INFLUENCE OF THE FIRST AND SECOND LANGUAGE ON THE PERCEPTION OF THAI TONES

Vance Schaefer

Submitted to the faculty of the University Graduate School

in partial fulfillment of the requirements for the degree

Doctor of Philosophy

in the Department of Second Language Studies of the College of Arts and Sciences,

Indiana University

October 2015

Accepted by the Graduate Faculty, Indiana University, in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

Doctoral Committee

[Chairperson's signature]

[Isabelle Darcy, Ph.D.]

[Second reader's signature]

[Rex Sprouse, Ph.D.]

[Third reader's signature]

[Laurent Dekydtspotter, Ph.D.]

[Fourth reader's signature]

[Charles Lin, Ph.D.]

[Fifth reader's signature]

[Öner Özçelik, Ph.D.]

[October 5, 2015]

Acknowledgements

Abstract

Convergent results from second language (L2) phonology literature suggest that when a phonetic or linguistic dimension is present or important in the first language (L1) of a learner, acquiring words that use this dimension in an L2 will be facilitated (Feature Hypothesis, McAllister, Flege & Piske, 2002). This study applies this hypothesis to the dimension of linguistic pitch by testing speakers from L1s which utilize lexically-contrastive pitch in varying manners and degrees – Mandarin Chinese, (using pitch as tone), Japanese (using pitch accent), English (using pitch in word stress), and finally Korean (not using pitch to distinguish words) – on their perception of Thai tones. Additionally, an L1-English-speaking group who have learned Mandarin as an L2 have also been tested on their perception of Thai tones to examine whether they can transfer their acquired perception of Mandarin tones to naive perception of Thai tones. This L2 Mandarin Learner groups has also been tested on a lexical decision task composed of Mandarin words and non-words to determine their acquisition of L2 Mandarin tone.

The current theoretical framework of L2 phonology models tends to largely focus on the perception of segmentals (i.e., vowels and consonants) (Dupoux, Peperkamp & Sebastian, 2001). Only a limited number of studies have examined the perception of tone by non-native speakers, and very few have considered speakers of pitch accent languages. My study attempts to examine whether the perception of suprasegmentals (i.e., tone, pitch accent, stress, intonation) follows the same mechanisms which have been put forward for segmentals. This type of analysis enables us to better understand L2 acquisition as it examines the "raw" perception of L2 tone and determines a baseline for how pitch is processed by non-learners of a tone language. This provides valuable information regarding what non-learners do, allowing us to posit how learners acquire tone in an L2. A comparative study of how foreign or L2 tones are perceived by speakers of languages differing in the linguistic use of pitch would more comprehensively illuminate the perception of

suprasegmentals, allowing for their better incorporation into L2 phonology theories and better informed pedagogical practices.

Research is guided by the following questions. Do speakers of various L1s differ in their perception of non-native tone? Does the presence of certain features in the L1, namely pitch to distinguish words, aid in the perception of non-native tone? Does the differing extent of lexically-contrastive pitch in various L1s result in a linguistic hierarchy of perceptual ability of a non-native tone language? Do learners of an L2 tone language perform at superior levels on the naïve perception of another tone language as compared to their counterparts who have not learned an L2 tone language? To evaluate listeners' naïve perception of Thai tones, a computer-based ABX task presents the tones of Thai in various combinations and records both accuracy rates and reaction times (RTs). The following hierarchy (from best to worst) based on the functionality of lexically-contrastive pitch in the L1 is predicted in the ability to discriminate tone: L1 Thai speakers > L1 Mandarin speakers > L1 English/L2 Mandarin > L1 Japanese speakers > L1 English speakers > L1 Korean speakers. Results from a preliminary pilot study suggest that the presence of pitch to distinguish words in the L1 shapes the perception of L2 tones in a positive manner. In short, the greater the functionality of the lexically-contrastive pitch in the L1 or in an L2, the greater the salience of pitch in the L1 or L2 and therefore, superior naïve perception of non-native features of lexically contrastive pitch.

Results have been able to answer to both a lesser and greater extent the research questions posed for this study. First, results show that lexically-contrastive pitch in the L1 does indeed influence the perception of a non-native lexically-contrastive pitch system, namely Thai tones. Furthermore, this influence differs in a hierarchical manner depending on the extent of the usage of lexically-contrastive pitch system. That is, L1 speakers of a tone language perform most accurately on the naïve tone perception of Thai tones, followed by L1 speakers of a pitch accent language, i.e., Japanese. In contrast, L1 speakers of English, a word stress language, perform at the same level of L1 speakers of a language which does not use pitch in a lexically-contrastive manner, i.e., Korean. Reaction times conform less to the expected hierarchy. The L1 Japanese and L1 Korean speakers are faster than the L1 Mandarin and L1 English speakers, respectively. Additionally, results for both accuracy rates and RTs on the perception of individual tone comparisons reveal both universal and language-specific tendencies, indicating that some tones are easier or more difficult to perception.

Second, results clearly demonstrate learning a tone language as an L2, i.e., Mandarin Chinese, by L1 speakers of English shapes the perception of a non-native tone language, i.e., Thai. Furthermore, results show that learners are generally able to encode tonal information to the extent that lexical access is constrained. Moreover, exposure and accordingly encoding of Mandarin tones appears to be sufficiently robust to shape the naïve perception of Thai tones. L2 Mandarin learners outperform non-learners of Mandarin who also speak English as an L1 and perform at comparable levels to L1 speakers of Mandarin on the naïve perception of Thai tones. However, performance between the L2 learners and L1 Mandarin speakers do not pattern in the exact same manner with a few discrepancies on the individual tone comparisons, most likely reflective of the respective use of lexically-contrastive pitch in the L1s. Also, production of Production of Mandarin tones by the L2 learners generally correlates with the encoding of Mandarin tones and superior naïve perception of Thai tones although not by all individual learners.

Lastly, results from both the naïve perception of Thai tones and encoding/production of Mandarin tones allow a framework for the perception of non-native tones to be forged. First, resulting data concerning individual tone comparisons can be characterized by both universal and language-specific tendencies among the target language groups. Furthermore, results can be interpreted within a comprehensive PIM-PAM model (PAM, Best, 1995; PIM, Brown, 1998) focusing on the features of lexically-contrastive pitch. That is, results of the current study can be more clearly explained by looking at the features of lexically-contrastive pitch. For example, the PAM model can explain the influence of lexically-contrastive pitch in languages that both use tone as it enables streamlined tone-to-tone mapping. Yet, this becomes more difficult when attempting to map pitch accent or word stress patterns to tone patterns. Additionally, the similarities or dissimilarities between the tones promoting or obstructing tone-to-tone mapping are not easily explained. One solution would be commandeer the PIM model (Brown, 1998) and Feature Hypothesis (McAllister et al., 2002) and look at the features of these tones to offer a more straightforward explanation. This approach would additionally offer a more insightful interpretation of pitch accent pattern- or word stress- to-tone mapping.

Table of Contents

List of Figures

List of Tables

1. Introduction

Most people who learn a second language (L2, i.e., foreign language) have an accent reflecting their first language (L1, i.e., native language). Looking at perhaps the most well-known example of an accent by non-native speakers may clearly illustrate the phenomenon of a foreign accent: the production of the /r/ and /l/ distinction in English generally by L1 speakers of Asian languages. How do we account for this inability to distinguish these two sounds? Simply put, it is said to be brought about by the lack of these two sounds in the L1 of the speakers. As learners of a foreign language, we often encounter new sounds for the first time which we cannot perceive, but are nevertheless asked to pronounce in the classroom. In such cases, when we attempt to produce the sound, we oftentimes substitute a sound from our native language which we perceive to be the closest approximation to the target sound. In other cases, we produce a sound not in our native language nor actually in the foreign language we are trying to mimic. Additionally, the notion of sound goes beyond isolated production of vowels or consonants, but extends to the position of these sounds within a word (e.g., initial, mid or final position, i.e., phonotactics), intonation patterns and more. It is then these discrepancies in the production of the sounds of a foreign language (i.e., difference between the targeted L1 sound and the sound produced by the L2 speaker) that create a foreign accent. Moreover, it cannot be overlooked that this non-target-like production of foreign sounds is often a result of non-target-like perception. This fact is perhaps very apparent to native speakers of English who are learning the tones of Mandarin Chinese, a language where non-target-like tone production also characterizes a foreign accent.

This common observation most likely lead to the creation of models for L2 phonology (e.g., Perceptual Assimilation Model, PAM, Best, 1995, and its L2 offshoot PAM-L2, Best & Tyler, 2007; Feature Hypothesis, McAllister, Piske & Flege, 2002) based upon such experiences of learners to explain and predict patterns of foreign accents. That is, such models intuitively rely on the similarity between the targeted sound in the foreign language and the nearest substituted approximation in the native language of the learner in an attempt to explain the resulting foreign accent. These models also almost exclusively look at sounds such as vowels and consonants, due perhaps to two reasons which may shed light on the situation. The first reason may merely be that somehow isolated vowels and consonants (i.e., segmentals) seem more “stable” and thus, easier to hear and work with while other characteristics of language (e.g., word stress, tone, intonation) to distinguish meaning between words may seem less “stable” and thus, more difficult to hear and work with. These other characteristics include vowel length and pitch (i.e., suprasegmentals), for example. Pitch is the use of the voice height to distinguish different words, which to native speakers of English may seem analogous to singing. Thus, just as two different vowels can distinguish the difference between two words such as ‘ship’ and ‘sheep’ in English, the use of pitch (oftentimes along with other features) may do so as well, e.g., the difference between words such as ‘an insert’ versus ‘to insert’ in English where the first syllable is stressed in the former word while the second syllable is generally stressed in the latter.

In English the use of pitch in word stress appears to be less salient than vowel length or intensity to most native speakers of English. This fact may reinforce the second reason as to why most models to explain the perception and production of non-native sounds tend to focus on consonants and vowels: there is most likely an overwhelming cultural bias toward Western languages, particularly toward English, where learning another Western language oftentimes limits the necessity of teaching or noting other pronunciation features beyond vowels and consonants due to assumed perceived similarities between most major Western European languages in terms of word stress and intonation, for example. Hence, native speakers of English when learning German or French or Spanish, perhaps the most commonly taught languages in the West in the last century until relatively recent times, were rarely taught the differences in stress usage and patterns.

However, globalization in the form of the political, economic and cultural rise of non-Western-European countries and their languages has expanded the selection of languages that native speakers of English now can and wish to learn, including Russian, Japanese, Arabic and perhaps the "golden child" of the moment, Mandarin Chinese. Yet, perhaps in the face of the continued historical bias in Western language education with its focus on vowels and consonants, models on the perception (and production) of second languages have tended to be generally insufficient to explain how learners adapt other dimensions of phonology in their perception and production patterns. One example is the use of pitch to contrast words in the form of tone. Tone is the use of pitch to distinguish two words. A famous Mandarin Chinese example for tone is the syllable /ma/ which, depending on the pitch pattern used, can mean 'mother' or 'horse' and a few other meanings. These pitch patterns sound similar to the intonation of English in a syllable such as "Well.", "Well?", or "Well!". If the common models to account for the discrepancies between L1 and L2 vowels and consonants rely on similarities between sounds in the first language and the second language, then, what sound is tone similar to in English or any other Western language? The difference between English and Mandarin is that with a change in intonational pitch lexical meaning does not change in English, only the syntactic or pragmatic/affective meaning does. Recently, advances in this direction were made with the development of a model that takes into account dimensions other than vowels and consonant categories, namely, vowel length. This model (McAllister, Flege & Piske, 2002) was developed for the acquisition of L2 Swedish, a language prominently featuring the use of pitch in word accents.

A model focusing on segmental categories (PAM, Best, 1995; detailed explanation to appear in Chapter 2) appears to differ from a model focusing on suprasegmentals (McAllister et al., 2002; cf., Eliasson, 1997). Whereas the segmental models use categories as perceptual entities that can be equated across phonological systems, For example, the inventory of vowels in a language may differ in their number or location within the vowel space or where the boundaries lay for each vowel lies within the vowel space. Thus, learning non-native vowels may merely require reorganizing these boundaries or suppressing a category. In contrast, suprasegmental categories are less well defined and perhaps accordingly, the suprasegmental model relies on the extent to which a certain feature is used in one's first language in addition to the similarities to the counterpart feature in a second language. Of note, the suprasegmental model appears to be based on the assumption that the feature in question, i.e., vowel length, is the same and only varies in the extent it is used in a language. However, with tone the picture is more complicated. Tone can be characterized by several patterns to contrast words as mentioned above in the case of Mandarin Chinese. Conversely, a segmental-based model such as PAM (Best, 1995) appears to allow for tone patterns in the manner of vowels and consonants, but has more difficulty in categorizing English word stress.

The suprasegmental model, however, combined with segmental models (Best, 1995) may offer the solution to explaining perceptual gaps between the first language and second language when it comes to tone, particularly if the gap between the two models is bridged by features rather than tone pattern (cf., Brown, 1998). That is, the PAM model (Best, 1995) explains the ease or difficulty of perceiving non-native sounds according to the perceived (dis)similarities between L1 and non-native sounds. Yet, this model never clearly defines what these (dis)similarities are. I propose for at least the perception of tone that tone features may provide the answer. In the case of languages with tonal categories, these categories serve as bundles of tone features. This streamlines the perceptual process through tone-to-tone mapping, but can be both a positive or negative influence depending on the (dis)similarities between tone patterns within a tonal space. For non-tone languages such as pitch-accent and word stress languages, rather than pitch patterns analogous to tone patterns, the features of pitch would come to the fore. For example, pitch features may be merely a matter of height or register. L1 speakers of Japanese and English may hear merely high and low pitches in their L1 and so, for non-native tones they may hear a series of pitch heights to reflect the tone patterns. They might possibly hear the fall or rise in pitch as well, i.e., direction. And of course, we cannot completely rule out the possibility that they hear pitch patterns, e.g., High+Low or Low+High, reflective of both pitch accent and word stress patterns of words in those languages. Thus, combining the PAM segmental model (Best, 1995) with the Feature Hypothesis suprasegmental model (McAllister, 2002) through the feature of pitch may offer an approach for interpreting non-native tone perception by speakers of a variety of languages using lexically-contrastive pitch to varying degrees. This is particularly the case where both the native language and foreign language are both tone languages, but vary in the inventory of tones in terms of physical type and number.

However, what do we do with the languages of Western Europe such as English with its word stress or a language like Japanese with pitch accent which are not considered to necessarily be tone languages? The term "word stress" in English is a catch-all term for what in actuality is a bundle of features that are used to perceive and produce stress in English. That is, when a syllable is stressed in English it is longer in length, higher in pitch, and louder as compared to an unstressed syllable which is "inversely" shorter in length, lower in pitch, and quieter relative to a stressed syllable. Additionally, an unstressed syllable is made shorter in part because the vowel quality of the stressed syllable is oftentimes reduced to a schwa /ə/ as in the 'a' in the word 'about' or to another such short sound as /ɪ/ or /ɛ/. In this respect, (and if we restrict the discussion to the same linguistic unit size – the syllable) we cannot compare word stress nor intonation for that matter to tone directly, but only through the pitch used to mark word stress. This approach enables us to compare the extent of the usage of pitch to distinguish words in English in word stress as compared to the extent of the usage of tone to distinguish words in Mandarin Chinese. Similarly, pitch accent languages like Japanese are comparable to English in that an "accent" is placed on one part of a word. However, unlike in English, pitch is the only feature of the accent used to distinguish accented or unaccented mora. Thus, we have a string of segmentals like /a.me/ meaning either 'rain' or 'candy' depending on the pitch level, High-Low or Low-High, respectively. In this respect, the greater salience of lexically-contrastive pitch in Japanese as compared to the other features used in English word stress may allow for Japanese to be classified as a tone language, albeit more limited in terms of pitch patterns like Mandarin Chinese.

Yet, when it comes to comparing the inventory of pitch patterns used in word stress or pitch accent to that of pitch patterns used in tone, again the picture becomes less clear. For English are there indeed pitch patterns? Can the mere difference in pitch register, i.e., high vs low, count as a pattern? This is dubious within a segmental-based framework like PAM (Best, 1995) as it relies on the existence and a comparison between such patterns or categories for segmentals. For word stress, is a stressed pattern level or does it vary in direction (i.e., contour)? Also, is the high pitch of stress the only one that exists whereas the low pitch is as the term "unstressed" implies merely a lack of high pitch? In this respect, English would therefore have only one pattern perhaps just as Mandarin Chinese does not count "no tone" as one of the four tone patterns. The same argument might also be made for a pitch accent language as Japanese where the fall in pitch on the mora (a unit smaller than a syllable) in a word is what is most saliently perceived as the accent and the other moras display relatively high or low pitch.

Additionally, there are languages which do not use pitch to distinguish words, but they like all languages use intonation. Again, the question begs whether intonation patterns would be comparable to tone patterns or even accessed as part of the representation of a word. It is commonly agreed that in English the intonational patterns for questions and statements differ in a fairly set form. But these patterns cross over utterances, albeit even one-word utterances (e.g., Done? Done.). Thus, this invites the question as to whether intonational patterns are robust enough to be considered "set patterns" or categories, given that they cannot define the lexical meaning of a word. The position taken by this study is that English speakers do not associate any particular intonation pattern with the meaning of a word, only a word stress pattern.

As a result, when investigating the perception of tones by listeners of non-tonal languages, we need to ask only whether the pitch used in English word stress or Japanese pitch accent to distinguish words (and not necessarily intonation) can be accessed and applied when listening to tones in a foreign language. In order to determine if this is the case, we need to compare speakers of these languages to speakers of languages which are either tonal or do not use pitch at all to distinguish words. In this respect, we can see whether or not speakers of English and Japanese apply the lexically-contrastive pitch in their respective languages when listening to foreign language tones. And as all languages use intonation, we could generally assume if there is a difference among the speakers of various languages in their perception of foreign language tones, then it is most likely not due to the use of intonation in the first language, but to the extent to which pitch is used to distinguish words in the first language. Additionally, we can see if speakers of English and Japanese can perceive tones at the same level of accuracy as speakers of another tonal language. And by further looking at which tones are easier or more difficult to perceive, we may be able to determine how pitch in the L1 is mapped onto the non-native tones, thus allowing us to determine which pitch patterns are possibly accessed in the L1 and therefore, which features of pitch usage in the L1 are similar to those in the non-native language. Thus, the foreign accent in Mandarin of L1 English speakers and that of L1 Japanese speakers would differ in terms of tones due to the realization of the differences in lexically-contrastive pitch in the respective L1s.

From the discussion above, we can assume that learning a tone language may therefore be easier for an L1 speaker of a tone language as this speaker would merely need to adjust the number of tones in their L1 inventory or modify the shape of the their L1 tones. Whereas for a speaker of a non-tone language, it would require much more: they would need to introduce this new dimension of tones into their phonological system at the lexical level. From this assumption, another question arises: If a speaker of a non-tonal language learns a tone language, would knowledge aid them in perceiving and therefore, learning the tones of another tone language? We might expect that learners of Mandarin tones display more accurate perception abilities than their counterpart monolingual speakers who have not learned Mandarin when listening to an unknown tone language, due to the former groups' experience with tone alone. However, the true underlying question is whether learners have acquired tone in Mandarin and whether they react as native Mandarin speakers in the perception of Mandarin tone. Thereupon, we can specifically ask whether learners of Mandarin would react in the same manner as native speakers of Mandarin upon hearing the tones of another tone language (in this study, we used Thai, a language they have never learned before). We would expect learners of Mandarin to be better at distinguishing Thai tones in perception than native speakers of English who have never learned Mandarin. Additionally, we might hazard that just as in the case of L1 speakers of Mandarin, the perception of such learners depends on the type of tones they have learned in Mandarin as compared to the type of tones they hear in Thai. Here, the mappings would be similar for Mandarin NSs, and English learners of Mandarin, when listening to Thai tones. A suprasegmental model based on "functional load" or the extent of pitch to distinguish words in the native language may only be able to predict overall performance of speakers of languages varying in the functional load and not performance on individual tone patterns. In opposition, similarities between the two different tone languages might be analyzed under a segmental model and account for why certain tones are easier or more difficult to perceive. The segmental model might further be reduced to analyzing tones as individual features (e.g., [+high]) in order to better determine what accounts for similarities between the L1 and non-native sounds.

In sum, the first language of a speaker must therefore be considered when determining a person's perception of a second language (i.e., foreign language) – expressly, as performance on the naïve perception of a non-native tone language is shaped by the presence of the functional load of pitch in the first language to distinguish words. Many native speakers of English view Mandarin Chinese as a very "alien" language due to their unfamiliarity with tones. Many learners of Mandarin wonder whether they will be able to perceive and produce Mandarin tones in order to truly master this language. Yet, English too utilizes pitch in differentiating words and actually has a very rich and complex system of using pitch to convey meaning in the language. As such, contrary to common assumptions, native English speakers do not approach Mandarin tones empty-handed in terms of familiarity in using pitch to convey meaning. The issue is rather one of learning to use pitch differences at the syllable level to convey lexical meaning. It is the goal of the present dissertation to dispel the notion that native speakers of many non-tonal languages such as English approach tones without any experience with lexically-contrastive pitch. Another goal is to further determine the extent to which learners of Mandarin can acquire Mandarin tone. This goal is reinforced by looking at the application of these Mandarin tone patterns or features to the naïve perception of Thai tones as compared to speakers of languages with varying degrees of lexically-contrastive pitch in their respective L1s. A third goal is to interpret the results of this study within a framework combining a suprasegmental and segmental model bridged by a focus on pitch features.

Chapter 2 first introduces the basic features of lexically-contrastive pitch in the five languages of relevance in this dissertation: Thai, Mandarin, Japanese, English and Korean. Then, the literature review sets the theoretical framework of the study by discussing the perception and processing of tone by both native speakers of tone and non-tone languages (including both learners and non-learners). This section provides an overview of lexical encoding and priming in the L1 and L2 in the context of naïve and L2 tone perception of a non-native tone language. In Chapter 3, I describe the methodology and the results of a pilot study examining the non-native perception of tones, which were used as a basis for the current study. Chapter 4 presents a series of experiments conducted to replicate in part the pilot study in order to examine non-native perception of tones in more detail, with different stimuli. In Chapter 5, I examine two aspects of perception of tones by L2 learners of Mandarin: First, I examine whether perception of an unknown tone language (i.e., Thai) is facilitated by having specific experience in acquiring a tonal L2 (i.e., Mandarin). Second, I examine directly whether L2 learners of Mandarin have acquired the tonal contrasts in their lexical representations of Mandarin words. In Chapter 6, I discuss the differences found between non-native and L2 perception of tones, and wrap up the study with conclusions concerning the theoretical framework.

2. Background

Just as languages differ in terms of vowels and consonant categories, they differ in terms of the way they use pitch (or fundamental frequency, abbreviated as F0, i.e., height of the voice) to signal linguistic functions. For English speakers, perhaps intonation first comes to mind. All languages appear to use intonation. We can express questions and statements and emotions such as surprise or anger for example by using intonation. Intonation crosses over phrases or entire sentences and can even be used for one word, but intonation does not change the lexical meaning of that word, but rather its pragmatic function, sometimes including its syntactic structure, for example whether it is a statement or question (e.g., the "well?" – "well." example in the introduction, or "Coming?" vs "Yeah, I'm coming."). In addition to intonation, many languages use pitch to differentiate the lexical meaning of words. Combined with intonation, this lexically-contrastive pitch marks each language with a unique "melody" which causes native speakers of English to often stereotypically (and somewhat derisively) characterize Swedish or Chinese as "sing-songy". For speakers of English the use of lexically-contrastive pitch in the form of tone in Mandarin Chinese seems vastly different from their own language system, and yet, English uses pitch in word stress to help distinguish words as well. Thus, using pitch to contrast words is actually a common linguistic phenomenon.

In this chapter we examine the characteristics and degree of usage of lexically-contrastive pitch in the languages of the participants of this dissertation to better understand how the respective L1s may factor into the naïve and L2 perception of a tone language. I review the literature on cross-linguistic perception of lexically-contrastive pitch within a theoretical framework grounded in second language phonology, touching upon issues of processing, encoding, and priming in both the L1 and L2.

2.1 Cross-linguistic differences in phonological pitch systems

We can distinguish four language types that differ typologically in the prominence attributed to pitch in order to distinguish words (ranging from maximal/exclusive use to none). I specifically address in each of the following sub-sections Mandarin, Thai, Japanese, English and Korean since these are the languages spoken by the participants in this research.

2.1.1 Tone languages

The majority of languages in the world are tone languages (e.g. Mandarin Chinese, Thai, Vietnamese, and languages in Africa and Central America) (Yip, 2002). They use various and sometimes complex pitch patterns (referred to as *tone*) to contrast words. Pitch patterns may be relatively flat and spoken at the high, mid or low end of the voice range, or they may have a simple contour, i.e., pitch movement, by falling or rising. Additionally, they may have more complex contours by falling and then, rising or conversely, rising and falling. In tone languages, each syllable generally has a set pitch pattern, perceived by listeners primarily through pitch height and pitch movement (Gandour, 1983). One well-known classic example is the Mandarin Chinese segmental string of /ma/ which depending on the tone can mean 'mother' (*mā*), 'hemp' (*má*), 'horse' (*mǎ*), 'to scold' (*mà*), or mark a question (*ma*) when toneless[1] (See Table 2.1). Additionally, tones in Mandarin have been said to vary in length. For example, in citation form, tone 3 is longer, but in running speech, tone 3 becomes a low falling tone without the rising part, a phenomenon which renders the differences in tone duration undependable in identifying the tones (Kratochvil,

[1] The four tone patterns can be indicated by the Chao numbers (Chao, 1948) and descriptions: 55 (high level), 35 (rising), 214 (low falling rising or dipping) and 51 (high falling), respectively. These tones are also correspondingly commonly referred to as tones 1, 2, 3, and 4. Yet, the Chao numbers serve as a convenient method to easily indicate and compare pitch levels and directions. However, it is important to keep in mind that the Chao numerical designations are relative, as the tones can vary physically depending on gender and individual variation (e.g. in larynx size), and on the height of the overall pattern increases for questions as opposed to statements (Lin, 2007).

1987, cited in Hallé et al., 2004)[2]. Table 2.1 offers an overview of the different tone patterns in the two tone languages relevant in this dissertation: Mandarin and Thai.

Table 2.1. Overview of tones in Mandarin and Thai

	Mandarin			Thai		
	e.g.	Chao	Label	e.g.	Chao	Label
Level	mā *mother*	55	high level Tone 1	ná: *aunt*	--	high
				na: *rice field*	--	mid
				nà: *custard apple*	--	low
Contour	má *hemp*	35	rising Tone 2	nǎ: *thick*	--	falling
	mǎ *horse*	214	low falling rising or dipping Tone 3			
				nâ: *face*	--	rising
	mà *to scold*	51	high falling Tone 4			

Based on Chao (1948: 24-25) for Mandarin, Zsiga and Nitisaroj (2007: 344) for Thai.

In contrast to Mandarin, Thai has five tones composed of three flat tones (i.e., low, mid, high) and two contour tones (i.e., falling, rising). For example, the segmental string /naa/ can mean 'rice field' (*naa*, mid tone), a nickname (*nàa*, low tone), 'face' or 'season' (*nâa*, falling tone), 'aunt/uncle (mother's younger sibling)' (*náa*, high tone) and 'thick' (*nǎa*, rising tone). However, there is a discrepancy between these phonological descriptions and phonetic descriptions. Literature concerning Thai traditionally describes Thai as having three level tones (Abramson, 1975). However, examining Figure 2.1 we see that the high tone (H) is not phonetically level but in fact almost parallels the rising tone (R) in shape. Indeed, the low and mid tones are not "truly flat" as the former falls while the latter rises at the end, albeit the two are relatively flatter than the

[2] Further discrepancies must be noted between different dialects such as Beijing Mandarin tones and Taiwanese Mandarin where for tone 2 the beginning part clearly falls in Taiwanese Mandarin while the end part rises slightly in opposition to Beijing Mandarin which clearly rises (Fon and Chiang, 1999, cited in Hallé, Chang, & Best, 2004).

so-called contour tones which clearly rise **and** fall or vice-versa. Yet, to a smaller or larger degree, listeners must follow the height and/or direction of pitch movement to distinguish tones from one another. For example, height may be more salient than direction as the only feature to distinguish the Low tone from the Mid tone. Likewise, direction alone may be sufficient to distinguish the falling from the rising tone but pitch height most likely aids in perception. However, the difference between Low and Mid or High and Rising may rely exclusively on height. Also, there are phonological descriptions of the contour tones being composed of Low, Mid and High level tones. For example, Yip (2002) refers to the five Thai tones as M, L, HL, H, LH[3] (i.e., Mid, Low, High+Low, High, Low+High) where HL refers to the falling tone and LH refers to the rising tone. This is a compositional position that is also taken by Morén and Zsiga (2006) in their analysis of Thai.

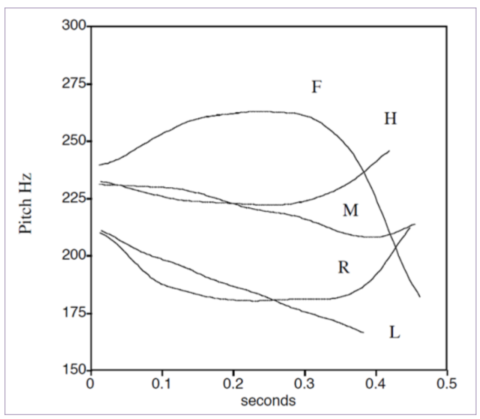


Figure 2.1. Contour shapes of Thai tones in citation form. Representative examples from one speaker. From Zsiga & Nitisaroj (2007: 347)

In sum, we can say that the two tone languages Mandarin and Thai use pitch contrasts maximally since pitch variations distinguish words lexically at the syllable level.

[3] Variation has been noted in the realization of Thai tones according to context exemplified by the contrast of the citation form with the connected speech form (Abramson, 1979). It is also the case that the so-called flat tones are not "completely" flat.

2.1.2 Pitch-accent languages

Pitch-accent languages also use pitch to distinguish words, but not to the same extent and manner as tone languages do. This diverse group of languages includes standard Japanese, Scandinavian languages, Serbo-Croatian, and a major dialect of Korean spoken in the Kyungsang region in the southeastern part of South Korea (Jun, Kim, Lee & Jun, 2006; Lee & Ramsey, 2000). Pitch accent languages are considered a sub-set of tone languages (Yip, 2002: 4) or a "mixture" of tone and stress languages (McCawley, 1978; Vance, 1987; Gussenhoven, 2004; Hyman, 2009; Ladefoged and Johnston, 2010). Yip (2002) adopts Hyman's definition of a tone language to encompass pitch accent languages: 'A language with tone is one which an indication of pitch enters into the lexical realization of at least some morphemes' (p. 4).

Thus, in a pitch accent language like Japanese, one mora of each word receives a high pitch (i.e., the "word accent") and determines the pitch on the preceding or subsequent moras (Kubozono, 1999), much like the placement of stress in English. As such, the pitch pattern of a word is predictable if the position of the word accent is known (Tsujimura, 2006: p. 74) and therefore, is considered lexical information needed to learn a word. In short, all the moras before the accented moras also have a high pitch while the moras after the accented mora have a low pitch. Also, the first mora of a word has a low pitch unless it is accented (Initial Lowering Rule, Haraguchi, 1977). One example is /ka.ki.ga/ which can mean 'oyster' (HLL or initial-accented where the first mora is marked by a high tone and the following two moras are marked by low tones), 'fence' (LHL or final-accented), or 'persimmon' (LHH or unaccented) where *ga* is a particle indicating the nominative case[4].

[4] Additionally, in the case of Japanese, it must be noted that some Japanese dialects vary from standard Japanese in their pitch accent systems having more or fewer patterns, including pitch accentless varieties, and/or having almost the opposite pattern for the "same" words as in the case of the Kyoto dialect (e.g., /ka.ki.ga/ 'oyster' LHL, 'hedge' HLL, and 'persimmon' HHH) (Yoshida & Zamma, 2001: 216).

Thus, pitch accent languages resemble tone languages as they employ pitch to contrast words, but with less pitch movement, or a more restricted inventory of contours that can be realized on one syllable/mora.

2.1.3 Stress-accent languages

A third group can be called "stress languages". In stress languages like English, pitch is never used alone to distinguish meaning or mark prominence. The stressed syllable generally features higher pitch, greater intensity, lengthened vowels and unreduced vowel quality (Fry, 1958). A typical contrast of stress in English would be the noun "IMpact" and the verb "imPACT" (where capital letters indicate the stressed syllable). Pitch may, however, play a lesser role than loudness and vowel duration in identifying stress among native speakers (Kochanski, Grabe, Coleman & Rosner, 2005) although pitch has also been shown to be the most salient feature (Fry, 1955; 1958; Lieberman, 1957, cited in Tuc, 2003). Otake and Cutler (1999) state that the role of stress alone is diminished in word activation perhaps due to the strong correlation of vowel quality with stress. For example, unlike the IMpact/imPACT minimal pair where the two are distinguished by both pitch and length, it is more common to distinguish such minimal pairs by pitch, length **and** vowel quality in English as in "DEsert" (noun) vs "deSERT" (verb).

Thus, stress-accent languages differ from both tone and pitch-accent languages as pitch is not necessarily used exclusively to mark prominence in words. Additionally, as seen in the case of English, pitch may not play the most salient role in marking the prominence of a syllable.

2.1.4 Languages with no use of lexical pitch ("Intonation languages")

Finally, as mentioned above some languages do not use tone, pitch accent or stress to distinguish words. We refer to these languages as "intonation" languages as they use pitch and the associated features of duration and intensity at the level of the phrase (even if the phrase is just one word long, such as “really?”), but they do not use pitch contrastively at the level of words to distinguish lexical meaning as is the case of other languages. Pitch is used only to signal meaning at the phrase level (i.e., intonation), and this is the case for the other language types mentioned above as well. Such languages include standard Korean (Kim-Renaud, 2009: 22), the Fukushima dialect of Japanese (Otake & Cutler, 1999; Otake & Higuchi, 2004; Utsugi, Koizumi & Mazuka, 2010), the Kumamoto dialect of Japanese (Otake & Cutler, 1999; Otake & Higuchi, 2004) or French (Dupoux, Sebastián-Gallés, Navarrete & Peperkamp, 2008: 685). However, like all languages, they use intonation which uses pitch height and pitch contour at a phrase or sentence level to communicate questions, statements, and emotional state but not to distinguish words (Hirst & Cristo, 1998).

Additionally, it must be noted for standard Korean that lexically contrastive usage of pitch may be appearing among younger speakers of the Seoul dialect born after 1965 (Silva, 2006). Korean possesses a three-way contrast for stops: lax, tense, and aspirated, and the VOT between lax and aspirated stops has decreased, creating a complete overlap between the two VOTs. To maintain the contrast, lax stops employ a lower F0 while aspirated stops have a higher F0, but only when these stops begin a word. Nonetheless, Korean does exhibit stress realized as various patterns of pitch differences in words such as /kal.bi/ which is stressed on the first syllable to mean ‘spare ribs’ or /a.chim/ which is stressed on the second syllable to mean ‘morning’ (Song, 2005: 40). Therefore, we cannot discount the existence of stress in Korean, but rather its lack of lexically contrastive usage in at least standard Korean and hence, the much lower prominence of pitch. However, in opposition to the lack of lexically-contrastive pitch usage in standard Korean, the Kyungsang Korean dialect (see section 2.1.2 above) does display lexically-contrastive pitch (Jun, Kim, Lee & Jun, 2006; Lee & Ramsey, 2000). As the Kyungsang Korean dialect is a major dialect spoken in the city of Busan and its surroundings, it must be kept in mind when examining speakers of Korean.

For standard Korean, pitch is not used at the lexical level but only at the phrasal level. Thus, “intonation-only” languages differ markedly from the other language types in their lack of lexically-contrastive pitch usage.

2.2 Perception and processing of tone

We now review the mechanisms underlying the perception, processing and lexical encoding of tone by native speakers.

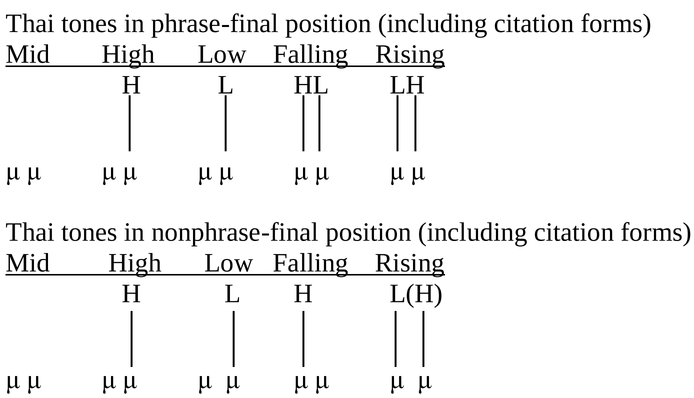
2.2.1 Perception of L1 tone

The two features thought to play the largest role in both L1 and L2 perception of tone are pitch height and pitch direction (Gandour, 1983; Wang, Jongman, Sereno, 2006; Vu, 1981), where direction refers to the movement of pitch whether up or down or level. In terms of pitch height, listeners evaluate whether the tone is high or low or, in some tone languages, in the middle area of the voice range. Even though height is relative and is influenced by gender and voice quality, Lee (2009) shows that native speakers of Mandarin are able to rapidly and accurately identify F0 height by gauging the voice range of multispeaker stimuli from both males and females with no previous exposure to the speakers and only the beginning of the syllable. In terms of pitch direction, listeners detect whether the tone moves up or down or both or whether it remains relatively flat. Thus, tone perception can be measured in terms of perception of pitch height and/or direction.

Across languages, listeners weigh pitch height more than pitch direction. However, the L1 influences the relative weight given to these two features. Gandour (1983) compared listeners of Thai, Mandarin, Cantonese, Taiwanese (and English) on the perception of 19 different synthetic tones reflecting pitch patterns found in tone languages (i.e., 5 level tones, 14 contour tones – 8 unidirectional tones of falling or rising tones and 6 bidirectional tones combining falling with rising tones – where beginning and end points varied within each subset). Participants were asked to differentiate pairs on an 11-point scale (*no difference* to *extreme difference*), after hearing the stimulus set twice. The findings show that every group used height more than direction to judge tone dissimilarity, even though tone language speakers were found to use direction to a larger extent than non-tone-language speakers (English). Variance in perceived differences was attributed to the two features of height and direction. This prominence can be ascribed to the fact that only pitch height allows listeners to discriminate between some tones (Tuc, 2003) such as the mid and high tone in Thai. The relative difference in weighting of pitch height and direction is thus shaped by the L1, influenced further by other factors, specific to the task, such as the number of voices used for the stimuli.

