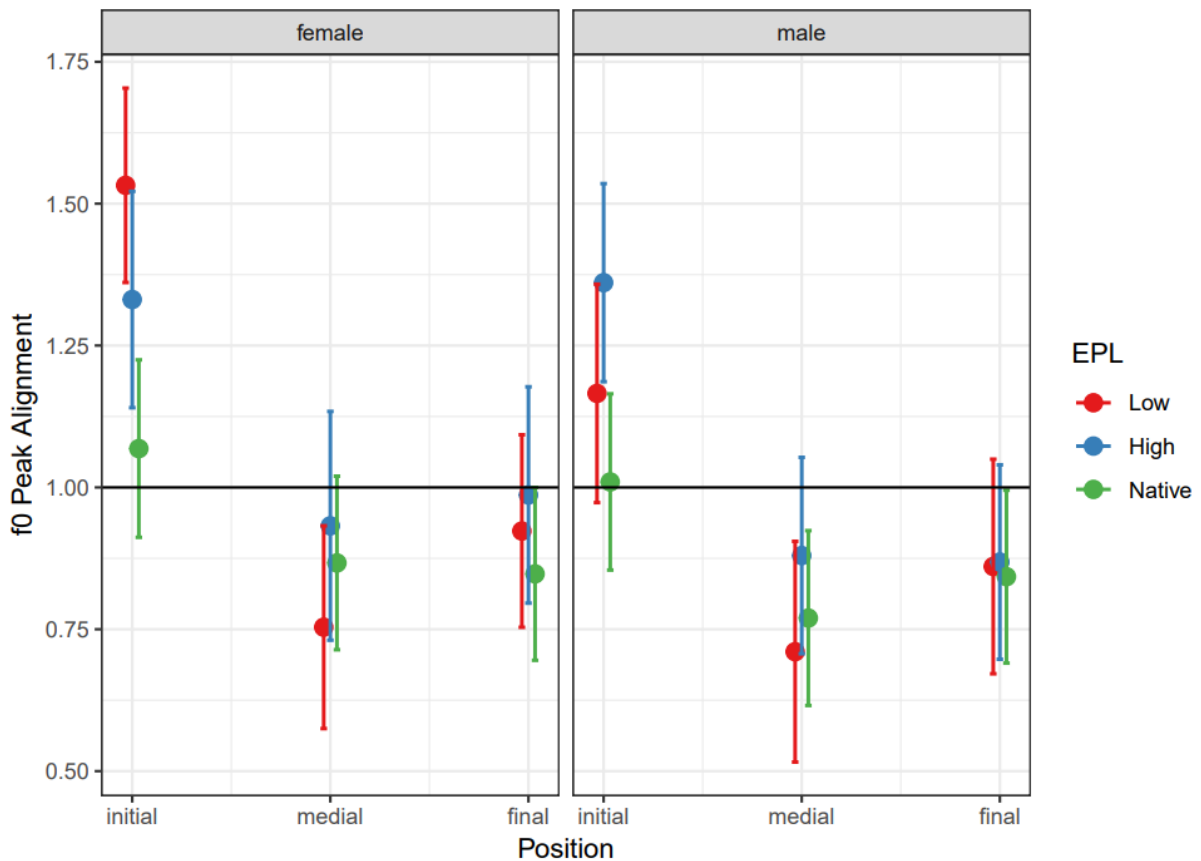
Figure 6. 4 Plot for interaction effects of positions (initial, medial, final) by English proficiency levels (Low, High and Native) for each gender

According to Tukey's HSD post-hoc test, the statistical differences are reported and grouped into sentence-initial, sentence-medial, and sentence-final as follows. The report here takes three-way interactions (Position, Gender, and English proficiency levels) into consideration.

193

1. Sentence-initial

    1.1 Sentence-initial and Thai females with low English proficiency

    1.1.1 Thai females with low English proficiency had different values of f0 peak alignment at the initial position compared to values at the medial position [t = 6.522, df = 30.4, p < 0.05] and final position [t = 5.260, df = 26.5, p < 0.05].

    1.1.2 Thai females with low English proficiency had different values of f0 peak alignment at the initial position compared to values at the medial position [t = 4.036, df = 61.5, p < 0.05] and final position [t = 3.771, df = 54.8, p < 0.05], produced by Thai females with high English proficiency.

    1.1.3 Thai females with low English proficiency had different values of f0 peak alignment at the initial position compared to values at the initial position [t = 5.564, df = 86.6, p < 0.05], medial [t = 5.079, df = 40.2, p < 0.05], and final position [t = 5.235, df = 40.2, p < 0.05] produced by native English females.

    1.1.4 Thai females with low English proficiency had different values of f0 peak alignment at the initial position compared to values at the initial position [t = 3.552, df = 83.4, p < 0.05], medial position [t = 5.626, df = 57.2, p < 0.05], and final position [t = 4.659, df = 54.0, p < 0.05] produced by Thai males with low English proficiency.

    1.1.5 Thai females with low English proficiency had different values of f0 peak alignment at the initial position compared to values at the medial position [t = 4.726, df = 48.1, p < 0.05], and final position [t = 4.829, df = 47.3, p < 0.05] produced by Thai males with high English proficiency.

    1.1.6 Thai females with low English proficiency had different values of f0 peak alignment at the initial position compared to values at the initial position [t = 6.307, df = 84.6, p < 0.05], medial position [t = 5.805, df = 40.6, p < 0.05], and final position [t = 5.272, df = 40.0, p < 0.05] produced by native English males.

    1.2 Sentence-initial and Thai females with high English proficiency

    1.2.1 Thai females with high English proficiency had different values of f0 peak alignment at the initial position compared to values at the medial position [t = 4.349, df = 18.7, p < 0.05], and final position [t = 5.716, df = 18.7, p < 0.05].

194

1.2.2   Thai females with high English proficiency had different values of f0 peak alignment at the initial position compared to values at the medial position [t = 3.349, df = 48.3, p < 0.05]. and final position [t = 3.492, df = 48.1, p < 0.05], produced by native English females.

1.2.3   Thai females with high English proficiency had different values of f0 peak alignment at the initial position compared to values at the medial position [t = 4.058, df = 64.0, p < 0.05] produced by Thai males with low English proficiency.

1.2.4   Thai females with high English proficiency had different values of f0 peak alignment at the initial position compared to values at the medial position [t = 3.105, df = 55.8, p < 0.05] and final position [t = 3.196, df = 55.0, p < 0.05] produced by Thai males with high English proficiency.

1.2.5   Thai females with high English proficiency had different values of f0 peak alignment at the initial position compared to values at the initial position [t = 3.400, df = 82.7, p < 0.05], medial position [t = 4.040, df = 48.8, p < 0.05] produced by native English males.

1.3 Sentence-initial and Thai males with low English proficiency

1.3.1   Thai males with low English proficiency had different values of f0 peak alignment at the initial position compared to values at the medial position [t = 3.648, df = 37.4, p < 0.05].

1.4 Sentence-initial and Thai males with high English proficiency

1.4.1   Thai males with high English proficiency had different values of f0 peak alignment at the initial position compared to values at the medial position [t = 4.061, df = 29.2, p < 0.05], and final position [t = 4.186, df = 28.5, p < 0.05].

1.4.2   Thai males with high English proficiency had different values of f0 peak alignment at the initial position compared to values at the final position [t = 3.444, df = 55.7, p < 0.05] produced by Thai males with low English proficiency

1.4.3   Thai males with high English proficiency had different values of f0 peak alignment at the initial position compared to values at the initial position [t = 4.164, df = 90.5, p < 0.05], medial position [t = 4.460, df = 42.2, p < 0.05],

and final position [t = 3.926, df = 41.6, p < 0.05] produced by native English males.

1.4.4 Thai males with high English proficiency had different values of f0 peak alignment at the initial position compared to values at the initial position [t = 3.450, df = 92.3, p < 0.05] produced by native English females.

2. Sentence-medial

2.1 Sentence-medial and Thai females with low English proficiency

2.1.1 Thai females with low English proficiency had different values of f0 peak alignment at the medial position compared to values at the initial position [t = -3.925, df = 58.9, p < 0.05] produced by Thai females with high English proficiency.

2.1.2 Thai females with low English proficiency had different values of f0 peak alignment at the medial position compared to values at the initial position [t = -4.305, df = 52.8, p < 0.05] produced by Thai males with high English proficiency.

2.2 Sentence-medial and Native English females

2.2.1 Native English females had different values of f0 peak alignment at the medial position compared to values at the initial position [t = -3.737, df = 41.7, p < 0.05] produced by Thai males with high English proficiency.

2.3 Sentence-medial and Thai males with low English proficiency

2.3.1 Thai males with low English proficiency had different values of f0 peak alignment at the medial position compared to values at the initial position [t = -4.421, df = 58.8, p < 0.05] produced by Thai males with high English proficiency.

3. Sentence-final

3.1 Sentence-final and Native English females

3.1.1 Native English females had different values of f0 peak alignment at the final position compared to values at the initial position [t = -3.889, df = 41.6, p < 0.05] produced by Thai males with high English proficiency.

6.4 Rising Slopes of the Rising Accents

To investigate the rising slope of rising accents, the f0 values of the L and H tonal targets were measured for the data collected from Thai learners and native English speakers. The rising slope values were calculated as the f0 difference between the peak (H) and the preceding L, or difference between maximum f0 and minimum f0 as a relative proportion of rising time or difference between low and high (LtoH). To examine the effects of such factors as levels of English proficiency, gender and focus positions, the hypothesis is:

Hypothesis 6.3:   The rising slope of rising accents produced by Thai learners differs from those of native English speakers due to different levels of English proficiency, gender and focus positions.

6.4.1 Rising slopes and effects of English proficiency, gender and focus positions

Table 6.7 shows the descriptive statistics of the relative rising slope for English native speakers or NSE (n = 217) and Thai learners or English L2 Learners or ELL (n = 207). The mean value of the rising slope for the former was 20.40, SD = 8.83 and that of the latter was 24.72, SD = 11.86. Thai learners with low proficiency (n = 105) have a mean rising slope value of 25.57, SD = 10.42, and those with high proficiency (n = 102) have a mean rising slope value of 23.84, SD = 13.17.

Table 6. 7 Descriptive statistics for the relative rising slope of rising accents

Subjects		n	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
NSE		217	20.40	8.83	7.21	13.63	26.40	48.60
ELL		207	24.72	11.86	0.68	15.98	33.04	56.99
	Low	105	25.57	10.42	4.95	17.62	32.94	53.19
	High	102	23.84	13.17	0.68	12.82	32.77	56.99

The main statistical manipulations were performed by a series of LMM analyses with the help of the 'lme4' package in R. In LMM, Speaker or Participant and Word were treated as random factors, while English Proficiency Level (EPL: low, high and native), Position (sentence-initial, sentence-medial and sentence-final), and Gender (female and male) were included as fixed

factors. When their interaction is significant, all group means from the interaction are compared by conducting pairwise comparisons, especially for 3-way interaction effect. That is, focusing on the groups in the interaction better describes the results of the analysis. For the relative rising slope in accented and focal syllables as in Table 6.2, the final full model with the smaller AIC value was chosen and displayed after running a series of model comparison. In this study, many maximal random effects models (e.g., both random intercept and slope) failed to converge or cannot find a solution. In this case, the researcher had to use a simpler model. For example:

model4.1 <-   lmer(RelativeRisingSlope*100 ~ EPL*Gender+Position+
(1+ EPL + Position |Subject)+(1 + EPL + Position|Word),
data = mydata1, REML=FALSE)

model5.1 <-   lmer(RelativeRisingSlope*100 ~ EPL+Gender+Position+
EPL*Gender+(1 + EPL + Position |Subject)+
(1+EPL+Position|Word), data = mydata1, REML=FALSE)

model9 <-   lmer(RelativeRisingSlope*100 ~ EPL+Gender+Position+
EPL*Gender*Position+EPL*Position+
Gender*Position+(1 + EPL + Position |Subject)+
(1 + EPL + Position|Word), data = mydata1, REML=FALSE)

In this case, the researcher had to use a simpler model (e.g., intercepts-only). One common means to test the model's fit is to rerun the analysis but include only the intercept terms which is often called the null model.

model_LHslope <- lmer(RelativeRisingSlope*100 ~ EPL + Gender + Position +
EPL * Gender * Position + EPL * Gender +
EPL * Position + Gender * Position +
(1|Subject) + (1|Word), data = mydata1, REML = FALSE)

nullmodel_ LHslope   <- lmer(RelativeRisingSlope*100 ~ 1+(1|Subject)+ (1|Word),
data = mydata1, REML=FALSE)

Table 6. 8 LMM results for the relative rising slope in target accented syllables

Random Effects					
s^2	32.69				
$\tau_{00 \ Subject}$	65.73				
$\tau_{00 \ Word}$	0.82				
ICC	0.67				
$n_{Subject}$	40				
n_{Word}	12				
Observations	424				
Marginal R^2 / Conditional R^2	0.145 / 0.718				
AIC	2854.578				
BIC	2939.6				
log-Likelihood	-1406.289				
Relative Rising Slope*100					
Predictors	*Estimates*	*std. Error*	*Statistic*	*p*	*df*
(Intercept)	29.24	3.56	8.22	**<0.001**	403.00
EPL[High]	3.12	5.57	0.56	0.576	403.00
EPL[Native]	-8.78	4.48	-1.96	**<0.050**	403.00
Gender[male]	-6.50	5.59	-1.16	0.245	403.00
Position[medial]	-3.52	2.00	-1.76	0.078	403.00
Position[final]	-0.14	1.81	-0.08	0.939	403.00
EPL[High]*Gender[male]	-4.48	7.91	-0.57	0.572	403.00
EPL[Native]*Gender[male]	7.23	6.81	1.06	0.289	403.00
EPL[High]*Position[medial]	-13.75	3.00	-4.58	**<0.001**	403.00
EPL[Native]*Position[medial]	5.14	2.35	2.18	**<0.029**	403.00
EPL[High]*Position[final]	-12.98	2.69	-4.83	**<0.001**	403.00
EPL[Native]*Position[final]	-0.67	2.18	-0.31	0.760	403.00
Gender[male]*Position[medial]	-5.74	2.90	-1.98	**<0.048**	403.00
Gender[male]*Position[final]	3.23	2.69	1.20	0.231	403.00
(EPL[High]*Gender[male])*Position[medial]	24.19	4.15	5.83	**<0.001**	403.00
(EPL[Native]*Gender[male])*Position[medial]	2.44	3.51	0.70	0.487	403.00
(EPL[High]*Gender[male])*Position[final]	9.70	3.85	2.52	**<0.012**	403.00
(EPL[Native]*Gender[male])*Position[final]	-5.24	3.31	-1.58	0.113	403.00

Table 6.8 reveals that there were significant effects of interaction between English proficiency, gender and position on the values of the relative rising slope in target accented syllables. The participants with low English proficiency levels had higher relative rising slope than other groups, for example, native English speakers with [t = -1.96, df = 403.00, p < 0.05]. Differences were also found between Thai learners with high English proficiency and medial position [t = -4.58, df = 403.00, p < 0.05] and final position [t = -4.83, df = 403.00, p < 0.05]. There were also effects of native English speakers and focus positions. That is an interaction of EPL[Native] and medial position [t = 2.18, df = 403.00, p < 0.05]. In addition, there was a statistically significant effect of Gender[male] and Position[medial] [t = -1.98, df = 403.00, p < 0.05]. In terms of 3-way interaction, the results show that there were effects of EPL[High], Gender[male] and Position[medial] [t = 5.83, df = 403.00, p < 0.05], as well as those of EPL[High], Gender[male] and Position[final] [t = 2.52, df = 403.00, p < 0.05] as in Figure 6.5.
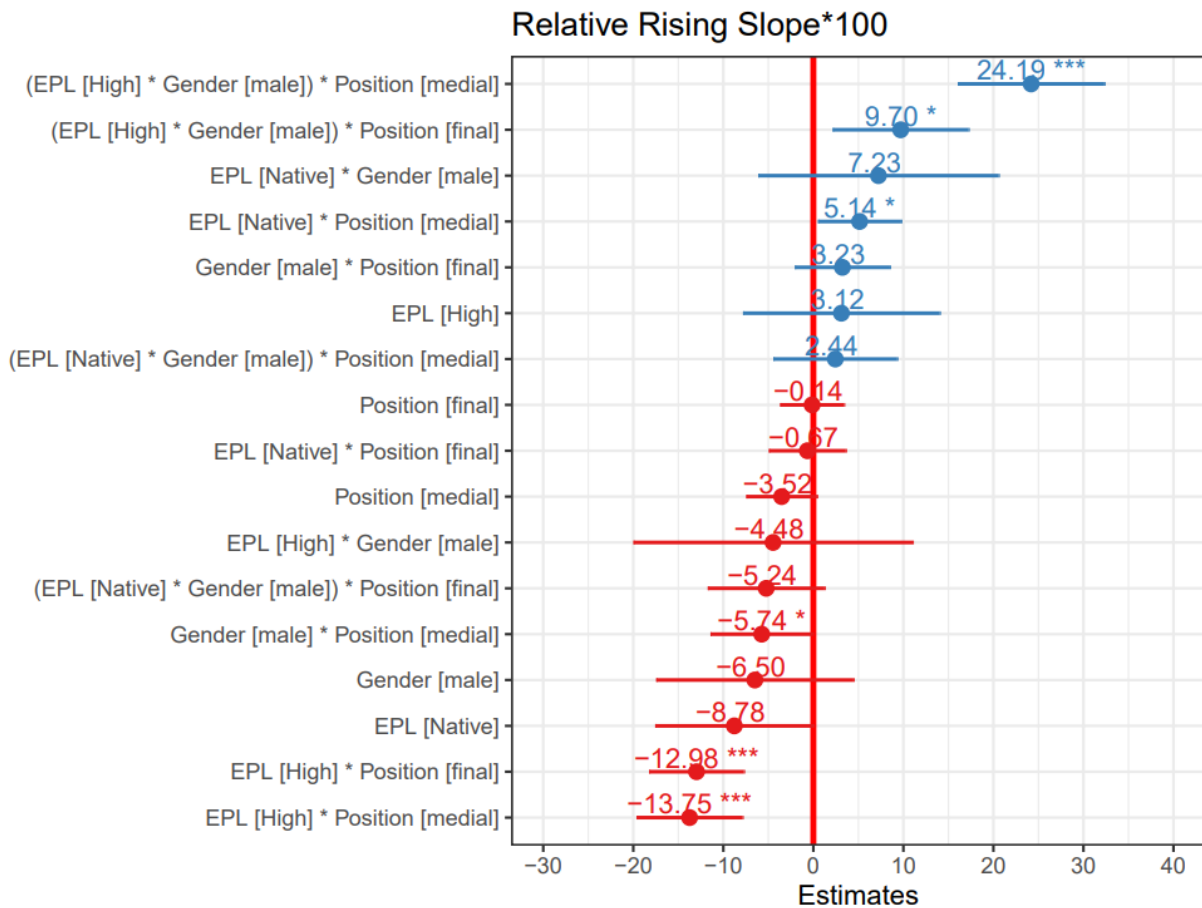
Figure 6. 5 Plot of estimates of model_LHslope from highest to lowest values and the red vertical line indicating no effect