Additionally, in opposition to the feature of height or direction in a unitary model, we must keep in mind an alternative approach of viewing direction as a sequence of high and low tones under a compositional model (Morén & Zsiga, 2006). Based on a perceptual study by Mixdorff, Luksaneeyanawin, Fujisaki, & Charnavit (2002) that asserts that rather than pitch contours or slopes, Thai tones consist of high and low pitches associated with certain points within a syllable, Morén and Zsiga (2006) put forth the Moraic Alignment Hypothesis to describe Thai tones. In this hypothesis, pitch heights are associated to certain moras in a word, and it is this alignment that serves as the key perceptual cues to distinguishing the five tones of Thai from one another. Each stressed syllable in Thai must be bimoraic (Bennet, 1994 cited in Zsiga and Nitisaroj, 2007) so that monomoraic CV syllables occur only in polysyllabic words and only then in an unstressed non-final position. When such syllables stand alone, there is actually a glottal stop after the vowel creating a bimoraic word (p. 346). Under the moraic alignment hypothesis, pitch height and its alignment with a specific mora serve as the primary cue to Thai tones in both citation and connected forms while pitch contour and slope are utilized in cases of ambiguity. Simply put, Morén and Zsiga’s (2006) autosegmental representation consisting of only High and Low tones can distinguish the five Thai tones, as illustrated in Figure 2.2. The mid tone does not bear either a High or Low tone while the high tone bears a High tone on the second mora and the low tone bears a Low tone also on the second mora. In contrast, the falling tone is composed of a High + Low tone with the high tone on the first mora and the low tone on the second mora while the rising tone is composed of a Low + High tone with the low tone on the first mora and the high tone on the second mora. In the case of these latter two tones, the tone on the second mora may or may not be elided.

Figure 2.2. Moraic alignment hypothesis (Morén & Zsiga, 2006)



(in Zsiga & Nitisaroj, 2007, pp. 344-345)

Thus, hearing the location of the high tone or low tone on the first or second mora of a word cues the listener into knowing the tone of the word and hence, the meaning. Under such a hypothesis, we might expect pitch accent speakers to be able to perceive Thai tones quite easily, especially if the L1 is also a mora-timed language like Japanese where each mora bears a high or low pitch determined by the placement of the accent and its high pitch. This possibility is further highlighted by the study of Cutler and Otake (1994) demonstrating that Japanese use a mora-based strategy when listening to L2 English words and thus showing the segmentation of L2 words to be language specific.

The syllable is considered the tone bearing unit (TBU) in Mandarin (Yip, 1995). However, the mora has also been posited as the TBU for Mandarin (Duanmu, 2004), which is the case for other Chinese languages. Thus, Mandarin speakers may actually perceive tones along the lines of the Moraic Alignment Hypothesis for Thai (Morén and Zsiga, 2006). For example, Lee, Bond, and Tao (2008) compared the identification of intact, silent-center, center-only and onset-only Mandarin tones by native speakers. They found that identification in all cases was above chance but in especially the onset-only cases, tones 2 and 3 were confused with one another as both have low F0 onsets while tones 1 and 4 were confused as both have high F0s. Confusion was asymmetrical with tone 1, which was identified more often as tone 4 than vice-versa, and with tone 2, identified as tone 3 more often than vice-versa. Such findings bolster the claims for a similar approach for Mandarin tones as the Moraic Alignment Hypothesis for Thai tones. That is, contour tones would be formed of sequences of high and low pitches and the location of the high or low pitch within a word might trigger identification of a tone. Thus, such an approach might be universal for tone perception and must therefore be kept in mind when examining the cross-linguistic perception of tones whether Thai or Mandarin tones. In sum, both the unitary and compositional models for tones appear plausible and indeed both may be valid. The unitary model may merely be language-specific "streamlined perceptual shorthand" which could both aid and hinder native speakers of Mandarin in mapping Thai tones onto Mandarin tones due to similarity or dissimilarity respectively. The compositional model on the other hand may be validated should Mandarin speakers perceive with high accuracy Thai tone patterns not present in Mandarin. The compositional approach may also be further validated should the Japanese with a mora-based pitch accent system be able to apply map moraically-aligned Thai tones onto Japanese high- and low-pitched moras. The position generally adopted for this study is a compositional approach with a unitary approach serving as a type of shorthand.

Additional features have been put forth to explain how native speakers perceive various tones. Gandour (1983) names three such features in addition to the two just reviewed: 1) F0 slope/magnitude of pitch change; 2) extreme endpoints; and 3) tone duration. This dissertation, however, focuses on the features of pitch height and pitch direction as these appear to be the most salient in tone perception for Mandarin, Cantonese, (Francis et al., 2008; Gandour, 1983), Taiwanese (Gandour, 1983), Thai (Gandour & Harshman, 1978; Gandour, 1983), and Vietnamese (Vũ Thanh Phương, 1981, cited in Tuc, 2003).

Similarities between tone patterns may also cause confusion (Leather, 1990) as in the case of tones with the same directional pattern (i.e., both flat or both rising) or in the same register (i.e. a low flat tone and low-to-mid rising tone). Thus, some tonal contrasts are easier to perceive than others for both native and non-native speakers of the target tone language (Abramson, 1975, 1978; Burnham et al., 1992). When comparing Thai tones in an AX task, L1 Thai listeners have the most difficulty comparing flat vs. contour tones (e.g., L vs. R) while the comparison of contour tones (R vs. F) with one another is the easiest (Burnham et al., 1992), suggesting that the perception of pitch height, specifically the "absolute initial pitch of the component tones" (p. 555), is the most salient for L1 Thai speakers when perceiving contrasting tones. Yet, even L1 Thai speakers frequently confuse the mid and low tones (Abramson, 1976), hinting that L1 Thai speakers generally require more than pitch height or absolute initial pitch when discriminating tones. That is, the low and mid tones have the same shape (i.e., direction) and differ only by height which additionally is minimal so that both tones possibly occur in the same register. In contrast, all other tone comparisons allow the listener to access the two differing features of direction and height.

Within a Moraic Alignment Hypothesis, we witness the same problem between these two tones: both have the mid tone on the first mora while the second mora is mid for the mid tone and low for the low tone. Thus, again listeners must rely on the only difference between the two tones: a mid mora vs a low mora. Interestingly, the high tone varies from the mid tone by only one difference: Both have a mid tone on the first mora while the second mora of the high tone is high. Thus, in order to account for the confusion between the low and mid tones for L1 Thai speakers, we must surmise under the Moraic Aligment Model that the height between the mid and high tone varies greatly compared to that between the low and mid tone. In contrast, under a compositional model the mid and high tone differ by both direction and height.

In a parallel situation for Mandarin, L1 speakers also confuse similar tones such as the rising and dipping tones (tones 2 and 3, respectively) (Li & Thompson, 1977; Leather, 1983, 1990; Wang, Spence, Jongman & Sereno, 1999). We also find a somewhat analogous situation in the acquisition of L1 Mandarin tone by children whereby tones 2 and 3 are not only acquired last but are also substituted for one another during acquisition (Li & Thompson, 1977). Also, as noted above L1 Mandarin speakers can generally place tones 1, 4 and tones 2, 3 into either the high or low registers respectively but confuse tones sharing similar heights (Lee, Tao & Bond, 2008).

In sum, we see that pitch height appears to be the most salient feature of tone for native speakers of both tonal and non-tonal languages (Gandour, 1983), but pitch direction also appears to be necessary to discriminate tones which have the same initial pitch level (Abramson, 1976) or occur in the same register. When tones have similar pitch heights and directions, they are easily confusable (e.g., Mandarin rising and dipping tones). Yet, the picture appears at times somewhat more complicated as pitch height and direction may not be sufficient to discriminate all tone comparisons (e.g., confusion among native Thai listeners in the perception of the mid and low tones).

2.2.2 Lexical encoding of suprasegmentals and tone in the L1

In this section, I review what is known about the manner in which tones are encoded into the lexical representation for words.

Just as for segmentals, suprasegmentals are encoded lexically and constrain word recognition (Soto-Faraco, Sebastián-Gallés, & Cutler, 2001), even though the strength of the effect is modulated by language. In Spanish, suprasegmental information can constrain lexical activation in the same way as segmental information does (cf., Dupoux, Sebastián-Gallés, Navarrete, & Peperkamp, 2008). That is, just as the differences between the consonants in "right" and "light" or between the vowels in "ship" and "sheep" cause L1 English speakers to hear (or activate) different words, the differences in suprasegmentals such as the use of pitch in word stress also activates different words (i.e., "an <u>in</u>sert" vs "to in<u>sert</u>" in English). However, in English, this effect is attenuated (as mentioned by the presence of other correlates of word stress such as vowel length etc.), suggesting that suprasegmental information alone does not constrain lexical activation to the same extent as segmentals (Cooper, Cutler, and Wales, 2002).

In Mandarin, in the case of minimal pairs differing in tone (e.g., *tang* [t^haŋ]: tone 1 "soup" 湯; tone 2 "candy" 糖), both words are initially activated, with tonal information being accessed soon after to disambiguate the pairs (Lee, 2007). Moreover, when classifying segmentals under the condition where irrelevant changes in the pitch height of the tone have been made, L1 speakers of Mandarin react more slowly than L1 English speakers, suggesting that Mandarin listeners have difficulties ignoring tonal information, unlike English listeners for whom it is not relevant (Lee and Nussbaum, 1993). However, L1 Cantonese speakers experience greater difficulty in their judgment of nonwords versus words on a lexical decision task when the target item differs in tone only as compared to differences in segmentals only, suggesting a greater reliance on segmental over tonal information in lexical representations (Cutler & Chen, 1997).

2.2.3 Priming in the L1

Lexical representations and the phonological information stored in them can also be investigated using priming. One such technique is repetition priming. Repetition priming is the effect where speakers hear a word that has the same phonological form (i.e., the same word in both sound and meaning) twice and react to hearing the same word (target) the second time more quickly than when hearing that word (prime) for the first time. In this respect, it is said that the first word primes the second word (i.e., causes a faster reaction) (cf., Zwitserlood, 1996). While facilitation is expected through faster reaction times (RT), inhibitory effects do occur. Facilitation is prelexical and concerns the "overlap" in form ("form-based facilitation") while inhibition is "a lexical-based effect" due to competition between activated words ("inhibitory lexical competition" or "competitive-lexicon?-based inhibition," Lee, 2007, p. 103). However, generally facilitation is the expected outcome for repetition priming for the same words while no facilitation is expected for minimal pairs. This effect has been witnessed both for the L1 for segments (e.g., /e/ vs /ɛ/ as in *Pere* 'Peter' vs *pera* 'pear' for Catalan speakers, Pallier et al. 2001) and for suprasegmentals such as pitch accent in Japanese (Cutler and Otake, 1999) where pitch-accent minimal pairs did not prime one another. Similarly, priming did not occur in word-stress minimal pairs in Dutch (Cutler and Donselaar, 2001) and in English (Cooper, Cutler, & Wales, 2002). That is, facilitative priming for minimal pairs is to be expected only if (word-stress) minimal pairs are perceived and treated as homophones during lexical processing (see Pallier et al., 2001). The fact that there is no priming indicates that they are not treated as homophones, and therefore, that the prosodic information is used to constrain lexical activation by reducing the number of activated candidates

(Cooper et al., 2002; Cutler & Donselaar, 2001; Cutler & Otake, 1999; Soto-Faraco, Sebastian-Galles, & Cutler, 2001, cf, Lee, 2007 for more details).

In the case of L1 encoding of tone, both facilitatory and inhibitory effects have been reported in the literature. For Cantonese, priming in monosyllabic words has been observed for tonal minimal pairs that overlap in onset and/or rhyme but not for tone (i.e., minimal tone pairs) (Yip, Leung, & Chen, 1998). That is, the researchers used four types of word pairs: 1) *cho* (tone 2) with *cho* (tone 1), 2) *cho* (tone 2) with *do* (tone 2), 3) *cho* (tone 2) with *chi* (tone 2) and 4) *cho* (tone 2) with *gwa* (tone 1). They found that only in the first case of tonal minimal pairs that priming occurred despite differences in tone. However, facilitation has been found for Cantonese when both rhyme and tone overlap as well as when both onset and rhyme overlap (Yip, 2001). Additionally, in another study (Cutler & Chen, 1995), both a facilitatory and inhibitory effect for both rhyme (i.e., vowel) and tone has been shown for Cantonese disyllabic words with a mismatch in rhyme and/or tone in either the first or the second syllable. When the first syllable of the target was identical to the prime in rhyme and tone while the second syllable varied in either rhyme or tone, lexical competition resulted, causing an inhibitory effect and slower reaction times. In the other case, when the second syllable of the target was identical to the prime in rhyme and tone while the first syllable varied in either rhyme or tone, a facilitatory effect emerged, resulting in faster reaction times (Cutler & Chen, 1995, cf., Lee, 2007). Thus, encoding of tone appears to depend on several factors, e.g., how much and/or what part of the syllable overlaps (i.e., onset, rhyme, tone), and whether the first or second syllable is the same for disyllabic words.

In the case of Mandarin tones, the picture is similar. Lee (2007) examined priming with Mandarin stimuli in four conditions: Tone overlap (*segmental mismatch*), Segment overlap (*tonal mismatch*), Tone+Segment overlap (*repetition*), and unrelated. Participants heard two words in succession with an inter-stimulus interval (ISI) of 50ms or 250ms and were asked to respond as quickly as possible to whether the second word was real or not (press YES or NO). Priming occurred in a hierarchical scale with statistical significance in both ISI lengths. That is, when both the segmental and tone information were the same between the two words, reactions were the fastest in the case of *repetition* (tone + segment overlap), followed by when segment overlap (*tonal mismatch*), followed next by tone overlap (*segmental mismatch*). In his study, Lee (2007) also examined mediated semantic priming, that is, the case where a word (e.g., ‘light’) primes a semantically-related one (e.g. ‘correct’) but through a third phonologically similar word creating a minimal pair (e.g., ‘right’) with the prime ‘light’. That is, using an English analogy, Lee tested whether using a prime such as “light” would prime the target of “correct” by the phonological similarity of “light” to “right”. However, Lee varied the words only in tone and not segments (e.g., *lou2* ‘hall’[5] versus *lou3* ‘hug’ priming *jian4zhu0* ‘building”). Again, participants heard two words in succession and were asked to decide whether the second item (target) was a real word or not. In the case of an ISI of 250ms between the two words, there was no priming, i.e., the prime word unrelated in meaning to the target word failed to activate its minimal counterpart (the phonologically similar word) which was semantically related to the target word. That is, when the participant heard, for example, the prime word “light,” it did not activate the word “right” and so, there would be no semantic priming of the target (second) word “correct”. However, in the case of an ISI of 50ms, there was priming, i.e., the prime word unrelated in meaning activated its minimal counterpart, thus priming the target second word. Lee concludes that tone is used to choose among activated candidates, but in the early stage of lexical activation tone does not prevent the activation of minimal tone pairs and priming effects are “rapid and short-lived” (Lee, 2007, p. 105). The

[5] The numbers in the examples (2, 3, 4, 0) refer to the tones applied to the syllables. For instance, lou2 refers to the word /lou/ with tone 2.

results of this study were corroborated by another study (Shuai, Li, & Gong, 2012) which found that segmentals initiate the activation of candidates under 200ms and that tonal information reduces the number of candidates afterwards though not appreciably after 400ms.

In sum, published research reveals a mixed picture of whether tones constrain lexical activation more or less so than segments in L1 speakers of a tone language. Clearly, however, tonal information is part of lexical representations in L1 tonal speakers and can be used to limit the number of activated candidates during word recognition. In this dissertation, I will investigate whether second language learners of Mandarin are able to encode tonal information in lexical representations to constrain lexical activation as well.

2.3 Second language phonology

Since the goal of this dissertation is to examine tonal perception and encoding in listeners whose L1 is not a tonal language, I will now turn to the specific challenges faced by naïve non-native listeners (= non-native listeners without any experience with the language) and second language learners (non-native listeners who have some experience with the language) during their processing of a non-native phonological system.

We do not perceive foreign sounds in a neutral way. Our L1 sound system acts as a filter when we hear foreign sounds. As a result, L2 learners experience difficulties distinguishing foreign words and often speak with an identifiable foreign accent. Indeed, when we first perceive a foreign/second language, we apply our "L1 ears" to the new sounds (Polivanov, 1931; Trubetzkoy, 1939; Weinreich, 1953) although it has been noted that this phonological filter may not apply to some sounds like Zulu clicks for certain L1 speakers (Best, McRoberts, & Goodell, 2001).

For L2-learners, the presence of two sound systems results in their interaction within one bilingual individual and guides the linguistic perception of sounds (Interaction Hypothesis, Flege, Yeni-Komshian & Liu, 1999; Weinreich, 1953), with the strength of influence depending on factors such as amount of L1/L2 usage and/or age of arrival (AOA) in the L2 environment (Flege et al., 1999). Interaction between two tonal inventories among bilingual speakers of two tone languages appears to affect tone production as well (Wu & Fon, 2010). Thus, the L1 plays a major role in the process of perceiving and acquiring the sounds of an L2 (Purcell & Suter, 1980) although other such factors as insufficient amount and type of language input (Bialystok & Hakuta, 1999) and other factors cannot be discounted.

2.3.1 Segmentals

The classic case of the L1 shaping perception of L2 segmentals, i.e., vowels and consonants, is the /l/~/r/ distinction in English for L1 speakers of Japanese (Goto, 1971). The lack of either /l/ or /r/ in the Japanese phonological inventory causes Japanese listeners to associate either to the Japanese flap /ɾ/ in perception. Japanese do this despite being able to produce the two sounds or to distinguish the two sounds in isolated cases at the level of L1 English speakers (Goto, 1971; Miyawaki, Strange, Verbrugge, Liberman, Jenkins, & Fujimura, 1975; Sheldon & Strange, 1982).

At the lexical level, a potential consequence of miscategorizing these English sounds is that L2 listeners judge non-words more frequently as words than L1 listeners when such a difficult segmental contrast is involved (Broersma, 2002; Broersma & Cutler, 2011, for Dutch learners of English). If segmental contrasts are perceived to be the same, minimal pairs that differ only in terms of that contrast may be perceived as homophones, and this can trigger positive priming between minimal pairs for these L2 learners (early or late) as evidenced by faster response times in a lexical decision task (Darcy, Dekydtspotter, Sprouse, Glover, Kaden, McGuire, & Scott, 2012; Pallier, Colomé, & Sebastian-Gallés, 2001), but not for native speakers.

Currently, several models attempt to account for L2 sound perception, predicting that L2 sound categories that are very different from the L1 sounds will eventually be accurately perceived and acquired, while L2 sound categories which are similar and only differ slightly from L1 sounds will remain difficult to perceive. While many other models have been put forth (Speech Learning Model, Flege, 1995; Native Language Magnet Model, Kuhl & Iverson, 1995; Kuhl, Conboy, Coffey-Corina, Padden, Rivera-Gaxiola, & Nelson, 2008), for the current study the Perceptual Assimilation Model (PAM: Best, 1995), which concerns naïve listeners, and the Perceptual Assimilation Model for Second Language Learners (PAM-L2: Best & Tyler, 2007) offer perhaps the optimal model applied to segmentals. They also seem to also best account for the perception of the suprasegmental of tone, particularly for tone to tone mapping. Under the PAM/PAM-L2 model, an L2 segmental category may or may not be assimilated to an L1 category depending on its perceived similarity or dissimilarity to an L1 segmental (see Figure 2.3). If the L2 sound is perceived to not be language (but rather some non-linguistic noise, for instance), then, it will be "unassimilable". If it is considered to be somehow a language sound, but a sound "very different" from any L1 segmental, the L2 segmental is considered to be "uncategorizable". In both these cases, discrimination between this sound and another sound categorized in one of these ways should be "very good" and "poor to moderate", respectively. Should the L2 segmental be perceived as similar to an L1 segmental and therefore, categorizable, it may be categorized in one of three ways. When contrasting two L2 segmentals, if both are assimilated to different L1 categories, then we have the case of a "two-category" pattern" with "great accuracy" in discrimination between the two. If both L2 segmentals are assimilated to the same L1 category, but one is considered a "better fit" to that category, then we have a "category goodness difference" pattern with "moderate to good" perception of that contrast. And if in this case both L2 segmentals are considered to be equally good, then we have a "single category" pattern with "poor to moderate" perception of that contrast. Additionally, contrasting one sound which is categorizable with one that is uncategorizable results in “very good” perception (PAM: Best, 1995; PAM-L2: Best & Tyler, 2007).

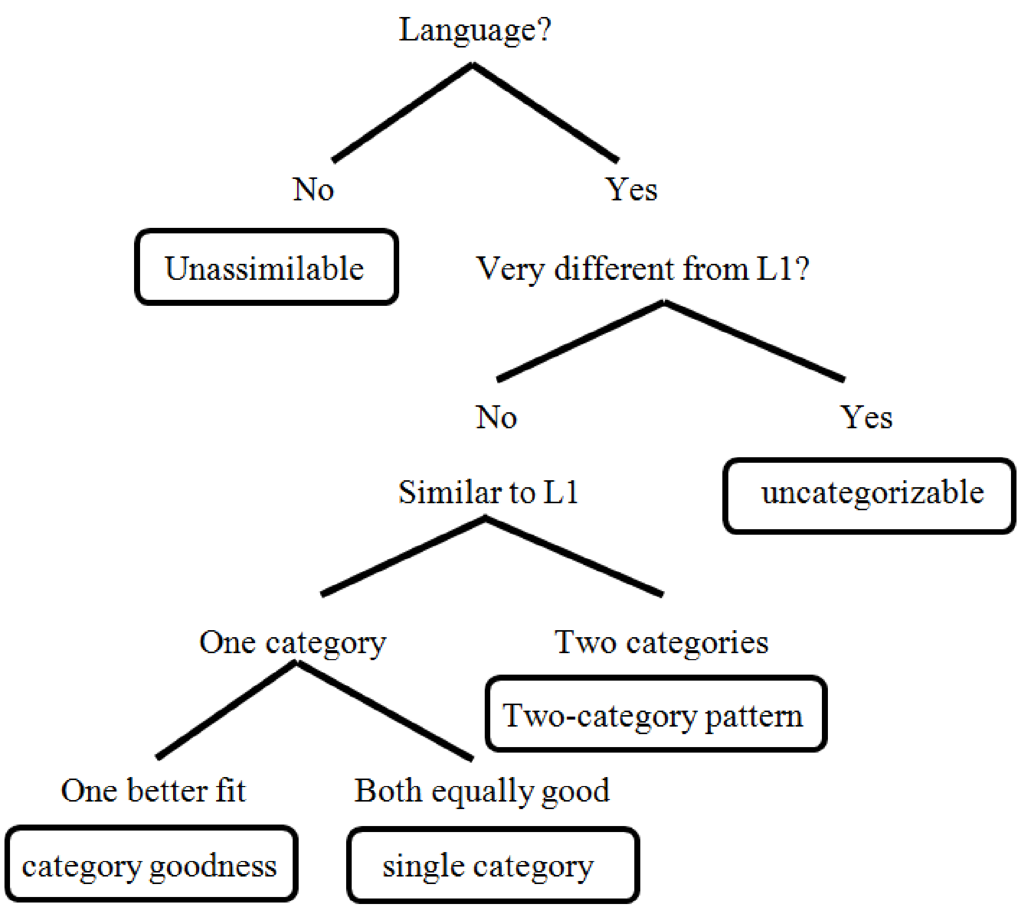


Figure 2.3 Categorization of a non-native or L2 sound in the Perceptual Assimilation Model (PAM) (Best, 1995)

Thus, we see that the PAM model revolves around the concept of similarity as a central predictor of perceptual inaccuracy. However, the PAM model does not offer a clear definition of what constitutes sufficient similarity or dissimilarity between sound categories. In response, within the framework of the PAM model, sounds could be further broken down into a combination of phonetic features, providing one possible way to determine this boundary between similar and different sounds (e.g., Phonological Interference Model, Brown, 1998). In addition to the PAM model, another approach adopted in this study which could be also applied to account for both naïve and L2 perception is the Feature Hypothesis (McAllister, Piske, & Flege, 2002) which unlike the PAM model was put forth to account for the L2 perception of suprasegmentals, namely vowel length (see next section 2.3.2 Suprasegmentals for a detailed description). However, as segmentals and suprasegmentals differ in nature, second language phonology for suprasegmentals must first be discussed in order to compare models for accounting for the perception of Thai tone.

2.3.2 Suprasegmentals

According to Polivanov (1931), the entire phonological system of the L1 will influence the perception and phonological processing of a non-native phonological system whether by L2 learners or non-native naïve listeners (i.e., second or a foreign language). It appears then straightforward to generalize findings from segmental perception to suprasegmental perception in that the suprasegmental perception may follow the same mechanisms that have been put forward for segmentals. We, however, cannot automatically infer from the perception of segmentals how the perception of suprasegmentals works. Especially when discussing pitch perception for the purpose of understanding cross-linguistic perception of tone, it is not straightforward as linguistic pitch exists in all languages in the form of intonation while for some to distinguish lexical items to varying degrees and manner. As Dupoux et al. (2008) point out, the current theoretical framework of L2 phonology models tends to largely focus on the perception of segmentals. A more refined understanding of the mechanisms underlying perception and processing of suprasegmental dimension is needed in order to be able to expand the segmental models to suprasegmental domains. Should it be found that suprasegmental perception does not work like segmental perception, a separate model will be required.

Research points to similarities between the segmental and suprasegmental domains in phonological processing: non-native suprasegmental dimensions such as word stress have been shown to be difficult to perceive for non-native (naïve) listeners who do not use this dimension in a lexically contrastive manner in their L1 (Dupoux, Pallier, Sebastián & Mehler, 1997). This phenomenon extends to the domain of non-native/naïve tone perception by L1 speakers of a non-tonal language (Burnham, Kirkwood, Luksaneeyanawin, & Pansottee, 1992). Similarly, previous studies show that naïve listeners with various L1s differ in the accuracy with which they identify tones due to their varying ability to attend to pitch height and/or direction possibly because of the influence of their L1 (Francis, Ciocca, Ma, & Fenn 2008; Gandour, 1983). L1speakers of a tone language are better able to attend to these cues and/or to map non-native tones onto their L1 tones and, therefore, outperform L1 speakers of non-tone languages in tone identification tasks (Wayland & Guion, 2004). Native speakers of non-tonal languages with no exposure to a tonal language, on the other hand, are less sensitive to tonal contrasts than speakers with previous tonal experience, i.e., L1 speakers of a tone language (Halle, Chang, & Best, for French listeners; Gandour & Harshman, 1978; Wang, Behne, Jongman, & Sereno, 2004). More details about the way non-tonal speakers perceive and encode tone are provided in section 2.4. For now, suffice it to say that suprasegmentals, including tone, appears to pattern like segmentals in terms of non-native or L2 phonological processing.

I turn now to a discussion of how the theoretical models manage to model suprasegmental perception in a non-native language. The PAM model (Best, 1995) is based on the absence or presence of a sound category that can distinguish meaning in the language. The case of suprasegmentals (vowel duration, tones or prosody in general, lexically-contrastive pitch) is somewhat problematic in that categories must exist in order for the PAM to apply. Any attempt to apply segmental models such as the PAM (Best, 1995) to suprasegmental dimensions therefore faces the challenge of conceiving suprasegmental dimensions as categories in the same way as segmental categories. In the case of tones, for instance, they are largely seen as lexical categories and are referred to as “tonemes,” parallel to the phoneme (Hallé, Chang, & Best, 2004; Francis et al., 2008 for tone**)**. Pitch accents are considered categories by So and Best (2010) but their status as categories is not fully clear yet. Lexical stress, on the other hand, is usually not looked at in terms of categories. In the domain of intonational patterns, several attempts at establishing their categorical nature indeed suggest that they (e.g. contrastive vs. neutral focus) are perceived categorically (Feldhausen, Pešková, Kireva, & Gabriel, 2011). Clearly however, more work is needed in this area.

Predictions made by the PAM model (Best, 1995) presupposes that assimilation (or equivalence classification) applies to both segmental and suprasegmentals in the same way. In short, if contrastive pitch is not used in L1 to the same extent, it might be perceived as very dissimilar and acquisition is predicted to be easier (see Dupoux et al., 2008 for an illustration of these predictions). Similarly, tone language speakers might experience “perceptual assimilation” from one non-native tone category to their native tone category. Recent evidence suggests that this is the case (So & Best, 2010). Also, attempts to model the mechanisms behind these effects in the same way as for segmentals suggest that Japanese listeners clearly assimilate Mandarin tones onto “Japanese pitch accent categories” (So, 2010).The PAM framework makes clear predictions about assimilation patterns between tone languages, but the predictions for other kinds of suprasegmental mappings are not clear (for instance between a tone and a pitch accent language), and crucially hinge on defining these dimensions as categories. In addition, these predictions are complicated by the possibility that intonational categories can also be represented in this way. Attempts at specifically developing assimilation maps between intonation and tones are underway (White, 1981; Hao, 2008).

As briefly mentioned above, another approach revolves around the central concept of prominence (or functionality). McAllister, Flege, and Piske (2002) demonstrated that the existence of vowel length in the L1 aided experienced learners in their perception of this same dimension in the target L2 of Swedish. L1 Estonian speakers performed at superior levels over both L1 English and L1 Spanish speakers, with the L1 English speakers performing better than the L1 Spanish speakers due to the limited usage – or prominence – of vowel length in English associated with lexical stress. As a result, the authors formulated the Feature Hypothesis: L2 phonetic features not used in a lexically contrastive manner in the L1 will be more difficult to acquire than those that are used to do so. According to this approach, the level of difficulty in L2 perception between the L1 and target L2 might be more accurately and precisely analyzed by incorporating the L2 phonetic-feature-based approach of prominence, as will be argued in the following section. Indeed, for suprasegmentals as stress such an approach has been deemed more explanatory (Dupoux, Sebastián-Gallés, Navarrete, & Peperkamp, 2008).

The predictions made by the Feature Hypothesis are more straightforwardly applicable to segmental and suprasegmental dimensions alike. Specifically, the Feature Hypothesis predicts that the more a dimension is prominent in the L1, the easier it might be to learn to discern and use that dimension for L2 phonological processing (see also Dupoux et al., 2008). This is attested to by a study on the acquisition of vowel length in Swedish which is lexically-contrastive by learners from three L1 languages which vary typologically in the use of lexically-contrastive vowel length. L1 speakers of Estonian where vowel length is lexically contrastive outperformed L1 speakers of English where vowel length is one correlate of word stress while the L1 English speakers in turn outperformed L1 speakers of Spanish which does not use vowel length to distinguish words. The predictions made by the Feature Hypothesis (McAllister et al., 2002), however, have not yet been tested for non-native, naïve tone perception.

Therefore, rather than the existence of pitch categories in the L1, we might consider prominence of lexically-contrastive pitch in the L1, and the resulting weighting of possible tone features such as pitch height and direction in the naïve perception of lexically-contrastive pitch such as tone (McAllister et al., 2002). However, within the Feature Hypothesis Model, in the case of clear-cut existence of pitch categories such as in Mandarin tone, we might be able to apply the PAM model (Best, 1995) to "streamline" the account of tone-to-tone mapping.

This concludes the section on L2 phonology with two subsections (segmentals and suprasegmentals). In sum, we can state that the L1 acts as a sieve in the perception of a non-native phonological system (Polivanov, 1931). However, the question remains whether the perception of non-native suprasegmentals patterns after that for non-native segmentals. The crux of the issue is whether lexically-contrastive pitch is categorical. If it is, the perception of non-native tone might be explained under the PAM model (Best, 1995) based upon categories and subsequent mapping between similar L1 and non-native categories. If not, the perception of non-native tone might be better explained under the Feature Hypothesis (McAllister et al., 2002) based upon the functionality of a feature in the L1, e.g., lexically-contrastive pitch.

I now turn to a more detailed review of what we know about the perception and the processing (including lexical encoding) of tonal contrasts in a non-native language. In the following two sections, I review the literature about tone perception for both naïve and L2 listeners (2.4.1), and then about the lexical encoding of tone in lexical representations (2.4.2). The goal of this section is to lay out predictions in terms of tone perception for native speakers of the four language types under consideration in this dissertation. I will also establish predictions for tone perception and tone encoding for L2 learners of a tone language.

2.4 Perception and processing of pitch and tone

A naïve listener who encounters a totally new sound system which utilizes pitch in a new manner does not approach the task without any "tools" but most likely employs the pitch patterns used in his or her L1. Accordingly, L1 linguistic pitch might shape perception of non-native and L2 tone (Wang, Jongman, & Sereno, 2006).

2.4.1 Perception of tone for naïve listeners and L2 learners

In this section, I will review the literature about tone perception in naïve and L2 listeners. This section is organized according to the four language types used in this dissertation and described in section 2.1 (Cross-linguistic differences in phonological pitch systems). That is, we look at the tonal languages of Mandarin and Thai, the pitch-accent language of Japanese, the word-stress language of English, and the "intonation-only-language" of standard Korean which does not feature lexically-contrastive pitch. Specifically, we look at the perception of non-native tone by speakers of these four languages. I consider in turn listeners who have never been exposed to the language (i.e., naïve listeners) and listeners who have learned a tone language (i.e., learners).

By speakers of a tone language. L1 tone language speakers transfer their L1 tone patterns onto the L2 tones or their ability from experience to track the pitch direction and/or height (Wayland & Guion, 2004). For example, L1 Mandarin speakers might map their falling tone [51] onto the Thai falling tone [51] or employ their ability to follow tone direction, enabling them to outperform non-tone language speakers. Yet, L1 tone language speakers also confuse L2 tones which are similar but not the same as their L1 tones. For example, L1 Mandarin speakers confuse the Thai mid [33] and Thai low [11] tones as Mandarin has no equivalent tone (Gandour, 1983), the closest being Mandarin tone 2 [35] and tone 3 [214] (Wayland & Guion, 2004).

Additionally, So and Best (2010) conclude that having tones in the L1 does not necessarily aid in the perception of L2 tones since as noted the L1 can also impede L2 perception both phonologically and phonetically. That is, experience in speaking a tone language not only strengthens the categorical perception range of sounds with similar frequencies, but may also impede perception of similar but sufficiently different tones as predicted under PAM (Best, 1995). Indeed, we find that L1 Mandarin speakers are more categorical in their responses than English L1 speakers to Mandarin tones (Wang, 1976; Leather, 1987; Stagray & Downs, 1993, cited in Sun, 1998) as L1 Mandarin speakers need to be less sensitive to pitch differences in order to account for the variation in tone production of a single tone pattern. For example, cross-linguistically we find that L1 Cantonese speakers are confused by similarities between their L1 Cantonese tones and target L2 Mandarin tones. They performed more poorly than the other two groups (i.e., Japanese and English) in distinguishing the target L2 Mandarin tone pairs 1 and 4 and tone pairs 2 and 3. Specifically, they mapped L2 Mandarin tone 1 [55] and 4 [51] onto their L1 Cantonese tone 1 [55] which has the corresponding allotone of 53 as well (i.e., Cantonese tone 1 is realized as either 55 or 53 just as English has the unaspirated /p/ as an allophone for the aspirated /p/, depending on context). Also, they mapped both L2 Mandarin tone 2 [35] and tone 3 [214] to Cantonese tone 2 [25 or 35] due to the similarity of their pitch contours. The authors classify these two cases respectively as “Single Category” and “Category Goodness Difference” assimilation patterns under Best’s Perceptual Assimilation Model (1995). These findings also suggest that L2 tonal categories can be assimilated to tonal categories used in the first language system[6].

[6] Also, it is important to note that bilingualism in two tone languages such as Mandarin Chinese and Taiwanese (i.e., Min) does not necessarily create differences in non-native tonal perception. That is, Gandour (1983)

By speakers of a pitch-accent language. The few studies to have looked at the perception of L2 tone by L1 pitch accent speakers have found that pitch accent speakers perform at comparable accuracy levels to L1 speakers of tone languages (Burnham et al., 1996; So, 2006). For example, L1 Swedish speakers mirrored both L1 Cantonese and L1 Thai speakers in their accuracy rates in the perception of Thai tones and only the L1 Cantonese speakers in terms of RTs but not the faster L1 Thai speakers (Burnham et al., 1996) which would be expected as Thais can access their lexicon. L1 speakers of pitch accent languages also outperform L1 English speakers in perception accuracy (Burnham et al., 1996) or show greater improvement when learning Mandarin tones (McGinnis, 1996), but both groups tend to notice pitch height (Guion and Pedersen, 2007).

Japanese listeners clearly assimilated Mandarin tones onto "Japanese pitch accent categories" (i.e., HH, LH, HL) in an identification task (So, 2010). In contrast, in another study employing an identification task (So and Best, 2010), Japanese had difficulty mapping L2 Mandarin tone 2[35] and tone 4 [51] onto Japanese LH and HL pitch patterns, respectively, although predicted to do so as a two-category assimilation pattern within a PAM framework (So & Best, 2010). In the first study, the Japanese participants listened to Mandarin tones containing the syllabic segment /fuu/ and chose a corresponding real Japanese word which also contained the bimoraic /fuu/ segment as part of the Japanese word [i.e., *fuutoo* 'envelope' (封筒), *fuufu* 'married

demonstrated that monolingual Mandarin speakers and bilingual Mandarin-Min speakers in Taiwan do not differ in how they process L2 Thai tones and that both groups use pitch height and direction to the same degree (cited in Wayland & Guion, 2004). However, how such speakers compared to monolingual Mandarin speakers on Mainland China perhaps needs to be examined to make a definitive conclusion on this point. The interaction of two tonal systems is an issue to keep in mind when examining many Mandarin speakers as many appear to exhibit varying degrees of bilingualism or at least exposure to a Chinese language or Mandarin dialect outside of standard Mandarin (i.e., *putonghua*).

couple' (夫婦) and *imafuu* 'modern style' (今風)] but with one of three different pitch patterns (i.e., LH, HL, or HH, respectively), or they chose "unknown". In the second study, the task was preceded by a familiarization section where participants were allowed to spend no more than two minutes to listen to the four tones by clicking on buttons for the four tones. They then proceeded to the identification task where they would choose one of four buttons labeled with the four tones upon hearing the target item. Two questions remain concerning the first study: the influence of the Japanese being able to access mental representations of lexical items and the notion of "Japanese pitch accent categories" (i.e., the validity of this interpretation). Additionally, the LH pattern of 'envelope' might actually be realized as an HH pattern as it is a a heavy syllable, so we must question whether the phonetic reality indeed overrides the phonological abstraction. As for the second study, one must question the memory load required for naïve listeners to "absorb" the four tones under two minutes and then, to label the items subsequently heard accordingly.