To examine differences among groups, all pairwise comparisons were done only on the optimal model. The results containing the means and SEs of English proficiency or EPL (low, high and native), gender (female and male) and positions (initial, medial and final) for the relative rising slope in target accented syllables are tabulated in Table 5.3. The formula for 'emmeans' is displayed below:

emmslope   <-   emmeans(model_LHslope, specs = pairwise ~ EPL : Gender : Position, adjust = "tukey")

Table 6. 9 Means and SEs of EPL x Gender x Position for the relative rising slope from pairwise comparison based on LMM

Subject groups	English Proficiency	Gender	Focus position	Mean	SE	df
Native English speakers (NSE)	Native			20.22	2.04	49.72
		female		20.74	2.87	48.47
			initial	20.47	3.04	61.10
			medial	22.08	3.00	57.91
			final	19.66	2.99	57.20
		male		19.70	2.87	48.48
			initial	21.20	3.03	60.01
			medial	19.51	3.01	59.08
			final	18.38	2.99	57.19
English L2 learner (ELL) or Thai learners	Low			24.36	2.93	48.56
		female		28.02	3.70	48.21
			initial	29.24	3.85	57.34
			medial	25.72	3.94	62.79
			final	29.11	3.83	56.16
		male		20.69	4.51	47.67
			initial	22.75	4.71	57.47
			medial	13.48	4.73	58.59
			final	25.83	4.68	55.62
	High			21.97	2.93	48.87
		female		22.23	4.53	48.22
			initial	32.36	4.69	56.37
			medial	15.09	4.82	63.34
			final	19.24	4.69	56.37
		male		21.71	3.70	48.19
			initial	21.39	3.89	59.66
			medial	22.57	3.87	58.75
			final	21.19	3.85	57.45

The results can be visualised in Figure 6.6, which shows the effects of independent variables and their interactions on the relative rising slope for the target accented syllables. The plot presents their interactions, wherein the y-axis represents the distribution of the rising slope values by the levels of English proficiency (Low, High and Native). The x-axis represents different prosodic position (initial, medial and final). The dot inside a line corresponds to the mean rising slope, and the plot is also grouped by gender.
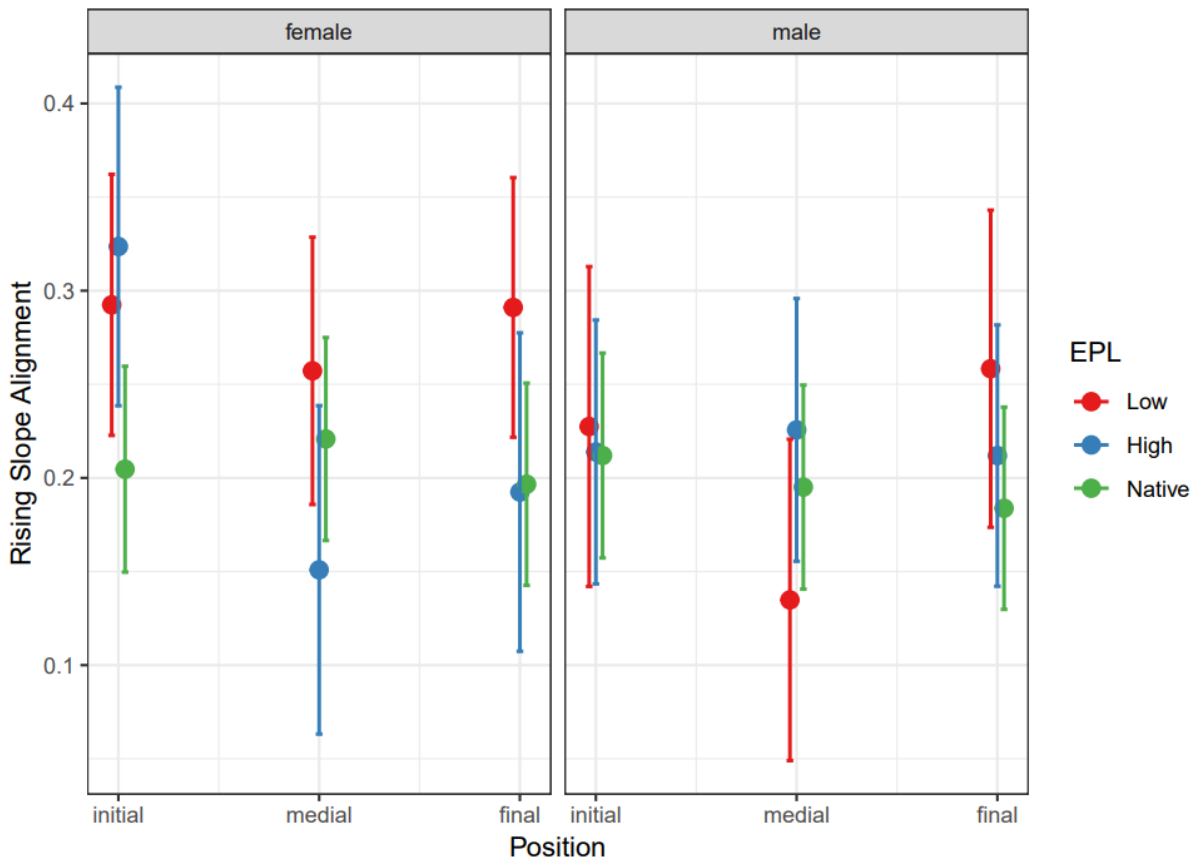
Figure 6. 6 Plot for interaction effects of positions (initial, medial, final) by English proficiency levels (Low, High and Native) for each gender

According to Tukey's HSD post-hoc test, the statistical differences of a significant pair only are reported and grouped into sentence-initial, sentence-medial as follows. The report here takes three-way interactions (Position, Gender, and English proficiency levels) into consideration.

1. Sentence-initial

    1.1 Sentence-initial and Thai females with high English proficiency

        1.1.1   Thai females with high English proficiency had different values of f0 slope alignment at the initial position compared to values at the medial position [t

= 6.936 , df = 281.2 , p < 0.05] and final position [t = 5.806 , df = 239.6 , p < 0.05].

    1.2 Sentence-initial and Thai males with low English proficiency

        1.2.1   Thai males with low English proficiency had different values of f0 slope alignment at the initial position compared to values at the medial position [t = 3.891 , df = 253.1 , p < 0.05].

2.   Sentence-medial

    2.1 Sentence-medial and Thai females with low English proficiency

        2.1.1   Thai females with low English proficiency had different values of f0 slope alignment at the medial position compared to values at the medial position [t = 2.007, df = 58.5, p < 0.05] produced by Thai males with low English proficiency.

    2.2 Sentence-medial and Thai males with low English proficiency

        2.2.1   Thai males with low English proficiency had different values of f0 slope alignment at the medial position compared to values at the final position [t = -5.337 , df = 245.5, p < 0.05].

6.5 Summary of the Results from Statistical analysis

The following section presents the results of investigating the alignment of (LH) tonal targets in rising accents produced by Thai learners and native speakers of English. The alignment of L and H onto segmental strings, as well as rising slopes were examined to see whether a rising accent produced by Thai learners differs from that of native English speakers due to different levels of English proficiency, gender and focus positions.

- *F0 valley alignment*

In terms of L target, the alignment of L was timed with the onset of the accented syllable. There were significant differences in the alignment of L between the two L1 groups. Due to levels of English proficiency, the low valley of Thai participants was aligned later than native English speakers, especially female groups. At the same time, the low valley of Thai female participants was aligned later than the male groups of native English speakers and Thai male participants.

Apart from final position, the low valley of Thai female participants regardless of English proficiency was timed later than native English females at sentence-medial position. Native English speakers (both males and females) had differences in the alignment of f0 valley even at the same focus position. Interestingly, the results showed that the participants in this study anchored the f0 valley with the onset of the target focused word later at the beginning than at the end of the sentences. The significant directions of differences in the alignment of the low valley in accented syllables can be seen in the table below.

Table 6. 10 The significant directions of differences in the alignment of the low valley in accented syllables

Position	Directions of Differences	Explanations
Initial	Native English f. > Native English m.	EPL is English proficiency. f. is for female.
Medial	Thai f. w/high EPL > Native English f. Thai f. w/high EPL > Thai m. w/high EPL	m. is for male. > is for directions of differences
Final	Thai f. w/low EPL > Native English f. Thai f. w/low EPL > Native English m. Thai f. w/high EPL > Thai m. w/high EPL Thai f. w/high EPL > Native English m. Thai f. w/high EPL > Native English f.	showing that the lefthand-side group of > takes more time to align f0 valley with accented syllables. Their f0 valley is still timed within the focused words.

- *F0 peak alignment*

In terms of H target at the sentence-initial position, the alignment of H was obviously affected by the later alignment of low valley, therefore, it reflected the phenomenon called 'the peak delay'. The alignment of f0 peak at sentence-medial and sentence-final produced by all the participants regardless of English proficiency level and gender, was timed well within the target focused syllables as the lower values of f0 peak alignment relative to the syllable duration. However, differences in the values and patterns of alignment of H can be observed within the group when taking English proficiency levels, gender and positions into account. The significant directions of differences in the alignment of the f0 peak in accented syllables can be seen in the table below.

Table 6. 11 The significant directions of differences in the alignment of the f0 peak in accented syllables

Position	Directions of Differences	Explanations
Initial	Thai f. w/low EPL > Native English f. Thai f. w/low EPL > Thai m. w/low EPL Thai f. w/low EPL > Native English m. Thai f. w/high EPL > Native English m. Thai m. w/high EPL > Native English m. Thai m. w/high EPL > Native English f.	EPL is English proficiency. f. is for female. m. is for male. > is for directions of differences showing that the lefthand-side group of > has their H target delayed into the focused words.
Medial	-	Their f0 peak alignment values are higher.
Final	-	

- *Rising slopes*

The results of rising slope presented the comparison of the rises in target focused words by English proficiency level, gender and position. It showed overall differences between the two groups in the realisation of rising accents. The L alignment was anchored to the beginning of focused words from most time-aligned final position, to less time-aligned medial position, and to least time-aligned initial position, for all participants regardless of English proficiency and gender, whereas the H alignment showed significant delay in the focused words from most peak-delayed initial position, to less peak-delayed final position, and to least peak-delayed medial position. The significant directions of differences in the alignment of rising slope in accented syllables can be seen in the table below.

Table 6. 12 The significant directions of differences in the alignment of rising slope in accented syllables

Position	Directions of Differences	Explanations
Initial	-	EPL is English proficiency. f. is for female. m. is for male. > is for directions of differences showing that the lefthand-side group of > uses a large pitch range (i.e., the rising accents with higher values) to mark the focused words.
Medial	Thai f. w/low EPL > Thai m. w/low EPL	
Final	-	

The values of rising slope for Thai females with low English proficiency regardless of English proficiency and sentence positions were higher than other groups of female participants and all male participants. Except for when looking closely at the values of f0, Thai females with high English proficiency across the participants had the significantly highest scaling of H target sentence-initially. This can reflect the use of large pitch range as well as large rising slope to mark narrow focus for this group. Noticeably, the values of rising slope for Thai males regardless of English proficiency and sentence positions were higher than native English males. Surprisingly, the scaling of f0 values of H target by Thai males with low English proficiency and Thai females with high English proficiency dropped the significantly lowest sentence-initially and then rose up higher again sentence-finally. This issue is open for future investigation.

Overall, the alignment of rising slope varied according to differences across participants. The rises in rising accents in narrow focused words produced by Thai participants were significantly longer or higher than the rises produced by native English speakers. Difference between both groups of participants suggests that they may be producing the different phonological category in rising accents.

Table 6.13 below summarises the results of post-hoc mean-separation tests for significant interaction effects between English proficiency levels, gender and positions on the alignment of the low valley from statistical analyses. Table 6.14 for the alignment of the high peak and 6.15 for relative rising slope in rising accents are produced by Thai learners and native English speakers. English proficiency levels include low (l), high (h) and native (n). Gender includes female (f) and male (m). Focus positions include sentence-initial (initial), sentence-medial (medial), and sentence-final (final). The asterisk * indicates statistically significant differences at $p < 0.05$.

Table 6. 13 Summary of the results of interaction effects between English proficiency levels, gender and positions on the alignment of the low valley in accented syllables

			Initial						Medial						Final					
			F			m			f			m			f			m		
			l	h	N	l	h	n	l	h	n	l	h	n	l	h	n	l	h	n
Initial	f	l																		
		h																		
		n						*		*							*			*
	m	l																		
		h																		
		n																		
Medial	f	l																		
		h								*			*							
		n																		
	m	l																		
		h																		
		n																		
Final	f	l															*			*
		h															*		*	*
		n																		
	m	l																		
		h																		
		n																		

Table 6. 14 Summary of the results of interaction effects between English proficiency levels, gender and positions on the alignment of the high peak in accented syllables

			Initial						Medial						Final					
			f			m			f			m			f			m		
			l	h	n	l	h	n	l	h	n	l	h	n	l	h	n	l	h	n
Initial	f	l			*	*		*	*	*	*	*	*	*	*	*	*	*	*	*
		h				*		*		*	*		*	*		*	*		*	
		n																		
	m	l										*								
		h			*			*					*	*				*	*	*
		n																		
Medial	f	l		*			*													
		h																		
		n					*													
	m	l					*													
		h																		
		n																		
Final	f	l																		
		h																		
		n				*														
	m	l																		
		h																		
		n																		

207

Table 6. 15 Summary of the results of interaction effects between English proficiency levels, gender and positions on the alignment of rising slope in accented syllables

			Initial						Medial						Final					
			f			m			f			m			f			m		
			l	h	n	l	h	n	l	h	n	l	h	n	l	h	n	l	h	n
Initial	f	l																		
		h								*							*				
		n																			
	m	l										*									
		h																			
		n																			
Medial	f	l										*									
		h																			
		n																			
	m	l																*			
		h																			
		n																			
Final	f	l																			
		h																			
		n																			
	m	l																			
		h																			
		n																			

6.6 Summary of the Chapter

The present chapter has reported the results of tonal alignment and scaling of rising accents in the expression of focus. To answer the research questions concerning factors affecting tonal alignment and scaling, the chapter has examined the alignment of the low valley and high peak, and the scaling of low and high tonal targets as the rising slope.

The study also conducted statistical analyses to examine the effects of different factors such as sentence positions (e.g., initial, medial, and final), gender and English proficiency levels on the rising accents. The results revealed that all of these factors influenced the phonetic realisation of rising accents used for the marking of narrow-focused words produced by Thai learners and native English speakers. These reflected differences in the later or earlier alignment of f0 valley or peak delay across participants and their effects also varied according to individual differences.

Chapter 7. Discussions

7.1 Introduction

This chapter begins with a discussion concerning the results as to the use of accent types to mark narrow focus as produced by native English speakers and native Thai learners (Chapter 4), as well as similarities and differences due to their levels of English proficiency. Section 7.3 considers the results for the use of acoustic cues for focus as produced by native English speakers and native Thai learners (Chapter 5), as well as the factors that affect the use of such acoustic cues. Section 7.4 concerns the tonal alignment and scaling produced by native English speakers and native Thai learners (Chapter 6), as well as the factors that affect the characteristics of tonal alignment and the scaling of rising accents. Section 7.5 then discusses to what extent L2 intonation learning theory (LILt) and the contrastive analysis hypothesis (CAH) account for the findings concerning the L2 English intonation for focus in this study. A summary of the chapter is provided in section 7.6.

7.2 Use of Accent Types in L1 and L2 English

This section discusses the findings of the study concerning the use of accent types for the marking of focus in L1 and L2 English. These findings are used to answer the following research questions as well.

RQ1:

What are the accent types that native speakers of English use to mark narrow informative focus?

RQ2:

What are the accent types that Thai learners of English use to mark narrow informative focus?

RQ3:

To what extent do Thai learners with different levels of English proficiency differ from native speakers in terms of the use of accent types to mark narrow informative focus?