There have been two studies looking at the reverse situation: the perception of L2 Norwegian word accent by speakers of L1 Mandarin (tone) and L1 German (word stress) (van Dommelen & Husby, 2008, 2009). In the follow-up study, they found that Mandarin speakers behaved as native speakers, both outperforming German speakers on the perception of Norwegian word accents on an ABX discrimination task. However, the Mandarin speakers unexpectedly performed worse than both native and L1 German speakers on the discrimination of tonal melodies. The authors speculate this result may be due to the greater number of participants with musical skills among the German group (8 of 11 vs 1 of 11 among the Mandarin speakers), adding another possible dimension to consider in the cross-linguistic perception lexically-contrastive pitch.

By speakers of a stress-accent language. Although English does not use lexical tone, we must consider the possibility that L1 English speakers might transfer the pitch used in word stress although they might experience difficulty in extracting only pitch from the other features of English word stress (i.e., vowel duration and intensity) as hinted to by their inability to do so in the production of L2 Japanese which uses only pitch (Kondo, 2007). Thus, English speakers might not transfer contrastive lexical pitch from their L1, but they do seem to transfer intonation patterns (e.g., question intonation to a rising tone in Mandarin, Francis et al., 2008, see below).

L1 English listeners tend to confuse similar tones in a comparable manner to L1 speakers of a tone language. When comparing Thai tones in an AX task, L1 English listeners have the most difficulty comparing flat vs. contour tones while the comparison of contour tones with one another is the easiest (Burnham et al., 1992), suggesting that the perception of pitch height is difficult for L1 English listeners as well. L1 English speakers also confuse the rising and dipping tones (tones 2 and 3, respectively) just as L1 Mandarin speakers do (Li & Thompson, 1977; Leather, 1983, 1990; Wang, Spence, Jongman & Sereno, 1999). Guion and Pedersen (2007) also showed that English speakers focus much more on pitch height rather than direction as compared to speakers of tone languages when perceiving synthetic Mandarin tones (as incidentally is the case for Japanese in their study as well).

By speakers of an "Intonation language". Hallé, et al. (2004) found that speakers of languages such as French which does not feature lexical pitch are not "deaf" to L2 tones, with their performance on an AXB discrimination task considered "not so bad" (p. 416). The researchers ascribe this performance level to that fact that within the framework of PAM whereby tones are considered analogous to phonemes or "tonemes", French speakers would find tonemes to be uncategorizable and therefore, when comparing two tonemes, performance should range from poor to good. That is, L1 French listeners do not perceive L2 tone categorically and therefore, have difficulty mapping tone onto any L1 French category (Hallé et al., 2004).This is in line with findings that L1 English speakers are more sensitive to small acoustic differences in pitch than L1 Mandarin speakers as only the latter group perceives tone categorically (Stagray and Downs, 1993; Wang, 1976).

Citing Beckman, Hirschberg, & Shattuck-Hufnagel (2005), Francis et al. (2008) note (in response to the Hallé et al., 2004 study) that intonational contours can be considered linguistic categories stored in long-term memory similar to categories for vowels, consonants, and tones, and so, should have the same influence on the perception of L2 tones, influencing in turn the weighting of certain features. Thus, French speakers should access their mental representations of intonational categories when perceiving L2 tones. This position is bolstered by So and Best (2010) who conclude within the framework of PAM (Best, 1995) that L2 prosodic categories (i.e., suprasegmentals such as tone) will be mapped onto all L1 prosodic categories which include tone categories, pitch accent categories, as well as intonational categories, an assertion further reinforced by the Hallé et al. (2004) and So (2010) study. Thus, So and Best (2010) make the interpretation within the framework of PAM that the English speakers they tested on the perception of tone may have mapped L2 Mandarin tone 2 [35] and tone 4 [51] onto English intonational categories for questions and statements, respectively. Francis et al. (2008) also showed that English speakers transfer intonation patterns in their perception of L2 Mandarin (e.g., question to a rising tone in Mandarin).

Studies examining the perception of non-native tone by Korean speakers have not been found to be published in English although there is a study in Japanese concerning the perception of Japanese pitch accent by Korean speakers of the Kyungsang dialect in an AXB task of nonce words, showing one possible advantage in non-native pitch perception over speakers of standard Korean (i.e., for CVH syllables where H indicates a long vowel, as opposed to CV-CV and CV-N syllables) (Sukekawa, Choi, Maekawa & Sato, 1995).

Comparing all four types of languages within the Feature Hypothesis (McAllister et al., 2002), we would expect differences in the perception of non-native tones by speakers of various non-tonal languages as non-tonal languages vary greatly in the degree and manner of usage of lexically-contrastive pitch while all employ intonation at the phrasal level. This can be attested to by a few studies which report that pitch accent language speakers (e.g., L1 Swedish, L1 Japanese) perform at comparable rates to L1 tone language speakers in their naïve perception of L2 tones (Burnham, Francis, Webster, Luksaneeyanawin, Attapaiboon, Lacerda, & Keller, 1996; So, 2006). Also, McGinnis (1997) shows that Japanese speakers show greater improvement compared to L1 English speakers when learning Mandarin tones, possibly suggesting that L1 English speakers have weaker categories or no categories to which they can assimilate tones. Conversely, we see that speakers of tone languages such as Mandarin appear to perceive the pitch accent of languages such as Swedish more accurately than speakers of a non-tone language such as Hindi (Eliasson, 1997). This study concludes that "the linguistic category of tone is cognitively more salient to learners if it is present in a comparable form in their native language" as "transfer is a viable factor in tonology" (p. 1274). This finding serves to further reinforce the hypothesis that accuracy in perceiving L2 lexically-contrastive pitch is commensurate to the functionality of lexically-contrastive pitch in the L1.

However, specifically modeling the perception of the suprasegmental of tone in parallel to segmental perception (i.e., PAM, Best, 1995) is complicated by the fact that pitch (F0) is used in all L1s in the form of intonation, unlike for segmental features, for which some L1s do not make use of certain features. As such, intonation has continued to be considered an influence on the perception of lexically contrastive pitch (Francis et al., 2008), i.e., tone, stress or pitch accent. However, while the case for the existence of intonational categories may be made, the question must be asked whether these intonational categories can be accessed to distinguish lexical items where the domain differs.

If only pitch used to distinguish lexical items can be accessed, then, the case must be made for the existence of lexically-contrastive pitch categories. While this is decidedly simple for the case of tone languages, this proves rather difficult for languages like standard Korean and French. That is, these "intonation-only" languages feature a pitch pattern on words, the pattern does not serve to distinguish that word from another word differing in pitch pattern (i.e., lexical pitch minimal pairs do not exist). Yet, the fact that pitch patterns do exist on the word level may afford the speakers of even these languages some perceptual ability to distinguish lexical pitch patterns. However, if speakers of "intonation-only" languages do not have lexically-contrastive pitch patterns, then they may be analogous to speakers of a language which lacks consonants in perceiving consonants. For word-stress languages it also is somewhat difficult to make the argument that categories exist. This is further exacerbated by the presence of more salient correlates of word stress than pitch, i.e., vowel length, vowel quality, and intensity. The case for categories for Japanese pitch accent is also not clear. Furthermore, if we consider the Moraic pitch alignment (Morén and Zsiga, 2006), then, the existence of categories for tones may be called into question or at least modified.

Thus, an alternative approach would be that lexically contrastive pitch usage be more minutely broken down into features in an attempt to analyze the perception of lexically contrastive pitch within current L2 phonology models. Therefore, rather than the actual lack of the pitch features or categories (i.e., PAM), we might rather consider the prominence, and the resulting weighting of possible features in the perception of L1 contrastive pitch accent in the form of tone (Francis et al., 2008; McAllister et al., 2002). However, as mentioned previously the two approaches are not necessarily exclusive and may be united by examining the similarity and dissimilarity between tone patterns by looking at features. Additionally, we cannot discount the possibility that tone patterns are merely streamlined bundles of features used by native speakers in perceiving tones in the L1 more efficiently.

Summary

To sum up the patterns just reviewed, naïve listeners most likely apply the pitch patterns from their L1 in the perception of a new lexical pitch system such as non-native tone (cf., Wang et al., 2006). In response, the languages in this study could be typologically organized by the function of lexical pitch in the respective languages as suggested by Van Lancker (1980). The pitch contrasts in these languages differ in both domain (small to large) and function (from tonal contrasts to marking focus structure or affect) as configured in a functional scale in Figure 2.4 (See section 2.1 *Cross-linguistic differences in phonological pitch systems* for details).

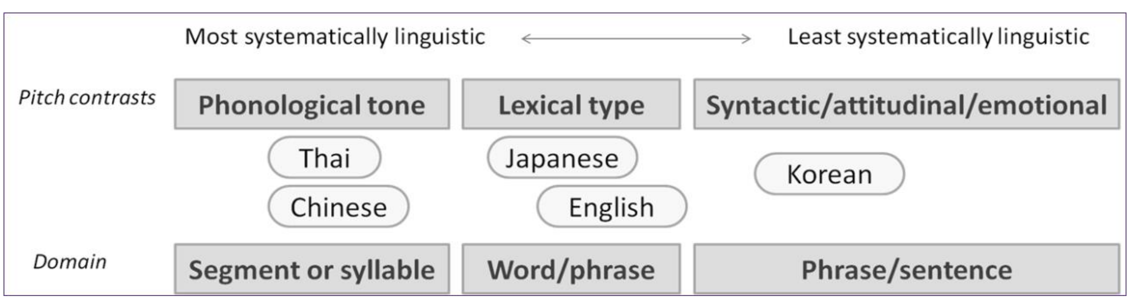


Figure 2.4. Functional scale of pitch contrasts (Adapted from Van Lancker (1980: 210)

If we assume that the L1 influences the perception and therefore, acquisition of the L2 phonological inventory, we might also assume that an acquired L2 might influence the naïve perception of another non-native phonological inventory. This situation would be expected as an extension of the Interaction Hypothesis (Flege et al., 1999; Weinreich, 1953) which claims that the L1 phonological system interacts with the acquisition of the L2 phonological system.

Additionally, an acquired L2 may influence the naïve perception of another language as extrapolated from research concerning L3 acquisition. For example, in the initial stages of L3 acquisition, the fact that an L2 is a foreign language, or “L2 status,” accounts for the heavy L2 influence on L3 acquisition vis-à-vis the L1 influence. Thereby, the more proficient one is in the L3, the less L2 influence there will be. Concurrently, the L1 influence will increase. As such, in the naïve perception of another non-native language, we would expect strong L2 influence under the caveat that the L2 is of a fairly proficient level, “at least an intermediate level” (Fernandes-Boëchat, 2007). Another factor, namely typological distance or psychotypology, i.e., perceived similarity**,** is also cited as the possible cause for greater L2 influence as compared to L1 influence on the acquisition of the L3 (Wrembel, 2010). Additionally, we must consider the interaction between the sound systems of an L1 and L2 and subsequent influence on the naive perception of non-native sounds (cf., Interaction Hypothesis, Flege, Yeni-Komshian & Liu, 1999; Weinreich, 1953). Therefore, it might also influence the naïve perception of another tone language (e.g., Thai) by L2 speakers of a tone language (e.g., Mandarin). With these research findings in mind, this dissertation will examine the influence of learning Mandarin as an L2 on the naïve perception of Thai tones.

2.4.2 Lexical encoding of tone for L2 learners

This section describes how learners of a tone language encode tone into their lexical representation of the words they learn in the L2. The connection between the perception of an L2 sound and the encoding of that sound are discussed [e.g., whether L2 learners can hear the difference between minimal pairs (e.g., /r/ vs /l/ in “rock” vs “lock”) differing by the target sound and how the L2 learners react to the meaning of such minimal pairs].

As described above, experience (i.e., learning) with a tone language has been shown to improve perception of tones by L1 speakers of a non-tonal language. For instance, in the case of Thai tones, Wayland and Guion (2003) have shown that experienced L1 English learners of Thai are better at perceiving mid and low tones than naïve L1 English listeners.

However, very few studies if any have examined the extent to which the learners have encoded tones in their lexical representations for L2 words. L1 speakers of a language featuring lexical stress (i.e., German) have been shown to lexicalize L2 tone (i.e., Mandarin) more efficiently than L1 speakers of a language which does not feature either lexical stress or lexical tone (i.e., French; Braun, Galts & Kabak, 2014). Braun et al. (2014) tested L1 speakers of three language groups (French, German, Japanese) varying by lexical stress on learning Mandarin tone. The study presented participants with three types of disyllabic non-words: same tones/different segments, different tones/same segments and same tones/same segments, and had the participants learn the non-words by matching them to pictures. The study concluded that lexical pitch in the L1 benefits the naïve perception of non-native tones (i.e., Mandarin Chinese) due to findings: on the condition of the different tones/same segments condition the L1 German group outperformed the L1 French group where the two languages differ by the presence/absence of lexical stress. However, the L1 Japanese group also underperformed the L1 German group, calling into question the functional load of pitch accent in Japanese (vis-à-vis word stress in German) as one possible reason for this group's poor performance. Yet, unlike the current study, the participants in the Braun et al. (2014) study were not L2 learners but non-learners who were taught non-words with L2 Mandarin tones. Nevertheless, the Braun et al study points to a possible link among these two types of suprasegmentals, namely lexically-contrastive pitch in the form of word stress and tone, and to an innate propensity to somehow apply one to the other in L2 perception via a mental representation.

This has been shown to be the case as mentioned in previous sections for the L2 perception of vowel length (McAllister et al., 2002).

One way to examine lexical encoding of tonal information in the representation of L2 words in learners would be to use a priming paradigm in addition to a tonal discrimination task. As touched upon above, positive priming between minimal pairs (i.e., faster reaction times on the target second word) occurs among learners but not native speakers when contrastive segmentals are perceived to be the same (Darcy et al., 2012; Pallier et al., 2001). That is, L1 speakers of Japanese who do not have a robust phonological representation of, for example, the /l/-/r/ distinction in their L2 English are highly likely to exhibit a priming reaction upon hearing the target of “right” after having heard the prime of “light” (i.e., faster reaction times as they perceive the two words to be the same word). Native speakers on the other hand would not experience any priming between the two, as the two are different words with different phonological representations. Therefore, for tone as well, if learners of a tonal language do not have a robust phonological representation of two words in a minimal pair varying by tone [e.g., /ma/ ‘mother’ (tone 1) vs ‘horse’ (tone 3) in Mandarin Chinese], they are expected to react faster on hearing the second word of the minimal word pair as they will perceive it to be the same word. We need to recall a few observations. First, for native speakers in the early stage of lexical activation tone does not prevent the activation of minimal tone pairs, resulting in competition between the activated words or “inhibitory lexical competition.” Additionally, a hierarchical pattern in reaction times appears to occur for Mandarin: when both the segmental and tone information are the same between two words, reactions are the fastest in the case of repetition (tone + segment overlap), followed by when segment overlap, followed next by tone overlap. Moreover, priming effects are “rapid and short-lived” (Lee, 2007).

Testing perception is not enough to investigate lexical encoding: We need to examine both processes as we cannot infer from L2 perception what the L2 encoding might be as shown by studies that have observed this disconnection between perception and encoding (Ota et al., 2009; Darcy et al., 2012). Generally, it has been a common assumption that accurate perceptual categorization of a minimal pair distinction such as that between /l/-/r/ must occur before the encoding of this phonological representation within the lexical representation is possible, an idea analogous to the assumption that L2 perception must precede L2 production (cf., Flege, 1993). However, both assumptions have been challenged by exceptions (Weber & Cutler, 2004; Hayes-Harb & Masuda, 2008; Darcy et al., 2012, where encoding of an L2 phonological representation occurred without accurate perceptual categorization of a minimal pair distinction; and Goto, 1971, where native-like L2 production occurred without native-like L2 perception although through training in the Goto case).

There are two examples where the lack of a phonetic category is accompanied by a lack of phonological/lexical representation. Specifically, when listeners cannot differentiate between two similar segmentals, they cannot differentiate between two words which are differentiated only by these two similar segmentals (e.g., the minimal pair of "beat" vs "bit" in English differentiated by /i/ vs /ɪ/). The first case is that of bilingual speakers of Spanish and Catalan, who were dominant in Spanish. They had difficulties in discriminating the minimal segmental pairs of /e/ and /ɛ/ (Pallier et al., 1997) and also minimal word pairs distinguished by these two segmentals (Pallier et al., 2001). Despite having learned Catalan at a young age Spanish-dominant bilinguals exhibited repetition priming between minimal pairs differing in /e/ and /ɛ/ unlike the Catalan-dominant bilinguals (Pallier et al., 2001). Another study (Ota et al., 2009) corroborated this finding with a different approach. In short, L1 Japanese speakers do not have a robust categorization of the /l/-/r/

distinction in their L2 English and so, associate the word “rock” with “key” as “rock” activates the word “lock” as well, despite the fact that the words in the study were presented visually and not aurally to the participants. These two studies would seem to support the assumption that phonetic categorization must precede phonological representation. However, this assumption would be more strongly supported by these two studies if they also had L2 participants who have both the phonetic categorization and the phonological representation in the same cases in order to definitively rule out other possible explanations (cf. Darcy et al., for detailed explanation).

Empirical evidence to turn this assumption on its head can be found in two studies showing instances where phonological representation has been encoded lexically while participants were shown to lack a robust phonetic categorization of the phones (Weber & Cutler, 2004; Hayes-Harb & Masuda, 2008). In an eye-tracking experiment L1 Dutch speakers showed difficulty distinguishing between the /æ/-/ɛ/ minimal segmental pair upon hearing the initial segment of a word (e.g., ‘pen’ vs ‘panda’), but only the /æ/ segment activated words containing either the /æ/ or /ɛ/ segmental while the /ɛ/ only activated words containing /ɛ/ and not those containing /æ/ (Weber & Cutler, 2004). This asymmetry indicates that the distinction is encoded – at least partially – in lexical representations, independently of phonetic perception. In a study involving suprasegmentals of geminates (long consonants) in Japanese, learners were able to distinguish the difference between geminates and singletons at the accuracy levels of native-speakers (i.e., not significantly different) when matching pictures to non-words (made-up brandnames of products). However, while the learners were able to produce a phone that clearly differed from a singleton, they did not produce native-like geminates when naming the same pictures. That is, learners may not have encoded the segment in the same manner as native speakers, but they encoded something that is different from both the singleton and the target geminate and stands in for the geminate in their phonological representation (Hayes-Harb & Masuda, 2008).

Additionally, Darcy et al. (2012) found a much clearer disconnect between phonetic categorization and phonological/lexical representation in case of the front back rounded vowels of French (/u/-/y/ and /œ/-/ɔ/) by L1 English speakers. On an ABX categorization task, broadly speaking intermediate and advanced learners performed significantly less accurately than the native speakers. However, results diverged when it came to the encoded phonological representations for these words, as witnessed on a lexical decision task with repetition priming: while advanced learners behaved like native speakers on the lexical task, they did not outperform the intermediate learners on the categorization task. More interestingly, intermediate learners displayed priming for the one contrast they were able to discriminate with high accuracy: the high vowel contrast. Conversely, for the contrast on which learners made most errors in discrimination (mid vowels), no sign of fuzzy lexical representations was observed, prompting the authors to conclude a dissociation between phonetic perception and lexical encoding. A similar divergence was found using the same methodologies (i.e., ABX and lexical decision task) for advanced learners of Japanese on their phonetic categorization and phonological representation of geminates (Kojima & Darcy, 2014). In short, the advanced learners did show "partial segmental deafness" but incongruously did not show "lexical deafness". Darcy et al. interpret this "lexical deafness" in a somewhat related vein to the interpretations of results in the Weber and Cutler (2004) and Hayes-Harb and Masuda (2008) studies when they propose that advanced learners have encoded "something" they term a fuzzy lexical representation which may not be a fully native-like representation of the targeted phone, but something that allows them to distinguish minimal word pairs distinguished by the minimal segmental pairs.

As such, we see that relationship between phonetic categorization and phonological representation is not clear-cut, with the former necessarily preceding the latter. However, as alluded to by Darcy et al. (2012), learners' lexical representations of L2 words can be fuzzy. If the methodology is not problematic, then, the question is of what nature is this fuzzy lexical representation. This current study may indirectly substantiate the results of Darcy et al. (2012) for segmentals and Darcy, Daidone & Kojima (2013) for geminates with results for tone using a similar methodology (i.e., ABX categorization task and lexical decision task). However, suprasegmentals do not behave like segmentals, and while geminates and tones are considered suprasegmentals, they seem to be of very different natures. In sum, the connection between phonetic categorization and phonological representation is still unclear. The fact that few studies have been conducted on the encoding of L2 tone only further clouds our understanding of this connection.

2.5 Research questions guiding this dissertation

As outlined in this background chapter, within the framework of the Feature Hypothesis (McAllister et al, 2002), PAM (Best, 1995) and L2 phonology of priming and encoding as discussed in this literature review, the primary aim of this dissertation is to ascertain how naïve listeners of a wide range of languages with varying levels of prominence and inventory of contrastive linguistic pitch patterns perceive the tones of a non-native tone language. This comparative approach of examining languages differing in the linguistic use of pitch should more comprehensively illuminate the perception of suprasegmentals: it should pinpoint whether the L1 inventory (i.e., type of pitch patterns) and the prominence of certain pitch features in the L1 influences the weighting and thereby, the non-native perception of certain tone features. Such an approach would provide valuable information regarding what non-learners do, allowing us to posit how learners acquire L2 tone. As such, this study represents a first step toward better understanding the L2 acquisition of tone. This study goes one step further by examining whether experience with a tone L2 facilitates the perception of another tone language by including a group of learners who have studied a tone language. To correlate the naïve perception of a tone language with the exposure through learning of another tone language, this study looks at the neglected area of the lexical encoding of suprasegmental dimensions in lexical representation by assessing the extent to which L2 learners store tonal information in lexical representations.

In sum, the research questions examined in this dissertation are the following:

RQ1: Does the varying presence of certain features in the L1, specifically lexically-contrastive pitch, aid in the naïve perception of non-native tones, thus resulting in a linguistic hierarchy of perceptual ability?

RQ1.a is: Are some tones or tonal comparisons easier to perceive than others (cf., pitch height or pitch direction) in the naïve perception of non-native tone?

RQ2: Does the learning of a tone language aid in the naïve perception of the tones of another non-native tone language?

RQ2.a is: Do L2 learners of a tone language differ in their naïve perception of another tone language as compared to participants with the same L1 who are not learning a tonal L2?

RQ2.b is: Do L2 learners of a tone language perform at levels comparable to L1 speakers of that tone language in a linguistic hierarchy of perceptual ability of another tone language?

RQ2.c is: Do L2 learners of a language perform in a similar manner to L1 speakers of that one language, i.e., Are some tones or tonal comparisons easier to perceive than others (cf., pitch height or pitch direction) in the naïve perception of another tone language? Do the two groups react in the same manner?

RQ3: Are L2 learners of a tone language able to lexically encode tonal information to constrain lexical access?

3. Pilot study: Non-native naïve perception of Thai tones

This pilot study specifically examined how Thai tones are perceived by naïve listeners (= non-learners of Thai) who speak a range of languages varying in the usage of lexically-contrastive pitch: tone language (Mandarin Chinese), a pitch accent language (Japanese, Kyungsang Korean dialect), a word-stress language (English), and a language which does not employ pitch in any way to distinguish the meaning of words (standard Korean). Pitch is generally defined as the variations in the height of the voice or fundamental frequency (abbreviated as F0) used to distinguish words.

Thai was adopted as the target language of this study as it has at least two level tones, providing a richer variety of level tones than Mandarin Chinese which given the nature of its tone inventory may cause listeners to rely on their perception of pitch direction. That is, Thai has a falling and rising tone as in Mandarin, but Thai also has three level tones (i.e., low, mid, high). However, the high level tone is phonetically not level as mentioned and more similar to the rising tone in shape. The similarities between the low and mid level tones and the rising and high tones in Thai provide a means for better understanding the perception of pitch height as opposed to the perception of pitch direction which is needed to perceive the contour tones. Also, logistically, there are many more learners of Mandarin than Thai, allowing the examination of the effect of learning a non-native tone language (i.e., Mandarin Chinese) on the naïve perception of a different (unknown) tone language (i.e., Thai).

The pilot study employed one methodology (AXB task) and referenced the Feature Hypothesis (McAllister et al., 2002). Results were expected to establish a baseline in an attempt to better understand the acquisition of L2 lexically contrastive pitch, specifically Thai tones, and provide a reference point to compare the acquisition of suprasegmentals such as tone with that of segmentals within other models such as PAM-2 (Best & Tyler, 2007). The pilot study was conducted with the goal of verifying that speakers of different L1s perform differently on tonal categorization as measured in overall accuracy rates and RTs on an AXB task (Chang, Hallé, Best & Abramson, 2008; Chiao, Kabak & Braun, 2011; Li & Shuai, 2011). Additionally, the pilot study attempted to ascertain whether these speakers attend to the features of pitch height and direction differently as measured in accuracy rates and RTs for each condition (i.e., comparison of each tone to every other tone). The results demonstrate that as a possible function of the functionality of L1 lexical pitch in the L1, speakers of the various languages perform in a hierarchical manner and attend to pitch height and direction differently. A number of experimental considerations will be discussed in Chapter 6.

The first research question of the pilot study[7] was exploratory in nature: "Does the functionality[8] of pitch in the L1 shape non-native tone perception?" Functionality here refers to the functional role of linguistic pitch to contrast lexical items. The hypothesis was that the greater the functionality of lexical pitch in the L1, the better the perception of non-native tones, resulting in a hierarchy of performance among the various L1s, relative to the functionality of lexical pitch in the L1. The notion of functionality was grafted onto that of the functional load of pitch from Van Lanker's model (1980) in order to predict a pitch functionality hierarchy of performance (see Table 3.1). The languages examined in this pilot study can be arranged according to language type on a functional scale, such as one proposed by Van Lancker. The languages use pitch contrasts in different domains (from small to large), and also in the function that the pitch fulfills, whether

[7] The results of this pilot study have been published in *Laboratory Phonology* in November 2014. This chapter was written independently of the published manuscript but contains the same data. See Schaefer, V. & Darcy, I. (2014). Lexical function of pitch in the first language shapes cross-linguistic perception of Thai tones. *Laboratory Phonology 5*(4), 489-522.
[8] In this study the term *functionality* is used instead of the term prominence which was used in the McAllister et al. (2002), especially as the term *functionality* serves as shorthand for the melding of prominence (McAllister et al., 2002) and functional scale (Van Lancker, 1980). Also, the term functionality began to be used instead of prominence to avoid any ambiguity due to the use of the term prominence to signify other meanings in the field of linguistics as pointed out by reviewers of publications of this pilot study.

lexical or non-lexical (i.e., from tonal contrasts to marking focus structure or affect, for example). When the domain and function of pitch overlap between the L1 and L2, we (generally) expect perception of the non-native linguistic pitch system to be more reliable.

In Table 3.1, specific predictions are put forth for tone perception accuracy reflecting the functionality of pitch to contrast words in the target languages.. Thus, we see that Mandarin, Japanese and English feature lexical pitch despite varying domains within which they do so while standard Korean does not employ lexically-contrastive pitch, but uses pitch at a domain larger than the word as do all of the languages. Based upon these characteristics, a scale of functionality from maximal to low can be established alongside a parallel hierarchy predicting accuracy of tone perception. Additionally, the condition of exclusivity is introduced to make a distinction between Japanese and English for both the functionality hierarchy and "Accuracy in tone perception" hierarchy, namely the exclusive reliance of L1 Japanese listeners on pitch to distinguish words as compared to L1 English listeners where pitch is one correlate among a few (i.e., vowel length, vowel quality, intensity) needed to distinguish words in most likelihood increases the Japanese listeners' sensitivity to lexical pitch usage. Thus, naïve listeners who encounter a new sound system which utilizes pitch in a new manner do not approach the task without any "tools": their perception will most likely rely on the pitch usage defined by their L1 phonological grammar. For example, the presence in the L1 of lexical pitch contrasts has been shown to facilitate perception of non-native tone (Burnham et al., 1996).

Table 3.1: Overview of pitch functionality typology and predictions for tone perception accuracy

Pitch pattern	Prosodic Domain	Lexical status of pitch use	Functionality for lexical distinction?	Predicted Sensitivity/ Accuracy in tone perception
Tone (Mandarin)	Syllable, word	Lexical	Maximal	Highest
Pitch-Accent (Japanese)	Word	Lexical	High-Intermediate (pitch is exclusive)	High-intermediate
Word Stress (English)	Word/Foot	Lexical	Low-intermediate (pitch is non-exclusive)	Low-intermediate
Intonation (Korean)	Intonational phrase, PPh	Non-lexical	Low	Lowest

The natural follow-up research question to the first research question (as to the influence of lexically-contrastive pitch in the L1 on the naïve perception of non-native tone) is to ask whether specific tone contrasts also influence accuracy. Referencing previous research (cf., Gandour, 1983 among many), the pilot study explored whether listeners are equally sensitive to the height or direction feature of tones or show a bias toward one or the other, and if this bias or lack thereof is influenced by the inventory of lexically-contrastive pitch usage in the L1. For example, English speakers process dynamic (contour) F0 variations more accurately than static F0 differences in disyllabic stimuli (Repp & Lin, 1990; Wood, 1974; Lee & Nusbaum, 1993). Also, Dutch listeners attend to F0 information when these correspond to contours having linguistic meaning in Dutch (disyllabic stimuli, question intonation) (Braun & Johnson, 2011). In response to such findings, it is natural to ask whether performance in non-native tone perception is affected by specific tonal shapes. The working hypothesis of the pilot study is that listeners will use height more than direction (e.g. Gandour, 1983) as listeners need to be first aware of tone height, even if merely a dual opposition of a high voice register versus a low voice register, to perceive movement in pitch, which is especially true if contour tones are indeed composed of a series of high and low tones (cf., compositional model of Morén & Zsiga, 2006). However, there could be various other factors that affect this prevalence, e.g., the use of multiple voices for the stimuli (see below).

3.1 Methodology

3.1.1 Participants

Participants were recruited from five language groups: Mandarin (n=10; females=6), Japanese (n=12, females=11), English (n=13; females=10), Korean (n=10; females=7) and Thai (n=2; males=2). The Thai speakers were recruited to ensure that the stimuli and AXB task itself were valid for native speakers. The participants were primarily graduate students or former graduate students who were involved in language studies (i.e., language education, linguistics, applied linguistics) with the exception of 11 participants who were undergraduate students (n=3) or not involved in language studies (n = 8) (i.e., Korean = 3, Japanese = 3, Mandarin = 3, English = 1, Thai = 1). Average ages ranged as follows: Japanese (25-50 years with an average age of 35.4); English (25-45 years with an average age of 31), Mandarin (24-31 years old with an average age of 27.1), and Korean (27-47 years old with an average age of 32.2). The two Thai listeners were 25 and 32 years old. The average time in an English-speaking country was 6.6 years for the Japanese, 4.5 years for the Koreans, 3.5 years for the Mandarin-speakers, and 2 years for the Thai speakers. The English speakers had spent an average of 1.7 years abroad in a non-English speaking environment.

The speakers of Mandarin, a tone language, were recruited to test whether L1 tone aids perception of L2 tones. These speakers included six speakers who also had various degrees of exposure to Taiwanese, another tonal language. Most had been exposed to another Chinese dialect even if they did not consider themselves a fluent speaker of that dialect. The speakers of Japanese, a pitch accent language, were recruited on the basis of speaking a dialect of Japanese which features pitch accent although not necessarily standard Japanese. Two speakers were from Tochigi and Ibaraki prefecture which are close to Fukushima prefecture, known for its accentless dialect. The English speakers were native speakers of American English who had no proficiency in Thai, Mandarin, Japanese, Korean or any tone language. Korean speakers were mainly from the Seoul area, but three were from the Kyungsang region where a pitch-accent dialect of Korean is spoken, and one speaker was from Cholla, an area abutting Kyungsang but with a dialect not featuring pitch accent (self-reported) (although these dialects do not correspond to their respective administrative borders, Lee and Ramsey, 2000).

3.2.1 Materials

Target stimuli included both words and nonwords in Thai. Target stimuli included words with the relatively easy-to-pronounce phonotactic structure of an open-syllable CV structure to avoid conflating other possible pronunciation problems with the tones. This is in line with Wayland and Guion (2003) who demonstrated that when comparing open syllable words with closed syllable words for mid and low tone items, the latter are distinguished by the onset and offset of the F0 while the former are distinguished by only F0 onset. Hence, the authors predict that closed syllables would be easier to perceive. As such all target words were open-syllable CV words considered more difficult to perceive than closed syllable words (Wayland and Guion, 2003), favoring a more faithful reflection of the participants' perception ability of Thai tone. A total of 80 items were selected (41 real words and 39 nonce words). (See Appendix A). Distracters were all real Thai words composed of similar vowels or consonants, including more difficult vowels such as [ɯ], [ɛ], and [ə] without regard to tone as well as some closed syllable items. The words/nonwords were presented in triplets for one trial: A, X, and B, where X=A or X=B. Additionally, as part of the design of an AXB task, four trials are presented for each comparison, to reduce bias induced by order: AAB, ABB, BAA, BBA. So, a comparison of the low tone (L) to the mid tone (M) would be presented four times (= four trials): LLM, LMM, MLL, and MML. A concrete example using the syllable [biː] would therefore look like the following: [biː]^{L} – [biː]^{L} – [biː]^{M}.

In total, ninety-six trials were created (48 test and 48 distracter trials). The test trials included three conditions: 1) height, comparing flat tones to test for a bias to height in perception, 2) direction, comparing contour tones for a direction bias, and 3) mixed, comparing flat tones with contour tones. Twelve triplets each for the direction and height conditions, and 24 triplets in the mixed condition were created (see Table 3.2). All trials were randomized and put into 3 blocks of 32 items, respectively.

Table 3.2: Overview of the tonal comparison(s) and number of trials used for each condition

Test Conditions (n=48)			Control Condition (n=48)
Direction (n=12)	Height (n=12)	Mixed (n=24)	Control (n=48)
	low-mid (LM)	low-rising (LR) low-falling (LF)	consonant
rising-falling (RF)	low-high (LH)	mid-rising (MR) mid-falling (MF)	vowel
	mid-high (MH)	high-rising (HR) high-falling (HF)	

The AXB stimuli were recorded by two native Thai speakers. Both spoke the Central Thai dialect. Sixteen different words were recorded, with three tokens of each, spoken without a carrier phrase. Another recording of distracters was made, with two tokens of each item.

The female voice was used for the A and B while the male voice was used for the X. The AXB task included a 500ms interstimulus interval (ISI) both between the A and X and between the X and B stimuli. The experiment was timed so that after the presentation of each trial, participants had 3000 milliseconds to make their answer. Reaction times were measured from the onset of the X stimulus.

The training task consisted of 16 trials with feedback indicating their accuracy and RT. The 16 tokens consisted of two trials comparing flat tones, two trials comparing contour tones, four comparing flat tones with contour tones, and eight distracter trials. None of these were used in the actual test task of 96 trials. These trials were randomized in their order within the training task section.

3.1.3 Procedure

Participants were tested individually. After signing the consent form, participants were given oral instructions about the task. They then read the instructions on the computer screen and did the training test. For each trial, participants heard and chose whether the middle sound (i.e., X) was more similar to the first sound (i.e., A) or the third sound heard (i.e., B), by pressing two clearly identified keys on the computer keyboard. After the training they were asked if they were fine with the task and understood what they were being asked to do. They then performed the task. The task required from 15-20 minutes in total. The task was then followed by a debriefing session, where participants were asked very broadly about their opinion on the difficulty of the task and whether they felt they had problems understanding which pairs were similar and what this may have been caused by. They also filled out a background questionnaire (see Appendix B).

3.2 Results

Accuracy rates and reaction times were obtained for each condition and for each participant and analyzed as follows. Reaction times shorter or longer than two standard deviations from the RT mean of each participant were replaced by the RT mean of each participant (5.2% of total RTs). Additionally, remaining reaction times lower than 300 ms were replaced by the mean RT of the participant (0.23% of total RTs). Data for three items in the height condition (one each for L vs. M, L vs. H, M vs. H) and for one item in the direction condition (R vs. F) were excluded from analysis as one Thai participant felt that the tones were not ideal models of the targeted tone. Means for individual participants and remaining items were screened for outliers with the result that no further item or participant was excluded. Mean accuracy in each condition was computed for each group. Similarly, the mean reaction time for correct trials was computed for each condition and each group. Linear mixed effects models to compare means were run in SPSS 20 by subjects on average accuracy comparing Condition (*Test* vs. *Control*), and then Subconditions, and Language Groups, as well as interactions. Unlike the other language groups, the Thai listeners were able to approach the task using lexical knowledge. Therefore, analyses were also run excluding this group. An alpha level of 0.05 was used for all statistical tests.

3.2.1 Analysis by condition (test vs control)

Overall accuracy rates in each condition are presented in Figure 3.1. Results show that overall accuracy in the test condition was slightly lower than on the control condition (74.8% vs. 79.1% correct). The analysis omitting Thai revealed a significant interaction ($F(3, 37) = 11.3$, $p < 0.001$) between group and condition. Whereas accuracy of all groups was comparable on the control condition (no effect of "group" on the control condition: $F(3, 67.3) = 1.5$, $p > 0.1$), there was a significant effect of group on the test condition, as suggested by the difference in accuracy rates ($F(3, 67.3) = 11.3$, $p < 0.001$). On the test condition, we observed that Mandarin participants outperformed other non-native groups (87% accuracy), followed by Japanese participants (77%), and by English and Korean (both at 67% accuracy). Mandarin listeners discriminated tonal contrasts with higher accuracy than the other groups, significantly outperforming both Korean and English participants ($p < 0.001$) but only marginally more accurate than the Japanese group ($p = 0.093$). Notably, Korean and English participants were not significantly different from each other ($p = 1$) on the test condition, as is visible in Figure 3.1.