After examinations of how Thai learners and native speakers of English express the focus on information, it is revealed that, in terms of pitch accent types, native English speakers predominantly employed rising accents such as L+H* bitonal accent, L*+H bitonal accents, or even simply H* monotonal accents when marking narrow focus, as reported in a number of other studies pertaining to focus marking prosodically in well-established English varieties (Arvaniti and Garding, (to appear); Pierrehumbert and Hirschberg, 1990; Ladd, 2008). In the case of narrow focus, those rising accents were placed on the nuclear syllables or words that were positioned at the beginning (Subject), the middle (Verb), and the end (Object) of the sentences. In the case of broad focus, native English speakers in this study generally accented the first and the last content word in relation to a common pattern for marking broad focus in English, as explained in the work of Ladd (2008). Thai learners in this study tended to mark focus prosodically. That is, in terms of accent types, they used the rising bitonal accents (L+H*, L*+H, L+<H*), the H* monotonal accent and even L* monotonal accent to signal focused words. However, they differed from native English speakers in terms of accent placement in broad focus. Thai learners placed accents or accentual prominence so as to mark this kind of focus on several words and at different positions in sentences. Otherwise, they placed accents on every content word. They used different accent types and different locations of accent placement across their own production of the same sentence (i.e., in the three repetitions). This observation is in line with the finding of Wennerstrom (1994) that Thai learners used intonation or accent placement inconsistently to signal meaning in their discourse. Notably, rising accents that were employed by native English speakers and native Thai learners appeared to have the same patterns phonologically. In the absence of a rising pitch movement or bitonal rising accent on accented words, native English speakers and Thai learners of English mostly used a monotonal high accent (H*). Thai learners also employed L accents in this study whereas native English speakers in this study did not. This L* accent can happen since the nuclear tone can be high or low tone represented by the diacritic *.

Rather than accent choices only (H*, L*+H, L+H*, and L+<H*), all of the speakers also employed other strategies in the expression of focus. When it comes to other strategies for marking focus, the speakers in this study differed widely. It seems that they employed several means of marking focus and were not limited to the use of a particular strategy. For example, native English speakers were mostly found to make use of both pre-focal and post-focal deaccenting, and pre-nuclear accenting, as well as compressed pitch range without deaccenting post-focal materials. Prosodic phrasing was another strategy that could be found as produced by native English speakers in this study. Thai learners were also found to employ various

strategies to accompany the accenting that they used to mark focus. These strategies included, but are not limited to, lengthening and phrasing with pauses due to native language interference (Abramson, 1979; Luksaneeyanawin, 1983, 1998). Phrasing or grouping words into small meaningful pieces of information was frequently used, as was pitch range compression without deaccenting. Thai learners rarely made the most of pre-focal or post-focal deaccenting which native English speakers normally do to indicate the focused elements.

The phenomenon of Thai learners using a rising accent and a high accent to mark focus in English can be considered a reflection of first language transfer, since Thai has an extra high lexical tone used for emphatic situations such as when expressing emotion and attitudes. The use of prosodic phrasing can also be considered a reflection of first language transfer, since in Thai a long piece of information tends to be divided into smaller meaningful chunks by means of pausing in speaking to signal the highlighted parts of information. The use of a compressed pitch range without deaccenting can also reflect the characteristics of Thai learners of L2 English in that Thais exhibited inconsistency in their accent placement in sentences. The use of deaccenting and prenuclear accenting can possibly be considered characteristics of L2 English produced by Thai learners in approximating L1 English, since both types of strategies were predominantly used by native English speakers.

At this point, one can see that marking of focus can be achieved by various different choices of accents, as well as using strategies such as pre-focal, or post-focal deaccenting. This reflects differences between languages in the systematic dimensions of intonation, as proposed by Ladd (2008) and as in Mennen's (2015) L2 intonational learning theory (LILt). There are also variations in the use of accent choices/types among individual speakers both within the same first language and across languages.

To examine whether or not Thai learners with different L2 proficiency employed accent types differently from native English speakers, levels of English proficiency were first taken into an account. The finding of this study suggested that Thai learners' levels of English proficiency are not the only factor that influence differences in the choices of accent types used to mark narrow focused words, but also the location or positions in the sentence where the speakers wanted to highlight.

As mentioned earlier, native speakers of English marked narrow focus at different positions in the same sentence with different accent types. In another way round, they also marked narrow

focus at the same positions in the same sentence with the same accent types. This happened from time to time, from sentence to sentence, across participants, since the task in this study was reading-aloud with three repetitions. The same was true for Thai learners of English. Therefore, regardless of first language group that the participants belonged to and different positions in the sentence that an accent occurred, the common set of accents for marking narrow focus produced by both groups consisted of the H* accent, the scooped accent (L*+H), the rising accent (L+H*), and the rising accent (L+<H*) with a delayed peak.

When considering another factor, such as English proficiency levels, regardless of positions, it goes without saying that there were differences between the two groups: native English speakers versus native Thai learners, to some extent in terms of accent types and accent placement. Thai learners with low and high English proficiency differed from native English speakers (as in Chapter 4). The point is that there was an interaction effect of English proficiency levels and focus positions found in this study. The discussion, therefore, should be based on this finding.

At the beginning of the sentence, native speakers of English used:

H*, L*+H, L+H*, and L+<H*

Thai learners of English employed the same set of accent types above to mark narrow focus. These accents included:

H*, L*+H, L+H*, and L+<H*

At the sentence-medial position, there were only three accent types that were used predominantly by native speakers of English:

H*, L+H*, L+<H*

As for Thai learners, they still used:

H*, L*+H, and L+H* with one additional accent; L*

At the sentence-final position, native speakers of English still employed:

H*, L+H*, and L+<H*

while the accents produced in the Thai learner data were:

H*, L+H*, and L+<H*

Overall, although such factors as English proficiency levels and position all together had effects on using accent types to highlight information, it was less clearly indicative from the statistical results that Thai learners had accent types different from native English speakers. This is also because qualitative analysis supported evidence that Thai learners, regardless of high and low English proficiency, tended to use the same phonological categories of tonal targets as the

categories used by native English speakers to mark narrow focus. It is likely to be only which accent type is more common for them. That led to an interesting question whether or not they produced the rising accents differently in phonetic-acoustic ways.

7.3 Use of Acoustic Cues to Focus in L1 and L2 English

This section discusses the findings of the study concerning the use of acoustic cues when marking focus in L1 and L2 English. These findings can then answer the following research question as well.

RQ 5:
To what extent do factors such as level of English proficiency, gender and focus position affect the acoustic characteristics of focus used by Thai learners for the marking of narrow informative focus in terms of f0, intensity and duration?

After investigating the effects of factors such as L2 proficiency level, gender and focus positions on the acoustic properties of narrow focus on target words, the findings show that, besides increasing the degree of f0, intensity and duration, marking focus this way was accomplished in different ways. This can be attributed to differences between individuals. This study shows that native English speakers and Thai learners all employed f0, intensity and duration as defined under the scope of the study. For example, they produced high peaks in focused words, but the height of the peak varies by gender, English proficiency level, and position. In particular, there was more often variation between groups than within groups. When considering gender, female subjects had higher f0 than male subjects, and when considering across language groups but the same gender, non-native females or Thai female subjects had higher f0 than native English female subjects. This was almost the same case between Thai male subjects and native English males.

Taking L2 proficiency into account, Thai learners with high levels of English proficiency did not produce intensity to mark focus more differently from Thai learners with low levels of English proficiency. However, Thai learners regardless of English proficiency levels produced intensity differently from native English speakers. This may be evidence that native and non-native speakers of English had different intensity, e.g., mostly higher for Thai learners.

Duration showed much greater variation across subjects (gender), across English proficiency and across position in this study. Meanwhile, Thai learners did mark focus by using f0 differently from each other and native speakers. In most cases of positions, they had longer duration than native English speakers. This may be because of word lengthening in Thai as explained below.

When it comes to linguistic context, position in the sentence plays a crucial role in how focus is marked. Values of f0, intensity and duration were higher when the speakers, and in particular the Thai learners, arrived at the point of focus. Degrees of f0, intensity, and duration also vary according to sentence position. For example, they produced a high peak in the target accented words in the initial positions and another peak with lower f0 in the middle or at the end of sentences and vice versa. The former creates a phenomenon like a stepping-down f0, and the latter creates a prenuclear accent. This is also a cause of variation in the use of f0, intensity and duration.

Duration was one of the more reliable acoustic cues to prominence as well as f0 and intensity in the primary stressed syllables of the focused accented words. These acoustic cues were found across native English speakers and Thai learners; however, in this study, Thai learners tended to produce much longer duration, greater loudness and higher f0 values than native English speakers. This manipulation of duration or lengthening of the stressed syllables of focused words as used by Thai learners is in line with the Thai literature which points out that the Thai language itself uses syllable duration as one of the most salient features to separate stressed from unstressed syllables (Hiranburana, 1971; Luangthongkum, 1977; Luksaneeyanawin, 1998)). Overall, these characteristics of duration can be attributed to L1 transfer. For instance, Korean L1 speakers in the work of McGory (1997) employed duration as a cue to accentual prominence, whereas Vietnamese L1 speakers did not use it due to the limited use of this acoustic cue in their native language.

At this point, one can see that, even though the marking of focus can be achieved by manipulating f0, intensity and duration, differences in the acoustic realisation of focus can also vary due to factors such as English proficiency, gender, and focus position. And in this study, it found that f0 could indicate differences in gender, intensity could indicate difference in the state of being native or non-native, and duration could indicate difference in position, and vice versa. Therefore, it supported the hypothesis stating that:

Thai learners use f0, duration and intensity as the acoustic parameters to mark focus. The use of f0, duration and intensity may differ from native speakers due to language groups, levels of English proficiency, gender and focus positions. Duration could be greater since in Thai duration is one of the most prominent features used to distinguish stressed and unstressed syllables (Luksaneeyanawin, 1998).

7.4 Tonal Alignment and Scaling of Rising Accents in L1 and L2 English

This section discusses the findings of the study concerning the phonetic realization of focus via the use of rising accents in terms of tonal alignment and scaling in L1 and L2 English. These findings can be used to answer the following research question as well.

RQ 6:

To what extent do factors such as level of English proficiency, gender and focus position affect the temporal relations between the tonal movement in rising pitch accents and segmental strings in narrow-focused words produced by Thai learners?

After investigating the effects of factors such as English proficiency level, gender and position on the tonal alignment and scaling of rising accents for narrow focus in target words, the findings show that these factors influenced variations in the characteristics of rising tonal targets. In terms of f0 valley (L) alignment and taking L2 proficiency levels into consideration, it found great variations within and across subject groups, but variation across groups was much greater. Although in this study native English females differed from each other, they also differed from native English males and Thai learners with high and low English proficiency levels, regardless of gender of Thai learners. The values of f0 valley alignment, therefore, varied resulting in either earlier or later alignment of f0 valley.

Similarly, Thai learners did perform differently from each other, especially for Thai females with high English proficiency. They not only differed from Thai males with high English proficiency, but also native English speakers (both men and women). Thai females with low English proficiency differed from native English speakers. Thai learners in this study tended to align f0 valley very late with the beginning of target accented words. In this sense, Thai learners did not perform differently from other Thai learners when not taking any factors into an account.

Definitely, differences in the alignment of f0 valley varied when English proficiency levels, gender and position were included.

With respect to f0 peak alignment, native English speakers did not perform differently from each other. That is, they mostly aligned f0 high peak (H*) with target accented syllables or words regardless of gender and positions. This showed no difference within subject groups of the same L1. In the case of Thai learners, the finding showed similarities and differences in the alignment of f0 peak which were influenced by English proficiency levels, gender and positions. Whereas Thai females with high and low English proficiency tended to align f0 peak within target focused words in the same way as Thai males with high and low English proficiency, they performed differently from each other when taking into consideration positions or locations of target accent words (or location of accent placement).

In this study, it was sentence-initial position that caused the most differences across subject groups, gender, levels of English proficiency and positions. For example, Thai females with high and low English proficiency had later alignment of f0 peak or peak delay at the beginning of sentence as did Thai males with high and low English proficiency. However, their alignment of f0 peak differed from the f0 alignment of native English speakers and those of the Thai learner group itself when considering positions for marking focus.

As in Chapter 4, Thai learners choose the right nuclear accent (rising accents: L+H*, L*+H, L*+H) for marking a particular accented syllable or word. However, they placed the right nuclear tone differently from native English speakers, such as somewhat of an earlier or later alignment of f0 peak, as in Chapter 6. This is known as a misalignment in intonation (Mennen, 2006, 2007). In this case of so-called misalignment in intonation, when looking closely at the issue about the effect of L2 proficiency level, such misalignment may be attributed to lack of sufficient knowledge in the matter of how to mark focus via pitch prominence, and this is the case for Thai learners. This does not mean that L2 learners with high proficiency will align f0 peak with accented syllables better because the finding of Chapter 6 did not confirm it. It just showed that there was variation in the f0 peak alignment due to other factors such as positions or location of accent placement.

As in Chapter 5, gender affected differences in the scaling or value of low and high tonal targets. Female subjects exhibited higher scaling of high tonal targets and males exhibited lower scaling of high tonal targets. When it comes to rising slope of the rising accents produced by native

English speakers and native Thai learners, positions in the sentence plays a crucial role in influencing the height of f0 slope. For the rising pitch accents (L+H*, L*+H, L*+H) in this study, L alignment was anchored to accented syllable onsets for most speakers. There was a significant peak delay in the alignment of H peak found in the data from Thai learners with low and high English proficiency, especially in sentence-initial positions. Difference in phonetic realisation of the rising accent like this might reflect that Thai learners were using another type of rising accent to mark narrow focus. Rising accent of this kind used by Thai learners was probably posited as an L+<H* accent (a rising accent with a delayed peak).

In addition, the H tone on focused and accented syllables produced by Thai learners was scaled and aligned higher, reflecting the use of pitch range as a strategy to mark this kind of focus. In this study, Thai learners showed higher values of f0, in particular at the initial position of either a minor intonational phrase (ip) or a full intonational phrase (IP). At this point referring back to the statistical analysis, even though the marking of focus can be achieved by using rising accents, differences in the tonal alignment and scaling of the rising accents can vary to some extent due to factors such as English proficiency, gender and focus position. The conclusions concerning L1 influence in terms of tonal alignment and scaling still need more support from Thai evidence. And it is likely to be the case when considering the fact that tonal alignment is language- or dialect-specific from the findings of recent studies (Grabe *et al.,* 2000; Atterer and Ladd, 2004; Fletcher *et al.,* 2005)

Overall, these findings of the present study are in line with those of previous studies of tonal alignment and segmental anchoring (Arvaniti *et al.,* 1998; Ladd, 2006; Ladd *et al.,* 2009; D'Imperio, 2012), in the sense that the L element of rising gesture was aligned with the onset of syllables and the H element was within duration of the accented syllables. It supported part of the hypothesis stating that:

> Thai learners may choose the right nuclear accent for a particular accented syllable or word, but they may place that right nuclear tone differently from native English speakers somewhere early or late, which is known as a misalignment in intonation (Mennen, 2006, 2007). However, it was less clearly indicative from the statistical results indicating that levels of English proficiency, gender and focus positions influences the tonal alignment and scaling of rising accents.

7.5 L2 Intonation Learning theory (LILt) and the Contrastive Analysis Hypothesis (CAH) account for the L2 English intonation for focus

This section would first attempt to discuss the findings of the study on the basis of the Contrastive Analysis Hypothesis (CAH), the Perceptual Assimilation Model (PAM), and finally, Intonation Learning theory (LILt). This would answer the following research question as well.

RQ4:
To what extent do the L2 Intonation Learning theory (LILt) and the Contrastive Analysis Hypothesis (CAH) account for the L2 English intonation for focus in this study?

Regarding the Contrastive Analysis Hypothesis (CAH), it predicts that L2 learners can encounter difficulties when L2 sounds do not exist in their L1 sound, and that L2 learners can find L2 sounds easy to produce when L2 sounds have shared characteristics with their L1 sounds. If this applies to learning L2 intonation or the use of accent types to mark focus, it follows that intonation patterns or accent types in their L1 would facilitate their production of L2 intonation. In other words, the patterns of intonation or types of accents in L1 are ready to use, and they simply recall them. This process can be accomplished by positive transfer (Major, 1987; Wells, 2006).

In this study, Thai learners have high tone, rising tone or even extra high tone, so they do not encounter problems in using H* accents or rising accents to mark the focus. Meanwhile, acoustic-phonetic characteristics of accentual and focal prominence are factors that cause more deviations in intonation patterns or choices of accents used for marking focus as produced by Thai learners. This is so because Thai learners would rely on duration or lengthening to distinguish sounds while native English speakers use pitch.

When it comes to the use of pitch, Thai learners would produce pitch with higher f0 values causing greater pitch size in some cases, as the result found in Chapter 5 about f0 and Chapter 6 about the alignment of f0 peak and rising slope. Possibly, the notion of CAH is supported partly in terms of L1 transfer, as in the case of the use of duration and high f0. That is, Thai learners attempted to manipulate longer duration and f0 height to mark narrow informative focus, and they expected to approximate the patterns produced by native English speakers. However, they might exaggerate the accented word with the use of such different duration and

f0. Otherwise, the Contrastive Analysis Hypothesis (CAH) failed to account for acoustic-phonetic phenomenon concerning this aspect.

Even though there are the Perceptual Assimilation Model (PAM) and the Second-Language Linguistic Perception (L2LP) model, they are out of interest for the study to make any connections. Apart from different objectives of each learning model, some reasons are mentioned in the literature review in 2.4.4 (Chapter 2), but it can be briefed again as follows.

The Perceptual Assimilation Model (PAM) is considered to be the model that indicates phonetic similarities and differences between the sounds of L1 and L2, in particular articulatory gestures and perception of L2 segments. The Perceptual Assimilation Model-L2 (PAM-L2) is the model used to predict linguistic behaviours of more advanced L2 learners, as well as to predict sound contrasts between the L1 and L2. PAM, PAM-L2 and L2LP take into account L1 influences on L2 speech acquisition and emphasise the description and prediction of issues of perception only, whereas the objective of this present study was based on investigating the issue of production of intonational phonology.

Referring back to the statement of Rasier and Hiligsmann (2007) and Mennen (2015), they said that there have been no models proposed to describe and predict difficulties in producing and perceiving L2 intonation in a direct way. Thus, Mennen proposed the L2 Intonation Learning theory (LILt). The purpose of this model is to explain difficulties that L2 learners have encountered when they are producing L2 intonation. The model builds on the basis of cross-language differences divided into four dimensions which were proposed by Ladd (2008) in the Autosegmental-Metrical theory. The first dimension is the systematic dimension concerning typological similarities and differences in the inventory of phonological categories. This dimension is, for example, about the categories of pitch accents and edge tones, as well as about how different phonological categories combine with one another. The second dimension is the realisational dimension concerning how phonological categories are phonetically realized; for example, in alignment and scaling of tonal targets. The third dimension is the semantic dimension concerning the use of phonological categories or tunes in expressing meaning. For instance, there are the use of a rising intonation to signal yes/no questions and a falling intonation to express ideas and give information. The fourth dimension is the frequency dimension concerning similarities and differences in the frequency of use and distribution of phonological categories. This dimension is, for example, about how frequently one accent type is used rather than other accents. It also includes the frequency and distribution of the use of

boundary tones. In addition, Mennen (2015) suggested that the L2 Intonation Learning theory (LILt) can allow for a systematic comparison between L2 learners with different levels of proficiency, different ages of arrival, different L1, or any other relevant variables.

In this study, only the first two dimensions are applied to give an account for intonation used by native English speakers and Thai learner to mark narrow focus. The systemic dimension is the first one since it concerns typological similarities and differences in the inventory of pitch accents and edge tones. Thus, it can account for accent choices that Thai learners and native English speakers used to mark focus in this study. From the fact that languages around the world differ in their intonational typology, using a wider range of accent types (L*+H, L+H*, L+<H*, H*, or even L*) to mark narrow focus as produced by Thai learners can be supported, whereas native speakers of English in this study used L*+H, L+H*, L+<H* and H* accents.

At this point, using intonation or rising accents to highlight information is not likely to be problematic for Thai learners. When they want to mark the focus of information, phonological categories of tonal targets in Thai are ready to manipulate. However, the realisational dimension may cause much more variation reflecting different characteristics of rising accents phonetically. From the results of Chapters 5 and 6, it is clear that English L2 learners, like Thais, could mark narrow focus with rising accents similar to native English speakers. The only difference is that Thai learners attempt to approximate the ways native English speakers emphasize information. To be near native, they may deliberately exaggerate native English speakers' patterns of intonation by lengthening words, and using higher degrees of f0 and intensity. The approximation or imitation like this can cause misalignment of f0 valley and f0 peak (peak delay) or in other words, mistakes in nuclear accent placement. Since languages differ in realisational dimension of intonation, later alignment of nuclear accents produced by Thai learners can be supported, just as the use of compressed pitch range or stepping-down accents.

As Mennen (2015) pointed out that the L2 Intonation Learning theory (LILt) can allow for a systematic comparison between L2 learners with different levels of proficiency, the results from statistical analysis are far from making firm conclusions. Even though deviations of L2 intonation can be diminished when L2 proficiency increases, Thai learners with high English proficiency did not perform differently from Thai learners with low English proficiency levels in several cases such as choices of accent types (Chapter 4), intensity and duration (Chapter 5) and alignment of L and H tonal targets (Chapter 6). Thus, the L2 intonation learning theory best

accounts for the findings in terms of systemic and realisational dimensions of intonation, but not in terms of English proficiency levels decreasing deviations of L2 intonation in this study.

In conclusion, if Ladd's (2008) taxonomy of cross-linguistic differences in intonation and Mennen's (2015) L2 intonational learning theory (LILt) are taken into account, this study gives evidence for systematic differences and realisational differences between L1 English and L2 English. That is, categorical and gradient prosodic strategies are employed in the expression of focus of this kind.

7.6 Summary of the Chapter

This chapter presents discussions of the findings concerning several issues. It considers similarities and differences in the use of accent types and deaccenting as produced by native English speakers and native Thai learners. It considers the results for the use of acoustic cues for focus as produced by native English speakers and native Thai learners. It discusses the findings concerning the phonetic realization of focus via the use of rising accents in terms of tonal alignment and scaling in L1 and L2 English. And it discusses the findings of the study on the basis of the Contrastive Analysis Hypothesis (CAH), and Intonation Learning theory (LILt) among others. Factors such as English proficiency, gender and focus position are included.

The discussions can be briefed as follows. First, choices of accents that Thai learners used to mark focus almost came from the same inventory as those of native English speakers. They differed only how often they are used to mark the focus. Accents used were also varied regardless of the factors such as English proficiency, gender and focus positions. Second, Thai leaners employed various acoustic cues to mark focus, but with a differing degree from native English speakers and from Thai learners themselves. Third, Thai learners could align the L element of rising gesture with the onset of syllables and the H element within duration of the accented syllables. They also produced rising accents with an earlier peak or a delayed peak, resulting from misalignment. And finally, the use of accent choices and tonal alignment were linked with Intonation Learning theory (LILt).

Chapter 8. Conclusions

8.1 Introduction

This chapter begins with the implications of the findings of the research for relevant areas in section 8.2, and the primary contributions that the present study has made are discussed in section 8.3. The chapter also considers the limitations of this study in section 8.4. Directions for further study and future research are then recommended in section 8.5.

8.2 Implications of the Findings

- Implications of the findings for syllabus design for TEFL in Thailand

The findings of the current study can shed some light on intonational features in L1/L2 English focus in instructional contexts. It does not attempt to judge which approaches are appropriate to teaching intonation, but it can offer ideas about what intonational elements should be included in L2 English classrooms. In other words, it has pedagogical implications, at least in terms of guidelines when designing and developing language syllabuses for teaching English pronunciation beyond articulatory phonetics, such as using intonation (e.g., H*, L*+H, L+H*, L+<H*) in the expression of focus and other prosodic focus-marking strategies such as deaccenting. Since the current study regards educational institutions and academia in Thailand as its most important audience groups, it aims to provide Thai teachers and linguists with knowledge of how to mark focus in English prosodically. More specifically, it can help teachers of English as a foreign language and researchers in the field understand more of the characteristics of L2 English. Using a compressed pitch range and prosodic phrasing are the main strategies Thai learners employed. Knowing that native speakers deaccenting post-focal materials can be helpful for learning how to highlight information.

'Thais already use high and rising accents to mark focus and it is the teacher's responsibility to make sure that students place accents on focused words to prevent misalignment.'

With a good understanding of which accent types are common for marking focus, and knowledge about mistaken placements of nuclear accents that can cause failures in communication, language teachers can teach their students to be aware of the possible

consequences of using such erroneous patterns. In addition, intonation training programmes with the help of computer-assisted instruction (CAI) can be developed for non-native learners, and especially advanced learners who aim to achieve near-native pronunciation or to develop better cross-cultural communication skills. Finally, comparing the use of accent types and focus-marking strategies in L2 English with L1 English in this way can help curriculum planners to detect differences and similarities among languages and to recognise what to incorporate in foreign language teaching.

- Implications of the findings for research into English as a lingua franca

Having examined L2 English prosodic features and the focus-marking strategies of native Thai speakers in comparison with those of native English speakers, the findings of the present study can be used for reference when conducting further research. They can be used for comparison with the L2 English characteristics produced by other groups of Thai learners or other groups of non-native speakers of English from different L1 backgrounds within the Autosegmental-Metrical (AM) framework and the use of Tones and Break Indices (ToBI) as a model for prosodic transcription. For instance, a comparative study concerning intonation and focus marking conducted in the context of English as the lingua franca spoken in the ASEAN community could exploit the same research methodology as that used in this study to search for universals or variations in prosody, intonation and focus marking among non-native speakers in this community. This could include investigations into differences and similarities in the phonological realisation of focus in terms of accent type, the acoustic cues to focal prominence, as well as phonetic realisation in terms of tonal alignment and scaling.

8.3 Research Contributions

- *General contributions to existing L2 research into intonation and focus marking*

There have been a handful of publications and studies which provide descriptive accounts of intonation and the marking of focus structure as produced by Thai learners, and such research has been conducted within the Autosegmental-Metrical (AM) framework. In addition, existing studies that provide comparative accounts of differences in the prosodic marking of focus between Thai learners and native speakers of English, or even native speakers of other languages, are scarce; especially those which would satisfy researchers interested in the field

223

of intonational phonology. Thus, this study has made general contributions to knowledge concerning the intonational phonology of L2 English in the first instance. It has also made contributions to existing literature in terms of adding descriptive and comparative accounts of intonational differences, such as using phonologically distinct tune types irrespective of semantic differences, as well as the use of focus-marking strategies. In particular, this study adds to the taxonomy of cross-language differences in intonation in the systematic dimensions proposed by Ladd (2008) and the L2 Intonation Learning theory (LILt) proposed by Mennen (2015).

- *Specific contributions to acoustic-phonetic studies on tonal alignment and scaling*

The second contribution is that this study investigates not only the phonological realisation of focus but also the acoustic-phonetic properties of focus in English intonation, including tonal alignment and scaling. The analysis and comparison of the characteristics of focus produced by native English speakers and Thai learners at this acoustic-phonetic level provide important information to L2 researchers who are interested in in-depth investigations to detect the extent to which realisations of focus differ phonologically and phonetically. Above all, the present study is the first to investigate the prosodic marking of focus produced by Thai learners at two different levels of analysis. This study also makes contributions to the AM theory of intonation analysis in terms of adding Thai learner data concerning tonal alignment and scaling, using ToBI as the model for prosodic transcription. The data obtained from this study can serve as empirical evidence to help develop a better understanding concerning the concepts of tonal alignment and scaling to a great extent. In particular, as previously mentioned, it contributes to a taxonomy of cross-language differences in intonation in the realisational dimension proposed by Ladd (2008) and the L2 Intonation Learning theory (LILt) proposed by Mennen (2015).

8.4 Limitations of the Present Study

This study has various different limitations. Factors that could relate to these limitations which were mentioned earlier in the introductory chapter are restated here, as well as some issues encountered during the investigation. Limitations in terms of the experimental design include the following.

1) Design of speech materials

Not all of the stimuli could be constructed with fully voiced words and put into carrier sentences of the same length. The carrier sentences were limited to short declarative sentences. The study was also limited to examinations of the expression of narrow informative focus in declarative sentences.

2) Design of tasks

The material used in this study was limited to question and answer tasks. The researcher used questions to prompt and elicit answers from the participants, but there is an inherent limitation in this strategy in that the participants only read aloud answers from lists.

3) Determining variables

Since a study of prosody involves several prosodic elements of speech, there are various factors which require a study by researchers who are interested in this area. In this present study, factors such as syllable structure, speaking rate, learner's age and length of residence in the UK of Thai learners were not included due to the limitations of material design and the time allowed for the investigation.

4) Acoustic analysis

The present study was limited to investigations into the fundamental frequency, intensity and duration of elements of speech. It did not examine vowel quality, as has been included in other studies of acoustic correlates of prominence in English (e.g., Peterson and Lehiste, 1960; Ladd and Silverman, 1984).

5) Line of research

This study conducted investigations into the use of (or the production of) accent types to mark focus and the relevant acoustic correlates of focus produced by Thai learners and native English speakers. Therefore, it lacks evidence relating to the results of perception experiments, which could examine whether or not such speakers perceive the marking of focus differently, which

would lead to the issue of intelligibility between native English speakers and non-native English speakers.

One problem arose from the participants themselves during the experiment. Some participants happened to move their mouths too close to or too far away from the microphone in front of them. This probably happened from time to time and could be reflected in differences in intensity in outlier data. A further problem occurred while conducting the phonological analysis of Thai learner data. As mentioned earlier, there has been no work published yet on the development of a Thai-ToBI, which would be useful when making cross-language comparisons between first and second languages. It is extremely difficult to detect L1 transfer in or interference with the characteristics of the L2 English produced by Thai learners in the areas under investigation. The present researcher could only make links between the findings obtained from the analysis of data and the relevant literature concerning the L2 English of learners from different L1 backgrounds and to the existing literature concerning Thai where applicable.

8.5 Directions for Further Research

This study has aimed to set the foundations for the study of the prosodic marking of focus in L2 English in general. It can also serve as a starting point for the further analysis of other prosodic features that can be cues to the expression of focus and its realisation. In particular, for Thai scholars and those who are interested in the field of intonational phonology, the following areas of future research would be recommended:

Under the same theoretical framework of the Autosegmental-Metrical (AM) approach and using the same method of prosodic transcription with the Tones and Break Indices (ToBI) system, further experimental investigations into the marking of focus could be conducted in many different ways. Some of the possibilities are listed below:

1) *Changing representative samples*

A future study on the same topic can be extended by collecting data from different groups of participants from the same L1. This could include, for example, Thai learners majoring in

English, and well-educated Thai learners from different disciplines. Alternatively, new variables as stated in 3) below can be considered.

2) Changing tasks

A future study can use other types of tasks and stimuli to collect data concerning the marking of focus. Depending on the research purposes, tasks could include collecting data from controlled or uncontrolled experiments, from reading passages to connected speech, or even short to long conversations. Stimuli could include a piece of writing, maps and pictures to facilitate experimentation.

3) Changing variables

Apart from L1 prosodic transfer, there are several factors such as L2 learners' learning experience and their motivation and goals for learning a target language which can be taken into consideration in future research. Narrow contrastive focus is also suggested for future research in terms of its phonological and phonetic realisation, which involves tonal alignment and scaling. Examining vowel formant patterns is another suggestion if a researcher is interested in acoustic differences in the vowel quality of accented syllables.

4) Conducting cross-language comparisons in L1 and L2 English context

A comparative study of the marking of focus could be conducted or extended between a group of Thai learners of English and either native speakers from other varieties of spoken English or non-native speakers of English from different first languages. English as *a lingua franca* in the ASEAN community is another research area that researchers can consider.

5) Conducting 3-way cross-language comparisons

This kind of research is promising but challenging. As long as there is no Thai-specific ToBI, it would be difficult to conduct a 3-way comparative study such as Thai versus L2 English versus L1 English. As stated in section 8.2, this study regards *'educational institutions and academia in Thailand as its most important audience groups among others'*. It thus aims to encourage Thai linguists and researchers to pay more attention to issues of prosodic elements of speech and consider the benefits of comparing cross-language differences using the standard

system for prosodic annotation such as ToBI as well as considering the development of Thai-ToBI. This would be a useful tool in exploring and investigating cross-language differences and similarities; for instance, between an intonation language such as English or German and Thai; between Thai and other tonal languages such as Chinese; or even between other different typological languages like Korean, Japanese, and French. A variety of topics are still left open for further research in this area.

Finally, when equipped with such a Thai-ToBI, Thai linguists and researchers can conduct in-depth investigations into cross-language differences concerning L1 Thai and L2 English intonation. As proposed in the AM theory, Ladd (2008, p.115) discussed a taxonomy of cross-language differences in four dimensions, including differences in the meaning or use of the same tune (semantic), differences in the inventory of phonologically distinct tune types irrespective of semantic differences (systematic), differences of detail in the phonetic realisation of the same tune (realisational), and differences in tune-text association and in the permitted structure of tunes (phonotactics). All of these dimensions can be linked to the L2 Intonation Learning theory (LILt) proposed by Mennen (2015). A better understanding of and further insight into such cross-language differences in intonation backed up by full empirical evidence may help scholars and researchers of intonational phonology to '*capture the potential prosodic features in English which may cause failure or success in communication between native English speakers and Thai learners of L2 English, as well as between Thai learners and non-native speakers of English around the world*'.

References

Abramson, A.S. (1962) *The vowels and tones of Standard Thai: acoustical measurements and experiments*. Bloomington: Indiana University Research Center in Anthropology, Folklore, and Linguistics, Publication 20.

Abramson, A.S. (1978) 'Static and dynamic acoustic cues in distinctive tones', *Language and Speech*, 21, pp. 319-325.

Abramson, A.S. (1979) 'Lexical tone and sentence prosody', *Haskins Laboratories*: *Status Report on Speech Research*, SR-59/60, pp. 153-160.

Abramson, A.S. and Svastikula, K. (1983) 'Intersections of tone and intonation in Thai', *Haskins Laboratories*: *Status Report on Speech Research*, SR-74/75, pp. 143-154.

Adamou, E. and Arvaniti, A. (2010) 'Language-specific and universal patterns in narrow focus marking in Romani', *Proceedings of Speech Prosody 2010*. Chicago, IL, USA, 10-14 May.

Adams, C. and Munro, R.R. (1978) 'In search of the acoustic correlates of stress: fundamental frequency, amplitude, and duration in the connected utterrances of some native and non-native speakers of English', *Phonetica*, 35, pp. 125-156.

Arvaniti, A. (2012) 'Segment-to-tone association', in Cohn, A.C., Fougeron, C. and Huffman, M.K. (eds.) *The Oxford handbook of laboratory phonology*, Oxford: Oxford University Press, pp. 265-274.

Arvaniti, A. and Baltazani, M. (2005) 'Intonational analysis and prosodic annotation of Greek spoken corpora', in Jun, S.-A. (ed.) *Prosodic typology: the phonology of intonation and phrasing*. Oxford: Oxford University Press, pp. 85-117.

Arvaniti, A. and Garding, G. (to appear) 'Dialectal variation in the rising accents of American English', in Cole, J. and Hualde, J.H. (eds.) *Papers in laboratory phonology 9: change in phonology*. Berlin: Mouton de Gruyter.

Arvaniti, A., Ladd, D.R., and Mennen, I. (1998) 'Stability of tonal alignment: the case of Greek prenuclear accents', *Journal of Phonetics*, 26, pp. 3-25.

Arvaniti, A., Ladd, D.R., and Mennen, I. (2006) 'Phonetic effects of focus and "tonal crowding" in intonation: evidence from Greek polar questions', *Speech Communication*, 48, pp. 667-696.