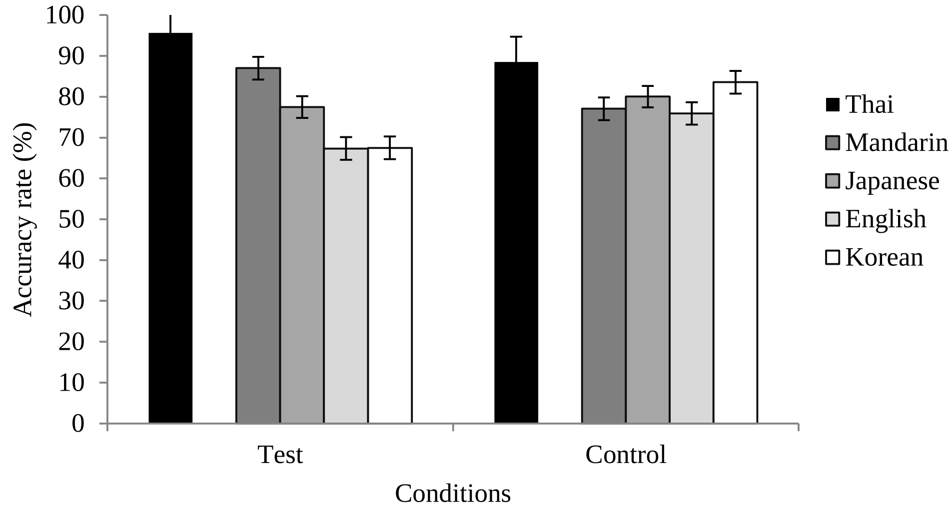


Figure 3.1: Accuracy rate (%) for each language group in the test vs. control condition. Error bars enclose +/- 1 SE. (Thai listeners are displayed in black for comparison purposes)

Reaction times on the test vs. control condition are shown in Figure 3.2 below. A main effect of condition (F (1, 37) = 31.4, p < 0.001) was observed. Overall, latencies in the test condition are about 120 ms slower than on the control condition (1265 ms vs. 1141 ms). Thai and Mandarin listeners are faster than the other non-native groups but the main effect of group was not significant (F (3, 37) = 1.7, p > 0.1). The interaction was not significant (F (3, 37) = 2.4, p = 0.08).

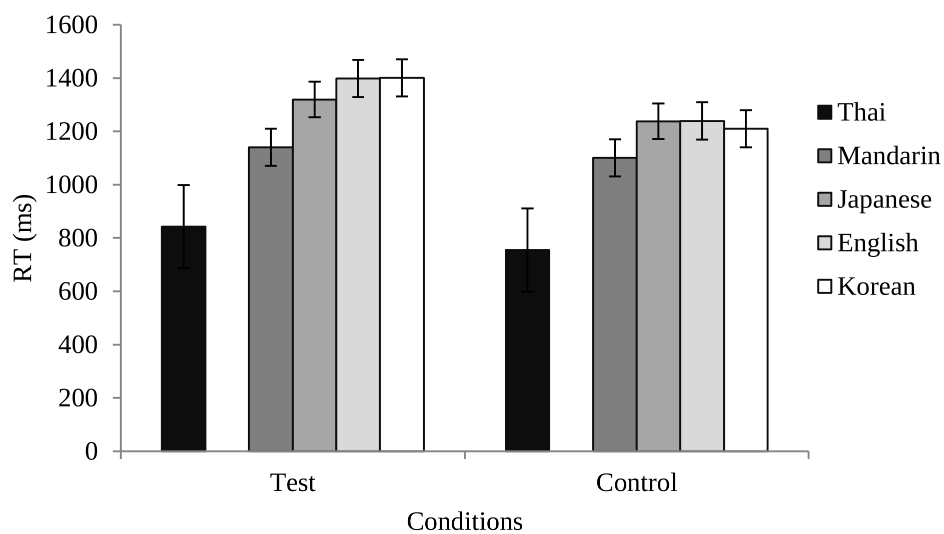


Figure 3.2: Reaction times (ms) for each language group in the test vs. control condition. Error bars enclose +/- 1 SE. (Thai listeners are displayed in black for comparison purposes)

The predicted hierarchy was in large part supported by the overall accuracy pattern that emerged from these accuracy and RT data. The functionality of pitch in the L1 determines accuracy in a phonological discrimination task. However, against the initial prediction, the data also revealed that English and Korean participants obtained very similar accuracy levels on the test condition, perhaps suggesting that F0 information is less readily accessible for phonological discrimination in these two groups (see discussion).

3.2.2 Analysis by subcondition (height, direction, mixed)

We now look at performance in each subcondition with accuracy and RT data for each group and subcondition shown in Table 3.3.

Table 3.3: Accuracy means (%) and reaction times (in ms) by language groups for each subcondition

	Accuracy				Reaction times			
Language group	height	direction	mixed	(SE)	height	direction	mixed	(SE)
Thai (n=2)	*88.9*	*100*	*97.9*	8.5	*887*	*846*	*800*	*164.8*
Mandarin (n=10)	76.7	97.3	87.1	3.8	1169	1041	1155	73.7
Japanese (n=11)	66.4	87.6	78.4	3.6	1303	1155	1347	70.3
English (n=10)	64.4	65.5	72.1	3.8	1288	1306	1389	73.7
Korean (n=10)	58.9	78.2	65.4	3.8	1370	1363	1301	73.7
Average (non-n.)	66.6	82.1	75.7		1282	1216	1298	

Note: SE = standard error; non-n. = non-native groups only

To compare the language groups, linear mixed models with repeated measures were run in SPSS 22 for the four conditions within each subject. The dependent variables were the binary accuracy data (1 vs. 0) and the continuous reaction times. The independent variables were language group (Thai, Mandarin, Japanese, English, Korean) and condition (height, direction, mixed). Subjects and items were entered as random effects in the model.

Looking at the Type III tests of fixed effects for the analysis on accuracy rates (not including Thai listeners), the analysis of accuracy rates on the F-tests showed an interaction between condition and language group ($F(6, 74) = 2.7$, $p < 0.05$). Univariate tests for the simple effect of condition within each group show that condition significantly impacts accuracy rates for all groups except for the English participants (English: $F(2, 74) = 1.58$, $p > 0.1$; for the other three, $p < 0.001$).

Results generally support the predicted trends regarding overall sensitivity to tonal contrasts. Against the initial working hypothesis, the results show that the height condition is overall the most difficult whereas the direction condition is the most accurate and incidentally has the fastest response times (RTs) as well. The one exception to this pattern is found in the English group: Their accuracy performance is flat (i.e., the same) on all of the conditions.

As suggested by the significant interaction between language group and condition, it is possible that certain language groups show a different pattern of responses across conditions. Univariate tests show that the effect of condition is significant for all groups except for the English native speakers (see Table 3.4).

Table 3.4: Accuracy means (%) and reaction times (in ms) by language groups for each subcondition

Language Group	F(2, 74) =	*p* <
English	1.581	.213
Korean	8.842	.000
Japanese	11.399	.000
Mandarin	9.747	.000

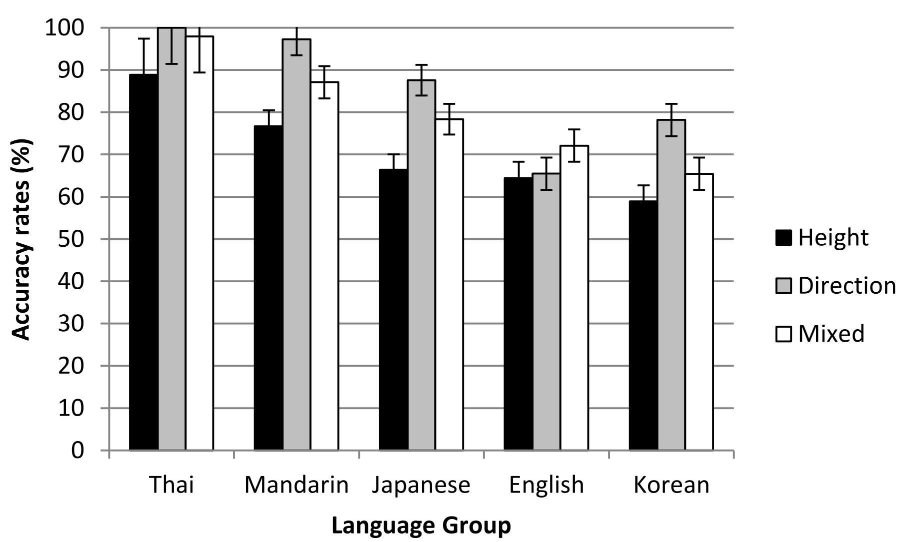


Figure 3.3. Accuracy rates for each language group and each condition. Error bars enclose +/- 1 SE.

More specifically, Mandarin listeners are significantly more accurate on the direction condition than on the height condition ($p < 0.000$), but only marginally more accurate than on the mixed condition ($p = 0.09$). Similarly, performance on the mixed condition is only marginally more accurate than on the height condition ($p = 0.08$). The Japanese have equal accuracy in both direction and mixed conditions (p > 0.1), and in both conditions they are more accurate than on the height condition (both $p < 0.03$). The Koreans are also more accurate on the direction than in both other conditions (both < 0.03). They are equally accurate on the mixed and on the height condition ($p > 0.4$).

Turning to RTs, as before, additional analyses were run for the RTs excluding the native speaker group. The average RTs in each condition for each language group are presented in Figure 3.4. A main effect of condition was revealed ($F(2, 74) = 3.8$, $p < 0.05$). Overall, the direction condition (e.g. raising-falling) showed faster reaction times in two out of four groups (average RT: height, 1282 ms; direction, 1216 ms; mixed, 1298 ms) which is significantly faster than for the mixed condition. The main effect of group is, however, not significant ($F(3, 37) = 2.1$, $p = 0.12$) while the interaction between condition and language group is marginal ($F(6, 74) = 2.0$, $p = 0.075$), indicating that groups display similar latencies to the same conditions. This and the lack of main effect of language group is likely due to the fact that performance among the groups is highly similar for the height and the mixed condition, especially for the Japanese, English and Korean groups. Univariate tests for the simple effect of group within each condition show that latencies differ by group only for the direction condition ($p < 0.05$), but not for the two other conditions ($p > 0.1$). Univariate tests verified that the effect of condition is really driven by the Japanese and Mandarin groups, as shown in Table 3.5.

Table 3.5: Statistical significance of the effect of condition for each language group (RT)

Language	F (2, 74) =	*p* <
English	1.455	.240
Korean	1.199	.307
Japanese	15.059	.000
Mandarin	5.917	.004

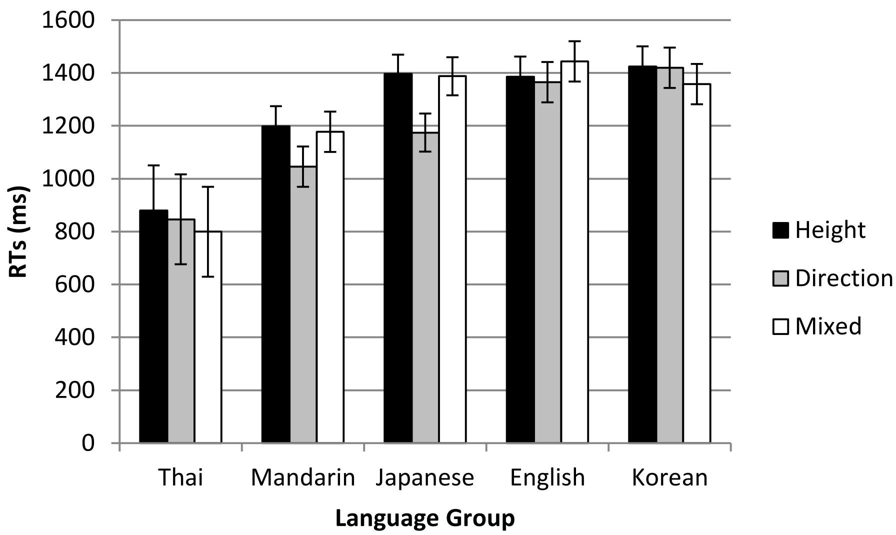


Figure 3.4. Reaction times for all five language groups and each condition. Error bars enclose +/- 1 SE.

In particular, Mandarin listeners were faster in the direction condition over both other conditions (both $p < 0.03$). Performance was equal in both other conditions ($p > 0.9$). The Japanese, similarly, were faster in the direction condition over both other conditions (both $p < 0.000$), and had equal performance in both other conditions ($p > 0.9$).

These results suggest that Japanese and Mandarin listeners are more able to track pitch direction, which is what was predicted for the Mandarin. While this finding was not specifically expected for the Japanese, it can be linked to the use of pitch accent in the Japanese phonological system, but more specific predictions need to be developed.

Taken together, the accuracy and reaction time data analyzed by condition reveal that performance is lowest in the height condition for all groups. This contradicts previous literature arguing that height might be more salient than direction for English and Japanese listeners (cf., Gandour & Harshman, 1978; Guion & Pedersen, 2007). In the data, both groups performed poorly on the height condition. We now look at height in detail.

Figures 3.5 and 3.6 illustrate the mean accuracy and reaction time on the height condition, split according to the specific comparison (low-mid, low-high, and mid-high). As can be seen, the overall lower performance on this condition might be due to the presence of a mid tone, as both accuracy and reaction times seem to indicate that this specific comparison of low and mid tone was the most difficult for the three non-tone language participants (English, Japanese and Korean). Indeed, performance on the low – mid comparison is lowest and might be statistically responsible for the effect. As a possible direction for future examination, we will explore the effect of the mid tone on performance in the height condition. The reasons for such difficulties can be looked for in acoustic salience of Low-Mid vs. Mid-High comparisons. For example, the Low-Mid comparison is often confused by native and non-native speakers of Thai (cf., Wayland & Guion, 2003) while such confusion has not been found to be the case for the Mid-High comparison. This may be due the fact that as noted in subsection 2.1.1 (page 24), the High flat tone is phonetically not flat but has a contour shape similar to the rising tone, providing two dimensions in difference between the Mid and High tone, namely height and contour.

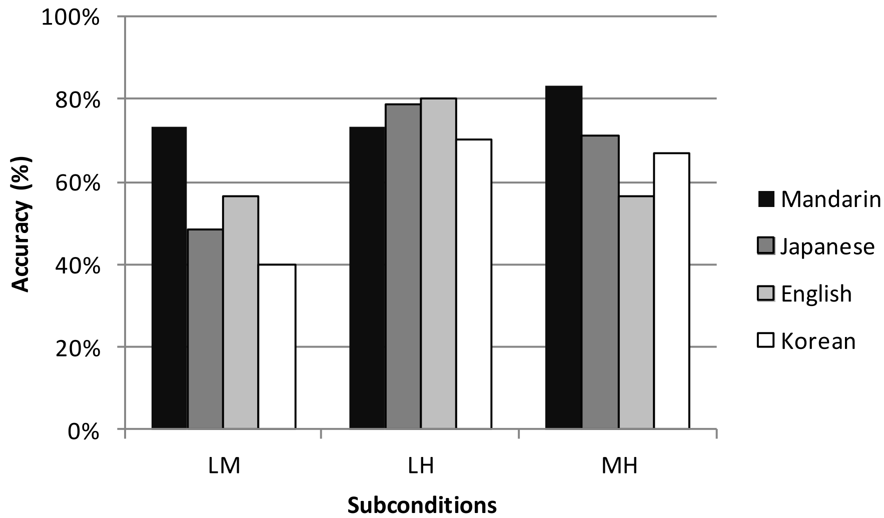


Figure 3.5. Accuracy rates for the Height conditions (L=low, M=mid, H=high) in each language group (excluding the Thai)

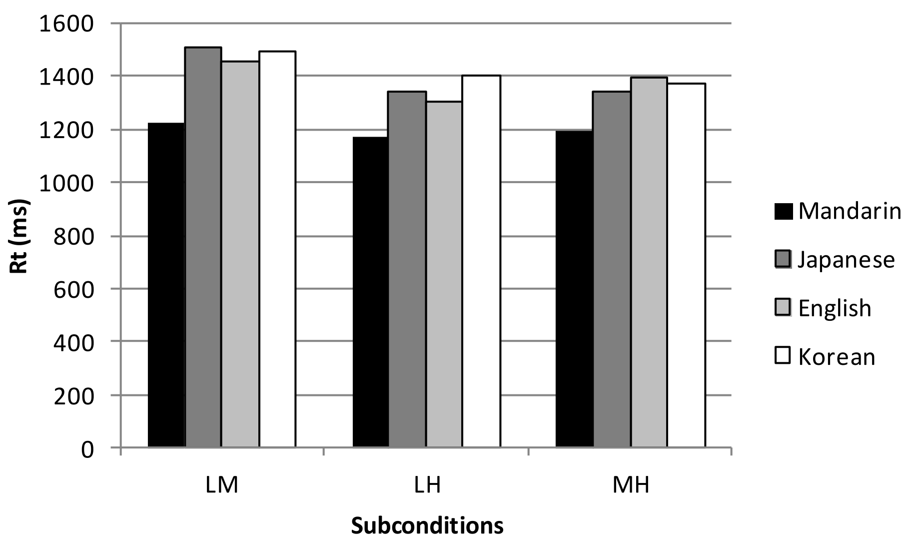


Figure 3.6. Reaction times for the Height conditions (L=low, M=mid, H=high) in each language group (excluding the Thai)

However, in reaction to the results in the above breakdown of specific tonal comparisons within the height condition, the data was reanalyzed according to the total 10 possible tonal comparisons rather than on the initially-planned three subcondition comparisons of *height*, *direction* and *mixed* (i.e., height vs direction). This was prompted by the observation that in fact the level high tone is not phonetically level but rather roughly parallels the contour of the rising tone. The accuracy and reaction times were therefore re-analyzed as a function of the ten tone subconditions (within subjects) for each non-native group (between subjects: Mandarin, Japanese, English and Korean).

A linear mixed effects model was conducted in SPSS 22 on the binary accuracy and continuous reaction times, excluding the Thai participants. *Language group* (Mandarin, Japanese, English, Korean) and *subcondition* (LH = Low-High, LM = Low-Mid, MH = Mid-High; HF = High-Falling, HR = High-Rising, MF = Mid-Falling, MR = Mid-Rising, LF = Low-Falling, LR = Low-Rising, RF = Rising-Falling) were declared as fixed effects. *Subjects* and *items* were entered as random effects in the model. When looking at the Type III tests of fixed effects on accuracy, the F-tests showed a main effect of subcondition ($F(9, 34) = 3.5$, $p < 0.01$), and a significant effect of language group ($F(3, 40.9) = 7.5$, $p < 0.001$), but no significant interaction between the two factors ($F(27, 1690.1) = 1.4$, $p = 0.071$). The same analysis on the RT showed a main effect of subcondition ($F(9, 31.3) = 2.5$, $p < 0.05$), no main effect of language group ($F(3, 38.1) = 1.9$, $p > 0.1$) and no interaction ($F(27, 1213.9) = 1.4$, $p = 0.076$). The natural conclusion to reach is that specific conditions differ in overall difficulty, as shown in Table 3.6.

Table 3.6. Mean accuracy and RT across all non-native language groups for each subcondition

Subcondition	Mean accuracy		Mean RT (ms)	
	Mean	Std. Error	Mean	Std. Error
LH	.76	.068	1244	65.3
LM	.55	.068	1342	71.1
MH	.70	.068	1315	67.7
HF	.83	.059	1273	58.5
HR	.59	.059	1466	62.6
MF	.69	.059	1284	60.7
MR	.85	.059	1277	58.0
LF	.73	.059	1365	60.0
LR	.85	.059	1215	58.2
FR	.82	.037	1218	43.9

Furthermore, the lack of significant interactions indicates that the difference between groups does not vary widely as a function of subcondition.

3.3 Summary of results

In sum, we see that the overall perceptual accuracy rates of Thai tones and reaction times resulted in a hierarchy of performance with the Thai native speakers outperforming the other language groups. Among the four non-native groups, the Mandarin speakers outperformed the other three groups followed by the Japanese who outperformed the remaining two groups of L1 English and L1 Korean. There appears to be a universal tendency for groups to react similarly to the L1 Thai group in performance but at lower levels of accuracy and RTs. Also, looking at individual comparisons we also see language-specific tendencies.

In sum, we see the following results:

- Accuracy rates for distracters (i.e., segmental comparisons) by all language groups were high, validating the task.
- Performance on the height condition was the worst for each L1 group, including the Thai group.
- As for the ten subconditions, groups generally performed poorly on the three height tones, especially the Low-Mid tone comparison and also on a fourth comparison, namely on the mixed condition comparing the High vs Rising tones, vis-à-vis other pairings. This last difficulty may be due to the Thai high flat tone resembling the rising tone in shape.
- Mandarin, Japanese and Korean patterns of accuracy in the various subconditions (i.e., height, direction, mixed) reflect Thai performances but at lower levels of accuracy and RTs. English speakers performed differently with flat results.

- The Mandarin group generally performed at accuracy levels significantly higher than the English and Korean groups but marginally higher as compared to the Japanese group.
- L1 Mandarin listeners performed best on the direction condition as expected (then the mixed and then, the height condition), reflecting their focus on direction over height. RTs are also fastest on the direction condition.
- L1 Mandarin listener RTs are faster than the other L1s possibly reflecting the higher functionality of pitch in Mandarin and experience with tone. Additionally, they may be able to map two out of five of the target Thai tones (i.e., falling and rising) onto Mandarin falling and rising tone. However, they maynot necessarily be able to map the Thai high flat tone onto the Mandarin high tone as the Thai high flat tone actually resembles the rising tone in shape. Japanese RTs are poorer than the RTs of the Mandarin group due to the lower functionality of pitch compared to Mandarin but Japanese RTs are higher than the English and Korean RTs.
- English listeners performed comparably on both the height and direction conditions as if neither intonational categories nor the high pitch correlate of lexical stress seem to be reflected in the scores. Again, the mid tone may have caused lower scores in the height condition. Again, English accuracy rates were flat across all three conditions.
- English and Korean speakers performed at similar levels of accuracy, which runs counter to the predicted hierarchy of performance set in the hypothesis.

3.4 Discussion

The results of this pilot study strongly indicate that the functionality of lexically-contrastive pitch use in different L1s shapes the cross-linguistic perception of non-native tone, indicated by the effect of the L1. The L1 Mandarin group performed significantly more accurately than the English and Korean groups overall and also the Japanese (but not significantly). The Japanese were significantly more accurate than the Korean group overall and the English listeners in terms of raw scores, suggesting that a higher L1 pitch functionality aids in non-native tone perception. That is, we clearly see performance on the perception of non-native Thai tones varies according to the functionality of L1 lexically-contrastive pitch in the L1. Thus, the predicted hierarchy of performance (of more accurate to less accurate) was generally confirmed as follows: L1 tone > L1 pitch accent > L1 stress = L1 without lexically-contrastive pitch. Globally these pilot findings confirm previous results obtained across studies and add strength by allowing a direct comparison among four language groups with the same methodology. The degree of functionality of pitch to signal lexical contrast appears to determine accuracy on this tonal categorization task.

The result showing the comparable levels of accuracy between the L1 English and L1 Korean groups was not predicted as it was predicted that the L1 English group with word stress would outperform the L1 Korean group which does not feature lexically-contrastive pitch. Several possibilities may account for this result. First, f0 is rarely used alone to distinguish words in English. This fact appears to yield the same performance in tone discrimination as if f0 was not used at all to signal lexical contrast (i.e., English = Korean). Additionally, the fact that f0 can be used exclusively to distinguish words in Japanese plays an important role as suggested by the different patterns obtained by the Japanese and English listeners. The flat performance by the English speakers is also consistent with findings showing that stress constrains lexical access only to a limited extent in English (Cooper, Cutler, & Wales, 2002). A second possible reason may be that Koreans are more accurate in this task because of their exposure to and acquisition of L2 English word stress. Alternatively, the effect could be due to exposure to a pitch-accent dialect (cf., Sukekawa, Choi, Maekawa, & Sato, 1995). Both these facts require a reanalysis of the data for the Korean speakers and to retest with a control group speaking only the Seoul dialect and who are also somewhat older, since lexically-contrastive pitch usage is appearing among younger speakers of the Seoul dialect (Silva, 2006). Such a group was indeed tested and will be discussed in section 3.5, pp. 89-95.

Concerning the research question as to whether different L1 language groups resort to different features in their L1 to perceive non-native tones (height or direction), we find the following results. Overall, most groups reflect a general trend that indicates a more robust tendency to rely on pitch direction, as observed by the fact that performance was less accurate on the height condition compared to other conditions. This result seems to contradict previous research that showed Japanese and English groups focusing more on pitch height (Guion and Pedersen, 2007). This result also conflicts somewhat with the previous results showing the mixed condition was the most difficult followed by the height condition (Burnham et al., 1992), requiring this issue to be re-examined. However, it is possible that these results are in part explained by the fact that two voices with different gender (male and female) were used, thus making the pitch height comparison more difficult than in other studies using only one voice, which while not confusing for L1 perception (Lee, 2009), may be for non-native, naïve perception. Another possible explanation could also be the fewer number of comparisons in the direction condition (only Rising = Falling) compared to height (Low=Mid, Low=High, Mid=High) or to mixed (see Bohn, 1995). Yet, this did not affect the English (who showed a flat performance in all conditions). One possible explanation for their performance is that the use of monosyllabic stimuli may have prevented them from applying intonational contours to tonal comparisons varying in tone contour, as suggested by similar findings obtained with disyllabic stimuli by Braun and Johnson (2011). This will be examined in more detail in chapter 4.

To conclude, this pilot study was conducted to further the understanding of cross-linguistic perception of tonal contrasts and expand current models of L2 phonology as a first step in defining naïve perception of lexically-contrastive pitch and a baseline for L2 tone acquisition.

3.5 Follow-up to the pilot study: Re-analysis of Korean speaker data and testing of additional Seoul-Korean speaker group

As explained above, an unexpected finding was the equal performance of the L1 English speakers and L1 Korean speakers. In this follow-up to the pilot study, we must specifically ask whether the L1 English speakers performed lower than expected or whether the L1 Korean speakers performed higher than expected. That is, we question whether F0 information is less readily accessible for phonological discrimination for the L1 English speakers and more so for the L1 Korean speakers. There are two possible scenarios. The first one is that the word stress correlate of pitch in English is not robust enough for the L1 English speakers to access/transfer to their perception of Thai tones. The second one is that the Koreans were able to use the presence of L1 pitch accent patterns (either from the Kyungsang dialect or "young-generation" Seoul dialect) to aid their perception of Thai tones. This paper explored the latter hypothesis and explicitly investigated whether dialectal exposure has influenced Korean performance.

The group of Koreans was not fully homogenous in terms of dialect as can be seen in Table 3.7. We see that four out of ten participants were not speakers of the Seoul dialect. Three were speakers of the pitch-accent Kyungsang dialect while one was a speaker from Cholla, a region in the southwest of South Korea abutting the Kyungsang dialect region. We know that Kyungsang listeners show categorical perception of pitch accent patterns (e.g., LH, HL, HH to distinguish minimal word pairs) (Kim & de Jong, 2007; Kim, 2011) and limited advantage in the naïve perception of Japanese pitch accent (Sukegawa, Choi, Maekawa, & Sato, 1995). If the L1 phonological system determines accuracy, then Kyungsang dialect speakers should outperform non-Kyungsang dialect speakers, unless emerging pitch accent patterns in Seoul Korean are robust enough to level performance across the whole group.

Table 3.7. Background of Korean participants

Code	Age	Dialect region
KRF1	35	Seoul (Standard)
KRF2	30	Seoul (Standard)
KRF3	29	Busan (Kyungsang)
KRF4	32	Seoul (Standard)
KRF5	30	Jinju (Kyungsang)
KRF6	27	Seoul (Standard)
KRF7	35	Seoul (Standard)
KRM1	27	Busan (Kyungsang)
KRM2	47	Cholla (near Kyungsang)
KRM3	30	Seoul (Standard)

We turn now to examining their performance on the test vs. control condition, in order to see if the people who come from pitch accent regions would perform more accurately than the others. Figure 3.6 shows the individual performance for each of the Korean participants on the combined test conditions, and Figure 3.7, on the control condition. Black bars represent Seoul dialect speakers while white bars represent Kyungsang dialect speakers; the lone grey bar represents the speaker from Cholla.

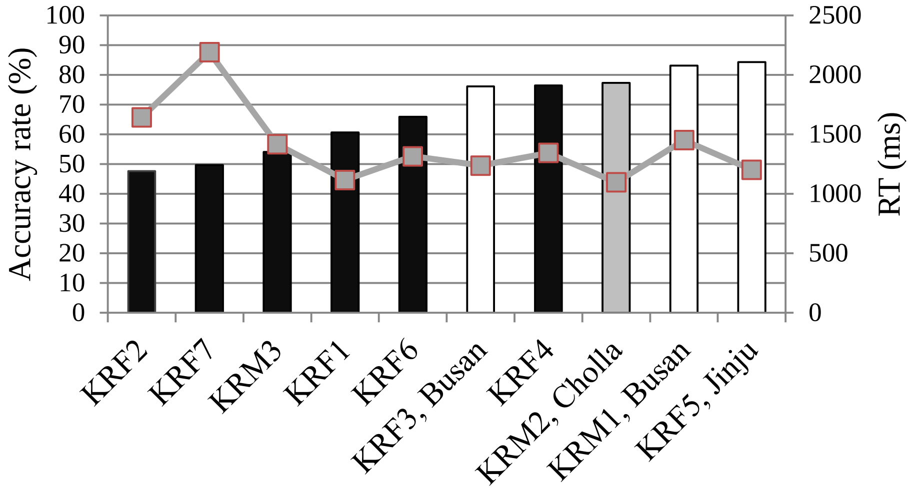


Figure 3.7: Mean individual accuracy rate (%) and RT (ms) for the Korean participants on the combined test conditions.

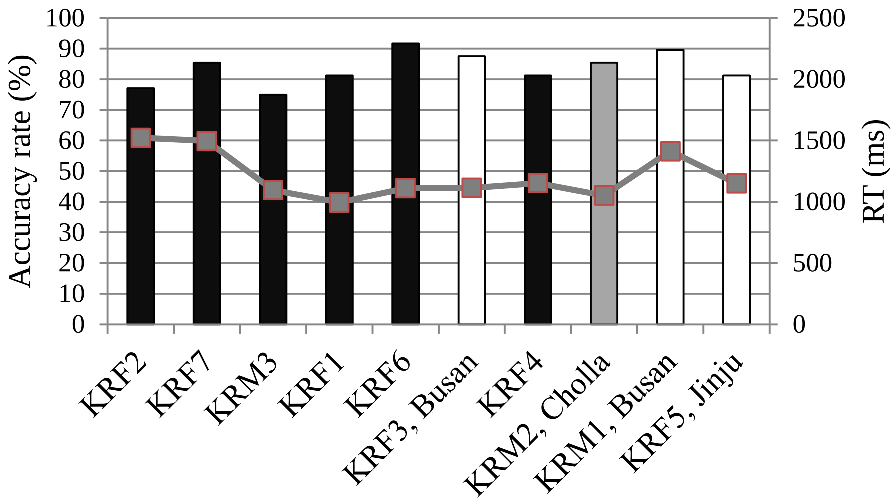


Figure 3.8: Mean individual accuracy rate (%) and RT (ms) for the Korean participants on the control condition.

In Figure 3.6, all the Kyungsang dialect speakers, i.e., the pitch accent speakers, are on the higher end of the accuracy spectrum, outperforming non-pitch-accent dialect speakers. In Figure 3.7 for the control condition, however, there is no clear relationship between dialectal group and performance. Figure 3.8 now compares the performance of the Korean group to the performance obtained by the other groups. The group of Koreans is split by dialect region. In Figure 3.6, Thai listeners are displayed in black for comparison purposes.

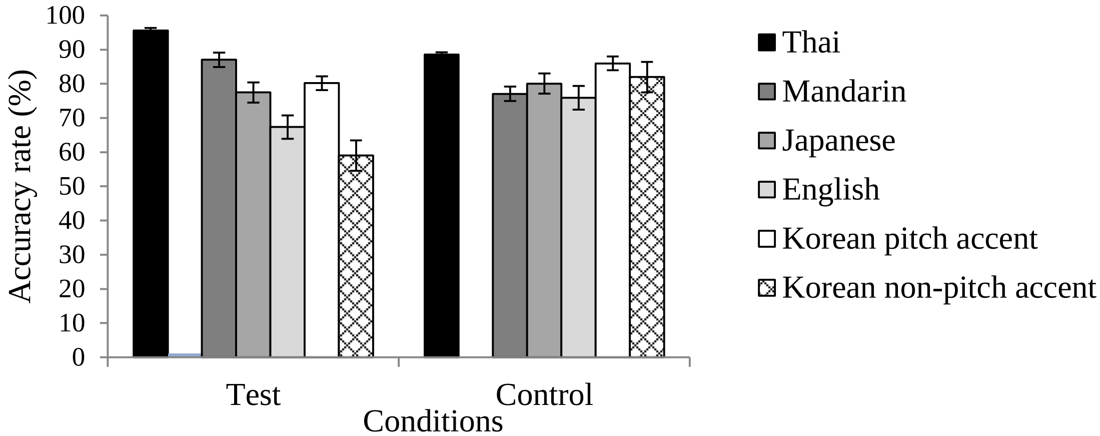


Figure 3.9: Mean accuracy rate (%) for each language group with the Koreans split into subgroups for the test vs. control condition. Error bars enclose +/- 1 SE.

We see a clear difference between the two Korean subgroups. The non-pitch accent speakers are in fact less accurate than the L1 English speakers, which conforms to the original prediction while the pitch-accent speakers in contrast outperform the English speaker group, and are more comparable to the Japanese speaker group. Additionally, in order to eliminate the influence of "native-speaker levels" of pitch-accent Kyungsang Korean dialect and to limit exposure to English, an additional 10 speakers of the Seoul-Korean dialect were tested onsite in Seoul Korea.

For the analysis of the group of Seoul Korean participants, a linear mixed effects model was conducted in SPSS 22 on the binary accuracy and continuous reaction times data. C*ondition* (test, control) was declared as fixed effect. *Subjects* and *items* were entered as random effects in the model. There was a main effect of condition on accuracy ($F(1, 90) = 16.5$, $p < 0.001$), indicating that performance on the test condition (M = 59.9% correct) was less accurate than on the control condition (M = 77.1% correct). Similarly, there was a marginal effect of condition on RTs ($F(1, 535.2) = 3.98$, $p = 0.05$): RT on the test condition trended to be slower (M = 1294 ms) than on the control condition (M = 1247 ms). The parameter estimates are presented in Tables 3.8 and 3.9.

Table 3.8: Parameter estimate, standard error, *t*-value, *p*-value, and 95% confidence interval of the predictors for the AXB accuracy

Fixed Effects	Estimate	Std. Error	df	t	Sig.	95% Confidence Interval lower	upper
Intercept	.77	.037	30.8	20.7	.000	.69	.85
[TestCondition=1]	-.17	.043	90	-4.1	.000	-.26	-.09
Covariance Parameters	Estimate	Std. Error					
Residual	.179	.009					
Subject	.005	.003					
Item	.025	.006					

Note: Control is the reference condition

Table 3.9: Parameter estimate, standard error, *t*-value, *p*-value, and 95% confidence interval of the predictors for the AXB reaction times

Fixed Effects	Estimate	Std. Error	df	t	Sig.	95% Confidence Interval lower	upper
Intercept	1247	57.3	9.6	21.8	.000	1119	1375
[TestCondition=1]	47.0	23.9	535.2	2.0	.050	.073	94.04
Covariance Parameters	Estimate	Std. Error					
Residual	74999	4590					
Subject	30314	14995					
Item	b	.0					

Note: Control is the reference condition. b. This parameter is set to zero because it is redundant.

To compare the results of the Seoul Korean participants to the other four non-native groups, a linear mixed effects model was conducted in SPSS 22 on the binary accuracy and continuous reaction times data. C*ondition* (test, control) and *Language group* (Mandarin, Japanese, English, Korean, Seoul Korean) were declared as fixed effect. *Subjects* and *items* were entered as random effects in the model. When looking at the Type III tests of fixed effects for accuracy, the F-tests showed no main effect of condition ($F(1, 90.0) = 2.7$, $p > 0.1$), a significant effect of language group ($F(4, 46.1) = 5.82$, $p < 0.002$), and a significant interaction between the two factors ($F(4, 4537.1) = 21.0$, $p < 0.001$). In other words, the exact same pattern of results was obtained with this new group added to the analysis. When looking at the Type III tests of fixed effects for the RT analysis, again, the exact same pattern of results as in the previous analysis emerges: the F-tests showed a main effect of *condition* ($F(1, 85.9) = 16.3$, $p < 0.001$), no significant effect of language group ($F(4, 46.0) = 1.6$, $p > 0.1$), and a significant interaction between the two factors ($F(4, 3256.7) = 6.4$, $p < 0.001$).

To more clearly show the effect of exposure to pitch-accent in the two groups of Korean participants, the participants of the first Korean group were reassigned according to whether they had exposure to pitch-accent in their native dialect. The six who did not were grouped with the new Seoul Korean participants (n = 16), and the four pitch-accent Korean speakers were in a separate group. Figure 3.9 displays the respective accuracy scores for these two Korean groups (white and dotted bars), in comparison to the other non-native groups. It becomes clear that the non-pitch-accent Korean participants in this reanalysis now corroborate the originally predicted hierarchy of performance. Interestingly, the four Korean speakers with exposure to pitch-accent perform at the same level as the Japanese speakers. A linear mixed effects model on the accuracy scores declared C*ondition* (test, control) and *Language group* (Thai, Mandarin, Japanese, English, Korean/pitch-accent, Seoul Korean/no pitch-accent) as fixed effect. *Subjects* and *items* were entered as random effects in the model. The Type III tests of fixed effects for accuracy were identical to the above analysis: there was no main effect of condition ($F(1, 98.7) < 1$), a significant effect of language group ($F(5, 47.0) = 9.8$, $p < 0.001$), and a significant interaction between the two factors ($F(5, 4718.1) = 20.6$, $p < 0.001$). Post-hoc pairwise comparisons also indicate that now, the Seoul Korean group was significantly less accurate than the English group on the test condition ($p = 0.41$). This group was also significantly less accurate than all other groups (all $p <$ 0.001) on the test condition – but not on the control condition.