Atterer, M. and Ladd, D.R. (2004) 'On the phonetics and phonology of "segmental anchoring" of F0: evidence from German', *Journal of Phonetics*, 32(2), pp. 177-197.

Baumann, S. Grice, M. and Steindamm, S. (2006) 'Prosodic marking of focus domains – categorical or gradient?', *Proceedings of speech prosody 2006*, Dresden, Germany, 2-5 May.

Beckman, M.E. and Ayers-Elam, G. (1997) *Guideline for ToBI labelling* (version 3.0) [manuscript and accompanying speech materials]. Available at: http://www.lin.ohio-state.edu/~tobi/.

Beckman, M.E. and Pierrehumbert, J.B. (1986) 'Intonational structure in Japanese and English', *Phonology Yearbook*, 3, pp. 255-309.

Beckman, M.E. and Venditti, J.J. (2011) 'Intonation', in Goldsmith, J., Riggle, J. and Yu, A.C.L. (eds.) *The handbook of phonological theory*. 2nd edn. Chichester, UK: Wiley-Blackwell, pp. 485-532.

Beckman, M.E., Hirschberg, J. and Shattuck-Hufnagel, S. (2005) 'The original ToBI system and the evolution of the ToBI framework', in Jun, S.-A. (ed.) *Prosodic typology: the phonology of intonation and phrasing*. Oxford: Oxford University Press, pp. 9-54.

Best, C.T. (1995) 'A direct realist view of cross-language speech perception', in Strange, W. (ed.) *Speech perception and linguistic experience: issues in cross-language research*. Timonium: York Press, pp. 171-232.

Best, C.T. and Tyler, M.D. (2007) 'Nonnative and second-language speech perception: commonalities and complementarities', in Bohn, O.-S. and Munro, M.J. (eds.) *Language experience in second language speech learning; in honor of James Emil Flege*. Amsterdam: John Benjamins, pp. 13-34.

Boersma, P. and Weenink, D. (2019) *Praat: doing phonetics by computer* (Version 6.1.03) [Computer program]. Available at: http://www.praat.org/ (Accessed: 2 September 2019)

Bolinger, D.L.M. (1958) 'A theory of pitch accents in English', *WORD*, 14, pp. 2-3.

Bolinger, D.L.M. (1972a) *Intonation: selected readings*. Harmondsworth: Penguin.

Bolinger, D.L.M. (1972b) 'Accent is predictable (if you're a mind reader)', *Language*, 48(3), pp, 633-644.

Bolinger, D.L.M. (1989) *Intonation and its uses: melody in grammar and discourse*. Stanford, California: Stanford University Press.

Boonkit, K. (2002) *Listening strategies with television texts: a study of Thai university students of English as a foreign language*. PhD thesis. University of Tasmania.

Brazil, D., Coulthard, M. and Johns, C. (1980) *Discourse intonation and language teaching*. London: Longman.

Breen, M., Fedorenko, E., Wagner, M., and Gibson, E. (2010) 'Acoustic correlates of information structure'. *Language and Cognitive Processes*, 25(7-9), pp. 1044-1098.

Büring, D. (2007) 'Semantics, intonation and information structure', in Ramchad, G. and Reiss, C. (eds.) *The Oxford handbook of linguistic interfaces*. Oxford University Press, pp. 446-474.

Büring, D. (2009) 'Towards a typology of focus realization', in Zimmermann, M. and Féry, C. (eds.) *Information structure: theoretical, typological, and experimental perspectives*. Oxford: Oxford University Press, pp. 177-205.

Büring, D. (20012) 'Focus and intonation', in Russell, G. and Fara, D.G. (eds.) *The Routledge companion to the philosophy of language*. London: Routledge, pp. 103-115.

Cangemi, F. (2015) mausmooth (Version 1.0) [Computer program]. Available at: http://phonetik.phil-fak.uni-koeln.de/fcangemi.html (Accessed: 23 May 2018).

Cauldwell, R. and Hewings, M. (1996) 'Intonation rules in ELT textbooks', *ELT Journal*, 50(4), pp. 327-334.

Chafe, W.L. (1974) 'Language and consciousness', *Language*, 50 (1), pp. 111-133.

Chafe, W.L. (1976) 'Givenness, contractiveness, definiteness, subjects, topics and points of view', in Li, C.N. (ed.) *Subject and topic*. New York: Academic Press, pp. 27-55.

Chen, H.C. (2015) 'Acoustical analyses and intelligibility assessments of timing patterns among Chinese English learners with different dialect background', *Journal of Psycholinguist Research*, 44, pp. 749-773.

Chen, S.W., Wang, B. and Xu, Y. (2009) 'Closely related languages, different ways of realizing focus', *Proceedings of interspeech 2009*, Brighton, UK, 6-10 September. pp. 1007-1010.

Chen, Y. (2006) 'Durational adjustment under corrective focus in Standard Chinese', *Journal of Phonetics*, 34, 176-201.

Chen, Y. (2010) 'Post-focus F0 compression: now you see it, now you don't', *Journal of Phonetics*, 38, pp. 517-525.

Chen, Y., Robb, M.P., Gilbert, H.R. and Lerman, J.W. (2001) 'A study of sentence stress production in Mandarin speakers of American English', The *Journal of the Acoustical Society of America*, 109(4), pp, 1681-1690.

Cheng, L. (1987) *Assessing Asian language performance: guidelines for evaluating limited-English-proficient students*. Aspen: Rockville Co.

Chomsky, N. and Halle, M. (1968) *The sound pattern of English*. New York: Harper and Row.

Clennell, C. (1997) 'Raising the pedagogic status of discourse intonation teaching', *ELT Journal*, 51(2), pp. 117-125.

Cohn, A.C., Fougeron, C., and Huffman, M.K. (2012) *The Oxford handbook of laboratory phonology*. Oxford: Oxford University Press.

Cole, J. (2015) 'Prosody in context: a review', *Language, Cognition and Neuroscience*, 30(1-2), pp. 1-31.

Connell, B. and Ladd, D.R. (1990) 'Aspects of pitch realisation in Yoruba', *Phonology*, 7, pp. 1-29.

Cooper, W.E., Eady, S.J. and Mueller, P.R. (1985) 'Acoustic aspects of contrastive stress in question answer contexts', *The Journal of the Acoustical Society of America*, 77(6), pp. 2142-2156.

Cruttenden, A. (1997). *Intonation*. 2nd edn. Cambridge: Cambridge University Press.

Cruz-Ferreira, M. (1987) 'Non-native interpretive strategies for intonational meaning: an experimental study', in James, A. and Leather, J. (eds.) *Sound patterns in second language acquisition*. Dordrecht: Foris, pp. 103-120.

Crystal, D. (1969) *Prosodic systems and intonation in English*. London: Cambridge University Press.

Cutler, A., Dahan, D. and van Donselaar, W. (1997) 'Prosody in the comprehension of spoken language: a literature review', *Language and Speech*, 40(2), pp. 141-201.

Daly, N. and Warren, P. (2001) 'Pitching it differently in New Zealand English: speaker sex and intonation patterns', *Journal of Sociolinguistics*, 5(1), pp. 85-96.

Derwing, T.M. and Munro, M.J. (1997) 'Accent, intelligibility, and comprehensibility: evidence from four L1s', *Studies in Second Language Acquisition*, 20, pp. 1-16.

Deterding, D. (1994) 'The intonation of Singapore English', *Journal of the International Phonetic Association*, 24(2), pp. 61-72.

Dilley, L.C. (2010) 'Pitch range variation in English tonal contrasts: continuous or categorical?' *Phonetica*, 67, pp. 63-81.

Dilley, L.C., Ladd, D.R., and Schepman, A. (2005) 'Alignment of L and H in bitonal pitch accents: testing two hypotheses', *Journal of Phonetics*, 33, pp. 115-119.

D'Imperio, M. (2000) *The role of perception in defining tonal targets and their alignment*. PhD thesis. Ohio State University, Columbus, USA.

D'Imperio, M. (2001) 'Focus and tonal structure in Neapolitan Italian', *Speech Communication*, 33, pp. 339-356.

D'Imperio, M. (2012) 'Tonal alignment', in Cohn, A.C., Fougeron, C. and Huffman, M.K. (eds.) *The Oxford handbook of laboratory phonology*, Oxford: Oxford University Press, pp. 275-287.

Eady, S.J. and Cooper, W.E. (1986) 'Speech intonation and focus location in matched statements and questions', *The Journal of the Acoustical Society of America*, 80, pp. 402-415.

Ellis, R. (2008) *The study of second language acquisition*. 2nd edn. Oxford: Oxford University Press.

Escudero, P. (2005) *Linguistic perception and second language acquisition: explaining the attainment of optimal phonological categorization*. Unpublished PhD thesis. Utrecht University.

Estebas-Vilaplana, E. (2000) *The use and realisation of accentual focus in central Catalan with a comparison to English*. Unpublished PhD thesis. University College London.

Estebas-Vilaplana, E. (2017) 'The teaching and learning of L2 English intonation in a distance education environment: TL_ToBI vs. the traditional models', *Linguistica*, 57(1), pp. 73-91.

Féry, C. and Kügler, F. (2008) 'Pitch accent scaling on given, new and focused constituents in German', *Journal of Phonetics*, 36, pp. 680-703.

Flege, J.E. (1987) 'The production of "new" and "similar" phones in a foreign language: evidence for the effect of equivalence classification', *Journal of Phonetics*, 15, pp. 47-65.

Flege, J.E. (1991) 'Perception and production: the relevance of phonetic input to L2 phonological learning', in Huebner, T. and Ferguson, C.A. (eds.) *Crosscurrents in second language acquisition and linguistic theories*. Amsterdam: John Benjamins, pp. 249-289.

Flege, J.E. (1995) 'Second language speech learning: theory, findings and problems', in Strange, W. (ed.) *Speech perception and linguistic experience: issues in cross-language research*. Timonium, MD: York Press, pp. 233-277.

Flege, J.E. (2009) 'Give input a chance!' in Piske, T. and Young-Scholten, M. (eds.) *Input matters in SLA*. Bristol: Multilingual Matters, pp. 175-190.

Flege, J.E. and Hillenbrand, J. (1984) 'Limits on phonetic accuracy in foreign language speech production', *The Journal of the Acoustical Society of America*, 76(3), pp. 708-721.

Flege, J.E., Birdsong, D., Bialytok, E., Mack, M., Sung, H. and Tsukada, K. (2006) 'Degree of foreign accent in English sentences produced by Korean children and adults', *Journal of Phonetics*, 34, pp. 153-175.

Flege, J.E., Munro, M.J. and MacKay, I.R.A. (1995) 'Effects of age of second-language learning on the production of English consonants', *Speech Communication*, 16, pp. 1-26.

Fletcher, J. and Evans, N. (2002) 'An acoustic phonetic analysis of intonational prominence in two Australian languages', *Journal of the International Phonetic Association*, 32(2), pp. 123-140.

Fletcher, J., Grabe, E. and Warren, P. (2005) 'Intonational variation in four dialects of English: the high rising tone', in S-A., Jun (ed.) *Prosody typology: the phonology of intonation and phrasing*. Oxford: Oxford University Press, pp. 390-409.

Fletcher, J., Stirling, L., Mushin, L. and Wales, R. (2002) 'Intonational rises and dialog acts in the Australian English map task', *Language and Speech*, 45(3), pp. 229-253.

Fokes, J. and Bond Z.S. (1989) 'The Vowels of Stressed and Unstressed Syllables in Nonnative English', *Language Learning*, 39(3), pp. 341-373.

Foley, J.A. (2005) 'English in Thailand', *Regional Language Centre Journal*, 36(2), pp, 223-234.

Frota, S. (2002) 'Tonal association and target alignment in European Portuguese nuclear falls', in Gussenhoven, C. and Warner, N. (eds.) *Laboratory phonology 7*. Berlin/New York: Mouton de Gruyter, pp. 387-418.

Fuchs, R. and Maxwell, O. (2015) 'The placement and acoustic realisation of primary and second stress in Indian English', *Proceedings of the 18th international congress of phonetic sciences (ICPhS)*, Glasgow, 10-14 August.

Gandour, J. (1974a) 'Consonant types and tone in Siamese', *University of California at Los Angeles Working Papers in Phonetics*, 27, p. 92-117.

Gandour, J. (1974b) 'On the representation of tone in Siamese', *University of California at Los Angeles Working Papers in Phonetics*, 27, p. 118-146.

Genzel, S., Ishihara, S. and Surányi, B. (2015) 'The prosodic expression of focus, contrast and givenness: a production study of Hungarian', *Lingua* 162, pp. 183-204.

German, J., Pierrehumbert, J.B. and Kaufmann, S. (2006) 'Evidence for phonological constraints on nuclear accent placement', *Language*, 82, pp. 151-168.

Goldsmith, J. (1976) *Autosegmental phonology*. PhD thesis. Massachusetts Institute of Technology.

Grabe, E. (1998) 'Pitch accent realization in English and German', *Journal of Phonetics*, 26, pp. 129-143.

Grabe, E. (2004) 'Intonational variation in urban dialects of English spoken in the British Isles', in Gilles, P. and Peters, J. (eds.) *Regional variation in intonation*. Linguistuche Arbeiten, Tuebigen: Niemeyer, pp. 9-31.

Grabe, E. and Post, B. (2002) 'Intonational variation in the British Isles', *Proceedings of speech prosody 2002*, Aix-en-Provence, France, 11-13 April, pp. 343-346.

Grabe, E., Kochanski, G. and Coleman, J. (2005) 'The intonation of native accent varieties in the British Isle: potential for miscommunication?', in Dziubalska-Kołaczyk, K. and Przedlacka, J. (eds.) *English pronunciation models: a changing scene*. Bern: Peter Lang, pp. 311-337.

Grabe, E., Kochanski, G. and Coleman, J. (2007) 'Connecting intonation labels to mathematical descriptions of fundamental frequency', *Language and Speech*, 50(3), pp. 281-310.

Grabe, E., Post, B. and Nolan, F. (2001) *The IViE Corpus*. Department of Linguistics, University of Cambridge. Available at: http://www.phon.ox.ac.uk/IViE.

Grabe, E., Post, B., Nolan, F. and Farrar, K. (2000) 'Pitch accent realization in four varieties of British English', *Journal of Phonetics*, 28, pp. 161-185.

Graham, C. and Post, B. (2018) 'Second language acquisition of intonation: peak alignment in American English', *Journal of Phonetics*, 66, pp. 1-14.

Grice, M. (1995) 'Leading tones and downstep in English', *Phonology*, 12, pp. 183-233.

Grice, M. and Baumann, S. (2007) 'An introduction to intonation: functions and models', in Trouvain, J. and Gut, U. (eds.) *Non-native prosody: phonetic description and teaching practice*. Berlin, Germany: Mouton de Gruyter, pp. 25-51.

Grice, M. and Savino, M. (2003) 'Map tasks in Italian: Asking questions about given, accessible and new information', *Catalan Journal of Linguistics*, 2, pp. 153-180.

Grice, M., Baumann, S. and Benzmuüller, R. (2005) 'German intonation in autosegmental-metrical phonology', in Jun, S.-A. (ed.) *Prosodic typology: the phonology of intonation and phrasing*. Oxford: Oxford University Press, pp. 56-83.

Grice, M., D'Imperio, M., Savino, M. and Avesani, C. (2005) 'Strategies for intonation labelling across varieties of Italian', in Jun, S.-A. (ed.) *Prosodic typology: the phonology of intonation and phrasing*. Oxford: Oxford University Press, pp. 363-389.

Gussenhoven, C. (1983) 'Focus, mode and the nucleus', *Journal of Linguistics*, 19(2), pp. 377-417.

Gussenhoven, C. (1984) *On the grammar and semantics of sentence accents*. Dordrecht: Foris.

Gussenhoven, C. (1985) 'Two views of accent: a reply', *Journal of Linguistics*, 21, pp. 125-138.

Gussenhoven, C. (1999) 'On the limits of focus projection in English', in Bosch, P. and van der Sandt, R. (eds.) *Focus: linguistic, cognitive, and computational perspectives*. Cambridge: Cambridge University Press, pp. 43-55.

Gussenhoven, C. (2004) *The phonology of tone and intonation*. Cambridge, UK: Cambridge University Press.

Gussenhoven, C. (2005) 'Transcription of Dutch intonation', in Jun, S.-A. (ed.) *Prosodic typology: the phonology of intonation and phrasing*. Oxford: Oxford University Press, pp. 119-145.

Gussenhoven, C. (2008) 'Types of focus in English', in Lee, C., Gordon, M. and Büring, D. (eds.) *Topic and focus: cross-linguistic perspectives on meaning and intonation*. Dordrecht: Springer, pp. 83-100.

Gussenhoven, C. and Rietveld, A.C.M. (1985) 'On the relation between pitch excursion size and prominence', *Journal of Phonetics*, 13, pp. 299-308.

Gut, U. and Pillai, S. (2014) 'Prosodic marking of information structure by Malaysian speakers of English', *Studies in Second Language Acquisition*, 36, pp. 283-302.

Halliday, M.A.K. (1967) *Intonation and grammar in British English*. The Hague, Paris: Mouton.

Halliday, M.A.K. (1970) *A course in spoken English: intonation*. Oxford: Oxford University Press.

Hartmann, K. (2008) 'Focus and tone', *Acta Linguistic Hungarica*, 55(3-4), pp. 415-426.

Hass, M.R. (1946) 'Techniques of intensifying in Thai', *WORD*, 2(2), pp. 127-130.

Hayes, D. (2008) 'Becoming a teacher of English in Thailand', *Language Teaching Research*, 12(4), p. 471-494.

Hayes, D. (2010) 'Language learning, teaching and educational reform in rural Thailand: an English teacher's perspective', *Asia Pacific Journal of Education*, 30(3), pp. 305-319.