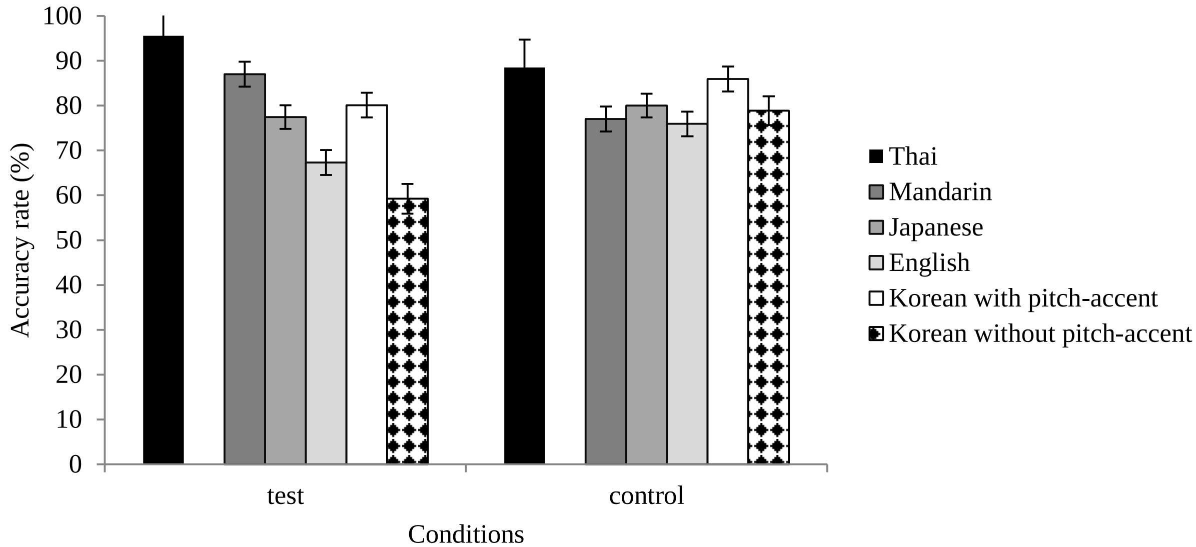


Figure 3.10: Accuracy rate (%) for each language group in the test vs. control condition. Error bars enclose +/- 1 SE. (Thai listeners are displayed in black for comparison purposes)

In sum, experiment 2 clearly demonstrates that the reason for the equal performance of the Korean and English groups in Experiment 1 was not due to the English performing less accurately than expected, but indeed to the Koreans performing more accurately. The presence of speakers of the Kyungsang pitch-accent Korean dialect in the original Korean group appears to have been the determining factor in the equal performance between the L1 Korean and L1 English group. Therefore, the originally predicted hierarchy of performance based on pitch functionality in L1 is further supported.

In addition to these points of consideration, L1 also must be defined in the narrow sense where one’s dialect (e.g., pitch-accented Kyungsang Korean dialect, non-pitch-accented Fukushima Japanese dialect) also impacts pitch functionality in a listener’s L1 phonological system (cf. Otake and Cutler 1999; Pallier, Bosch, and Sebastian-Gallés 1997; Weinreich 1953).

3.6 Conclusions and limitations

The previous analysis of the pilot study offered evidence that performance on the perception of non-native Thai tones varies according to the functionality of lexically-contrastive pitch in the L1. The predicted hierarchy of performance (of more accurate to less accurate) was first partially supported as follows: L1 tone > L1 pitch accent > L1 stress = L1 without lexically-contrastive pitch. The results from the follow-up analysis provided further evidence for the interpretation obtained from the findings of the pilot study, confirming more fully the initial hypothesis: L1 tone > L1 pitch accent > L1 stress > L1 without lexically-contrastive pitch. In conclusion, the functionality of lexically-contrastive pitch use in different L1s shapes the cross-linguistic perception of non-native tone. This interpretation is supported here even more specifically by the effect of the native dialect (D1) in the Korean group.

It has also been pointed out that English listeners may have performed less accurately due to the fact that F0 is rarely used alone to distinguish words in English and may be only a weak cue to word stress (Sluijter & van Heuven, 1996). Yet, the results of the second pilot study appear to demonstrate that pitch in English word stress may indeed be robust enough to differentiate these speakers' performance in Thai tonal perception accuracy from that of the L1 standard Korean speakers. Therefore, the question as to whether the English performed less accurately than expected may have also been answered: it appears that they have not, as they outperform the speakers of L1 standard Korean. Additionally, it appears that exposure to the pitch correlate of L2 English word stress by the speakers of standard Korean was not sufficient enough to boost their performance to the levels of the L1 speakers of English, even though this must remain speculative at this point.

While the findings in the pilot study seem to lend support to this hypothesis, the small sample size accessible in this pilot study did not allow for definite conclusions and alternative explanations cannot be fully ruled out. For example, perhaps differences in English proficiency among participants indirectly contribute to the observed differences independently of the dialectal origin. To conclude, these findings in the pilot study highlight the need to more thoroughly examine the L1 of participants in terms of dialect and possibly other factors. The influence of the acoustic correlates of word stress in English on the naïve and/or L2 perception of other linguistic pitch systems must be further studied.

While the pilot study was exploratory in nature, it was nevertheless able to delineate the outlines of the influence of the functionality of lexically-contrastive pitch on the naïve perception of Thai tones. The very exploratory nature of the pilot, however, leads to several limitations in design.

First, concerning the instrument design and stimuli, the ABX design was "non-traditional" in its composition: rather than using the same set of segments with all five tones (i.e., [no:i] for the low, mid, high, falling, rising tones), the pilot study used a different set of segments for each tone, to reduce monotony. A drawback of this choice is that it remains somewhat unclear to what extent the different sets of segments may have influenced the difficulty of the various trials.

Second, the control segment items were of a wide variety, presenting a few possible difficult sounds for the non-native participants. Also, the control items were not set up in a complete paradigm for the AXB task, i.e., ABB/AAB/BAA/BAB comparing each minimal pair.

Third, real words along with nonce words had been used in the pilot study. While this is not a problem for participants who do not know Thai, it allows the L1 Thai speakers to access their mental lexicon and not merely their L1 phonology, making the comparison between groups difficult.

Fourth, exposure to the five tones and to tone comparisons was not strictly balanced, allowing for a possible bias. Due to the design examining the height, direction and mixed tone combinations, there was a bias toward the direction condition as the comparison only consisted of the falling vs rising tone which was presented with the same number of trials as for the other two subconditions. As a result, for the direction condition there were twelve trials comparing the falling vs rising tones compared to three trials for each possible comparison in the height condition (i.e., low vs mid, mid vs high, low vs high x 3 each).

In the following experiments, these limitations were taken into account during the design, such that the number of times participants hear a particular tone have been balanced against each other; The control segmental stimuli have been made simpler so that almost no sound might be difficult to perceive for any one language group; No real words from Thai were used, and the items were vetted to ensure they were not words in any of the languages of the participants, etc. Also, the subconditions of height, direction and mixed were replaced with a total of the ten possible tone comparisons in the monosyllabic ABX task. The specific design will be described in detail in each section.

In addition, several choices regarding the stimuli, while not limitations per se, raise additional questions. In the pilot study, only monosyllabic stimuli were used. It is possible that the use of disyllabic stimuli would make a difference – particularly for English participants, as suggested by findings such as Braun and Johnson (2011). In the following chapter, a series of experiments will be presented where both monosyllabic and disyllabic tonal stimuli reflecting English word stress are used.

Further, the pilot study employed an AXB task, rather than an ABX, which has been shown to exhibit conflicting outcomes, resulting from for example the X being ignored by participants (cf., Gerrits & Schouten, 2004). Since an ABX sequence also allows measuring RTs more straightforwardly, this version will be used in the following experiments.

An interesting aspect of the pilot study was also the background of participants, both in terms of dialectal/regional variant, and in terms of other languages learnt, as these aspects were shown to have measurable effects on performance, and therefore need to be carefully considered. Almost all were assumed to have had superior proficiency in English as an L2. This issue is tied into another issue of the participants having been mostly graduate students studying language areas (e.g., linguistics, second language studies, foreign language). Similarly, a few of the Koreans were native speakers of the Kyungsang Korean pitch-accent dialect and a few of the Mandarin speakers exhibited exposure to various Mandarin dialects and/or “Chinese languages” (Taiwanese, Cantonese). In the following experiments, care will be taken to describe participants’ backgrounds in detail to allow for a cautious interpretation of results.

Finally, a second language learning element has been added: a group of native speakers of English who have learned or are still currently learning Mandarin Chinese as an L2 has been added to the original five participant groups of the pilot study.

Thus, the above-mentioned limitations in the pilot study have been generally addressed in the following experiments as will be outlined and discussed in more detail in the following chapters.

4. Naïve perception of tones

4.1 Introduction

The goal of the current experiments is to replicate the pilot study discussed in Chapter 3, and to examine the perception of Thai tones in the form of non-words by listeners who speak Thai, Mandarin, Japanese, English, and Korean as their L1. Unlike the pilot study, the current study employs two tasks: one for monosyllabic stimuli as in the pilot study and an additional one for disyllabic stimuli. The task with disyllabic stimuli was added as the English group in the pilot study was thought to require disyllabic stimuli to better "apply" the word stress of English to the perception of the task stimuli.

An additional difference concerns the order of presentation of the stimuli: the pilot study features an AXB order while that of the current study features an ABX order. The ABX task is considered more conventional; furthermore, the AXB task can produce "contradictory results," e.g., the third stimuli of B in an AXB task may be ignored by the listeners (Gerrits & Schouten, 2004: 364). The cognitive load of the ABX tasks using non-words and with a 500ms interstimulus interval (ISI) between the three stimuli (i.e., A, B, X) allows access to a participant's perception of the similarity between three stimuli as shaped by the use of lexically-contrastive pitch in the L1.

These tasks were devised in order to answer the first of the research questions:

1. RQ1: Does the varying presence of certain features in the L1, specifically lexically-contrastive pitch, aid in the naïve perception of non-native tone, thus resulting in a linguistic hierarchy of perceptual ability?

Two sub-questions to this research question emerged from the results of the pilot study in Chapter 3:

1.1 RQ1.a: Are some tones or tonal comparisons easier to perceive than others (cf., pitch height or pitch direction) in the naïve perception of non-native tone?

1.2 RQ1.b: Are tones presented in disyllabic stimuli easier or more difficult to perceive compared to monosyllabic stimuli?

4.2 Experiment 1a: ABX Thai tones, monosyllabic

We now turn to the methodology and results for a monosyllabic ABX task which tested several language groups differing in lexically-contrastive pitch.

4.2.1 Methodology

4.2.1.1 Participants

Participants were recruited from five L1 language groups: Mandarin, Japanese, English, Korean, and Thai. The Thai speakers were recruited as a native-speaker control group. In total, 116 participants were recruited (Mandarin = 31, females = 25; Japanese = 23, females = 11; English = 24, female = 14; Korean = 29, females = 19; Thai = 9, females = 8). None of them knew Thai nor had any knowledge of a tone language (except the Mandarin participants). The age ranges of each group were as follows: Thai: 32-41 (average age: 31.7; SD = 7), Mandarin: 18-37 (average age: 26.1; SD = 6.3), Japanese: 18-37 (average age: 28.8; SD = 5.2), English: 18-50 (average age: 31.2; SD = 9.1), Korean: 20-54 (average age: 31.3; SD = 7.9). On the whole, participants were either undergraduate or graduate students at Indiana University with a few exceptions. Demographic and language background information are presented in Appendix H (pp. 305-314). None of them participated in the pilot study described in chapter 3. Each participant was recruited either through flyers distributed on campus or through English language support courses, Mandarin as a foreign language courses or Second Language Studies courses at Indiana University or by word of mouth (i.e., participants would contact their friends or classmates). All participants were paid $10 for participation in the monosyllabic ABX task and disyllabic ABX task discussed in section 4.3. All procedures were approved by the Indiana University Institutional Review Board.

4.2.1.2 Stimuli and conditions

The design of this experiment was similar to the pilot study. There were two types of stimuli: Target and Control. The target non-word stimuli varied by tone but not segment while the control non-word stimuli varied by segment but not tone. Non-words were checked by native speakers of all the languages of the participants, i.e., Thai, Mandarin, Japanese, English and Korean, to ensure that the stimuli were indeed non-words in all of these languages. However, a few stimuli were deemed to be words in other Mandarin dialects or Taiwanese. Additional distracters were not included.

For the target stimuli, an equal number of stimuli from each of the five Thai tones was used: low (L), mid (M), high (H), rising (R), and falling (F). The monosyllabic control items differed by only one segment, either a vowel or consonant, and were equally divided among the five possible Thai tones. Open syllable items were created as open-syllable CV words are considered more difficult to perceive than closed syllable words (Wayland and Guion, 2003).

Two target tone syllables and eight control syllables (two pairs varying by vowels and two varying by consonants) were selected as stimuli. Each of the two tone target syllables carried each of the five Thai tones; these five tonal items were then paired to form ten possible tone comparisons: F-R, H-F, H-R, L-F, L-H, L-M, L-R, M-F, M-H and M-R. This resulted in 20 tonal pairs (10 with syllable one, 10 with syllable two). For control items, the four syllable pairs also carried each of the five tones (both members of the pair always carried the same tone), resulting in 20 pairs of items. Table 1 presents the overview of the experimental items.

Each of these pairs was then arranged in an experimental trial (a triplet) where one member of the pair (A, B) was repeated. This produced four triplets for each pair: ABA, ABB, BAB, BAA

(The third token of the triplet is the X token). All four possible combinations of ABA/ABB/BAA/BAB were used to balance presentation and prevent bias. Thus, 40 pairs of items produced 160 experimental trials. Before completing the experimental blocks of 160 trials, participants were given 10 trials as a “training” or familiarization phase. Each trials contained two syllable pairs carrying the five tones, resulting in 10 pairs of items, arranged as above into 10 triplets. The members of each pair in each trial carried the same tone. Thus, only differences in segments were used in the trials and not differences in tones as the purpose of the training was to ensure that the participants understand how to do the task. Feedback was only provided during the training session. Each tone was heard 6 times each in the training session, and 96 times in the experimental trials. Therefore, in total in this task, each tone was heard 102 times.

Table 4.1 Stimuli, conditions and number of trials in the ABX monosyllabic task

Test (tone) (80 trials**)		Control (segmental) (80 trials**)		Training (segmental only) (10 trials**)	
Tone types	segments	Tone types	Segment pairs	Tone types	Segment pairs
Falling-Rising Low-Falling Low-Rising Mid-Falling Mid-Rising High-Falling High-Rising Low-Mid Low-High Mid-High	[no:i] โนย* [p^{h}uai] พวย	Falling Rising Low Mid High	[bɛ:o] เบว - [tɛ:o] เตว [wia] เวีย - [t^{h}ia] เทีย [ua] อัวะ - [ia] เอีย [ria] เรีย - [rɯ:i] รืย	Falling Rising Low Mid High	[wuːi] วูย - [p^{h}uːi] พูย [dua] ดัว - [ŋiːn] งีน

*Spellings in the Thai script may not reflect the actual tones[9].
** Test condition = 10 tone types x 2 segments pairs x 4 orders (AAB, ABB, BBA, BAA) = 80 trials; Control condition = 5 tone types x 4 segment pairs x 4 orders (AAB, ABB, BBA, BAA) = 80 trials; training = 5 tone types x 2 segment pairs = 10 trials

4.2.1.3 Speakers for Thai stimuli, and elicitation method

Two female voices of the Central Thai dialect recorded the stimuli. The speaker for the A and B tokens was a 28-year old speaker from the Bang Phlat (บางพลัด) district of Bangkok, Thailand, while the speaker for the X tokens was a 25-year old speaker from the Min Buri (มีนบุรี) district of Bangkok, Thailand. This method allows us to use only two voices rather than three, and furthermore, using one voice for the X token and another voice for the A and B tokens adds more

[9] A complex combination of spelling conventions determines the tone of any given syllable in Thai. First, a vowel within a syllable may be marked above by one of four diacritics denoting tone: ◌่,◌้, ◌๊,◌๋. These diacritics are referred to as one (ไม้เอก), two (ไม้โท), three (ไม้ตรี) and four (ไม้จัตวา), respectively. A syllable may also not be marked by any of the four diacritics. This is, however, where the similarity to the markings of Mandarin tones in *pinyin* ends. The tones of syllables in Thai are further determined in conjunction with three other factors. The first factor is the class of the consonant (i.e., low, mid, high *which do not correlate to the tones with the same designations*) preceding the vowel. The consonant does not necessarily represent phonetically different consonants. For example, there are several letters that represent [k^h] such as ข (‘high’ consonant) and ค (‘low’ consonant), e.g., ไข่ [k^{h}ai] low tone ‘egg’ ควาย [k^{h}waai] mid tone ‘water buffalo’. The second factor is whether the syllable is “alive” or “dead” (i.e., long vowel/open syllable or short vowel/closed syllable, respectively). The third factor is whether one of two silent letters (ห [h] or อ [ɔ]) precedes the initial consonant of the syllable. This third factor is somewhat analogous to English where adding a final ‘e’ to the word ‘fin’ changes the pronunciation of the ‘i’ as in the case of ‘fine.’ While students are taught the rules in school, anecdotally many Thai appear to spell and read by rote memorization. Again, this is analogous to English where generally spelling is memorized despite being taught a few rules like “’i’ after ‘e’ except after ‘c’” etc.

difficulty to the task than using one voice for all three tokens. Unlike in the pilot experiment, two female voices were used to prevent any possible difficulty in comparing voice height between a male (low) and female (high) voice. One of the speakers had just completed her PhD at her home institution in Thailand and was studying English in the Intensive English Program at Indiana University while the other speaker was a graduate student (LLM degree) in the law program at Indiana University.

When creating the monosyllabic items, the first speaker recited the stimuli in the five-tone paradigm order that Thai speakers learn for reciting the tones: M, L, F, H, R. She repeated this three times for one item using this 5-tone paradigm. The second speaker (for the X stimuli) mimicked the first speaker and recited the target item with the same five-tone paradigm order, speaking into a second microphone. After recording one item by both speakers, they then moved onto another stimuli item. The two speakers monitored each other's pronunciation. This ensured a greater consistency and accuracy in producing the items with the correct tones.

Recordings were made in a soundproof room with the aid of a professional sound technician. The Thai stimuli were both recorded using a Shure SM7B microphone. A Shure M267 4-channel mixer was used to sum all the signals or in this case amplify signals from a dynamic microphone. A Motu 828 (audio interface) converter was used to convert analog to digital signals. A G5 Mac using Peak LE 5.2 digital audio workstation (i.e., a computer plus Peak LE audio software) was used to record and make edits on the stimuli. All files were recorded at 48kHz/24bit and left unprocessed. Thai stimuli were recorded on a 2-track stereo channel with 2 microphones, and then the two tracks were split into separate mono files. A Peak LE 6 audio editor was used to export dual mono tracks from the original stereo file. They then were spliced into individual wav. files using a program in Praat (Boersma & Weenink, 2015; Boersma, 2001). Finally, the wav.

soundfiles from both speakers were normalized for amplitude so that all items were comparable in loudness. The Thai monosyllabic stimuli were normalized using Audition software at 13.3Hz.

Tone patterns and their labels were checked using Praat to determine by visual inspection and by ear whether the tones were indeed the targeted tones. Soundfiles were also checked for clear splicing to ensure that final or initial sounds of the items were not cut off.

4.2.1.4 Procedure

Participants were tested individually in a quiet psycholinguistics laboratory on the campus of Indiana University. After arriving in the lab, participants first read and signed the consent form. Each participant sat in front of a personal computer, wearing high-quality headphones. They listened to the auditory stimuli at a self-selected comfortable volume level.

The experimental session started with the instructions presented on the screen. Participants were instructed to listen to the triplets to decide whether the last token (X) was more similar to the first (A) or the second one (B). They indicated their response by pressing a clearly labelled key (A or B) on the computer keyboard. The researcher sat near the participant as they did the training session, and asked whether they understood what was expected of them. Participants were allowed to ask questions to clarify. They then started the experimental blocks. The 160 experimental triplets were randomly presented to participants. Trial presentation was controlled by DMDX (Forster & Forster, 2003). An interstimulus interval (ISI) of 500ms was inserted between each stimulus within a triplet[10]. A break was inserted every 40 trial, resulting in four blocks of 40 trials.

---

[10] Listeners process pitch differently as witnessed in different findings of how different uses of pitch in the L1 influence non-native tone perception. In certain cases, this difference could be equivalent to linguistic vs. nonlinguistic processing of pitch. Therefore, we might expect different mental representations of tone between L1 and L2 speakers. That is, naïve L1 English speakers may rely on acoustics, or the physical properties of sound, and hear Thai tones not as language but as music or sound, for example, and thus, perform well on perception tasks based on their musical ability or ability to hear pitch (i.e., acoustic mode of processing).

In opposition to such processing, sounds may be processed as language under two other modes of processing: phonetic and phonological modes. Strange (2010) describes the phonetic mode as being perhaps necessary to adjust to the context-dependent allophonic details of a very different dialect in one's language and thus, also used in the

Response time-out was set at 2500ms to speed up the task. Reaction times were measured from the start of the X token which is the last of the three tokens. In total, duration of this task was approximately 20 minutes. The DMDX (Forster & Forster, 2003) script for this task is presented in Appendix I (pp. 315-322).

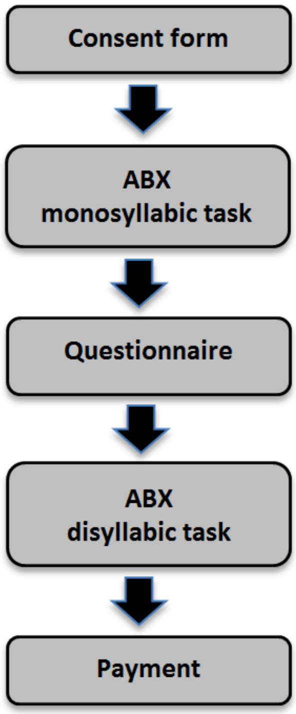
This task was administered as part of a larger testing session containing the two ABX tasks (monosyllabic and disyllabic). After finishing the monosyllabic ABX task, the participant was asked to fill out a questionnaire. After this, the participant then proceeded to the disyllabic ABX task (with a procedure highly similar to the monosyllabic AXB task). When the participants finished the disyllabic AXB task, they were debriefed, i.e., asked if they had any comments or questions about the two AXB tasks. This generally required between 45-60 minutes in total.

---

beginning of acquiring an L2 with its non-native segments. In the phonetic mode, the cognitive load is greater due to the need for greater attentional focus and thus, perception is less optimal (i.e., slower and unclear in noisy conditions) than in the phonological mode. In contrast, the phonological mode is used when listening to the same or similar dialect where context-dependent phonetic variations (i.e., speech rate, differences in prosodic structures, minor dialectal variations of familiar speakers) are not compensated for. This mode is more automatic and therefore, faster.

In order to tease apart whether listeners interpret tonal information phonologically, it seems crucial to increase the task demands in order to tax cognitive resources. Not doing that (keeping the demand of the task low) might allow listeners to perform a perceptual task at an acoustic level, therefore making claims about the phonological processing of tonal information inappropriate. Greater cognitive load in the form of more demanding perceptual tasks (e.g., longer interstimulus intervals) and complex stimuli (Strange, 2010) would allow the teasing apart of the modes of processing and be reflected in accuracy scores and reaction times (RTs). Several methods are possible, including stimulus uncertainty and memory load. Specific task features are the use of different voices for the stimuli, the use of natural stimuli, nonsense words, continuous speech and stimuli in sentence context. Also, embedding the stimuli in noise is a possibility to increase task demand as is the use of specific Inter stimulus interval durations (Strange & Shafer, 2008; Wayland and Guion, 2004). It should also be stated that the presence of memory load along with phonetic variability is what matters and not the size of the memory load when teasing out an acoustic-based approach to L2 stimuli (Dupoux et al. 2001). As a result, I have chosen an ISI of 500 ms, but more specifically increase stimulus complexity with difference voices, longer stimuli or noise, in order to enhance the use of the phonological mode of processing for tone (see also Højen and Flege, 2006).

Figure 4.1: Flowchart of general procedure



4.2.2 Results

Data from the ABX task and from the questionnaire were first converted to an Excel spreadsheet and then, entered into SPSS for statistical analyses. The data from the language background questionnaire was examined for the purpose of screening participants with non-target profiles. One Korean native speaker was removed because he had attended four years of college in China. Also, one English native speaker was removed because he had been exposed to Cantonese as a child and later experienced a short exposure to Mandarin while living in China. Additionally, participants who performed beyond two standard deviations of the accuracy on the control condition were excluded from the final analysis. This included one Korean speaker and one Mandarin speaker. Their data was also excluded from the RT analysis as well. The final number of participants in each group is as follows: total (n = 112) Thai (n = 9, female = 8), Mandarin (n = 30, female = 24), Japanese (n = 23, female = 11), English (n = 23, female = 14) and Korean (n = 27, female = 18).

A few Korean speakers noted in the questionnaire that they had studied some Chinese/Mandarin or Japanese, but they did not have substantial exposure to either language. That is, they had not lived in those countries nor did they rate their listening skills highly. Additionally, the scores of these individuals on the ABX varied: a few were in the higher range among the Korean speaker group and many more were in the mid or lower range. Hence, they were retained for analysis.

Reaction times (RT) that were 300ms and below and any over 2500ms were removed. In total, only 0.00014% of the RT datapoints were removed (N = 3). The RT data were examined and were found to be skewed. Hence, a log-transformation was applied to the RT (Log-RT) to obtain a normal distribution.

After cleaning the data, language groups were compared in terms of accuracy and reaction times on their overall performance on the tone (test) condition versus the segmental (control) condition. Additionally, performance for the ten individual tonal comparison types was also analyzed for accuracy and reaction times.

Both the Log RTs and the unconverted RTs in ms are reported with their lower and upper confidence intervals (CI) in all the following tables for results in this chapter. For accuracy data, since the data structure is categorical (1 vs. 0 for correct vs. incorrect answer), a Generalized Estimating Equations (GEE) model for a binary response was fitted to the data, declaring subjects as a random factor. For RT data which are continuous, a Linear Mixed Effects model was fitted to the data. For most analyses, Language (Thai, Mandarin, Japanese, English, Korean) and Condition (test vs control) or Subcondition (the 10 comparisons) are declared as fixed factors, and Subjects are declared as a random factor. Sidak correction for multiple comparisons was used over Bonferroni as it is a less conservative method to calculate significance when conducting multiple comparisons; otherwise, considering the multiple comparisons needed over several groups and many tonal comparisons, the chances of revealing true significance would be greatly diminished and therefore, not accurately reflect the true situation.

4.2.2.1 Analysis by condition: Test and control

For each group, a mean accuracy score and mean reaction times (RT) were computed for each condition and sub-condition (i.e., individual tonal comparisons). Table 4.2 below presents the mean accuracy rates and RTs for the test and control conditions in each group. Mean accuracy rates and RTs for the subconditions are presented in Table 4.4 and Table 4.6, respectively.

Raw scores for accuracy rates on the tone test condition of the monosyllabic ABX task in Table 4.2 and Figure 4.2 show that the Thai group perform the most accurate (M = 91.3%) and the Korean group perform the least accurate (M = 79.3%). The three language groups for Mandarin, Japanese and English perform at "in-between" scores (M = 85%, M = 82.1%, M = 79.5%, respectively). In contrast, accuracy scores for the segment control condition range from the Thai group at the highest (M = 96%) to the English group at the lowest (M = 93.5%).

In Table 4.2 and Figure 4.3, we see that for RTs on the tone test condition, the Japanese are the fastest (M = 951 ms) while the English group is the slowest (M = 1170 ms). The three groups of Korean, Thai and Mandarin perform at in-between rates (M = 1000 ms, M = 1059 ms, M = 1081 ms, respectively). As for the RTs on the segment control condition, the Japanese (M = 824 ms) and Korean (M = 851 ms) are the fastest with the English (M = 991 ms) the slowest.

Table 4.2: Mean accuracy (%), mean RT (ms) and mean log RT for each group on test and control conditions

	Accuracy (%)		RT (ms)		Log RT	
Language group	Test (SE)	Control (SE)	Test (CI)	Control (CI)	Test (CI)	Control (CI)
Thai (n=9)	91.3 (2.5)	96.0 (1.4)	1059 (934; 1199)	938 (826; 1062)	3.03 (2.97; 3.08)	2.97 (2.92; 3.03)
Mandarin (n=30)	85.0 (1.6)	93.6 (1.0)	1081 (1034; 1131)	946 (901; 992)	3.03 (3.01; 3.05)	2.98 (2.95; 3.0)
Japanese (n=23)	82.1 (2.0)	94.5 (0.9)	951 (897; 1007)	824 (774; 876)	2.98 (2.95; 3.0)	2.92 (2.89; 2.94)
English (n=23)	79.5 (2.6)	93.5 (1.4)	1170 (1108; 1237)	991 (923; 1063)	3.07 (3.04; 3.09)	3.0 (2.97; 3.03)
Korean (n=27)	79.3 (1.7)	93.8 (1.0)	1000 (935; 1068)	851 (794; 913)	3.0 (2.97; 3.03)	2.93 (2.9; 2.96)

Note: SE = standard error. CI = confidence interval (lower; upper).

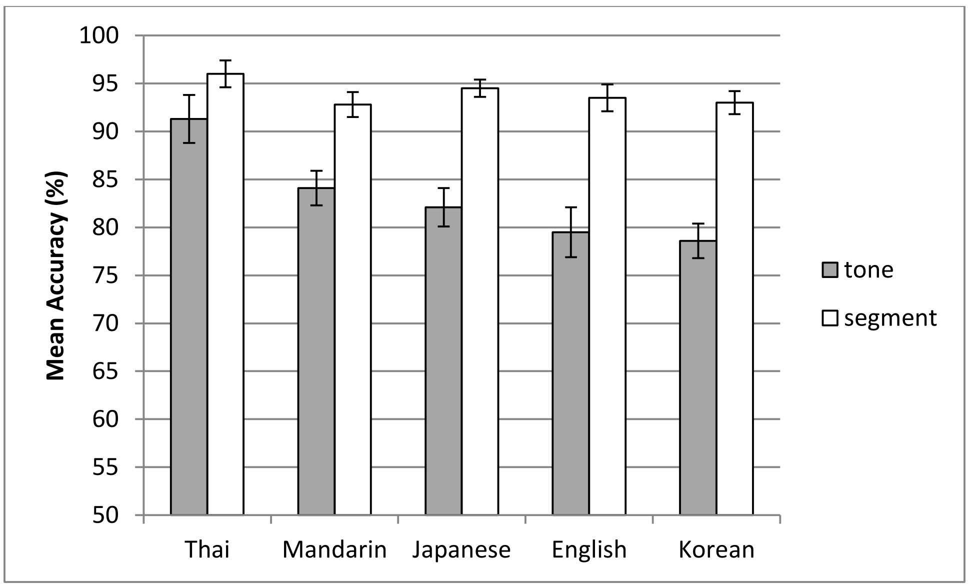


Figure 4.2: Mean accuracy rate (%) for each language group on test and control conditions. Error bars represent +/-1 SE.

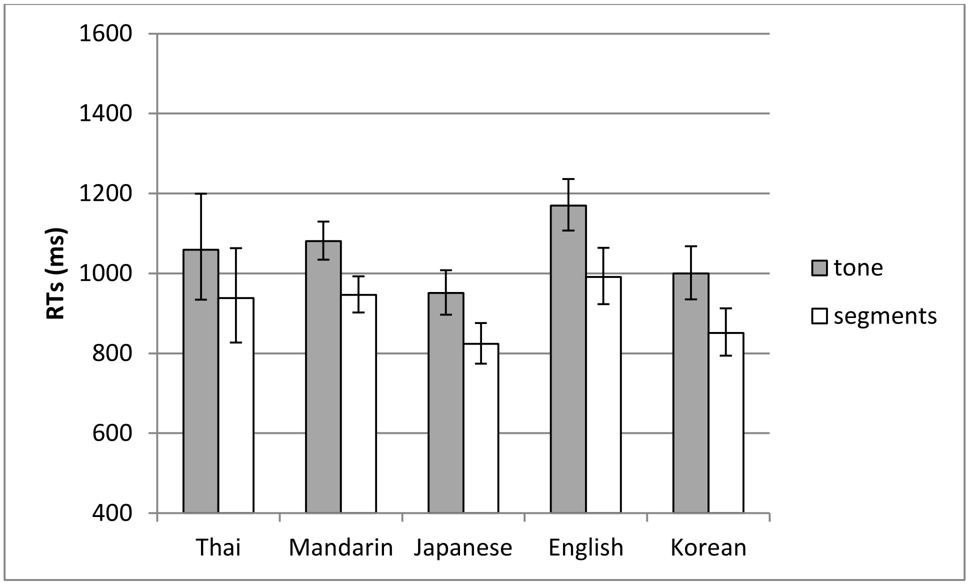


Figure 4.3: Mean reaction times (ms) for each language group on test and control conditions. Error bars represent the upper and lower CI from RTs from a Linear Mixed Effects model.

The un-aggregated data were used for statistical analysis. Accuracy data will be analyzed first, RT analysis will follow.

Accuracy Data

To examine the accuracy data, a GEE model was fitted with the fixed factors Language (Thai, Mandarin, Japanese, English, Korean) and Condition (test vs control), with Subjects as a random factor. The Type III tests of fixed effects showed no main effect of Language (χ^2 (4) = 5.34, p = 0.25), a significant effect of Condition (χ^2 (1) = 268.7, p < 0.001), and a significant interaction between the two factors (χ^2 (4) = 9.85, p = 0.043).

The interaction shows that while groups do not differ on the Control condition, their performance varied on the Test condition (See Table 4.5). There were no statistical differences for the control condition, i.e., perception of segmentals. However, overall test results revealed that performance differed significantly between groups on the test condition (χ^2(4) = 19.24, p = 0.001). On the test condition for tones, the native Thai speakers performed significantly better than the English (p = 0.011), Korean (p < 0.001) and Japanese (p = 0.038). The Thai group did not perform statistically better than the Mandarin group ($p = 0.278$). In turn the Mandarin group did not perform statistically better than the other language groups.

RT Data

For RTs, a Linear Mixed Effects model was run on the LogRTs with fixed factors as Language (Thai, Mandarin, Japanese, English, Korean) and Condition (test vs control), declaring Subjects as random factors. When looking at the Type III tests of fixed effects, the F-tests showed a main effect of Condition ($F(1, 15550) = 867.4$, $p < 0.001$), a significant effect of Language ($F(4, 106) = 6.1$, $p < 0.001$), and a significant interaction between the two factors ($F(4, 15550) = 2.67$, $p = 0.030$).

In this case, examining the interaction shows that groups differed on the test condition and the control condition (See Table 4.11). Univariate tests revealed that performance differed significantly between groups on both the test condition ($F(4, 115.07) = 6.0$, $p < 0.001$) and control condition ($F(4, 113.7) = 5.32$, $p = 0.001$). On the tone test condition, the Japanese were significantly faster than the English group ($p < 0.001$) and the Mandarin speakers ($p = 0.006$). Additionally, the Koreans were significantly faster than the English group as well ($p = 0.003$). Reflective of the RTs on the test tone condition, on the control segmental condition, the Japanese were significantly faster than the English ($p = 0.001$) and Mandarin speakers ($p = 0.005$) as well. The Korean group was statistically faster than the English group ($p = 0.028$).

Analysis by Korean Subgroup

Analyses on the Korean group were run to examine the possible differences between a group of speakers of standard Korean and a group of speakers of the Kyungsang Korean dialect which features pitch accent. In Chapter 3 on the pilot study, it was noted that Korean speakers performed at equal accuracy rates to the English speakers in their naïve perception of Thai tones on an ABX task. This was unexpected as it was predicted that the Korean group would perform less accurately than the English group due to the lower prominence of lexically-contrastive pitch in Korean as compared to English. In order to understand the possible cause behind this result, individual results for the Korean participants were examined. The speakers of the Kyungsang Korean dialect which features lexically-contrastive pitch performed more accurately than the speakers of Standard Korean. When the data for the dialect speakers was removed, the Standard Korean group performed less accurately than the English group as predicted while the Kyungsang Korean group performed more accurately than the English group and in fact at the level of the Japanese group. In the pilot experiment, this difference in dialect explained the equal performance in accuracy between the English group and the Korean group as a whole.

Hence, for the current study the Korean group was divided into the two groups of Standard Korean and Kyungsang Korean to examine whether the findings in the pilot study apply to the current study. Overall accuracy rates and reaction times are very similar in the two dialect groups. Table 4.3 presents the Korean group split by L1 dialects (grey-highlighted cells).

Table 4.3: Mean accuracy (%) and mean RT (ms) for each group and Korean dialect subgroups on test and control conditions

	Accuracy		RT		Log RT	
Language group	Test (SE)	Control (SE)	Test (CI)	Control (CI)	Test (CI)	Control (CI)
Thai (n=9)	91.3 (2.5)	96.0 (1.4)	1059 (934; 1199)	938 (826; 1062)	3.03 (2.97; 3.08)	2.97 (2.92; 3.03)
Mandarin (n=30)	85.0 (1.6)	93.6 (1.0)	1081 (1034; 1131)	946 (901; 992)	3.03 (3.01; 3.05)	2.98 (2.95; 3.0)
Japanese (n=23)	82.1 (2.0)	94.5 (0.9)	951 (897; 1007)	824 (774; 876)	2.98 (2.95; 3.0)	2.92 (2.89; 2.94)
English (n=23)	79.5 (2.6)	93.5 (1.4)	1170 (1108; 1237)	991 (923; 1063)	3.07 (3.04; 3.09)	3.0 (2.97; 3.03)
Korean (n=27) (overall)	79.3 (1.7)	93.8 (1.0)	1000 (935; 1068)	851 (794; 913)	3.0 (2.97; 3.03)	2.93 (2.9; 2.96)
Kyungsang Korean (n=10)	79.8 (2.7)	92.9 (1.5)	1016 (922; 1088)	927 (746; 914)	3.01 (2.96; 3.04)	2.97 (2.87; 2.96)
Seoul Korean (n=17)	78.9 (2.3)	94.4 (1.3)	997 (906; 1097)	869 (792; 954)	3.0 (2.96; 3.04)	2.94 (2.90; 2.98)

Note: SE = standard error. CI = confidence interval (lower; upper).

The two groups were divided and examined against each other only in the same type of analyses as was conducted on the language groups above.