Hedberg, N. and Sosa, J.M. (2008) 'The prosody of topic and focus in spontaneous English dialogue', in Lee, C., Gordon, M. and Büring, D. (eds.) *Topic and focus: cross-linguistic perspectives on meaning and intonation*. Dordrecht: Springer, pp. 101-120.

Heldner, M. (2003) 'On the reliability of overall intensity and spectral emphasis as acoustic correlates of focal accents in Swedish', *Journal of Phonetics*, 31, pp. 39-62.

Hellmuth, S. (2006) *Intonational pitch accent distribution in Egyptian Arabic*. PhD thesis. University of London.

Hellmuth, S. (2007) 'The relationship between prosodic structure and pitch accent distribution: evidence from Egyptian Arabic', *The Linguistic Review*, 24, pp. 291-316.

Henderson, E.J.A. (1949) 'Prosodies in Siamese: a study in synthesis', *Asia Major*, New Series 1, pp. 189-215.

Herschensohn, J. (2007) *Language development and age*. Cambridge: Cambridge University Press.

Hewings, M. (1995) 'Tone choice in the English intonation of non-native speakers', *International Review of Applied Linguistics in Language Teaching*, 33(3), pp. 251-265.

Hiranburana, K. (2017) 'Use of English in the Thai workplace', *Kasetsart Journal of Social Sciences*, 38, pp, 31-38.

Hiranburana, S. (1971) *The role of accent in Thai grammar*. PhD thesis. University of London.

Hirschberg, J. (2002) 'Communication and prosody: functional aspects of prosody', *Speech Communication*, 36, pp. 31-43.

Hirst, D. and Di Cristo, A. (1998) *Intonation systems: a survey of twenty languages*. Cambridge: Cambridge University Press.

Hirst, D.J. (2005) 'Form and function in the representation of speech prosody', *Speech Communication*, 46, pp. 334-347.

Hualde, J.I. and Prieto, P. (2016) 'Towards an International Prosodic Alphabet (IPrA)', *Laboratory Phonology: Journal of the Association for Laboratory Phonology*, 7(1), 5, pp. 1-25.

Huang, B.H. and Jun, S.-A. (2011) 'The effect of age on the acquisition of second language prosody', *Language and Speech*, 54(3), pp. 387-414.

Hughes, A. and Trudgill, P. (1987) *English accents and dialects: an introduction to social and regional varieties of British English*. 2nd edn. London: Edward Arnold.

Ioup, G. (2008) 'Exploring the role of age in the acquisition of a second language phonology', in Edwards, J.G.H. and Zampini, M.L. (eds.) *Phonology and second language acquisition*. Amsterdam: John Benjamins Publishing Company, pp, 41-62.

Ioup, G., Boustagui, E., Tigi, M.E. and Moselle, M. (1994) 'Reexamining the critical period hypothesis: a case study of successful adult SLA in a naturalistic environment', *Studies in Second Language Acquisition*, 16, pp. 73-98.

Ito, K., Speer, S.R. and Beckman, M.E. (2004) 'Informational status and pitch accent distribution in spontaneous dialogues in English', *Proceedings of speech prosody 2004*, Nara, Japan, 23-26 March, pp. 279-282.

Iwasaki, S. and Ingkaphirom, P. (2005) *A reference grammar of Thai*. Cambridge: Cambridge University Press.

Jackendoff, R.S. (1972) *Semantic interpretation in generative grammar*. Cambridge, Mass.: MIT Press.

Jenkins, J. (2002) 'A sociolinguistically based, empirically researched pronunciation syllabus for English as an international language', *Applied Linguistics*, 23, pp. 43-103.

Jeon, H.-S. and Nolan, F. (2017) 'Prosodic marking of narrow focus in Seoul Korean', *Laboratory Phonology*, 8(1), pp. 1-30.

Jindapitak, N. (2010) *An attitudinal study of varieties of English: voices from Thai university English learners*. MA thesis. Prince of Songkla University.

Jun, S.-A. (2005) 'Korean intonational phonology and prosodic transcription', in Jun, S.-A. (ed.) *Prosodic typology: the phonology of intonation and phrasing*. Oxford: Oxford University Press, pp. 202-229.

Jun, S.-A. and Fletcher, J. (2014) 'Methodology of studying intonation; from data collection to data analysis', in Jun, S.-A. (ed.) *Prosodic typology II: the phonology of intonation and phrasing*. Oxford: Oxford University Press.

Jun, S.-A. and Oh, M. (2000) 'Acquisition of second language intonation', *Proceedings of the 6th international conference on spoken language processing (ICSLP 2000)*, Beijing, China, 16-20 October, pp. 73-76.

Kainada, E. and Lengeris, A. (2015) 'Native language influences on the production of second-language prosody', *Journal of the International Phonetic Association*, 45(3), pp. 269-287.

Kallayanamit, S. (2004) *Intonation in Standard Thai: contours, registers and boundary tones*. PhD thesis. Georgetown University.

Kamphikul, N. (2015) 'Reduplication in Thai: an OT account', *SEL8205: Issues in phonological theory II*. Newcastle University. Unpublished essay.

Kawai, M. and Wignaraja, G. (2007) 'ASEAN+3 or ASEAN+6: which way forward?', *Paper presented at the conference on multilateralising regionalism*. Geneva, Switzerland, 10-12 September, pp. 1-53.

Khamkien, A. (2010) 'Teaching English speaking and English speaking tests in the Thai context: a reflection from Thai perspective', *English Language Teaching*, 3(1), pp. 184-190.

Kirkpatrick, A. (2008) 'English as the official working language of the Association of Southeast Asian Nations (ASEAN): features and strategies', *English Today 94*, 24(2), pp. 27-34.

Kochanski, G., Grabe, E., Coleman, J., and Rosner, B. (2005) 'Loudness predicts prominence: fundamental frequency lends little', *The Journal of the Acoustical Society of America*, 118, pp. 1038-1054.

Koreman, J., Andreeva, B. and Barry, W. (2008) 'Accentuation cues in French and German', *Proceedings of speech prosody 2008*, Campinas, Brazil, 6-9 May, pp. 613-616.

Krifka, M. (2008) 'Basic notions of information structure', *Acta Linguistica Hungarica*, 55(3-4), pp. 243-276.

Kuronen, M. and Tergujeff, E. (2018) 'Second language prosody and its development: connection between different aspects', *The Language Learning Journal*, pp. 1-15.

Ladd, D.R. (1984) 'Declination: a review and some hypotheses', *Phonology Yearbook*, 1, pp.53-74.

Ladd, D.R. (2006) 'Segmental anchoring of pitch movements autosegmental association or gestural coordination?', *Rivista di Linguistica*, 18(1), pp. 19-38.

Ladd, D.R. (2008) *Intonational phonology*. 2nd edn. Cambridge: Cambridge University Press.

Ladd, D.R. and Schepman, A. (2003) 'Sagging transition between high pitch accents in English: experimental evidence', *Journal of Phonetics*, 31, pp. 81-112.

Ladd, D.R., Scheman, A., White, L., Quarmby, L., and Stackhouse, R. (2009) 'Structural and dialectal effects on pitch peak alignment in two varieties of British English', *Journal of Phonetics*, 37, pp. 145-161.

Ladefoged, P. and Johnson, K. (2011) *A course in phonetics*. 6th edn. Boston, MA: Wadsworth, Cengage Leaning.

Lado, R. (1975) *Linguistic across cultures: applied linguistics for language teachers*. Ann Arbor: University of Michigan Press.

Lange, C. (2007) 'Focus marking in Indian English', *English World-Wide*, 28(1), pp. 89-118.

Lehiste, I. (1970) *Suprasegmentals*. Cambridge, MA: MIT Press.

Lenneberg, E.H. (1967) *Biological foundations of language*. New York: Wiley.

Lepetit, D. (1989) 'Cross-linguistic influence in intonation: French/Japanese and French/English', *Language Learning*, 39(3), pp. 397-431.

Liberman, M.Y. (1975) *The intonation system of English*. Unpublished PhD thesis. Massachusetts Institute of Technology, Cambridge, USA.

Lickley, R., Schepman, A. and Ladd, D.R. (2005) 'Alignment of "phrase accent" low in Dutch falling rising questions: theoretical and methodological implications', *Language and Speech*, 48(2), pp. 157-183.

Lieberman, P. (1960) 'Some acoustic correlates of word stress in American English', *The Journal of the Acoustical Society of America*, 32, pp. 451-454.

Long, M. (1990) 'Maturational constraints on language development', *Studies in Second Language acquisition*, 12, pp. 251-285.

Low, E.L. and Grabe, E. (1999) 'A contrastive study of prosody and lexical stress placement in Singapore English and British English', *Language and Speech*, 42(1), pp. 39-56.

Lu, Y.-A. and Kim, M. (2016) 'Prosody transfer in second language acquisition: tonal alignment in the production of English pitch accent by Mandarin native speakers', *Tsing Hua Journal of Chinese Studies*, 46(4), pp. 785-816.

Luangthongkum, T. (1977) *Rhythm in standard Thai*. Unpublished PhD thesis. University of Edinburgh.

Luksaneeyanawin, S. (1983) *Intonation in Thai*. Unpublished PhD thesis. University of Edinburgh.

Luksaneeyanawin, S. (1998) 'Intonation in Thai', in Hirst, D. and Di Cristo, A. (eds.) *Intonation systems: a survey of twenty languages*. Cambridge: Cambridge University Press, pp. 376-394.

MacKay, I.R.A., Flege, J.E. and Imai, S. (2006) 'Evaluating the effects of chronological age and sentence duration on degree of perceived foreign accent', *Applied Psycholinguistics*, 27, pp. 157-183.

Major, C.A. (2014) *The effect of age on second language acquisition in older adults*. MA thesis. Brigham Young University.

Major, R.C. (2008) 'Transfer in second language phonology: a review', in Edwards, J.G.H. and Zampini, M.L. (eds.) *Phonology and second language acquisition*. Amsterdam: John Benjamins Publishing Company, pp, 63-94.

Mayo, C.J. (1996) *Prosodic transcription of Glasgow English: an evaluation study of GlaToB I*. Master's thesis. University of Edinburgh.

McGory, J.T. (1997) *Acquisition of intonational prominence in English by Seoul Korean and Mandarin Chinese speakers*. PhD thesis. Ohio State University.

Mennen, I. (1999) *Second language acquisition of intonation: the case of Dutch near-native speakers of Greek*. PhD thesis. University of Edinburgh.

Mennen, I. (2006) 'Phonetic and phonological influences in non-native intonation: an overview for language teachers', *QMUC Speech Science Research Centre Working Paper*, WP-9, 1-18.

Mennen, I. (2007) 'Phonological and phonetic influences in non-native intonation', in Trouvain, J. and Gut, U. (eds.) *Non-native prosody: phonetic description and teaching practice*. Berlin, Germany: Mouton de Gruyter, pp. 53-76.

Mennen, I. (2015) 'Beyond segments: towards a L2 intonation learning theory', in Delais-Roussarie, E., Avanzi, M. and Herment, S. (eds.) *Prosody and language in contact: L2 acquisition, attrition and languages in multilingual situations.* Berlin: Springer, pp. 171-188.

Mennen, I. and De Leeuw, E. (2014) 'Beyond segments: prosody in SLA', *Studies in Second Language Acquisition*, 36, pp. 183-194.

Morén, B. and Zsiga, E. (2006) 'The lexical and post-lexical phonology of Thai tones', *Natural Language and Linguistic Theory*, 24, pp. 113-178.

Moyer, A. (2009) 'Input as a critical means to an end: quantity and quality of experience in L2 phonological attainment', in Piske, T. and Young-Scholten, M. (eds.) *Input matters in SLA*. Bristol: Multilingual Matters, pp. 159-174.

Munro, M.J. and Derwing, T.M. (1995) 'Foreign accent, comprehensibility, and intelligibility in the speech of second language learners', *Language Learning*, 45(1), pp. 73-97.

Muntendam, A. and Torreira, F. (2016) 'Focus and prosody in Spanish and Quechua: insights from an interactive task', in Armstrong, M.E., Henriksen, N. and Vanrell, M.M. (eds.) *Intonational grammar in Ibero-Romance: approaches across linguistic subfields*. Amsterdam: John Benjamins, pp. 69-89.

Nagy, J. (2015) 'The use of intonational cues marking new information in non-native speech', *EduLingua*, 1(2), pp. 1-14.

National Identity Board (2000) *Thailand into the 2000s*. Bangkok: Office of the Prime Minister, Kingdom of Thailand.

Nava, E. (2008) 'Prosody in L2 acquisition', in Slabakova, R., Rothman, J., Kemchinsky, P. and Gavruseva, E. (eds.), *Proceedings of the 9th generative approaches to second language acquisition conference (GASLA 2007)*, Somerville, MA: Cascadilla Proceedings Project, pp. 155-164.

Ng, M.L. and Chen, Y. (2011) 'Proficiency in English sentence stress production by Cantonese speakers who speak English as a second language (ESL)', *International Journal of Speech-Language Pathology*, 13(6), pp. 526-535.

Nguyễn, T.A.-T., Ingram, C.L.J. and Pensalfini, J.R. (2008) 'Prosodic transfer in Vietnamese acquisition of English contrastive stress patterns', *Journal of Phonetics*, 36, pp. 158-190.

Nolan, F. and Farrar, K. (1999) 'Timing of f0 peaks and peak lag', *Proceedings of the 14th international congress of phonetic sciences (ICPhS99)*, San Francisco, CA, pp, 961-967.

Noss, R.B. (1964) *Thai reference grammar*. Washington, D.C.: Foreign Service Institute.

O'Brien, M. and Gut, U. (2010) 'Phonological and phonetic realisation of different types of focus in L2 speech', in Dziubalska-Kolaczyk, K., Wrembel, M. and Kul, M. (eds.) *Proceedings of the 6th international symposium on the acquisition of second language speech, New Sounds 2010*. Poznań, Poland, 1-3 May, pp. 331-336.

O'Connor, J.D. and Arnold, G.F. (1973) *Intonation of colloquial English: a practical handbook*. 2nd edn. London: Longmans.

Office of the Education Council, Thailand. (2017) *Education in Thailand*. Bangkok: Office of the Education Council, Ministry of Education.

Oliver, D. and Andreeva, B. (2004) 'Peak alignment in broad and narrow focus in Polish and Bulgarian: a cross-language study', *Proceedings of the 5th conference on formal description of Slavic languages (FDSL-5, 2003)*, Leipzig, 26-28 November.

Ortega-Llebaria, M. and Colantoni, C. (2014) 'Relations between form-meaning associations, access to meaning, and L1 transfer', *Studies in Second Language Acquisition*, 36, pp. 331-353.

Pell, M.D. (2001) 'Influence of emotion and focus location on prosody in matched statements and questions', *The Journal of the Acoustical Society of America*, 109(4), pp. 1668-1680.

Peng, S.-H., Chan, M.K.M., Tseng, C.-Y., Huang, T., Lee, O.J. and Beckman, M.E. (2005) 'Towards a Pan-Mandarin system for prosodic transcription', in Jun, S.-A. (ed.) *Prosodic typology: the phonology of intonation and phrasing*. Oxford: Oxford University Press, pp. 231-270.

Peters, J., Hanssen, J. and Gussenhoven, C. (2014) 'The Phonetic realization of focus in West Frisian, Low Saxon, High German, and three varieties of Dutch', *Journal of Phonetics*, 46, 185-209.

Peterson, G.E. and Lehiste, I. (1960) 'Duration of syllable nuclei in English', *The Journal of the Acoustical Society of America*, 32, pp. 693-703.

Pickering, L. (2001) 'The role of tone choice in improving ITA communication in the classroom', *TESOL Quarterly*, 35(2), pp. 233-255.

Pickering, L. (2009) 'Intonation as a pragmatic resource in ELF interaction', *Intercultural Pragmatics*, 6(2), pp. 236-255.

Pierrehumbert, J.B. (1980) *The phonology and phonetics of English intonation*. Unpublished PhD dissertation. Massachusetts Institute of Technology.

Pierrehumbert, J.B. and Hirschberg, J. (1990) 'The meaning of intonational contours in the interpretation of discourse', in Cohen, P., Morgan, J. and Pollack, M. (eds.) *Intentions in Communication*. Cambridge, MA: MIT Press, pp. 271-311.

Pierrehumbert, J.B. and Steele, S. (1989) 'Categories of tonal alignment in English', *Phonetica*, 46, pp. 181-196.

Piske, T., Mackay, I.R.A. and Flege, J.E. (2001) 'Factors affecting degree of foreign accent in an L2: a review', *Journal of Phonetics*, 29, pp. 191-215.

Pitrelli, J.F, Beckman, M.E. and Hirschberg, J. (1994) 'Evaluation of prosodic transcription labelling reliability in the ToBI framework', *Proceeding of the 3rd international conference on spoken language processing (ICSLP 94)*, Yokohama, Japan, 18-12 September, pp. 123-126.

Pittayaporn, P. (2007) *Prosody of final particles in Thai: interaction between lexical tones and boundary tones* [Poster]. Exhibited at International Workshop on Intonation Phonology: Understudied or Fieldwork Languages. Saarbrucken, Germany. 5 August 2007.

Potisuk, S., Gandour, J.T. and Harper, M.P. (1994) 'F0 correlates of stress in Thai', *Linguistics of the TiBeto-Burman Area*, 17(2), pp. 1-27.

Prieto, P. (2009) 'Tonal alignment patterns in Catalan nuclear falls', *Lingua*, 119, pp. 865-880.

Prieto, P. (2011) 'Tonal alignment', in van Oostendorp, M., Ewen, C.J., Hume, E.V. and Rice, K. (eds.) *The Blackwell Companion to Phonology*. Oxford, UK: Wiley-Blackwell, pp. 1185-1203.