To examine the accuracy data, a GEE model restricted to these two Korean groups was fitted with the fixed factors Language (i.e., Dialect: standard Korean, Kyungsang Korean) and Condition (test vs control), with Subjects as a random factor. The Type III tests of fixed effects showed no main effect of Dialect ($\chi^2(1) = 0.20$, $p = 0.658$), a significant effect of Condition ($\chi^2(1) = 91.90$, $p < 0.001$), and no significant interaction between the two factors ($\chi^2(1) = 1.23$, $p = 0.268$). There was no significant difference between the two dialects on the accuracy rates of the test ($p = 0.814$) and control conditions ($p = 0.424$) on the ABX monosyllabic task. However, there were significant within-group differences between the performance on the test and control conditions. For RTs, a Linear Mixed Effects model was run on the LogRTs with Language (i.e., Dialect: standard Korean, Kyungsang Korean) and Condition (test vs control) as fixed factors and with Subjects as random factors. When looking at the Type III tests of fixed effects, the F-tests showed a main effect of Condition ($F(1, 3432) = 298.5$, $p < 0.001$), no significant effect of Language ($F(1, 23) = 0.111$, $p = 0.742$), and a significant interaction between the two factors ($F(1, 3432) = 8.69$, $p = 0.003$). There were no significant differences between the RTs of the two Korean dialects on the test ($p = 0.944$) and control conditions ($p = 0.461$). However, Univariate tests revealed that performance did not differ significantly between the test and control conditions within group (For Kyungsang Korean group: $F(1, 3715.2) = 169.37$, $p < 0.001$; for Standard Korean group: $F(1, 3715.5) = 150.1$, $p < 0.001$). Therefore, interaction between the two factors of Language and Condition may be attributed to the difference in mean RTs between the two dialect groups on the control condition (Kyungsang dialect: M = 927 ms vs Seoul Korean dialect: M = 869 ms) as opposed to the much smaller difference between the mean RTs on the test condition (Kyungsang dialect: M = 1016 ms vs Seoul Kyungsang dialect: M = 997 ms).

In sum, separating the two Korean dialect groups did not show any significant difference in accuracy (p = 0.814) or reaction times (p = 0.944) on the test condition between these two dialectal groups, unlike the results described in Chapter 3 (see section 3.5).

4.2.2.2 Analysis by subcondition

We now turn to the comparison of individual tones, for both accuracy and RTs. We examine the individual tonal comparisons to determine whether the perception of one tone type may be more difficult than perceiving another tone type by analyzing the comparisons of the various tone types. Similar behavior among all languages to various tone types may suggest a universal tendency inherent to that specific tone type despite the differences in the prominence of lexically-contrastive pitch in the native language. On the other hand, differences among languages in behavior may indicate language-specific tendencies due to the prominence of lexically-contrastive pitch in the native language. (Table 4.4 presents the individual accuracy means, and Table 4.6 presents the RT means for the test and control items in each subcondition for each group). Mean scores on accuracy rates and RTs for all the test tone and control segment subconditions across the language groups are as follows. Mean accuracy rates across all language groups for the control segment subconditions range from a low of 89.9% by the Korean group on the control subcondition using a rising tone on all three items (i.e., A, X, B) to a high of 97.9% by the Thai group on the control subcondition using a high tone on all three items. On all the segment control subconditions, accuracy rates were relatively similar across groups.

In comparison, the test tone subconditions showed greater variety: the accuracy rates varied from 62.5% by the Korean group on the Low vs Rising comparison to a high of 97.2% by the Thai group on the Mid vs Rising comparison. The overall trend for the tone test subconditions across all groups was as follows: Thai group performed at the highest accuracy rates for almost all the tone test subconditions while the English and Korean groups performed with the least accuracy for almost all the tone test subconditions with the Koreans generally less accurate than the English group.

Accuracy Data

Looking at the accuracy rates for each tone subcondition for each language group (See Table 4.6 and Figures 4.4a-4.4e), we are able to discern two possible trends. First, perceptual accuracy appears to be influenced by the type of the two tones compared. For example, “direction” tone comparisons (i.e., comparing two direction tones) appear to be the easiest (e.g., Falling vs Rising) while “height” tone comparisons appear to be the most difficult (e.g., Low vs Mid). The “mixed” tone comparisons comparing tone height (e.g., low tone, mid tone) with tone direction (e.g., falling, rising) appear to fall between the “direction” and “height” tone comparisons in terms of accuracy rates. In contrast, accuracy rates on the control segment subconditions do not show any hierarchical trend: accuracy rates on the control segment subconditions are fairly flat across all subconditions and languages (Figure 4.6). Second, certain tone comparisons were less accurate universally. For example, the Low vs Mid and Low vs High tone comparisons were the least accurate for the Thai (*M* = 87.5%, *M* = 88.9%, respectively). The other groups performed less accurately on these two comparisons as well, but not necessarily as their least accurate tone comparison score. However, on some tone comparisons such as the Low vs Rising tone the Thai performed at a fairly high rate (*M* = 90.3%) while on the same tone comparison the other groups performed at one of their lowest rates of accuracy.

Table 4.4: Mean accuracy rates for each group on each tonal comparison

Test subconditions	Mean accuracy (SE)				
	T	M	J	E	K
F vs R	91.7 (3.4)	94.2 (1.4)	90.8 (2.6)	90.6 (2.6)	87.0 (2.7)
L vs F	91.7 (3.4)	90.4 (2.2)	92.9 (2.0)	86.5 (3.7)	88.5 (2.3)
L vs R	90.3 (3.8)	75.0 (3.2)	67.9 (3.4)	64.6 (4.3)	63.0 (4.2)
M vs F	91.7 (3.9)	86.3 (2.5)	90.2 (2.0)	83.9 (3.4)	82.5 (3.2)
M vs R	97.2 (2.6)	92.5 (1.7)	88.6 (3.0)	84.9 (3.0)	86.0 (2.0)
H vs F	93.1 (2.1)	92.1 (2.1)	84.8 (3.2)	80.7 (4.1)	82.5 (3.5)
H vs R	90.3 (3.8)	85.0 (2.9)	80.4 (3.7)	77.6 (4.2)	80.0 (3.3)
L vs M	87.5 (3.4)	71.7 (3.6)	75.5 (3.7)	72.4 (4.4)	69.0 (3.9)
L vs H	88.9 (4.6)	84.6 (2.8)	85.3 (2.9)	79.7 (4.4)	79.5 (3.7)
M vs H	90.3 (3.8)	77.9(3.2)	64.1 (3.8)	74.5 (3.2)	74.5 (3.0)
Control subconditions					
F	95.1 (1.3)	94.2 (1.4)	94.6 (1.3)	93.5 (1.7)	96.5 (0.9)
R	96.5 (2.2)	92.3 (1.4)	93.2 (1.4)	93.5 (1.5)	91.0 (1.5)
L	94.4 (2.5)	93.3 (1.3)	95.1 (0.9)	93.2 (1.6)	93.3 (1.2)
M	95.8 (1.7)	94.4 (1.0)	94.8 (1.2)	92.7 (2.0)	96.0 (1.1)
H	97.9 (1.0)	94.0 (1.4)	94.8 (1.4)	94.8 (1.4)	92.2 (1.7)

Note: SE = standard error.

T = Thai; M = Mandarin, J = Japanese, E = English, K = Korean; L = low tone, M = mid tone, H = high tone, F = falling tone, R = rising tone.

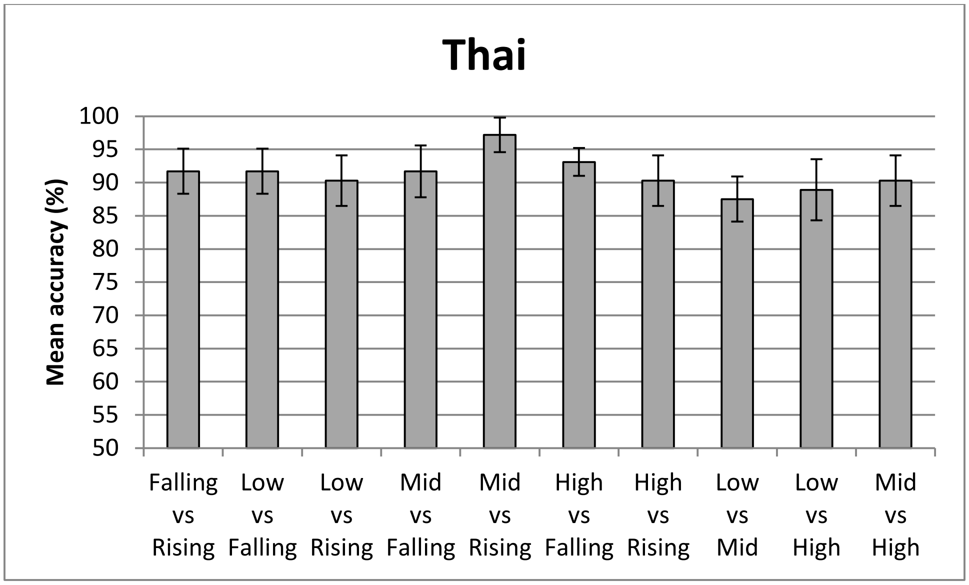


Figure 4.4a: Accuracy rates (%) for the Thai group on each test subcondition. Error bars represent +/- 1 SE.

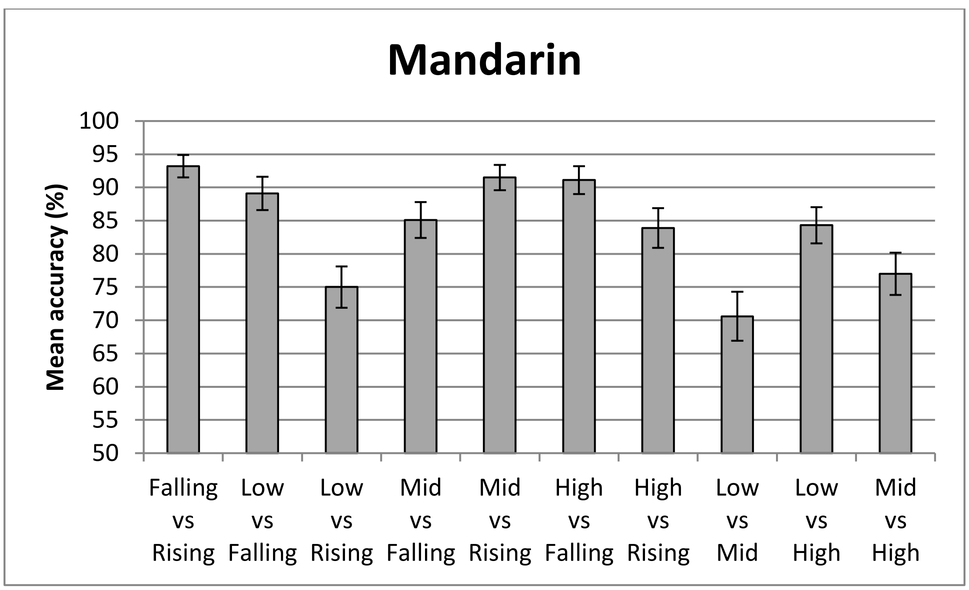


Figure 4.4b: Accuracy rates (%) for the Mandarin group on each test subcondition. Error bars represent +/- 1 SE.

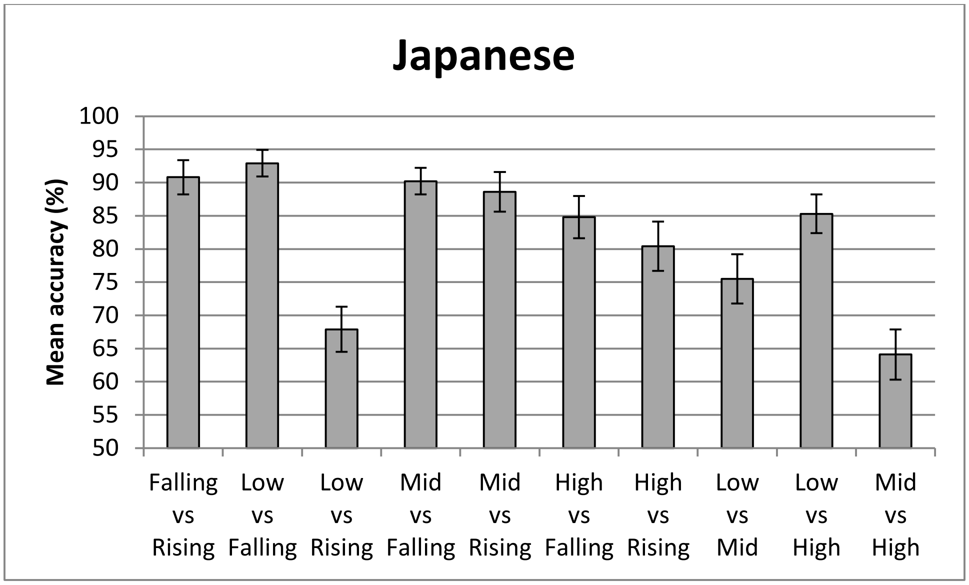


Figure 4.4c: Accuracy rates (%) for the Japanese group on each test subcondition. Error bars represent +/- 1 SE.

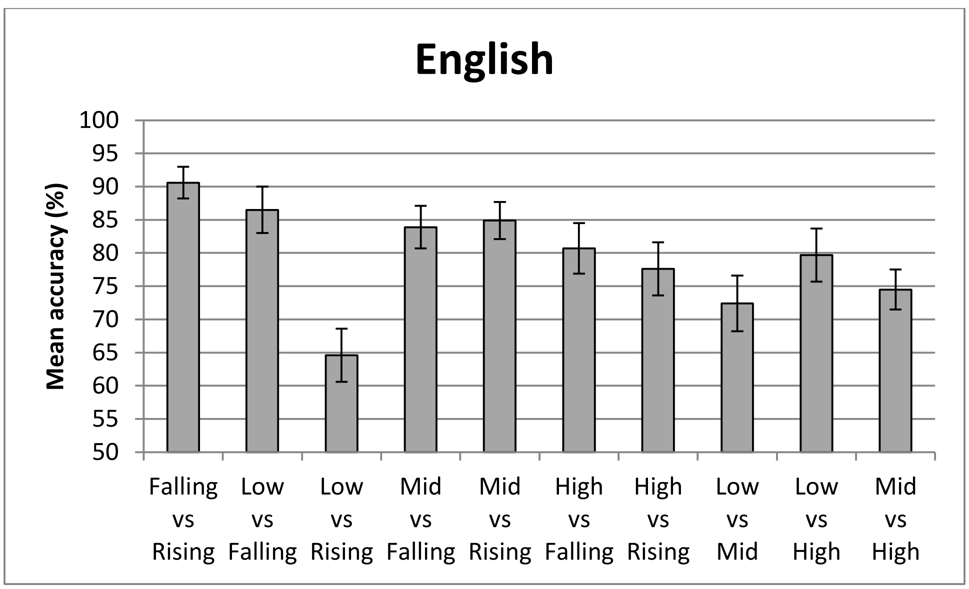


Figure 4.4d: Accuracy rates (%) for the English group on each test subcondition. Error bars represent +/- 1 SE.

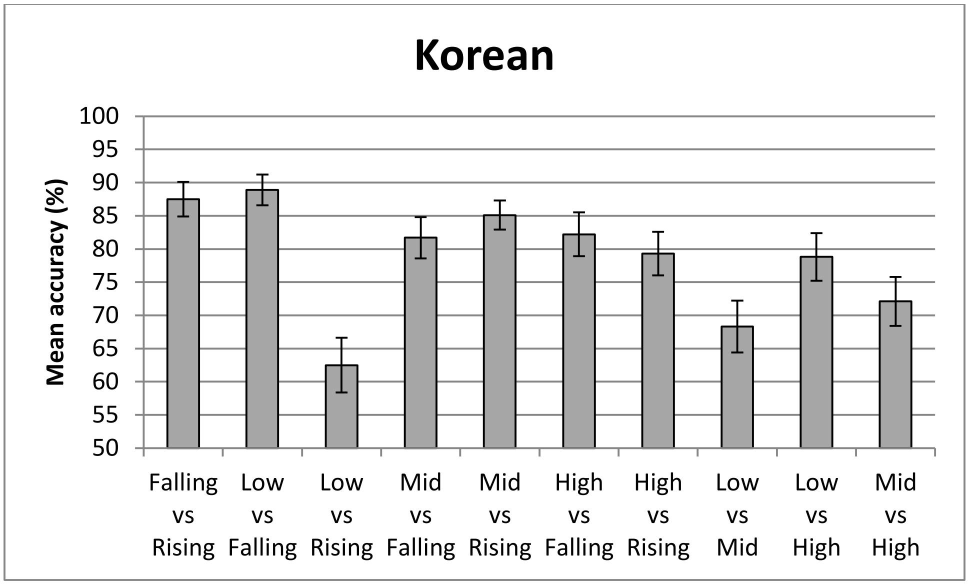


Figure 4.4e: Accuracy rates (%) for the Korean group on each test subcondition. Error bars represent +/- 1 SE.

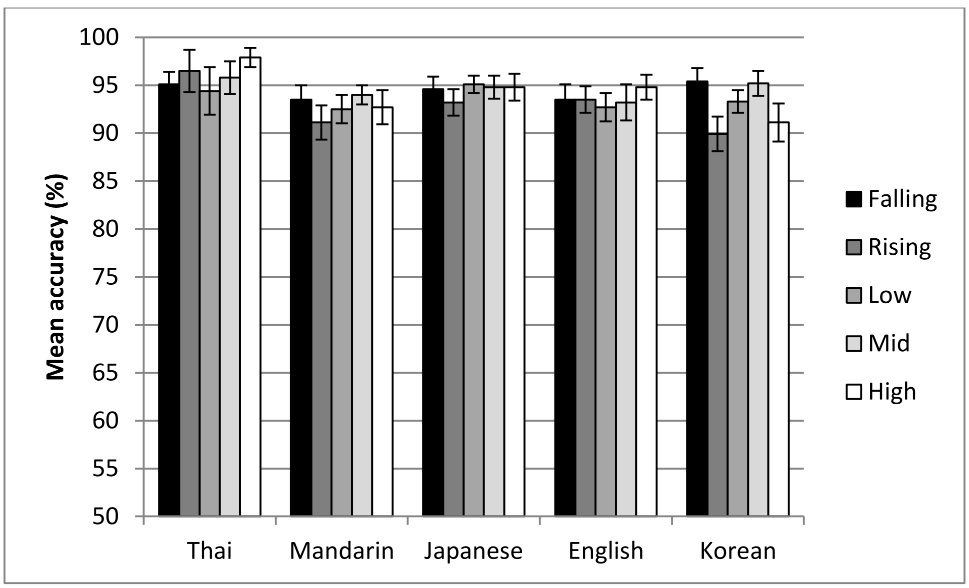


Figure 4.5: Accuracy rates (%) for each group on each control subcondition. Error bars represent +/- 1 SE.

To examine the accuracy data on individual tone comparisons, a GEE model was fitted with the fixed factors Language (Thai, Mandarin, Japanese, English, Korean) and Subcondition (individual tonal comparisons for both the tone and segment subconditions), with Subjects as a random factor. Examining individual tonal comparisons, the Type III tests of fixed effects showed a marginal effect of Language ($\chi^2(4) = 7.95$, $p = 0.094$), a significant effect of subcondition ($\chi^2(14)$

= 689.56, $p < 0.001$), and a significant interaction between the two factors ($\chi^2(56) = 262.69$, $p < 0.001$).

The interaction suggests that the performance of the groups varies on the different subconditions (See Table 4.5). The Thai group performed significantly better than the other groups on four of the ten individual tone comparisons: Low vs Mid, L vs Rising, Mid vs High, and Mid vs Rising. The Thai group performed significantly better than all the groups on the Low tone vs Rising tone comparison ($p < 0.001$ for English, Korean and Japanese; $p = 0.022$ for Mandarin). The Thai group performed significantly better than the English ($p = 0.014$), L1 Korean ($p = 0.012$) and Japanese ($p < 0.001$) groups on the Mid tone vs High tone comparison. The Thai group performed significantly better than the English ($p = 0.021$) and Korean ($p = 0.007$) groups on the Mid tone vs Rising tone comparison. The Thai group performed more accurately than the Korean ($p = 0.004$) and Mandarin ($p = 0.014$) groups on the Low vs Mid tone comparison. However, overall test results revealed that performance differed significantly between groups on the some of the test subconditions ($\chi^2(4) = 11.622{\sim}32.555$, $p < 0.05$), i.e., High vs Falling ($p = 0.004$), Low vs Mid ($p = 0.002$), Low vs Rising ($p < 0.001$), M vs High ($p < 0.001$), M vs Rising ($p = 0.002$) and the segment condition bearing the High tone ($p = 0.02$).

In Table 4.5 we can see on which subconditions the groups' accuracy scores are statistically different. Among the control subconditions, the Thai group is more accurate than the Korean group on the segment trials using the High tone. Otherwise, there is no significant difference in accuracy on the control subconditions between the language groups.

Table 4.5: Statistically significant accuracy rates on each tonal comparison (i.e., subcondition) for each group

Comparison	Accuracy
Tone overall	Thai > English (*p* = 0.011) Thai > Japanese (*p* = 0.038) Thai > Korean (*p* < 0.001)
FR	--
LF	--
LR	Thai > English (*p* < 0.001) Thai > Korean (*p* < 0.001) Thai > Japanese (p < 0.001) Thai > Mandarin (*p* = 0.022)
MF	--
MR	Thai > English (*p* = 0.021) Thai > Korean (*p* = 0.007)
HF	--
HR	--
LM	Thai > Korean (*p* = 0.004) Thai > Mandarin (*p* = 0.014)
LH	--
MH	Thai > English (*p* = 0.014) Thai > Japanese (*p* < 0.001) Thai > Korean (*p* = 0.012)
Segment overall	--
F (control)	--
R (control)	--
L (control)	--
M (control)	--
H (control)	Thai > Korean (*p* = 0.033)

RT Data

Moving onto the examination of RTs across language groups, we see in Table 4.6 that all but one RT were below 1,000 ms on the segment control subconditions (fastest RT: *M* = 815 ms) while the majority of RTs on the tone test subconditions were above 1,000 ms (slowest RT: *M* = 1,373 ms). In general, the Japanese and Korean groups were the fastest on both the test tone and control segment subconditions. The English group was generally the slowest followed by the Mandarin and Thai groups. Across all groups, the RTs for the low vs mid, mid vs high and low vs rising comparisons tend to be universally slower while the falling vs rising and low vs falling comparisons were generally the fastest for most groups. There appear to be no language-specific tone comparisons where some groups were slower or faster than other groups.

Looking at the RTs for each tone subcondition for each language group (Figures 4.8a-4.8e), we are able to discern a possible trend: perception appears to be influenced by the type of the two tones compared. For example, RTs for "direction" tone comparisons comparing direction appear to be the fastest (i.e., Falling vs Rising) while RTs for "height" tone comparisons appear to be the slowest (i.e., Low vs Mid). The RTs for "mixed" tone comparisons comparing tone height (e.g., low tone, mid tone) with tone direction (e.g., falling, rising) appear to fall between the "direction" and "height" tone comparisons in terms of RTs. In contrast, RTs on the control segment subconditions do not show any hierarchical trend: RTs on the control segment subconditions are fairly flat across both all subconditions and languages (Figure 4.7).

Table 4.6: Mean RTs and logged RTs for each group on each tonal comparison

Test subconditions	RT (CI)					Log RT (CI)				
	T	M	J	E	K	T	M	J	E	K
F vs R	983 (828;1168)	948 (896; 1003)	867 (810; 929)	1064 (1007; 1124	872 (815; 934)	2.99 (2.92; 3.07)	2.98 (2.95; 3.0)	2.94 (2.91; 2.97)	3.03 (3.0; 3.05)	2.94 (2.91; 2.97)
L vs F	1060 (894; 1257)	989 (931; 1050)	860 (798; 927)	1071 (994; 1155)	892 (827; 962)	3.03 (2.95; 3.1)	3.0 (2.97; 3.02)	2.93 (2.9; 2.97)	3.03 (3.0; 3.06)	2.95 (2.92; 2.98)
L vs R	1156 (997; 1341)	1169 (1120; 1222)	1058 (979; 1144)	1303 (1212; 1402)	1150 (1059; 1249)	3.06 (3.0; 3.13)	3.07 (3.05; 3.09)	3.02 (2.99; 3.06)	3.12 (3.08; 3.15)	3.06 (3.02; 3.1)
M vs F	1019 (878; 1183)	1091 (1021; 1167)	919 (852; 991)	1146 (1062; 1237)	971 (896; 1052)	3.01 (2.94; 3.07)	3.04 (3.01; 3.07)	2.96 (2.93; 3.0)	3.06 (3.03; 3.09)	2.99 (2.95; 3.02)
M vs R	1044 (919; 1186)	1121 (1060; 1186)	992 (936; 1051)	1211 (1126; 1302)	983 (922; 1049)	3.02 (2.96; 3.07)	3.05 (3.03; 3.07)	3.0 (2.97; 3.02)	3.08 (3.05; 3.11)	2.99 (2.95; 3.03)
H vs F	1030 (909; 1166)	1050 (986; 1119)	890 (836; 949)	1102 (1034; 1174)	981 (895; 1075)	3.01 (2.96; 3.07)	3.02 (2.99; 3.05)	2.95 (2.92; 2.98)	3.04 (3.01; 3.07)	2.99 (2.96; 3.03)
H vs R	1010 (894; 1141)	1046 (988; 1108)	947 (880; 1018)	1139 (1059; 1225)	983 (910; 1061)	3.0 (2.95; 3.06)	3.02 (2.99; 3.04)	2.98 (2.94; 3.01)	3.06 (3.02; 3.09)	2.99 (2.96; 3.03)
L vs M	1070 (936; 1223)	1203 (1131;1279)	1004 (947; 1064)	1254 (1168; 1348)	1047 (964; 1138)	3.03 (2.97; 3.09)	3.08 (3.05; 3.11)	3.0 (2.98; 3.03)	3.1 (3.07; 3.13)	3.02 (2.98; 3.04)
L vs H	1144 (1008; 1300)	1100 (1045; 1158)	1000 (922; 1085)	1146 (1072; 1225)	1028 (960; 1102)	3.06 (3.0; 3.11)	3.04 (3.02; 3.06)	3.0 (2.96; 3.04)	3.06 (3.03; 3.09)	3.01 (2.98; 3.04)
M vs H	1087 (983; 1203)	1180 (1124; 1240)	1063 (995; 1135)	1373 (1295; 1455)	1205 (1131; 1284)	3.04 (2.99; 3.08)	3.07 (3.05; 3.09)	3.03 (3.0; 3.06)	3.14 (3.11; 3.16)	3.08 (3.05; 3.11)
Control subconditions										
F	924 (804; 1062)	950 (897; 1005)	816 (763; 872)	985 (909; 1068)	840 (785; 899)	2.97 (2.91; 3.03)	2.98 (2.95; 3.0)	2.91 (2.88; 2.94)	2.99 (2.96; 3.03)	2.92 (2.9; 2.95)
R	928 (817; 1054)	939 (891; 990)	826 (776; 879)	999 (919; 1085)	850 (786; 919)	2.97 (2.91; 3.02)	2.97 (2.95; 3.0)	2.92 (2.89; 2.94)	3.0 (2.96; 3.04)	2.93 (2.9; 2.96)
L	933 (809; 1075)	938 (890; 989)	820 (768; 876)	963 (896; 1036)	854 (792; 920)	2.97 (2.91; 3.03)	2.97 (2.95; 3.0)	2.91 (2.89; 2.94)	2.98 (2.95; 3.02)	2.93 (2.9; 2.96)
M	958 (845; 1085)	952 (905; 1001)	841 (783; 903)	1029 (947; 1118)	873 (814; 935)	2.98 (2.93; 3.04)	2.98 (2.96; 3.0)	2.92 (2.89; 2.96)	3.01 (2.98; 3.05)	2.94 (2.91; 2.97)
H	942 (826; 1074)	950 (901; 1000)	815 (762; 872)	979 (922; 1039)	841 (781; 907)	2.97 (2.92; 3.03)	2.98 (2.95; 3.0)	2.91 (2.88; 2.94)	2.99 (2.96; 3.02)	2.93 (2.89; 2.96)

Note: CI = confidence interval (lower; upper).

T = Thai; M = Mandarin, J = Japanese, E = English, K = Korean; L = low tone, M = mid tone, H = high tone, F = falling tone, R = rising tone.

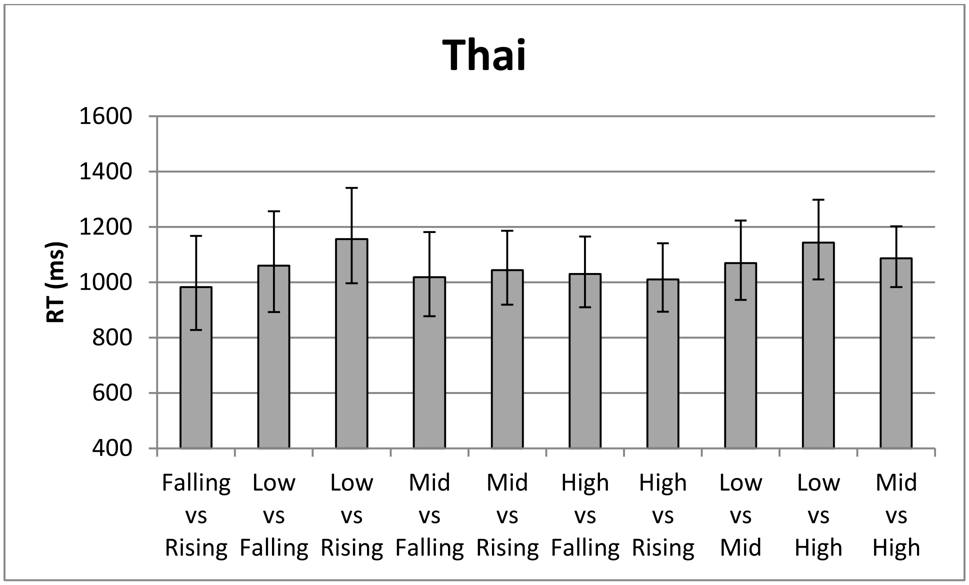


Figure 4.6a: Reaction times (ms) for the Thai group on each test subcondition. Error bars represent the upper and lower CI from RTs from a Linear Mixed Effects model.

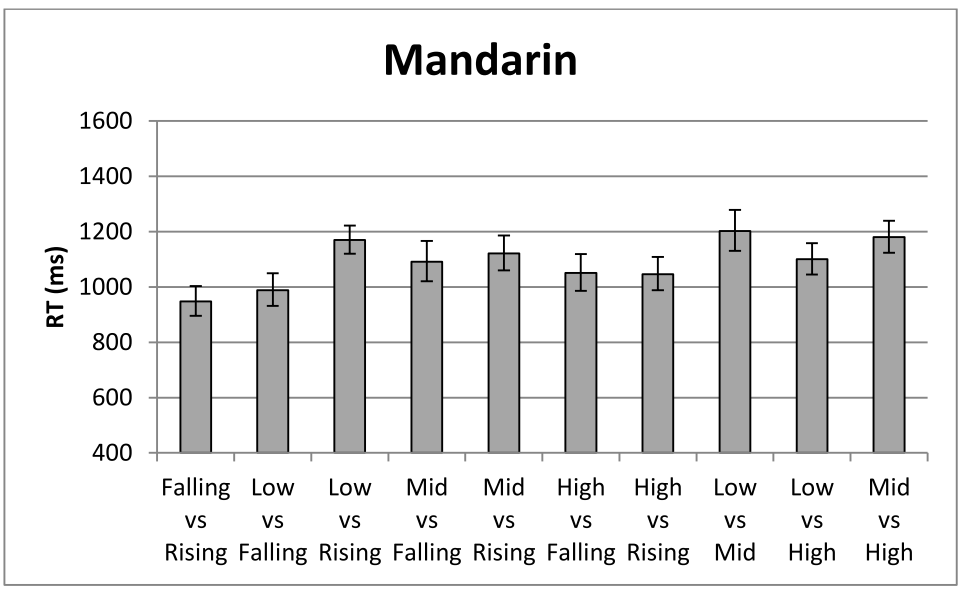


Figure 4.6b: Reaction times (ms) for the Mandarin group on each test subcondition. Error bars represent the upper and lower CI from RTs from a Linear Mixed Effects model.

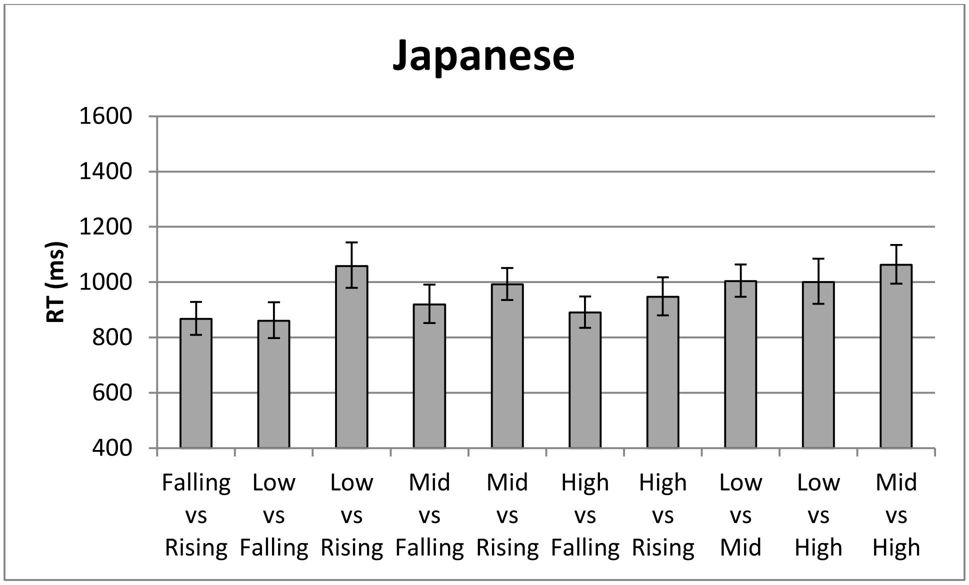


Figure 4.6c: Reaction times (ms) for the Japanese group on each test subcondition. Error bars represent the upper and lower CI from RTs from a Linear Mixed Effects model.

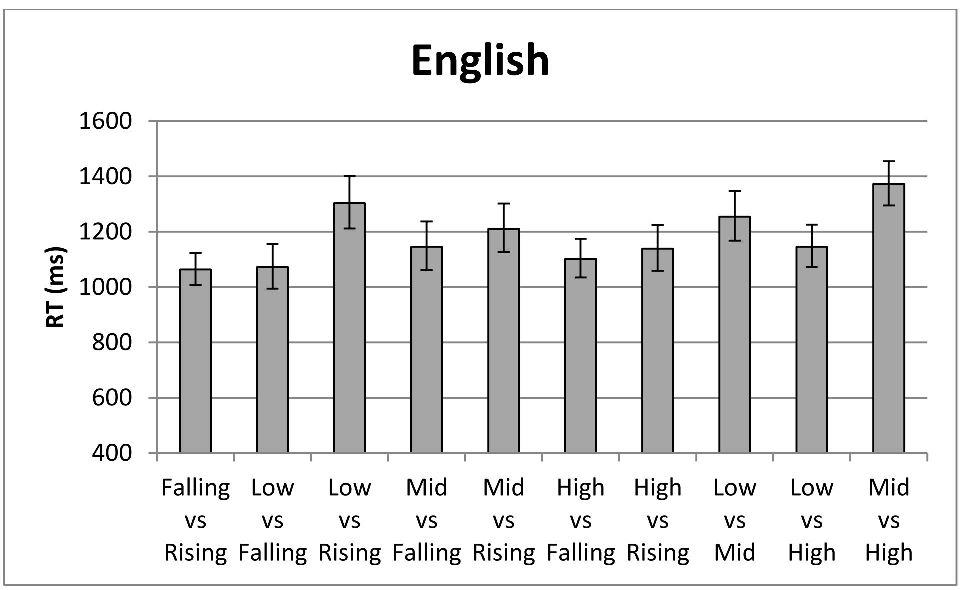


Figure 4.6d: Reaction times (ms) for the English group on each test subcondition. Error bars represent the upper and lower CI from RTs from a Linear Mixed Effects model.

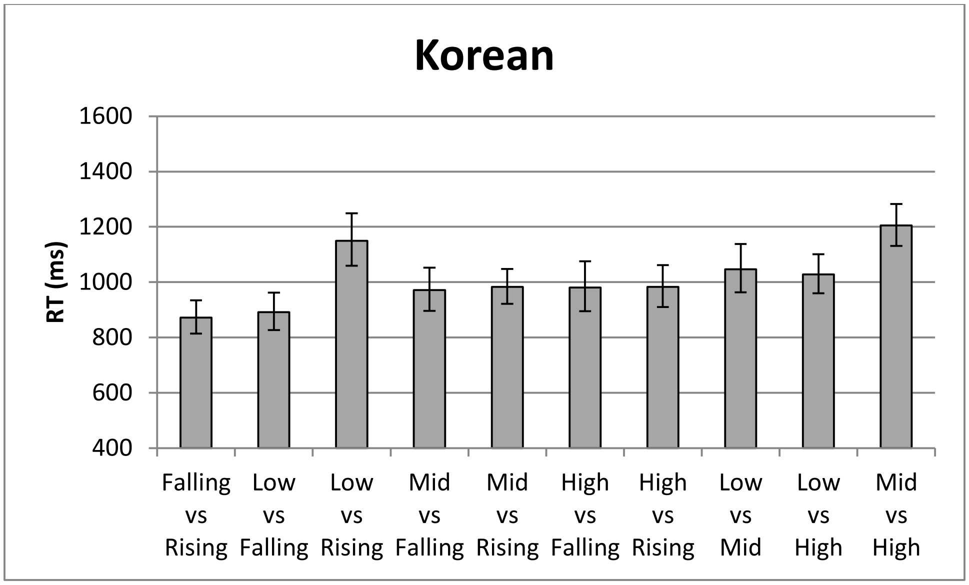


Figure 4.6e: Reaction times (ms) for the Korean group on each test subcondition. Error bars represent the upper and lower CI from RTs from a Linear Mixed Effects model.