Prieto, P. and Torreira, F. (2007) 'The segmental anchoring hypothesis revised: syllable structure and speech rate effects on peak timing in Spanish', *Journal of Phonetics*, 35, pp. 473-500.

Prieto, P., van Santen, J. and Hirschberg, J. (1995) 'Tonal alignment patterns in Spanish', *Journal of Phonetics*, 23, pp. 429-451.

R Development Core team (2019) *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN: 3-900051-07-0, Available at: http://www.R-project.org.

Ramiŕez Verdugo, D. (2002) 'Non-native interlanguage intonation systems: a study based on a computerized corpus of Spanish learners of English', *ICAME Journal*, 26, pp. 115-132.

Ramiŕez Verdugo, D. (2005) 'The nature and patterning of native and non-native intonation in the expression of certainty and uncertainty: pragmatic effects', *Journal of Pragmatics*, 37, pp. 2086-2115.

Ramiŕez Verdugo, D. (2006) 'Prosodic realization of focus in the discourse of Spanish learners and English native speakers', *Estudios Ingleses de la Universidad Complutense*, 14, pp. 9-32.

Rasier, L. and Hiligsmann, P. (2007) 'Prosodic transfer from L1 to L2: theoretical and methodological issues', *Nouveaux cahiers de linguistique française*, 28, pp. 41-66.

Rietveld, A.C.M. and Gussenhoven, C. (1985) 'On the relation between pitch excursion size and prominence', *Journal of Phonetics*, 13, pp. 299-308.

Roach, P. (2004) 'British English: Received Pronunciation', *Journal of the International Phonetic Association*, 34(2), pp. 239-245.

Roach, P. (2009) *English phonetics and phonology: a practical course*. 4th edn. Cambridge: Cambridge University Press.

RStudio Team (2017) *RStudio: integrated development environment for R*. RStudio, Inc., Boston, MA. Available at: http://www.rstudio.com.

Selkirk, E. (1995) 'Sentence prosody: intonation, stress, and phrasing', in Goldsmith, J. (ed.) *The handbook of phonological theory*. Cambridge, MA: Blackwell Publishers, pp. 550-569.

Silpachai, A. (2013) 'Prosodic characteristics of two focus types in emphatic context in Thai', *Proceedings of meetings on acoustics,* Montreal, Canada, 2-7 June. doi: 10.1121/1.4798795.

Silverman, K., Beckman, M., Pitrelli, J., Ostendorf, M., Wightman, C., Price, P., Pierrehumbert, J. and Hirschberg, J. (1992) 'ToBI: a standard for labeling English prosody'*, Proceedings of the 2nd international conference on spoken language processing (ICSLP 92)*, Banff, Alberta, Canada, 12-16 October.

Silverman, K.E.A. and Pierrehumbert, J.B. (1990) 'The timing of prenuclear high accents in English', in Kingston, J. and Beckman, M.E. (eds.) *Papers in laboratory phonology I: between the grammar and physics of speech*, Cambridge: Cambridge University Press, pp. 72-106.

Sityaev, D. and House, J. (2003) 'Phonetic and phonological correlates of broad, narrow and contrastive focus in English', *Proceedings of the 15th international congress of phonetic sciences,* Barcelona, pp. 1819-1822. ISBN 1-876346-48-5.

Sluijter, A.M.C. and van Heuven, V.J. (1996) 'Spectral balance as an acoustic correlate of linguistic stress', *The Journal of the Acoustic Society of America*, 100(4), pp. 2471-2485.

Sumdangdej, S. (2007) *Input and the acquisition of suprasegmental phonology in English by Thai school children*. PhD thesis. Durham University.

Swerts, M., Krahmer, E. and Avesani, C. (2002) 'Prosodic marking of information status in Dutch and Italian: a comparative analysis', *Journal of Phonetics*, 30, pp. 629-654.

Tahta, S., Wood, M. and Loewenthal, K. (1981) 'Foreign accents: factors relating to transfer of accent from the first language to a second language', *Language and Speech*, 24, pp. 265-272.

Tingsabadh, M.R.K. and Abramson, A.S. (1993) 'Thai', *Journal of the International Phonetic Association*, 23, pp. 24-28.

Trofimovich, P. and Baker, W. (2006) 'Learning second language suprasegmentals: effect of L2 experience on prosody and fluency characteristics of L2 speech', *Studies in Second Language Acquisition*, 28, pp. 1-30.

Tumtavitikul, A. (1997) 'The metrical structure of Thai in a non-linear perspective', in Abramson, A.S. (ed.) *Southeast Asian linguistic studies in honour of Vichin Panupong*. Bangkok: Chulalongkorn University, pp. 309-324.

Tumtavitikul, A. and Thitikannara, K. (2006) 'The intonation of Thai emotional speech', in Warren, P., Watson, C.I. (eds.) *Proceedings of the 11th Australian international conference on speech science and technology*, University of Auckland, New Zealand, 6-8 December. Australian Speech Science and Technology Association Inc., pp. 287-291.

Turk, A., Nakai, S. and Sugahara, M. (2006) 'Acoustic segment duration in prosodic research: a practical guide', in Sudhoff, S., Lenertová, D., Meyer, R., Pappert, S., Augurzky, P., Mleibek, I, Richter, N. and Schließer, J. (eds.) *Methods in empirical prosody research*. Berlin: Mouton de Gruyter, pp. 1-27.

Turk, A.E. and Sawusch, J.R. (1996) 'The processing of duration and intensity cues to prominence', *The Journal of the Acoustical Society of America*, 99, pp. 3782-3790.

Ueyama, M. (2000) *Prosodic transfer: an acoustic study of L2 English vs. L2 Japanese.* PhD thesis. University of California, Los Angeles.

Ueyama, M. and Jun, S.-A. (1998) 'Focus realization in Japanese English and Korean English intonation', *Japanese and Korean Linguistics*, 7, pp. 629-645.

van Els, T. and de Bot, K. (1987) 'The role of intonation in foreign accent', *The Modern Language Journal*, 71(2), pp. 147-155.

van Maastricht, L., Krahmer, E. and Swerts, M. (2016) 'Prominence patterns in a second language: intonational transfer from Dutch to Spanish and vice versa', *Language Learning*, 66, pp. 124-158.

van Santen, J.P.H. and Hirschberg, J. (1994) 'Segmental effects on timing and height of pitch contours', *Proceedings of the 3rd international conference on spoken language processing (ICSLP 94)*, Yokohama, Japan, 18-22 September.

Venditti, J.J. (2005) 'The J_ToBI model of Japanese intonation', in Jun, S.-A. (ed.) *Prosodic typology: the phonology of intonation and phrasing*. Oxford: Oxford University Press, pp. 173-200.

Warotamasikkhadit, U. (2000) 'The accentual system of Thai polysyllabic words', *MANUSYA: Journal of Humanities*, 3(2), pp. 32-44.

Wayland, R., Guion, S.G., Landfair, D. and Li, B. (2006) 'Native Thai speakers' acquisition of English word stress patterns', *Journal of Psycholinguist Research*, 35, pp. 285-304.

Weinreich, U. (1953) *Languages in contact, findings and problems*. New York: Linguistic Circle of New York.

Welby, P. (2003) 'Effects of pitch accent position, type, and status on focus projection', *Language and Speech*, 46(1), pp. 53-81.

Wells, J.C. (2006) *English intonation: an introduction*. Cambridge: Cambridge University Press.

Wells, W.H.G. (1986) 'An experimental approach to the interpretation of focus in spoken English', in Johns-Lewis, C. (ed.) *Intonation in discourse*. London: Croom Helm, pp. 53-75.

Wennerstrom, A. (1994) 'Intonational meaning in English discourse: a study of non-native speakers', *Applied Linguistics*, 15(4), pp.399-420.

Wichmann, A. (2000) *Intonation in text and discourse: beginnings, middles, and ends*. Harlow, England: Longman.

Wickham, H. (2016) *ggplot2: Elegant Graphics for Data Analysis*. 2nd edn. Springer-Verlag New York. ISBN 978-3-319-24277-4.

Wickham, H. (2017) *reshape2: Flexibly reshape Data: A Reboot of the Reshape Package*. R package (Version 1.4.3). Available at: http://CRAN.R-project.org/package=reshape2.

Wickham, H., Chang, W., Henry, L., Pedersen, T.L., Takahashi, K., Wilke, C., Woo, K., Yutani, H. and RStudio (2019) *ggplot2: Create elegant data visualisations using the grammar of graphic*. R package (Version 3.2.1). Available at: http://CRAN.R-project.org/package=ggplot2.

Wickham, H., François, R., Henry, L., Müller, K. and RStudio. (2018) *dplyr: a grammar of data manipulation*. R package (Version 0.7.6). Available at: http://CRAN.R-project.org/package=dplyr.

Wiriyachitra, A. (2001) 'A Thai university English scenario in the coming decade', *Thai TESOL Newsletter*, 14, pp. 4-7.

Wong, W.Y.P., Chan, M.K.M. and Beckman, M.E. (2005) 'An autosegmental-metrical analysis and prosodic annotation conventions for Cantonese', in Jun, S.-A. (ed.) *Prosodic typology: the phonology of intonation and phrasing*. Oxford: Oxford University Press, pp. 272-300.

Wongsothorn, A., Hiranburana, K. and Chinnawongs, S. (2002) 'English language teaching in Thailand today', *Asia Pacific Journal of Education*, 22(2), pp. 107-116.

Xu, Y. (1999) 'Effects of tone and focus on the formation and alignment of f0 contours', *Journal of Phonetics*, 27, pp. 55-105.

Xu, Y. (2010) 'In defense of lab speech', *Journal of Phonetics*, 38, pp. 329-336.

Xu, Y. (2013) 'ProsodyPro – A tool for large-scale systematic prosody analysis', *Proceedings of tools and resources for the analysis of speech prosody (TRASP 2013)*, Aix-en-Provence, France, pp. 7-10.

Xu, Y. (2017) 'Intonation', in Behr, W., Gu, Y., Handel, Z., Huang, C.-T.J. and Myers, J. (eds.) *Encyclopedia of Chinese language and linguistics*. Boston: Brill, pp. 458-466.

Xu, Y. and Wang, Q.E. (2001) 'Pitch targets and their realization: evidence from Mandarin Chinese', *Speech Communication*, 33, pp. 319-337.

Xu, Y. and Xu, C.X. (2005) 'Phonetic realization of focus in English declarative intonation', *Journal of Phonetics,* 33, pp. 159-197.

Yimngam, S., Premchaisawadi, W. and Kreesuradej, W. (2011) 'Prosody analysis of Thai emotion utterances', in Munoz, R., Montoyo, A. and Metais, E. (eds.) *Natural language processing and information systems*. New York: Springer, pp. 177-184.

Young-Scholten, M. (1994) 'On positive evidence and ultimate attainment in L2 phonology', *Second Language Research*, 10(3), pp. 193-214.

Young-Scholten, M. (1995) 'The negative effects of 'positive' evidence on L2 phonology', in Eubank, L., Selinker, L. and Sharwood Smith, M. (eds.) *The current state of interlanguage*. Amsterdam: Benjamins, pp. 107-122.

Young-Scholten, M. (2013) 'Great expectations in phonology? Second language acquisition research and its relation to the teaching of older and younger learners', in Whong, M., Gil, K.-H. and Marsden, H. (eds.) *Universal grammar and the second language classroom*, New York: Springer, pp. 207-229.

Zerbian, S. (2013) 'Prosodic marking of narrow focus across varieties of South African English'*, English World-Wide*, 34(1), pp. 26-47.

Zimmermann, M. and Onea, E. (2011) 'Focus marking and focus interpretation', *Lingua*, 121, pp. 1651-1670.

Zsiga, E. and Nitisaroj, R. (2007) 'Tone features, tone perception, and peak alignment in Thai', *Language and Speech*, 50(3), pp. 343-383.

Appendices

Appendix A: Consent form and participant information sheet for native English speakers and native Thai speakers

<div style="border:1px solid">

Newcastle University

CONSENT FORM

SCHOOL OF EDUCATION, COMMUNICATION AND LANGUAGE SCIENCES

Title of project:	The Intonation Patterns of Thai Learners of English
Researcher name:	Nisit Kamphikul (n.kamphikul@ncl.ac.uk)
Contact details:	School of Education, Communication and Language Sciences, King George VI Building, Newcastle University, Newcastle upon Tyne, Tyne and Wear, NE1 7RU, United Kingdom.
Primary supervisor:	Prof. Martha Young-Scholten (martha.young-scholten@ncl.ac.uk)
Secondary supervisor:	Dr SJ Hannahs (s.j.hannahs@ncl.ac.uk)
Contact details:	School of English Literature, Language and Linguistics, Percy Building, Newcastle University, Newcastle upon Tyne, Tyne and Wear, NE1 7RU, United Kingdom.

I, the undersigned participant, confirm that (please tick/check box appropriately):

1	I have read and understood the information about the project in the information sheet.	☐
2	I have been given the chance to ask questions about the project and my participation.	☐
3	I agree to voluntarily participate in this research project.	☐
4	I understand that I can withdraw at any time without giving reasons or being penalised or questioned for withdrawing or any consequences of any kind.	☐
5	I understand that an audio recorder will be used to collect data and I agree to my voice being recorded for the purpose of this research project	☐
6	The procedures on confidentiality and anonymity have been clearly explained to me.	☐
7	I understand that the recording of my voice and transcriptions will be stored in coded files on computers and in locked cabinets.	☐
8	I understand that anonymised parts of my data may be used in language research, conference presentations and written report/publications.	☐
9	Storage and usage of data has been clearly explained to me.	☐
10	I understand that I will receive payment as incentives for my participation in this project.	☐

Name of participant: _____

Signature of participant: _____

Date: _____

Name of researcher: _____Mr Nisit Kamphikul_____

Signature of researcher: _____

Date: _____

</div>

PARTICIPANT INFORMATION SHEET

Title of project:	The Intonation Patterns of Thai Learners of English
Researcher name:	Nisit Kamphikul (n.kamphikul@ncl.ac.uk)
Contact details:	School of Education, Communication and Language Sciences, King George VI Building, Newcastle University, Newcastle upon Tyne, Tyne and Wear, NE1 7RU, United Kingdom.
Primary supervisor:	Prof. Martha Young-Scholten (martha.young-scholten@ncl.ac.uk)
Secondary supervisor:	Dr SJ Hannahs (s.j.hannahs@ncl.ac.uk)
Contact details:	School of English Literature, Language and Linguistics, Percy Building, Newcastle University, Newcastle upon Tyne, Tyne and Wear, NE1 7RU, United Kingdom.

We invite you to participate in this research project which is being carried out by *Nisit Kamphikul* as part of his IPhD/PhD under the supervision of *Prof. Martha Young-Scholten* and *Dr SJ Hannahs*. Before you agree to take part in it, you need to understand some basic information on why the research is being conducted. Take your time to read the information here before deciding to participate in the project. If you have questions or concerns, or if you want more information, please feel free to contact *Nisit Kamphikul*, or any of the other people involved in this research.

Purpose and aims of the research: We are looking at how English intonation of native and non-native speakers deviate when they are producing an utterance.

Participation selection: We are inviting you to participate in this project because you are a native speaker of the languages we are interested in (i.e., English and Thai).

Voluntary participation: Your participation in this project is entirely voluntary. If you agree to participate, we will ask you to sign a consent form. You have the right to withdraw your consent and participation at any time. You only need to notify the researcher beforehand. If you withdraw, your data will be used only if you permit us to use it. Otherwise it will be destroyed.

What is involved: If you agree to participate, we will ask you to visit the Phonetics Laboratory (Room 2.12) at King George VI Building, where you will be audio-recorded reading aloud a series of sentences, asking and answering questions, and a passage. We will also ask you some questions about your background. Each recording should take no longer than 45 minutes. Financial reward as incentives will be given to you, and we also thank you very much for your participation.

Benefits and risks: Participation involves no known or anticipated risks but it may cause inconvenience as it will require around an hour of your time per recording. The potential benefits associated with your participation are that your data will have contribution to an understanding of the English intonation systems and thus will help teachers of ESL/EFL understand more of L2 intonation patterns, and more of pitch variation in English as an L1/L2 across population.

Anonymity and confidentiality: As for signature, we will ask you to give your first name and just your surname initial. We will keep the information we ask you for separate from the data from the tasks we audio-record. The audio data will be transcribed and connected to a code instead of your name. Only the Principal Investigator/the researchers will have access to information that connects your identity and data which will be protected during and after the data collection in a locked cabinet.

Dissemination of results: The results of the study will be shared by the researcher in publications and presentations. These will be available from the Principal Investigator on request.

Appendix B: Questionnaire for native English speakers

Questionnaire for Native Speakers of English (NSE)

Please provide some background information about yourself. It will be used for research purposes only.

1. Name: _____

2. Age: _____

3. Gender: _____ male   _____ female

4. Education level at the time of recording:

    _____ Bachelor's degree

    _____ Master's degree

    _____ IPhD/PhD

5. Your programme of study   _____

6. What year are you in?   _____

7. Hometown (Birthplace) :   _____

8. If you are not a student, what is your career? _____

9. Does your English/accent/pronunciation belong to Southern British English or RP (Received Pronunciation) English?   _____ Yes   _____ No

10. How would you rate your speaking skill in English? (check one)

    _____ good   _____ very good   _____ excellent

- -

Thank you very much.

Appendix C: Questionnaire for native Thai speakers

<table>
<tr><td>

Questionnaire for Thai Learners of English (TLE)

Please provide some background information about yourself. It will be used for research purposes only.

1. Name: _____

2. Age: _____

3. Gender: _____ male _____ female

4. Education level at the time of recording:

    _____ Bachelor's degree

    _____ Master's degree

    _____ IPhD/PhD

5. Your programme of study _____

6. What year are you in? _____

7. Hometown (Birthplace): _____

8. Number of months for your staying in the UK: _____ months

9. Average length of exposure to English per week: _____ hours

10. Your IELTS scores for speaking skill: _____

Thank you very much.