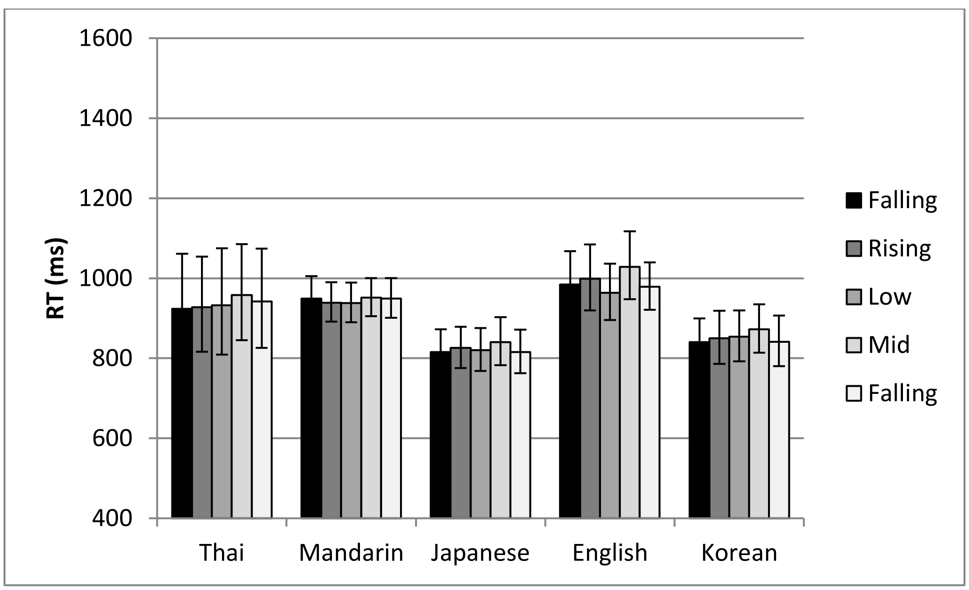


Figure 4.7: Reaction times (ms) for each group on each control condition. Error bars represent the upper and lower CI from RTs from a Linear Mixed Effects model.

A Linear Mixed Effects model was run on the LogRT with Language (Thai, Mandarin, Japanese, English, Korean) and Subcondition (individual tonal comparisons for both the tone and segment subconditions) as fixed factors and Subjects as a random factor. When looking at the Type III tests of fixed effects, the F-tests showed a main effect of Subcondition (F(14, 15485) = 93.166,

$p < 0.001$), a significant effect of Language ($F(4, 107) = 6.033$, $p < 0.001$), and a significant interaction between the two factors ($F(56, 15485) = 1.572$, $p = 0.004$).

The interaction again demonstrates the performance of the groups varies on both the Test and Control subconditions (See Table 4.11). As for the reaction times for the tonal comparisons, the Japanese and Korean groups were faster than the English and Mandarin groups on the whole. The Japanese group was significantly faster than the English group on every subcondition, whether test or control. The Korean group was significantly faster than the English group on half the test subconditions: the Falling vs Rising ($p = 0.001$), Low vs Falling ($p = 0.013$), Low vs Mid ($p < 0.001$), Mid vs High ($p = 0.039$) and Mid vs Rising ($p < 0.001$) comparisons. The Japanese were also faster than the Mandarin speakers on the High vs Falling ($p = 0.002$), High vs Falling ($p = 0.002$), Low vs Mid ($p < 0.001$), Low vs Falling ($p = 0.037$), Low vs Mid ($p < 0.001$) Mid vs Falling ($p = 0.005$), Mid vs Rising ($p = 0.019$). In comparison, the Japanese were significantly faster than the English and Mandarin groups on the all the control subconditions as well. Additionally, all the groups showed significant within-group differences on the RTs between all the subconditions except for the Low vs High comparison: ($F(4, 236.4) = 2.175$, $p = 0.072$).

In Table 4.7, we can see for which subconditions groups obtained significantly different RTs. The Mandarin speakers were faster than the English speakers on the Mid vs High comparison ($p = 0.003$). The Thai were faster than the English group on the Mid vs High comparison ($p = 0.001$). The Korean were significantly faster than the Mandarin on the Mid vs Rising subcondition ($p = 0.035$).

Table 4.7: Statistically significant differences between groups for LogRTs on each tonal comparison (i.e., subcondition)

Comparison	Reaction times
Tone overall	Japanese > English (p < 0.001) Korean > English (p = 0.003) Japanese > Mandarin (p = 0.006)
FR	Japanese > English (p < 0.001) Korean > English (p < 0.001) Mandarin > English (p = 0.41)
LF	Japanese > English (p = 0.001) Korean > English (p = 0.008) Japanese > Mandarin (p = 0.044)
LR	Japanese > English (p = 0.001)
MF	Japanese > English (p = 0.001) Korean > English (p = 0.033) Japanese > Mandarin (p = 0.008)
MR	Japanese > English (p < 0.001) Korean > English (p < 0.001) Japanese > Mandarin (p = 0.030) Korean > Mandarin (p = 0.026)
HF	Japanese > English (p < 0.001) Japanese > Mandarin (p = 0.003)
HR	Japanese > English (p = 0.005)
LM	Japanese > English (p < 0.001) Korean > English (p = 0.013) Japanese > Mandarin (p < 0.001)
LH	--
MH	Japanese > English (p < 0.001) Korean > English (p = 0.030) Mandarin > English (p = 0.001) Thai > English (p = 0.001)
Segment overall	Japanese > English (p = 0.001) Korean > English (p = 0.028) Japanese > Mandarin (p = 0.005)
F (control)	Japanese > English (p = 0.004) Korean > English (p = 0.030) Japanese > Mandarin (p = 0.007)
R (control)	Japanese > English (p = 0.003) Mandarin > Japanese (p = 0.019)
L (control)	Japanese > English (p = 0.013) Japanese > Mandarin (p = 0.017)
M (control)	Japanese > English (p = 0.003) Korean > English (p = 0.027) Japanese > Mandarin (p = 0.051)
H (control)	Japanese > English (p = 0.001) Korean > English (p = 0.021) Japanese > Mandarin (p = 0.004)

Analysis by Korean Subgroup

Again, the Korean group was divided into the two groups of standard Korean speakers and Kyungsang Korean speakers to scrutinize more closely the individual tone comparisons for any differences between the two groups. To examine the accuracy data for individual tone comparisons of the two Korean dialects, a GEE model was fitted with the fixed factors Language (i.e., Dialect: standard Korean, Kyungsang Korean) and Subcondition (individual tonal comparisons for both the tone and segment subconditions), with subjects as a random factor. There was no significant effect of Language between the two dialects ($\chi^2(1) = 0.112$, $p = 0.738$), a significant effect of Subcondition ($\chi^2(14) = 424.3$, $p < 0.001$), and a significant interaction between the two factors ($\chi^2(14) = 38.24$, $p < 0.001$). There were no significant differences on the accuracy rates of the individual tone comparisons between the two Korean dialects. Additionally, the Kyungsang Korean group showed almost no significant within-group differences on the accuracy rates between the test subconditions while the Seoul Korean groups showed many more.

For RTs of the two Korean dialects on the individual tone comparisons, a Linear Mixed Effects model was run on the LogRTs with Language (i.e., Dialects: standard Korean, Kyungsang Korean) and Condition ((individual tonal comparisons for both the tone and segment subconditions) as fixed factors and Subjects as random factors. When looking at the Type III tests of fixed effects, the F-tests showed a main effect of Subcondition ($F(14, 3406) = 36.0$, $p < 0.001$), no significant effect of Language ($F(1, 23) = 0.056$, $p = 0.815$), and no significant interaction between the two factors ($F(14, 3406) = 1.37$, $p = 0.155$). There were no significant differences between the RTs of the two Korean dialects on the individual tone comparisons of the ABX monosyllabic task. However, both groups showed significant within-group differences on the RTs between the test subconditions.

4.2.3 Discussion

Overall results demonstrate that different L1s do indeed result in different outcomes in the naïve perceptual accuracy of non-native tones. This is most likely due to the varied functionality of lexically-contrastive pitch in the various L1s. In contrast to the varied performance on the test condition, all groups performed at comparable levels on the control condition. Thus, we might assume that influence of intonation being relatively equal among the languages (i.e., all languages use intonation, and intonation falls under a different domain than lexically-contrastive pitch), it is indeed the varying functionality of lexically-contrastive pitch in the L1 that plays the deciding role in influencing the naïve perception of Thai tone to a corresponding degree. Furthermore, a hierarchy in performance emerges commensurate to the functionality of lexically-contrastive pitch in the L1.

A few unexpected results have emerged. First, the hierarchical prediction made based on findings from Chapter 3 does not bear out in a robust manner. Second, RTs do not reflect the accuracy results in the same hierarchical manner and seem to defy the predicted hierarchy based on the functionality of lexically-contrastive pitch. This section will address these two unexpected results along with those on the subconditions in terms of L2 phonology models and universal and language-specific trends.

Overall accuracy performance

Regarding the first question, possible explanations for the equal performance of the English and Korean groups are explored. Analyses first separated the Korean group into two Korean dialect groups due to possible influence of lexically-contrastive pitch in the Korean Kyungsang dialect as was previously suggested by the pilot study analyses (cf., Chapter 3 Section 3.5). However, in the current study analyses did not show any significant difference between these two groups Thus, pitch accent appears not to be sufficiently salient in Kyungsang Korean. Another possible reason behind this lack of difference may be that standard Korean somehow overrides the role of dialectal pitch accent (e.g., exposure in education, mass media or greater conformity to standard Korean), particularly among the younger generation. Another conceivable reason for the equal performance for these dialectal groups cannot discount the possibility of the emergence of lexically-contrastive pitch in standard Korean (cf., Silva, 2006) feasibly neutralizing the difference between standard Korean and Kyungsang Korean. We might also suspect the influence of word stress of the L2 English of many of the L1 Korean-speaking participants. In contrast, the English group may be performing poorly as pitch is not the most salient correlate of word stress, rendering the functionality of lexical pitch in English word stress equal to that of having no lexical pitch as in standard Korean. In sum, the statistically "equal" accuracy rates between the English and Korean groups in this study continue to defy the predicted hierarchy based on the functionality of the lexically-contrastive pitch in the L1 and therefore, need further analyses and/or testing.

A difference in cognitive load between the two different task designs of the pilot study and current study must be explored in order to account for the overall less clear differences between L1 groups found in the current experiment. The instrument (i.e., ABX task, stimuli) employed to test participants in the current experiment was perhaps less cognitively demanding, obscuring a possibly more nuanced picture of the perceptual abilities of the various L1 groups. For example, the statistically weaker results hinting at a performance hierarchy could be due to several design differences. Let's consider the stimuli in the two tasks of the pilot study and current study. In the pilot study, the four-paradigm trials (i.e., ABB, ABA, BAA, BAB) used different segments for each of the four parts of the paradigm while in the current task the same segments were used for each of the four parts. Thus, in the pilot study AXB task, we have [ba:] HHR, [pu:] RRH, [su:] HRR, and [t^ha:] RHH, for example, while in the current ABX task, we have [no:i] used for all four trials of the four-part paradigm. This may have caused two differences: greater awareness to listen for tonal differences and an inadvertent introduction of a "short training" practice. The fact that the current study used only two different segment strings (i.e., [no:i], [p^{h}uai]) for all the tonal comparisons may have caused the participants to focus on tone whenever the two types of segments were heard rather than having to expend attentional resources on determining whether vowels or consonants were different. Additionally, the participants may have also been able to be trained somewhat on tones by being aware that they were to listen for tones upon hearing these two types of segments. In short, the cognitive load may have been less in the current study than in the pilot study.

Overall reaction time results

Looking at reaction times for overall performance, we see a less clear correlation between the functionality of lexically-contrastive pitch in the L1 and a hierarchy of naïve perceptual accuracy of non-native tone. In fact, the reaction times on the tone test condition simultaneously both illuminate and confound interpretation of the overall results. To illustrate, predictably the Japanese were statistically faster than the English group and unexpectedly faster than the Mandarin group. We also see another unexpected trend: the L1 Korean group was faster than the L1 English group. Thus, we see that the RT data do not follow the predicted hierarchy of performance based upon the functionality of lexical pitch in the L1. This tendency for RTs was also witnessed in the pilot study in Chapter 3: RTs do not necessarily follow accuracy rates. This may be attributed to the possibility that as a response strategy certain groups favor speed over accuracy (e.g., Japanese group). Crucially, we see RTs on the control segmental condition reflect those on the test tone condition: The Japanese group was statistically faster than the English and Mandarin groups, and the Korean group was faster than the English group. This observation appears to point to an overall tendency in RTs among all groups, regardless of whether reacting to the test or control condition.

Therefore, the question is whether the Japanese performed faster or the Mandarin speakers performed slower or a combination of the two although performance on the control segmental condition appears to bias interpretation toward the first explanation, i.e., the Japanese prioritize speed over accuracy as a strategy. The case for the Korean group's faster speed over the English group may result from the same "speed-over-accuracy priority".

However, combining the accuracy results with the RT data teases apart the language groups and thus, corroborates a more detailed hierarchy of performance. That is, the English group would be at the lower end of the hierarchy and the Japanese and Korean groups in the middle along with the Mandarin group. If we prize accuracy over RT, then the Mandarin group would still occur higher in the hierarchy above the Japanese and Korean group. Thus we would see Thai, Mandarin > Japanese, Korean > English. A combination of accuracy and RT data, thus, supports a Feature Hypothesis explanation for overall performance by the various language groups as determined by the functionality of lexically-contrastive pitch in the L1. Results on the individual tone comparisons further provide a more nuanced understanding of the perception of non-native tone.

Performance on the subconditions

This experiment analyzes the perception of each of the five different individual Thai tones (i.e., individual tonal subconditions). This analysis varies from that of the data analysis in the pilot study where tones were grouped into types (e.g., level, direction) and then, analyzed as types. We see both universal and language-specific tendencies which do not necessarily coincide with the overall results for accuracy and RTs. First, it should be noted that six out of the ten tone comparisons did not show any significant differences between the language groups in terms of accuracy. As such, barring problems with the task design and stimuli, it can be said that all groups perceive certain tones more easily, thus indicating a universal tendency. In contrast, statistical differences did emerge between the language groups on the accuracy rates for the remaining four of the ten comparisons: Low vs Mid, Low vs Rising, Mid vs High, and Mid vs Rising. These differences do not necessarily reflect the perceptual differences in the same hierarchical order as that of the overall accuracy performance. That is, the Thai group is again statistically more accurate than the Japanese, English and Korean groups, but the English group did not perform statistically less accurate than the Japanese and Korean groups. There are other detailed differences on the four tone comparisons as well which diverge from the overall accuracy rate pattern. On two of the comparisons (i.e., Low vs Mid, Mid vs Rising), the L1 Thai group was not more accurate than the Japanese group while on two of the comparisons (i.e., Low vs Rising and Low vs Mid), the Thai group was more accurate than even the Mandarin group. Thus, it is among the accuracy results for these four tonal comparisons that language-specific tendencies emerge.

As for RTs on the individual tone comparisons, most reflect the trend witnessed in the overall performance of the language groups. Again, the Japanese and Korean groups were faster than expected. Thus, RTs for individual tone comparisons generally do not shed much light on the perception of tone. However, it is the tonal comparisons where this trend did not bear out (i.e., language-specific tendencies) that illuminate perception of tone. As a result, we again see both universal and language-specific tendencies.

Interpretation of results within the Feature Hypothesis and Perceptual Assimilation Model

Results on individual tone comparisons also point out both universal and language-specific trends, where the latter provides hints to how tones are processed by each L1. For example, not surprisingly we would expect confusion on the four tonal comparisons (i.e., Low vs Mid, Low vs Rising, Low vs Mid, Low vs Rising) that participants had difficulty with due to the strong phonetic similarities between the shapes of the two tones compared in these subconditions (See Figure 2. Contour shapes of Thai tones in citation form in Section 2.1.1). However, some language groups perform better or worse than other language groups on these comparisons against the Thai group.

This would indicate that something in the L1 of these languages is coming into play to either hinder or promote perception of at least one of the two tones being compared. It is precisely here that we begin to see glimpses of the nature of lexically-contrastive pitch. This influence on Thai tone perception of this varying functionality of lexically-contrastive pitch in the L1 may be explained by applying a combination of models for second-language perception, namely the Feature Hypothesis (McAllister et al., 2002) and the Perceptual Assimilation Model (Best, 1995).

First, in contrast to overall tone perception, it appears difficult to reconcile a Feature Hypothesis (McAllister et al., 2002) approach with individual tone perception. That is, the original McAllister et al. study looked at the presence and absence of vowel length without any necessary variety to the type of vowel length, but in the case of tones, this approach must also take into account the many types of tone patterns. Yet, it is precisely by looking at the individual tone comparisons that we may find insight into how tones are perceived and how to apply a model of perception to these results. So saying, it seems that rather than a Feature Hypothesis Approach, a PAM (Best, 1995) approach may better account for the results although the two models might be reconciled based on features smaller than tone patterns as will be discussed more in detail below after a few initial interpretations of the results.

We now turn to addressing these results within the framework of several models, namely the Feature Hypothesis (McAllister et al., 2002) and the Perceptual Assimilation Model (PAM, Best, 1995). Within a PAM model framework, we first assume that the Mandarin group views the Thai stimuli as language and thus, not as “unassimilable”. Hence, the most likely question is what the Mandarin group finds similar or dissimilar between Mandarin and Thai tones. For example, they might conjecture that the Thai Falling and Rising tones being most likely the best fits (i.e., category goodness) would be mapped onto the Mandarin Falling and Rising tone. The Thai Low and Mid tones might either be "uncategorizable" if viewed as sufficiently different from L1 Mandarin tones. If these two tones are viewed as sufficiently similar to L1 Mandarin tones, they may be "equally good" for the Mandarin Rising tone. Thus, it is the phonetic similarity of L1 and non-native tones that account for non-native tone perception. We see this for both native speakers and non-tonal language speakers as well: Confusion among the Thai and English groups has been noted for the Low vs Mid tone comparison (Abramson, 1976 for Thai; Wayland & Guion, 2004 for English), this is not the case for the Mid vs High tone.

Furthermore, the asymmetrical performance of the Mandarin speakers on the level tones comparisons appears to merely be a matter of tone-to-tone mapping between phonetically similar tones that are additionally without counterparts in the L1. For example, the Mandarin group performed well on the Mid vs High tone comparisons, but poorly on the Low vs Mid Comparisons. In short, Mandarin has a High tone and again, the phonetic contour of the High tone differs from that of the Mid tone while again both the Mid and Low tone are similar in contour and do not exist in Mandarin. Thus, the Mandarin group appears to rely on two factors influenced by the L1 inventory of Mandarin tones: Mandarin tones (i.e., High or Level, Rising, Dipping, Falling) differ by direction rather than merely height. As a result, we could use a PAM approach based on a unitary model of tone patterns or possibly a "feature-based" approach if we consider [+height] or [+direction] or more precisely a compositional model using [+ high pitch]. However, the tone patterns in Mandarin and Thai differ as can be seen in Figure 4.8. We can clearly see that while both languages have a falling and high tone that they differ in both shape and duration, which may present issues for the Mandarin group when attempting to map tones between the Thai and Mandarin.

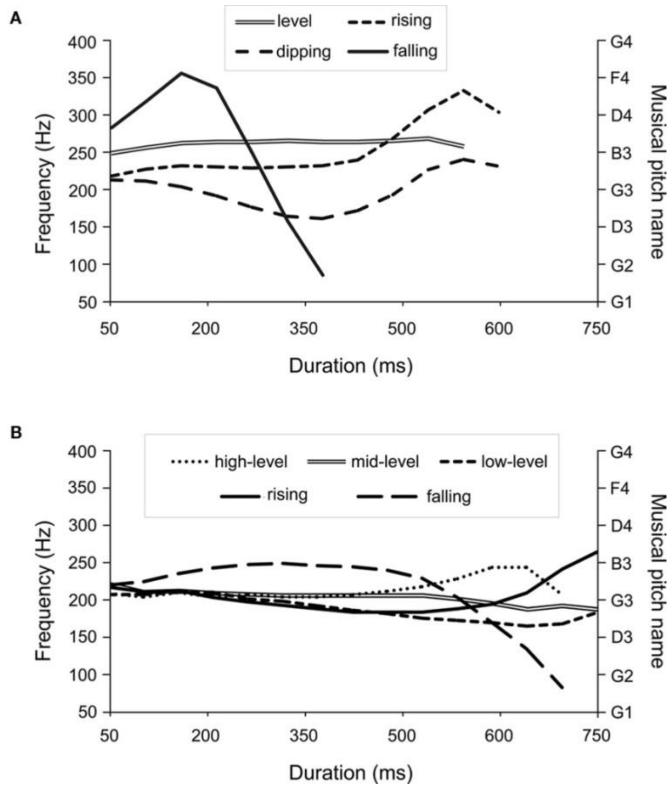


Figure 4.8 Comparison of Mandarin and Thai tones recorded from female speakers (Tillman et al., 2011).

As for the Japanese, English and Korean speaker groups, within a PAM framework we might consider all three groups to have LH and HL categories albeit to different levels of robustness (i.e., varying levels of lexical pitch functionality strongest for Japanese and weakest for Korean where pitch is present in words but not used contrastively). Additionally, as Japanese uses Low and High pitches on moras, this group may exhibit an ability to track Thai low and high level pitches with great accuracy. Indeed, superior performance overall and on individual tonal comparisons by the Japanese group may validate a moraic alignment interpretation of Thai tone (cf., Morén & Zsiga, 2006) as they may merely be hearing a High pitch on one mora combining with a Low pitch on another mora for even the contour tones. Thus, for the Japanese the Thai Rising tone which is syllabic may be merely "decomposed as a low-pitched mora followed by a high-pitched mora, for example.

Closely related to the idea of perceiving low- and high-pitched moras, it may simply be that similarity in register allows listeners of the non-tonal languages to gauge the similarity of tones. For example, most groups performed at high levels of accuracy on the Low vs High tone comparison, but less accurately on the Low vs Mid tone comparison and Mid vs High tone comparison where the Mid tone might be perceived to be in the same register of either the Low or High tone, resulting in confusion. Also, the English group performed poorly on the Mid vs High tone comparison with both tones being perceived as belonging to the upper register. Additionally, the Thai group outperformed these groups on the Low vs Rising tone comparison where both tones occur in the lower register.

As for the level tones, the Thai group performed significantly better than the English, Korean and Japanese groups on the Mid vs High comparison and marginally so against the Mandarin group on the same comparison. In contrast, on the Low vs Mid tone comparison, the Thai group performed statistically more accurately than the Korean and Mandarin groups and marginally so against the English group. We need to examine why there are differences in performance between the Thai on the one hand and the Japanese on the other hand and less so the English and Mandarin as well. First, the Japanese performed at a much lower level of accuracy on the Mid vs High tone comparison as compared to the Low vs Mid tone comparison. This may point to a possible categorization of the Mid tone as being a High tone so that the Thai Mid and High tone seem perceptually similar while the Low and Mid tone seem perceptually different. Again, it may mean that in Japanese phonology there is only the High tone with the Low tone being merely the lack of any tone (i.e., [+high] and [-high] and not [+low]); This feature-based approach may also explain why the Mid tone is categorized as a High tone (i.e., pitch is present even if lower). This may be further reinforced by the fact that phonetically mid-level pitch accents in Japanese are essentially “allotones” (or perhaps more appropriately “allo-pitch-accents”) of high pitches due to downstepping. Yet, as the High tone is phonetically different in its contour from the Mid tone it is still a question as to why the performance by the Japanese group on the Mid vs High tone comparison appears to be worse than that on the more confusing Low vs Mid tone comparison where the shape of these two tones are similar. This may have to do with categorizing the slight rise in the High tone shape within the phonological range of a high pitch in Japanese phonology (cf., Stagray and Downs, 1993). The reason for the Thai performing only marginally better than the English group on the Low vs Mid is simply due to the poorer performance by the Thai on this comparison vis-à-vis their performance on the Mid vs High comparison.

This same categorization may be the mechanism behind poor performance by the English group who exhibited their lowest accuracy rates on the Mid vs High tone comparison. Again, the pitch correlate of word stress in English may merely be equivalent to the presence of pitch (i.e., [+High]). Results on the Low vs Mid tone comparison where the English and Korean groups pattern differently may further bolster this interpretation. The English group performed better than the Korean group. Thus, while it has been noted that the English group has difficulty with this comparison (Wayland & Guion, 2004), the Korean group appears to have had more difficulty. Yet, it is precisely on such a tonal comparison that English word stress may aid the English group. That is, the high and low pitch of the word stress system of English would allow the English group to perform better on this comparison.

In sum, overall performance actually obscures the minute details and thereby, the “true mechanism(s)” operating behind tonal perception which are conversely brought to light by examining performance on individual tone comparisons.

4.3 Experiment 1b: ABX Thai tones, disyllabic

We turn now to the disyllabic ABX task, which not only may put a greater cognitive load onto the same participants of the monosyllabic ABX task, possibly teasing out more statistically significant results, but also may conversely provide “greater help” to L1 English speakers who may need two syllables to better hear the difference in tones reflective of word stress in English. To be exact, a high pitch accompanies the other correlates of vowel length and intensity to mark a stressed syllable in English while a low pitch combines with a lack of intensity, shorter vowel and oftentimes reduced vowel to characterize an unstressed syllable (cf., “a record” vs “to record”). Thus, as noted in Chapter 3 on the pilot study in Section 3.6 (p. 21), monosyllabic stimuli might not reflect the situation of English word stress. That is, the difference between the noun and verb of “record” in terms of word stress becomes salient as there is an unstressed syllable next to a stressed syllable. Therefore, English speakers may perceive some Thai tones better when they are presented in disyllabic words, creating an environment comparable to English word stress albeit characterized by other more salient correlates of word stress such as vowel length (cf., Kochanski et al., 2005). Additionally, this use of the disyllabic stimuli may aid the Japanese group as well due to common lexical pitch accent patterns (e.g., [a.me] HL ‘rain’ vs LH ‘candy’), particularly as pitch is the only correlate of Japanese word accent.

4.3.1 Methodology

4.3.1.1 Participants

Participants were the same as in Experiment 1a.

4.3.1.2 Stimuli and conditions

There were two types of stimuli: Target and Control. The target non-word stimuli varied by tone but not segment while the control non-word stimuli varied by segment but not tone. Non-words were checked by native speakers of all the languages of the participants, i.e., Thai, Mandarin, Japanese, English and Korean, to ensure that the stimuli were indeed non-words in all of these languages, and that they were possible non-words in Thai. Additional distracters were not included.

For disyllabic stimuli, each syllable carries a tone. The full matrix of pairings would be a possible 20 pairs of tones on each disyllabic item (i.e., LM, LH, LF, LR, ML, MH, etc.). Thus, considering the large number of possible combinations of pairings of the five tones, only five disyllabic combinations were selected: LM, LH, MH, HF, HL. These patterns were determined by several factors thought to affect their perception. Specifically, this selection allowed the testing of the potential mapping of L1 English word stress patterns on the target items (e.g., LM, LH, MH, HL) (cf., Brugos, Shattuck-Hufnagel, & Veilleux, 2006). Second, disyllabic items that were felt to reflect English pitch patterns were chosen. The five combinations were: LH-HL, LH-MH, LM-HL, LM-LH, and LM-MH.

- The LH-HL comparison was thought to reflect not only a difference in pitch direction but also the two possible disyllabic word stress patterns in English as in INsert vs inSERT, making it relatively easy for L1 English speakers to map onto their English and therefore, easy for them to perceive.
- The LH-MH and LM-LH were also thought to reflect word stress but differ as can be seen in the magnitude of pitch change when comparing an “extreme change” from the lower point to higher point versus a “moderate change” from a lower point to a mid-point.
- The LM-HL comparison combines a change in both pitch direction and magnitude.
- The LM-MH comparison was included to determine if participants could easily gauge the register of tones. That is, LM is in the lower register of the voice while

MH is in the higher register. This would determine whether participants could at least differentiate tones in either the lower or upper register.

The disyllabic experiment, however, did not feature a comparison between flat and contour tones, but this condition was included in the monosyllabic experiment. Also, “redundant” comparisons such as LM-HF which might be similar to either LM-HL in the task were not included. This resulted for each disyllabic non-word in five pairings; In total, for two non-words and four presentation orders, there were 40 trials in the test (tone) condition. Forty trials for this test condition were created by using the five tone comparisons on two syllables (i.e., [dua.pʰuui], ตัวพูย[kiːŋ.kæ] กึงแก) with four different orderings (i.e., ABA, ABB, BAA, and BAB). Thus, for one of the LH-HL trials the disyllabic ABX task looks as follows:

[dua.pʰuui]	[dua.pʰuui]	[dua.pʰuui]
L H	H L	L H
A	B	A

For the disyllabic control stimuli, the same patterns of HF, HL, LH, LM and MH were used as these were the only patterns present in the disyllabic tone comparisons. The stimuli varied in segments, either by only one vowel or one consonant as can be seen in Table 4.11 below. Again, the total number of trials included two pairs of disyllabic stimuli presented in the four possible ABX orderings of ABA, ABB, BAA, and BAB for five comparisons. This resulted in 40 possible trials (2 pairs of items x 4 orderings x 5 tone combinations). In total, there were 80 experimental trials. A trial for the control stimuli on the disyllabic ABX task looks as follows:

[teo.fiːŋ]	[teo.foːŋ]	[teo.fiːŋ]
M H	M H	M H
A	B	A

Table 4.8: Stimuli, conditions and number of trials in the ABX disyllabic task

Test (=tone) (40 trials**)		Control (=segmentals) (40 trials**)		Training (=segmentals only) (10 trials**)	
Tone types	Segments	Tone types	Segment pairs	Tone types	Segment pairs
Low+High-High+Low	[dua.pʰuui] ดัวพูย* [kiːŋ.kæ] กึงแก	High+Falling	[pʰuːi.wuːi] พูยวูย - [ruːi.wuːi] รูยวูย [teo.fiːŋ] เตวฟีง - [teo.foːŋ] เตวโฟง	High+Falling	[no:bɛ:o] โนเบว - [luai.pʰuai] ลูยพวย [ro:bɛ:o] โรเบว - [lu:i.pʰuai] ลวยพวย
Low+Mid-High+Low		High+Low		High+Low	
Low+High-Mid+High		Low+High		Low+High	
Low+Mid-Low+High		Low+Mid		Low+Mid	
Low+Mid-Mid+High		Mid+High		Mid+High	

*Spellings in the Thai script may not reflect the actual tones.
** The tone test condition = 5 tone types x 2 segments x 4 orders (AAB, ABB, BBA, BAA) = 40 trials; the control test condition = 5 tone types x 2 segments x 4 orders (AAB, ABB, BBA, BAA) = 40 trials; training = 5 tone types x 2 segments = 10 trials

In the training session, the trials compared segments with the same tone combination for the A, X and B stimuli. The training items basically mirrored those of the control condition using the same tone combinations of HF, HL, LH, LM and MH, but with the A and B item differing greatly in terms of segments, e.g., [no:bɛ:o] vs [luai.pʰuai]. Each of the five tone combinations was heard 6 times each in a total of 10 trials. In sum, the total number of times a tone combination was heard in the disyllabic task was LM x 66, LH x 66, MH x 54, HL x 54, HF x 30 (total = 270). Thus, participants heard 540 tone samples. They heard the low tone 186 times; the mid tone for 120 times; the high tone for204 times; and the falling tone for 30 times. The times the listeners heard each tone was controlled as best as possible to prevent possible bias.

4.3.1.3 Speakers for Thai stimuli/Elicitation method

The same two female voices of the Central Thai dialect were used to create the disyllabic Thai stimuli (See Section 4.2.1.3 for details). In the case of the disyllabic items, speaker one read the first test (i.e., tone) non-word [dua.pʰu:i] with the Low-Mid tone pattern three times. Speaker two repeated this on a second microphone. Then, speaker one read the second test non-word [kiːŋ.kæː] with the same Low-Mid tone pattern and speaker two repeated this. This ensured that all the stimuli were consistently recorded using the same pattern. Care was also taken to ensure the native-speaker recorders pronounced the disyllabic words as one word. Care was also taken to make sure there was some space between the same or different words for splicing purposes.

Stimuli were recorded in the same session and with the same technical settings as for the monosyllabic stimuli. The disyllabic stimuli were normalized to an average value of 13Hz. Similar checks were also performed on the stimuli. In particular, Praat (Boersma & Weenink, 2015) was used to examine the spectrographs of the tones to ensure accurate labelling. For example, it was checked whether a rising tone had not been confused with a low tone or vice-versa.

4.3.1.4 Procedure

The procedure was the same as in Experiment 1a; Participants first completed experiment 1a, and then the background questionnaire, followed by this experiment. The overall duration of this part was 20 minutes.

The DMDX (Forster & Forster, 2003) program for the ABX disyllabic task was overall similar to the DMDX monosyllabic ABX task (see 4.2.1.4), except for the following point. For the disyllabic Thai ABX task in DMDX, only one break was inserted again every 40 items for only one break in the 80-trial disyllabic Thai ABX task. The 40 trials were randomized within each block. Blocks were randomized with each other as well. The RT was measured from the onset of the X stimulus. The script is reprinted in Appendix J (pp. 323-329).

4.3.2 Results

Data from the ABX task and from the questionnaire were first converted to an Excel spreadsheet and then, entered into SPSS for statistical analyses.

The same participants excluded from the analysis of the monosyllabic task were also excluded here. The final number of participants in each group is as follows: total (n = 110) Thai (n = 9, female = 8), Mandarin (n = 29, female = 23), Japanese (n = 22, female = 10), English (n = 23, female = 14) and Korean (n = 27, female = 19).

Reaction times (RT) that were 300ms and below and any over 2500ms were removed. In total, only 0.00009% of the RT datapoints were removed (N = 1). As a log-transformation was applied to the RT (LogRT) to obtain a normal distribution for the monosyllabic ABX task, the same was done to the data for the disyllabic ABX task to make the data between the two tasks more comparable. After cleaning the data, a Generalized Estimating for a binary response was run for overall accuracy and a Linear Mixed Effects Model was run for the LogRT with Language (Thai, Mandarin, Japanese, English, Korean) and Condition (test vs control) as fixed factors and declaring Subjects as a random factor.

4.3.2.1 Analysis by condition: Test and control

For each group, a mean accuracy score and mean reaction times (RT) were computed for each condition and sub-condition (i.e., individual tonal comparisons) and are presented in the tables below. First, we will look at the test and control conditions for overall accuracy rates and RTs.

Mean scores for accuracy rates and RT for each group on the disyllabic ABX task are presented in Table 4.9 and Figure 4.8. They show the Thai performed the most accurately on the test tone condition (*M* = 91.7%) while the English group performed at the least accurate level (*M* = 77.5%). The Japanese, Mandarin and Korean groups performed at levels in between the Thai and English groups at *M* = 85.8%, *M* = 85.9%, *M* = 83.9%, respectively. On the control (segment) condition, accuracy rates are similarly high for all the language groups (*M* = 90 percentile). The range of accuracy rates for the control condition (all scores were approximately 96%) is smaller than that for the test tone condition (from a low of 77.5% to a high of 91.7%).

In Table 4.9 and Figure 4.9, we see that for RTs on the tone test condition, the Japanese are the fastest (*M* = 1106 ms) while the English group is the slowest (*M* = 1368 ms). The three groups of Korean, Thai and Mandarin perform at in-between rates (*M* = 1194 ms, *M* = 1262 ms, *M* = 1156 ms, respectively). As for the RTs on the segment control condition, the Japanese (*M* = 945 ms) and Korean (*M* = 983 ms) are the fastest with the Mandarin (*M* = 1156 ms) the slowest although similar to both the Thai (*M* = 1140) and English (*M* = 1139) groups.

Table 4.9: Mean accuracy (%), mean RT (ms) and mean log RT for each group on test and control conditions

	Accuracy		RT		Log RT	
Language group	Test (SE)	Control (SE)	Test (CI)	Control (CI)	Test (CI)	Control (CI)
Thai (n=9)	91.7 (2.6)	96.1 (1.4)	1262 (1133; 1406)	1140 (987; 1316)	3.1 (3.05; 3.15)	3.06 (2.99; 3.12)
Mandarin (n=29)	85.9 (2.0)	96.7 (0.7)	1281 (1237; 1327)	1156 (1113; 1200)	3.11 (3.09; 3.12)	3.06 (3.05; 3.08)
Japanese (n=22)	85.8 (1.4)	96.5 (0.9)	1106 (1040; 1177)	945 (886; 1008)	3.04 (3.02; 3.07)	2.98 (2.95; 3.0)
English (n=23)	77.5 (3.1)	96.6 (0.7)	1368 (1305; 1434)	1139 (1057; 1228)	3.14 (3.12; 3.16)	3.06 (3.02; 3.09)
Korean (n=27)	83.9 (1.8)	96.9 (0.6)	1194 (1131; 1260)	983 (921; 1050)	3.08 (3.05; 3.1)	2.99 (2.96; 3.02)

Note: SE = standard error.CI = Confidence interval (lower; upper).

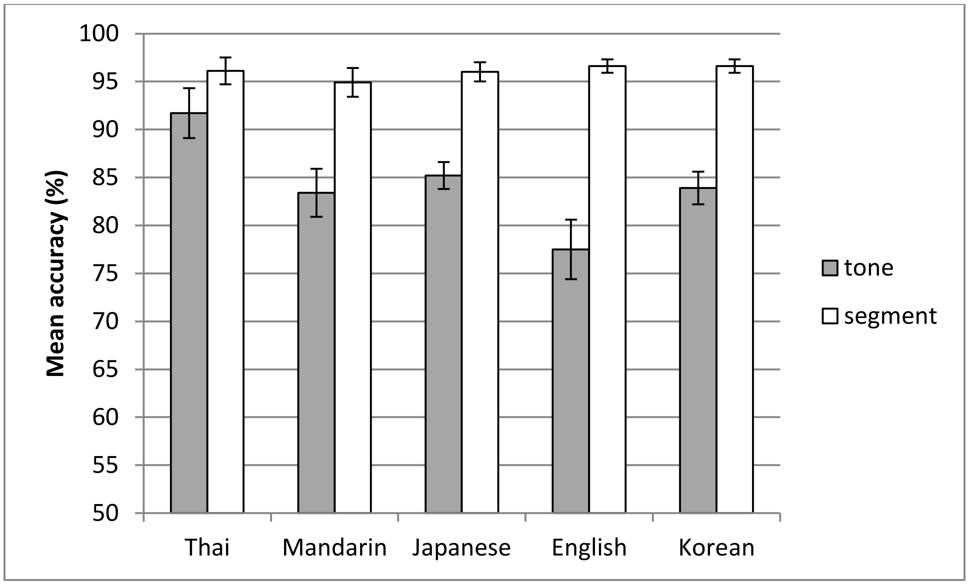


Figure 4.8: Mean accuracy rate (%) for each group on test and control conditions. Error bars represent +/-1 SE.