</td></tr>
</table>

Appendix D: Speech material for question and answer task

1. Test sentences with prompt questions

| No.	Prompt questions	Test sentences with target words	Focus size/ type
1	Who read a novel?	**NANNY** read a novel.	subject focus
2	What did Nanny read?	Nanny read a **NOVEL**.	object focus
3	What did Nanny do with a novel?	Nanny **READ** a novel.	verb focus
4	What did you say?	Nanny read a novel.	broad focus
5	Who won a medal?	**RONNEY** won a medal.	subject focus
6	What did Ronney win?	Ronney won a **MEDAL**.	object focus
7	What did Ronney do with a medal?	Ronney **WON** a medal.	verb focus
8	What did you say?	Ronney won a medal.	broad focus
9	Who had a melon?	**JIMMY** had a melon.	subject focus
10	What did Jimmy have?	Jimmy had a **MELON**.	object focus
11	What did Jimmy do with a melon?	Jimmy **HAD** a melon.	verb focus
12	What did you say?	Jimmy had a melon.	broad focus
13	Who lent the money?	**MANNY** lent the money.	subject focus
14	What did Manny lend?	Manny lent the **MONEY**.	object focus
15	What did Manny do with the money?	Manny **LENT** the money.	verb focus
16	What did you say?	Manny lent the money.	broad focus

2. Filler sentences

| No.	Prompt questions	Test sentences with target words	Focus size/ type
1	Who moved on Monday?	**HENRY** moved on Monday.	subject focus
2	When did Henry move?	Henry moved on **MONDAY**.	object focus
3	What did Henry do on Monday?	Henry **MOVED** on Monday.	verb focus
4	What did you say?	Henry moved on Monday.	broad focus
5	Who lived in London?	**JENNY** lived in London.	subject focus
6	Where did Jenny live?	Jenny lived in **LONDON**.	object focus
7	What did Jenny do in London?	Jenny **LIVED** in London.	verb focus
8	What did you say?	Jenny lived in London.	broad focus

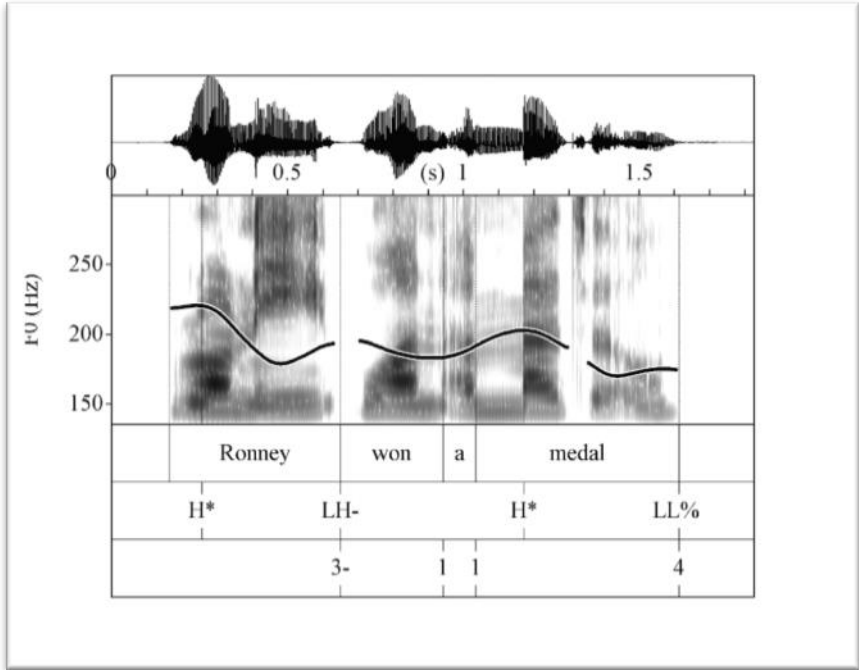**Appendix E: Examples of prosodic marking for broad focus as produced by native English speakers**
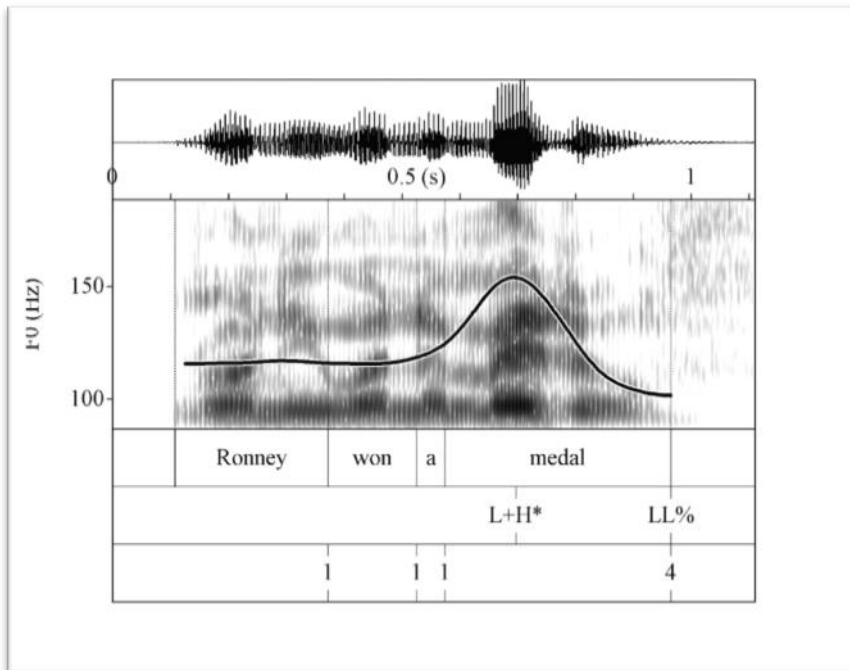
Rising accent (L+H*) with pre-focal deaccenting as uttered by the male native English speaker en1_m10_q44023
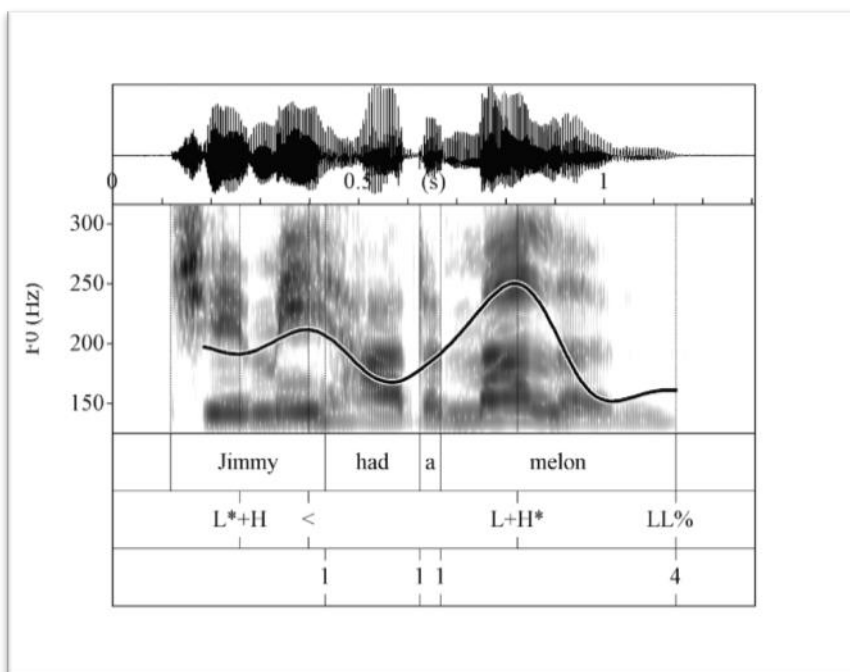
Rising accent (L+H*) with pre-focal deaccenting as uttered by the male native English speaker en1_m10_q44023
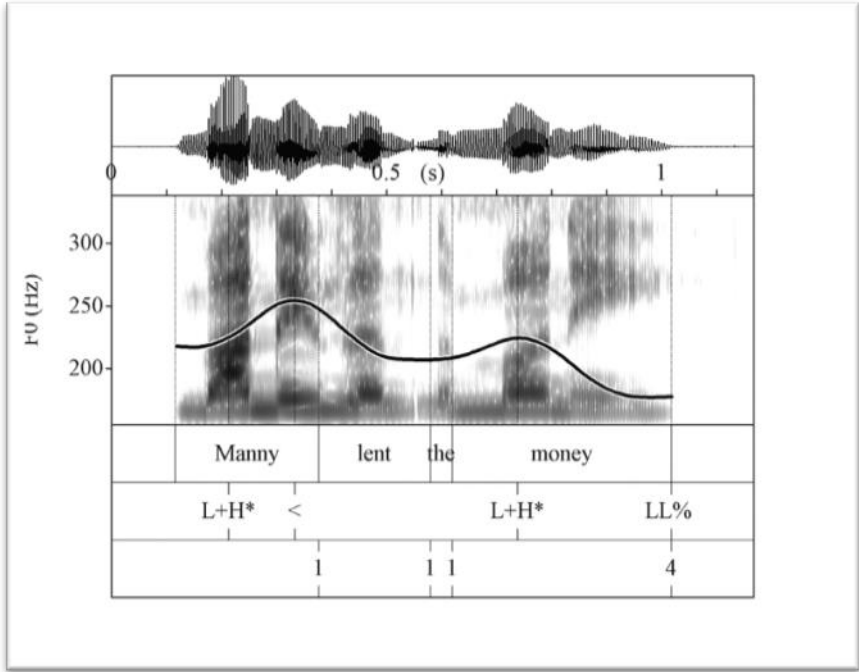
Rising accent (L+H*) accompanied by compressed pitch range as produced by the female native English speaker en1_f03_q44063
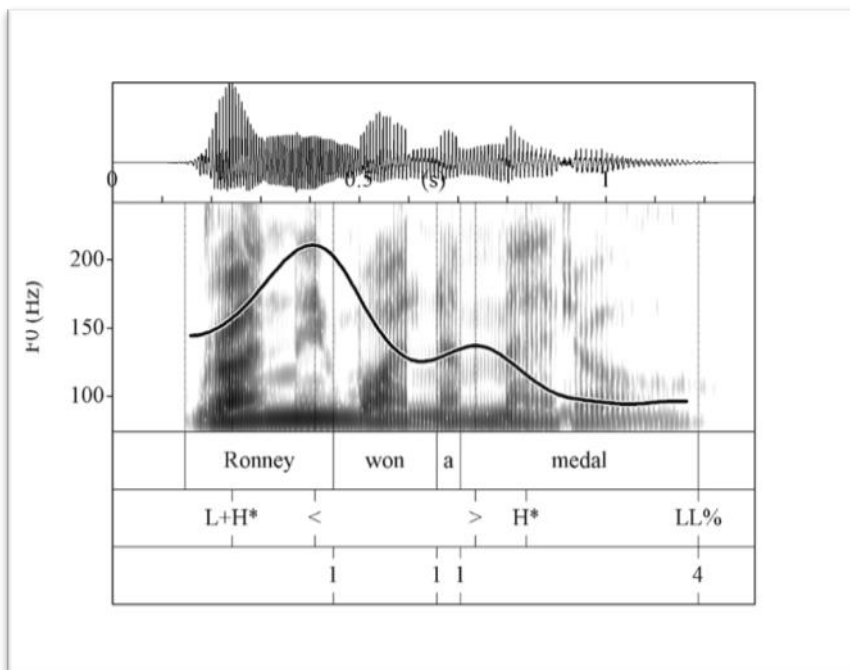

Characteristics of prosodic phrasing or re-phrasing as produced by the female native English speaker en1_f07_q44023
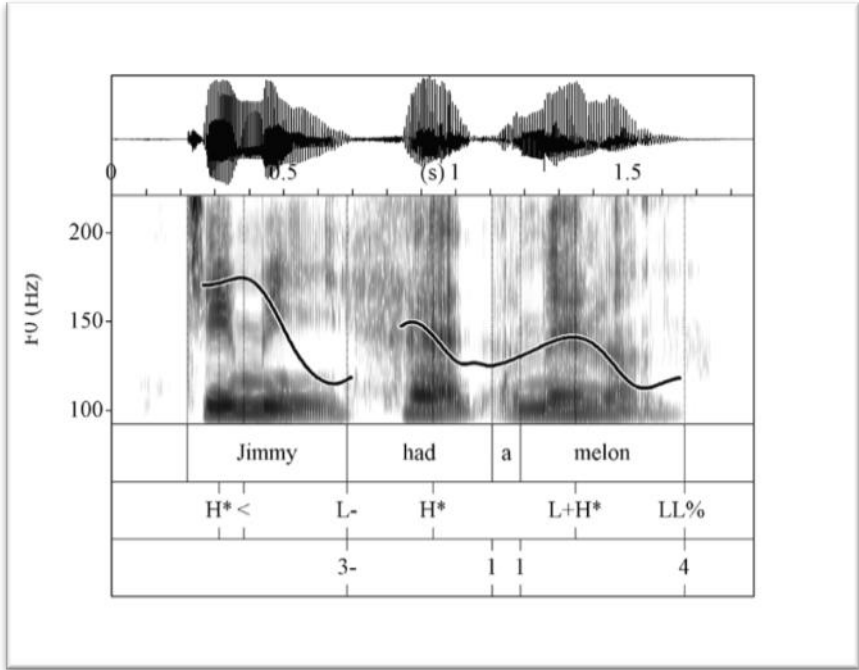
**Appendix F: Prosodic marking of broad focus as produced by native Thai speakers**
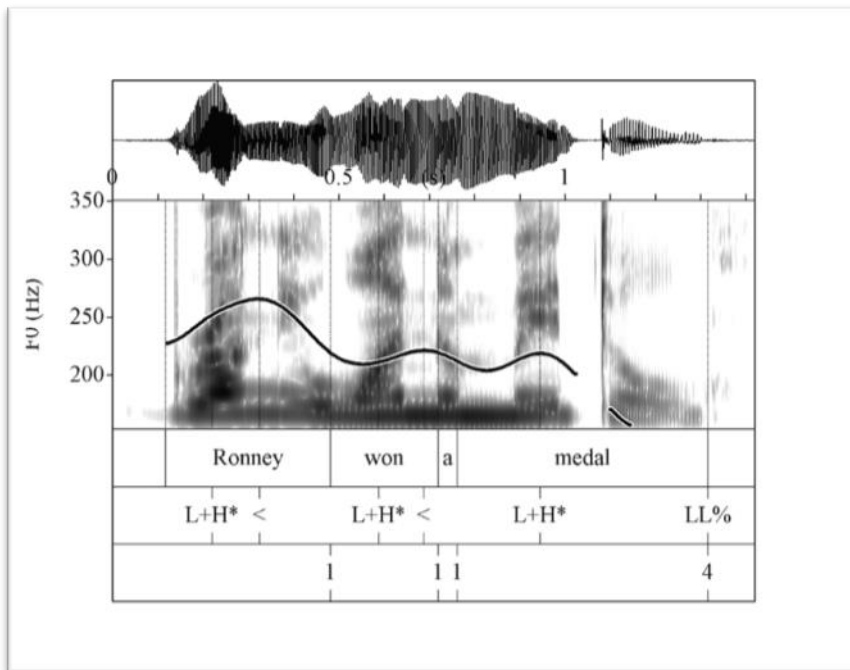
Characteristics of pitch range compression as uttered by the female native Thai speaker en2_f01_q44022

Accent placement on Subject and Object positions as produced by the male native Thai speaker en2_m01_q44023

Characteristics of prosodic phrasing or re-phrasing as uttered by the male native Thai speaker en2_m08_q44033
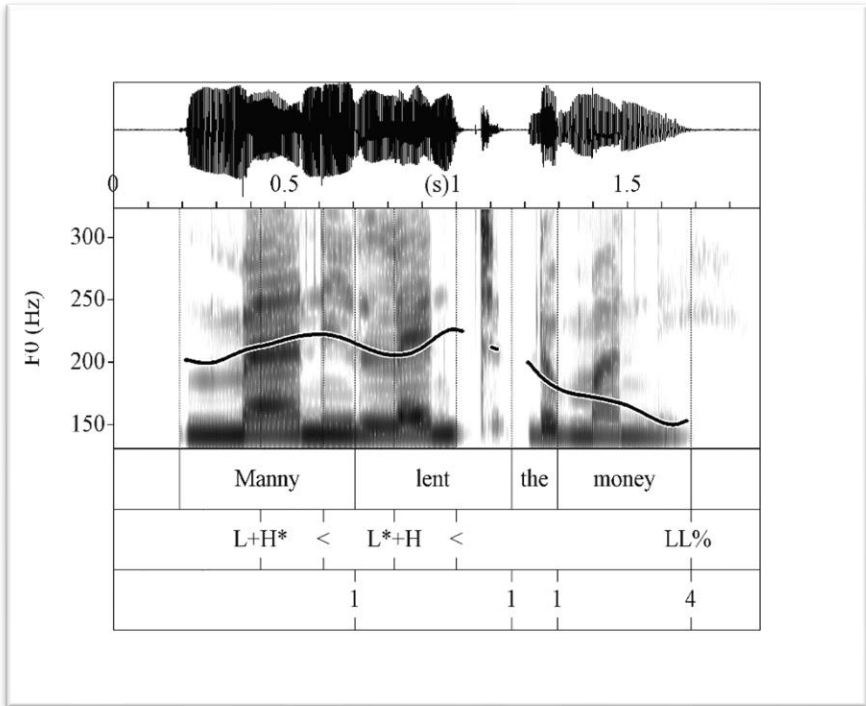
Accent placement on Subject and Verb positions as produced by the female native Thai speaker en2_f04_q44062