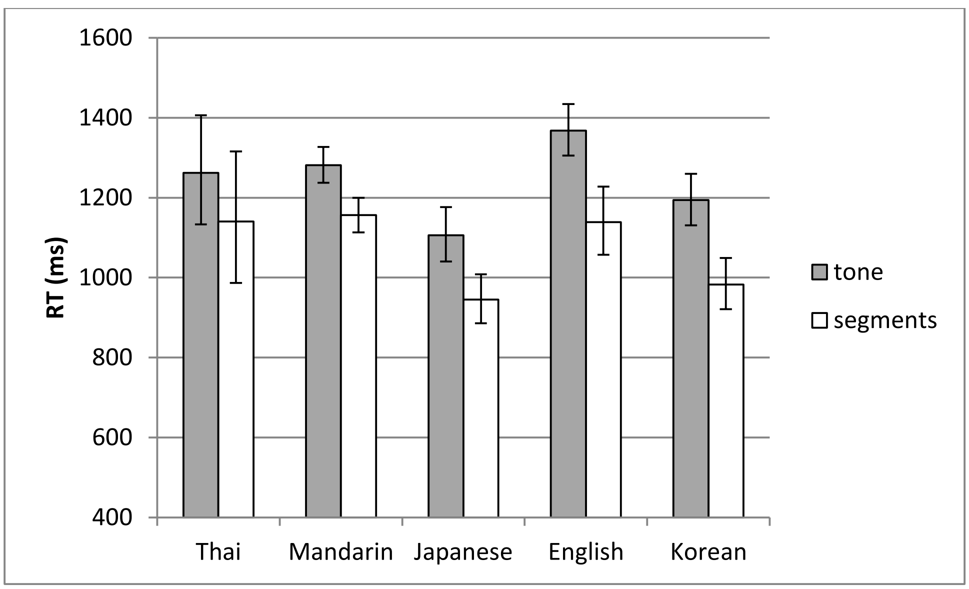


Figure 4.9: Mean RT (ms) for each group on test and control conditions. Error bars represent the upper and lower CI from RTs from a Linear Mixed Effects model.

The un-aggregated data were used for statistical analysis. Accuracy data will be analyzed first, RT analysis will follow.

Accuracy Data

To examine the accuracy data, a GEE model was fitted, declaring the fixed factors Language (Thai, Mandarin, Japanese, English, Korean) and Condition (test vs control), with Subjects as a random factor. The Type III tests of fixed effects showed no main effect of Language ($\chi^2(4) = 3.02$, $p = 0.55$), a significant effect of Condition ($\chi^2(1) = 165.2$, $p < 0.001$), and a significant interaction between the two factors ($\chi^2(4) = 10.01$, $p = 0.040$).

The interaction shows that while groups do not differ on the Control condition, their performance varies on the Test condition where there is a main effect of Language (See Table 4.15). Overall test results revealed that performance differed significantly between groups on the test condition ($\chi^2(4) = 13.33$, $p < 0.010$), where the native Thai speakers performed significantly better than only the English group ($p = 0.004$) but not significantly better than the other groups. In contrast, all groups performed at statistically comparable levels on the control condition (See Table 4.17).

RT Data

For RTs, the Japanese and Korean were the fastest on both the test and control conditions. While the English group was the slowest on the tone test condition, it was not the slowest on the control segment condition. The Thai and Mandarin performed the slowest on this condition with the English group closer to their RTs than to those of the Japanese and Korean groups.

A Linear Mixed Effects model was fitted, to examine the LogRTs, declaring Language (Thai, Mandarin, Japanese, English, Korean) and Condition (test vs control) as fixed factors with Subjects as a random factor. When looking at the Type III tests of fixed effects, the F-tests showed a main effect of Condition ($F(1, 7864) = 580.80$, $p < 0.001$), a significant effect of Language ($F(4, 104) = 8.55$, $p < 0.001$), and a significant interaction between the two factors ($F(4, 7965) = 11.56$, $p < 0.001$).

The interaction shows that the performance of the various groups differed on both the Test and Control conditions (See Table 4.12). Univariate tests revealed that performance differed significantly between groups on the test condition ($F(4, 118.0) = 6.79$, $p < 0.001$) and control condition ($F(4, 115.4) = 9.11$, $p < 0.001$).On the test condition, the Thai speakers were not faster than any other language groups. However, the English were slower than both the Japanese ($p < 0.001$) and the Korean ($p = 0.002$) groups while the Mandarin group was slower than the Japanese ($p = 0.001$) (See Table 4.12). The reaction times on the control condition generally reflect those on the test conditions with both the Japanese and Korean group reacting significantly faster than the English ($p = 0.002$, $p = 0.036$, respectively) and Mandarin groups ($p > 0.001$ by both Japanese and Korean groups).

Analysis by Korean Subgroup

Analyses on the Korean group were run examining the possible differences between a group of speakers of standard Korean and a group of speakers of the Kyungsang Korean dialect which features pitch accent, in the same way as was done in the previous experiment and in Chapter 3. Table 4.10 presents the Korean group split by L1 dialects (grey-highlighted cells).

Table 4.10: Mean accuracy (%) and mean RT (ms) for each group and Korean dialect subgroups on test and control conditions

	Accuracy		RT		Log RT	
Language group	Test (SE)	Control (SE)	Test (CI)	Control (CI)	Test (CI)	Control (CI)
Thai (n=9)	91.7 (2.5)	96.1 (1.4)	1262 (1133; 1406)	1140 (987; 1316)	3.1 (3.05; 3.15)	3.06 (2.99; 3.12)
Mandarin (n=29)	85.7 (2.0)	96.7 (0.7)	1281 (1237; 1327)	1156 (1113; 1200)	3.11 (3.09; 3.12)	3.06 (3.05; 3.08)
Japanese (n=22)	85.8 (1.4)	96.5 (0.9)	1106 (1040; 1177)	945 (886; 1008)	3.04 (3.02; 3.07)	2.98 (2.95; 3.0)
English (n=23)	77.5 (3.1)	96.6 (0.7)	1368 (1305; 1434)	1139 (1057; 1228)	3.14 (3.12; 3.16)	3.06 (3.02; 3.09)
Korean (n=27)	83.9 (1.8)	96.9 (0.6)	1194 (1131; 1260)	983 (921; 1050)	3.08 (3.05; 3.1)	2.99 (2.96; 3.02)
Kyungsang Korean (n=10)	83.0 (2.7)	97.0 (1.0)	1180 (1103; 1264)	940 (859; 1028)	3.07 (3.04; 3.1)	2.97 (2.93; 3.01)
Seoul Korean (n=17)	84.5 (2.4)	96.8 (0.8)	1203 (1114; 1300)	1013 (928; 1106)	3.08 (3.05; 3.11)	3.01 (2.97; 3.04)

Note: SE = standard error.CI = Confidence interval (lower;upper)

To examine the accuracy data, a GEE model with Language (i.e., Dialects: standard Korean, Kyungsang Korean) and Condition (test vs control) as fixed factors and Subjects as a random factor was fitted to the data. The Type III tests of fixed effects showed no main effect of Language ($\chi^2(1) = 0.007$, $p = 0.931$), a significant effect of Condition ($\chi^2(1) = 85.36$, $p < 0.001$), and no significant interaction between the two factors ($\chi^2(1) = 0.186$, $p = 0.666$). Overall test results revealed that performance did not differ significantly between groups on the test condition ($\chi^2(1) = 0.177$, $p = 0.674$). Similarly, there was no significant difference in accuracy rates on the control conditions ($p = 0.893$) for the two Korean subgroups.

A Linear Mixed Effects model was run on the LogRTs with Language (i.e., Dialect: standard Korean, Kyungsang Korean) and Condition (test vs control) as fixed factors and Subjects as random factors. When looking at the Type III tests of fixed effects, the F-tests showed a main effect of condition ($F(1, 1780) = 270.95$, $p < 0.001$), no significant effect of language ($F(1, 23.05) = 0.646$, $p = 0.430$), and a significant interaction between the two factors ($F(1, 1780) = 5.306$, $p = 0.021$). There was no significant difference in RTs between the two Korean dialects on the test ($p = 0.718$) and control condition ($p = 0.242$).

Thus, the analysis shows a significant interaction between the factors of Language (i.e., dialect) and Condition. Yet, there is no statistically significant difference on either the test or control condition. This contradiction could be the effect of a difference between the two dialect groups on the mean RT of the two groups on the test condition (Kyungsang Korean: $M = 1180$ ms vs Seoul Korean: $M = 1203$ ms) versus the difference between the mean RTs of the two groups on the control condition (Kyungsang Korean: $M = 940$ ms vs Seoul Korean: $M = 1013$ ms).

4.3.2.2 Analysis by subcondition

We now turn to the comparison of individual tones, for both accuracy and RTs. We examine the individual tonal comparisons to determine whether the perception of one tone type may be more difficult than perceiving another tone type by analyzing the comparisons of the various tone types. All language groups patterned similarly on certain tone types, suggesting a universal tendency despite the differences in the prominence of lexically-contrastive pitch in the L1. On the other hand, differences among groups in performance may indicate language-specific tendencies due to the prominence of lexically-contrastive pitch in the native language. Table 4.11 and Table 4.13 present the individual accuracy means and RT means, respectively, for the test and control items in each subcondition for each group.

The mean accuracy rates and RTs for both the test and control conditions across the language groups are as follows. Accuracy rates for the control condition are very high, ranging in the mid to high 90-percentile for all groups. In comparison, accuracy rates for the test condition range from a low of 63% to a high of 97.2%. There are some universal tendencies. For example, the Low+Mid vs High+Low comparison is the highest score for almost all the groups. On the other hand, the Low+High vs Mid+High comparison is the lowest score for all the groups. In contrast, some accuracy rates are language specific, such as the Low+Mid vs Mid+High comparison. For the Korean group, their performance on the Low+Mid vs Mid+High comparison is high relative to their performance on the other tone conditions. In contrast, the Thai performance on the Low+Mid vs Mid+High comparions is more middling than high relative to their performance on the other tone comparisons. Also, we see that the Japanese performed least accurately among the groups on the Low+Mid vs Low+High comparison (See Table 4.11 and Figures 4.10a-4.10e).

Table 4.11: Mean accuracy (%) for each group on each tonal comparison

Test subconditions	Mean accuracy (SE)				
	T	M	J	E	K
LH vs HL	95.8 (2.0)	92.2 (2.4)	92.6 (1.9)	80.7 (3.5)	85.0 (2.5)
LM vs HL	97.2 (1.7)	95.3 (1.5)	93.2 (1.8)	81.8 (3.5)	91.0 (2.1)
LH vs MH	81.9 (6.2)	66.4 (4.3)	72.7 (3.1)	63.0 (5.1)	70.5 (2.7)
LM vs LH	91.7 (2.8)	85.3 (3.3)	77.8 (3.6)	80.2 (3.5)	81.5 (3.3)
LM vs MH	91.7 (2.0)	90.1 (2.1)	92.6 (1.7)	81.8 (3.6)	91.5 (2.5)
Control subconditions					
HF	95.8 (2.0)	97.4 (1.1)	96.6 (1.2)	94.8 (1.8)	99.5 (0.5)
HL	97.2 (2.6)	97.0 (1.2)	98.3 (0.9)	96.9 (1.1)	94.5 (1.2)
LH	95.8 (2.0)	97.4 (0.9)	93.8 (2.1)	97.9 (1.0)	96.5 (1.4)
LM	94.4 (2.1)	96.6 (1.2)	96.0 (1.7)	95.8 (1.4)	97.5 (1.0)
MH	97.2 (1.7)	95.3 (1.3)	97.7 (1.3)	97.4 (1.0)	96.5 (1.5)

Note: SE=standard error.
T = Thai; M = Mandarin, J = Japanese, E = English, K = Korean ; L = low tone, M = mid tone, H = high tone, F = falling tone, R = rising tone. Hence, LM = Low tone + Mid tone disyllabic word, LH = Low tone + High tone disyllabic word, MH = Mid tone + High tone disyllabic word, HL = High tone + Low tone disyllabic word, HF = High tone + Falling tone disyllabic word.

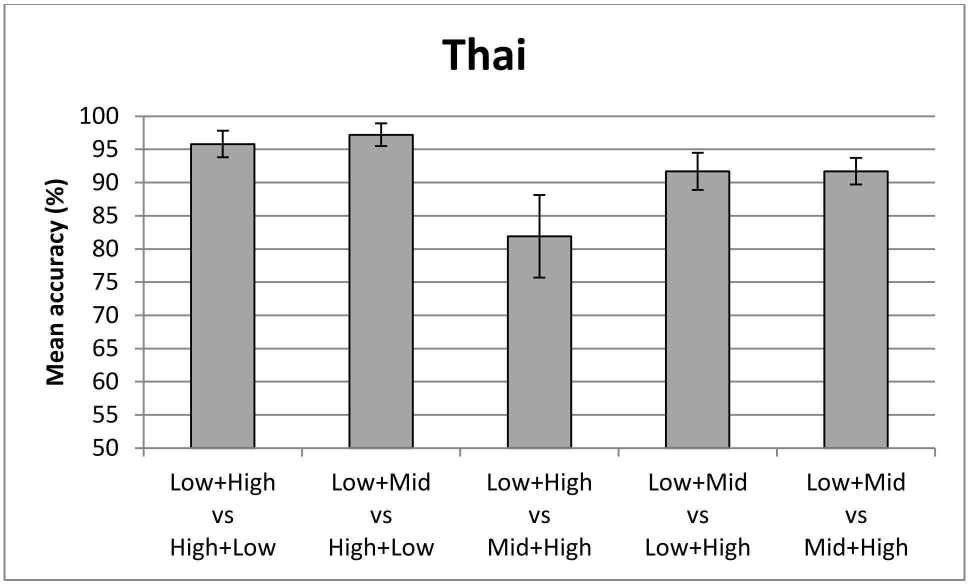


Figure 4.10a: Mean accuracy (%) for the Thai group on each test subcondition. Error bars represent +/-1 SE.

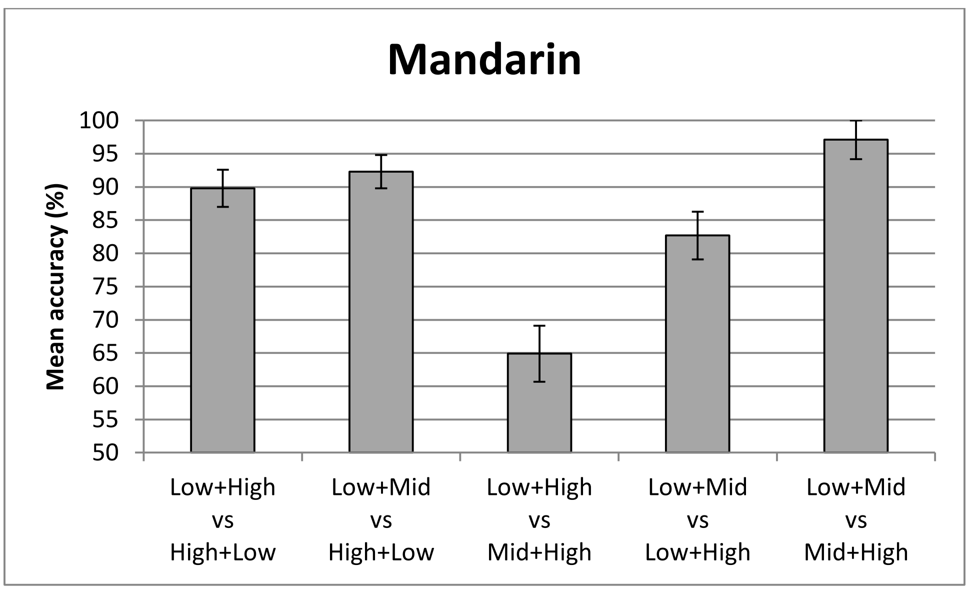


Figure 4.10b: Mean accuracy (%) for the Mandarin group on each test subcondition. Error bars represent +/-1 SE.

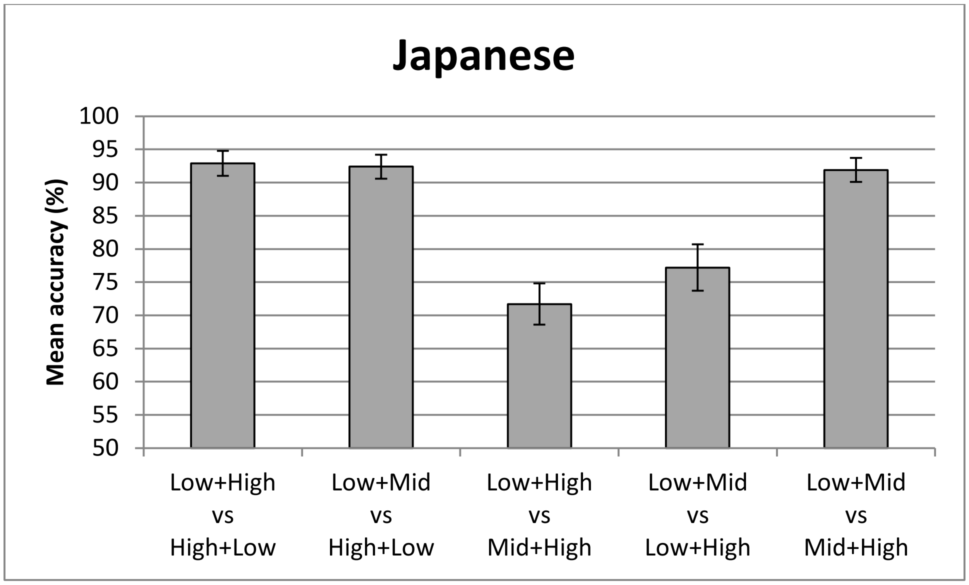


Figure 4.10c: Mean accuracy (%) for the Japanese group on each test subcondition. Error bars represent +/-1 SE.

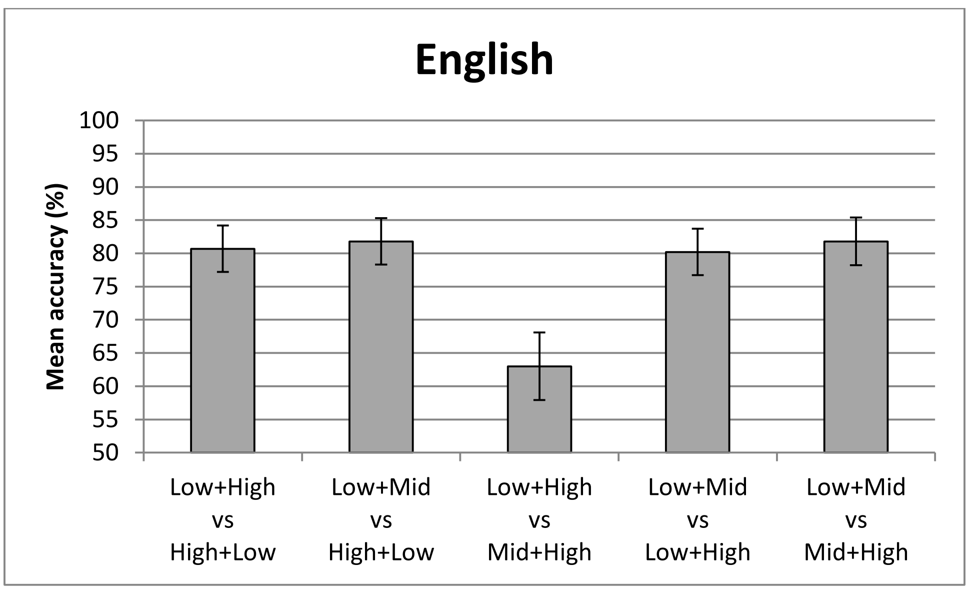


Figure 4.10d: Mean accuracy (%) for the English group on each test subcondition. Error bars represent +/-1 SE.

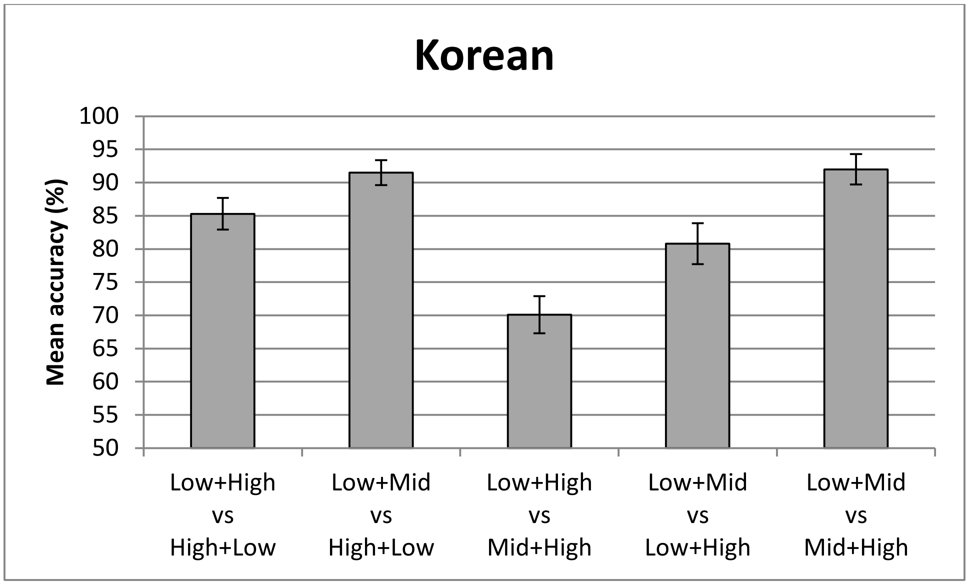


Figure 4.10e: Mean accuracy (%) for the Korean group on each test subcondition. Error bars represent +/-1 SE.

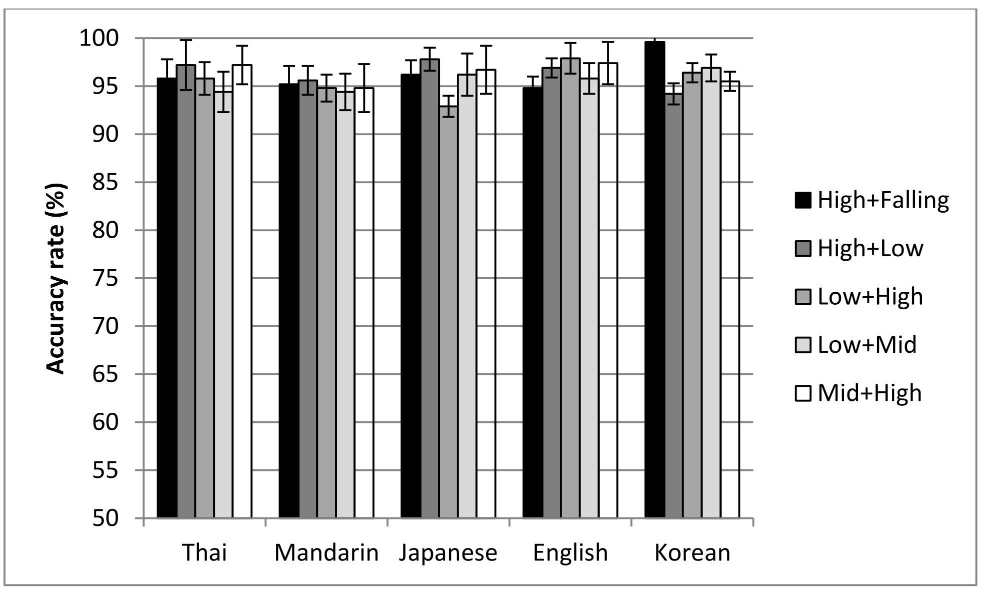


Figure 4.11: Mean accuracy (%) for each group on each control condition. Error bars represent +/-1 SE.

Accuracy Data

To examine the accuracy data on individual tone comparisons, a GEE model for a binary response was fitted with the fixed factors Language (Thai, Mandarin, Japanese, English, Korean) and Subcondition (individual tonal comparisons for both the tone and segment subconditions), with Subjects as a random factor. Examining individual tonal comparisons, we see that the Type III tests of fixed effects showed no main effect of Language ($\chi^2(4) = 3.884$, $p = 0.422$), a significant effect of Subcondition ($\chi^2(9) = 443.771$, $p < 0.001$), and a significant interaction between the two factors ($\chi^2(36) = 155.482$, $p < 0.001$). In Table 4.12 we can see in which subconditions there was a statistically significant difference among the groups.

The interaction again shows that the performance of the groups varies on the Test subconditions but not the Control subconditions. For accuracy rates on individual tone comparisons, we see in Table 4.12 that the Low+High vs High+Low and the Low+Mid vs High+Low show the greatest differences between language groups. We see that the Thai and Japanese significantly outperform the English on both these comparisons. Additionally, the Korean group was outperformed by the Thai group ($p = 0.008$) on the Low+High vs High+Low. Also, the Mandarin group outperformed the English on the Low+Mid vs High+Low comparison ($p = 0.004$). Another point to note is the Low+Mid vs Low+High comparison where the Thai outperformed the Japanese ($p = 0.023$).

On the control segment conditions, none of the groups performed significantly more accurate than the other groups.

Table 4.12: Statistically significant accuracy rates on each tonal comparison (i.e., subcondition) for each group

Comparison	Accuracy
Tone overall	Thai > English ($p = 0.004$)
LH-HL	Thai > English ($p = 0.002$) Japanese > English ($p = 0.030$) Thai > Korean ($p = 0.008$)
LM-HL	Thai > English ($p = 0.001$) Mandarin > English ($p = 0.004$) Japanese > English ($p = 0.031$)
LH-MH	--
LM-LH	Thai > Japanese ($p = 0.023$)
LM-MH	--
Segment overall	--
HF (control)	--
HL (control)	--
LH (control)	--
LM (control)	--
MH (control)	--

RT Data

Turning now to the RT data, we observe that the RTs for the control condition were generally faster than those for the test condition for all groups (See Table 4.13 and Figures 4.12a-4.12e). In both conditions, the Japanese and Korean groups were the fastest but not uniformly so, i.e., the RT on the Low+Mid vs Low+High comparison appears to be much higher than the RTs on the other test comparisons. The English group was the slowest on the test condition but not on the control condition where the Thai and Mandarin group were generally somewhat slower. Again, there were some universal tendencies across all groups: RTs were the slowest on the Low+Mid vs Low+High comparison while the RTs on the Low+High vs High+Low and Low+Mid vs High+Low were generally the fastest. There appear to be no language-specific tendencies.

Table 4.13: Mean RTs and logged RTs for each group on each tonal comparison

Test subconditions	RT (CI)					Log RT (CI)				
	T	M	J	E	K	T	M	J	E	K
LH vs HL	1203 (1043; 1387)	1180 (1128; 1236)	1037 (970; 1111)	1301 (1228; 1379)	1132 (1062; 1207)	3.08 (3.02; 3.14)	3.07 (3.05; 3.09)	3.02 (2.99; 3.05)	3.11 (3.09; 3.14)	3.05 (3.03; 3.08)
LM vs HL	1150 (1017; 1300)	1179 (1118; 1243)	1008 (926; 1100)	1297 (1226; 1373)	1088 (1014; 1168)	3.06 (3.01; 3.11)	3.07 (3.05; 3.09)	3.0 (2.97; 3.04)	3.11 (3.09; 3.14)	3.04 (3.01; 3.07)
LH vs MH	1292 (1173; 1423)	1362 (1308; 1419)	1141 (1082; 1204)	1458 (1364; 1560)	1276 (1204; 1352)	3.11 (3.07; 3.15)	3.13 (3.12; 3.15)	3.06 (3.03; 3.08)	3.16 (3.13; 3.19)	3.1 (3.08; 3.13)
LM vs LH	1448 (1318; 1591)	1439 (1386; 1495)	1282 (1213; 1356)	1484 (1416; 1555)	1345 (1284; 1409)	3.16 (3.12; 3.2)	3.16 (3.14; 3.17)	3.11 (3.08; 3.13)	3.17 (3.15; 3.19)	3.13 (3.11; 3.15)
LM vs MH	1250 (1100; 1421)	1301 (1250; 1355)	1114 (1039; 1194)	1332 (1257; 1412)	1176 (1101; 1256)	3.1 (3.04; 3.15)	3.11 (3.1; 3.13)	3.05 (3.02; 3.08)	3.12 (3.1; 3.15)	3.07 (3.04; 3.1)
Control subconditions										
HF	1156 (1004; 1330)	1152 (1108; 1198)	953 (895; 1015)	1120 (1038; 1209)	990 (924; 1061)	3.06 (3.0; 3.12)	3.06 (3.04; 3.08)	2.98 (2.95; 3.01)	3.05 (3.02; 3.08)	3.0 (2.97; 3.03)
HL	1162 (991; 1362)	1178 (1123; 1237)	952 (885; 1024)	1152 (1055; 1258)	998 (934; 1066)	3.07 (3.0; 3.13)	3.07 (3.05; 3.09)	2.98 (2.95; 3.01)	3.06 (3.02; 3.1)	3.0 (2.97; 3.03)
LH	1120 (955; 1313)	1139 (1087; 1193)	956 (891; 1025)	1127 (1042; 1219)	968 (904; 1035)	3.05 (2.98; 3.12)	3.06 (3.04; 3.08)	2.98 (2.95; 3.01)	3.05 (3.02; 3.09)	2.99 (2.96; 3.02)
LM	1138 (988; 1310)	1152 (1102; 1205)	918 (854; 987)	1143 (1054; 1239)	971 (901; 1046)	3.06 (2.99; 3.12)	3.06 (3.04; 3.08)	2.96 (2.93; 2.99)	3.06 (3.02; 3.09)	2.99 (2.95; 3.02)
MH	1124 (968; 1306)	1157 (1113; 1204)	948 (886; 1014)	1155 (1064; 1253)	990 (924; 1060)	3.05 (2.99; 3.12)	3.06 (3.04; 3.08)	2.98 2.95; 3.01)	3.06 (3.03; 3.1)	3.0 (2.97; 3.03)

Note: CI=Confidence interval (lower; upper).
T = Thai; M = Mandarin, J = Japanese, E = English, K = Korean; L = low tone, M = mid tone, H = high tone, F = falling tone, R = rising tone. Hence, LM = Low tone + Mid tone disyllabic word, LH = Low tone + High tone disyllabic word, MH = Mid tone + High tone disyllabic word, HL = High tone + Low tone disyllabic word, HF = High tone + Falling tone disyllabic word.

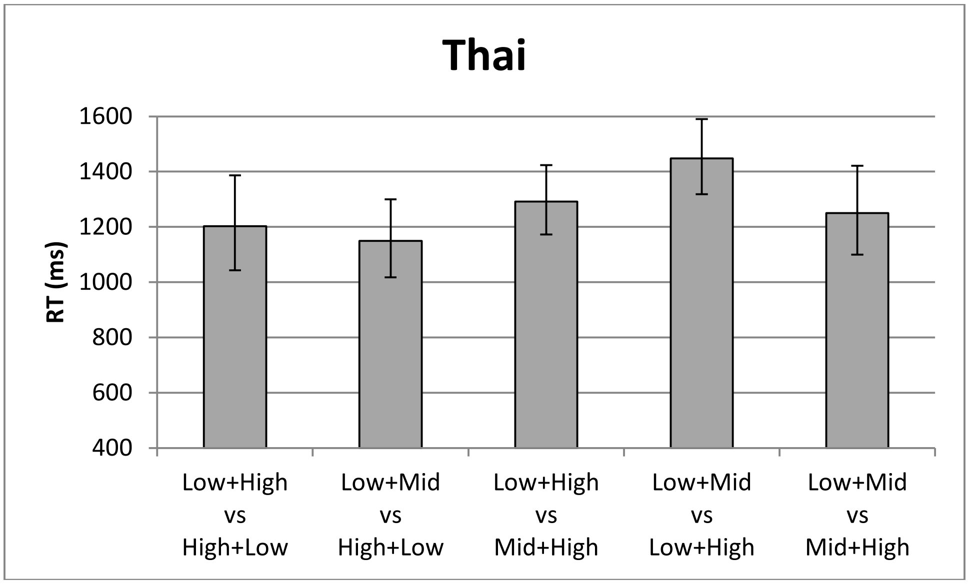


Figure 4.12a: Reaction times (ms) for the Thai on each test subcondition. Error bars represent the upper and lower CI from RTs from a Linear Mixed Effects model.

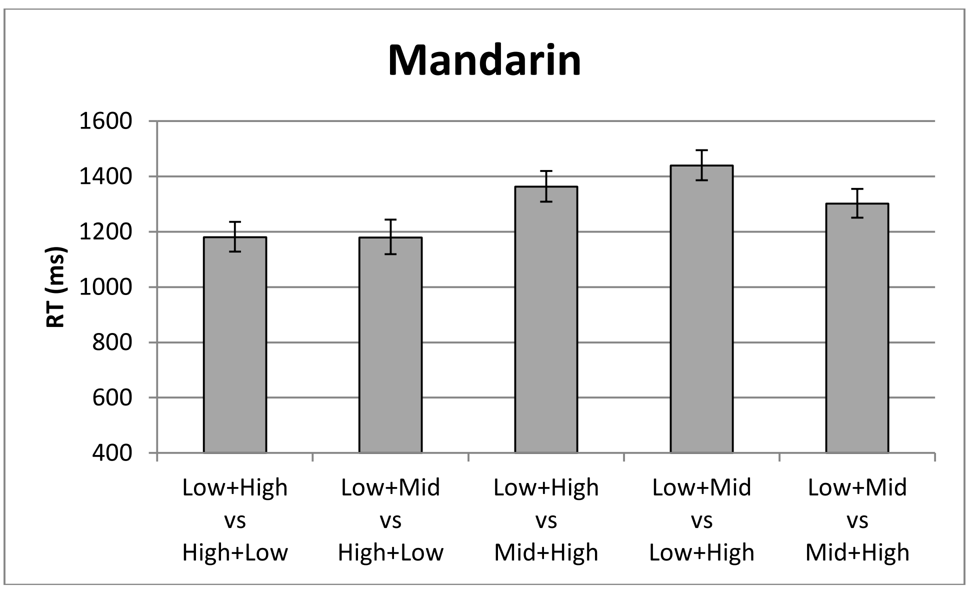


Figure 4.12b: Reaction times (ms) for the Mandarin on each test subcondition. Error bars represent the upper and lower CI from RTs from a Linear Mixed Effects model.

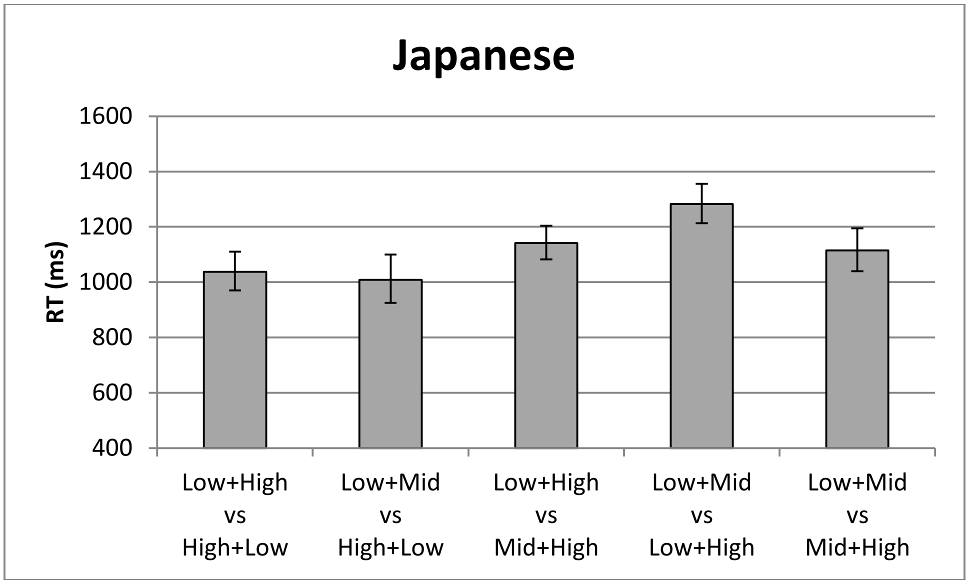


Figure 4.12c: Reaction times (ms) for the Japanese on each test subcondition. Error bars represent the upper and lower CI from RTs from a Linear Mixed Effects model.

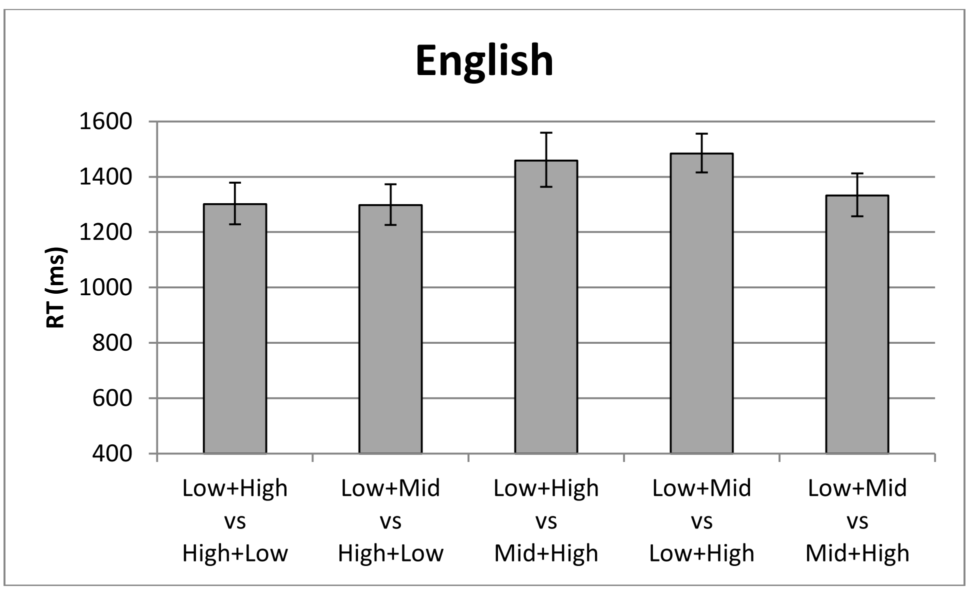


Figure 4.12d: Reaction times (ms) for the English on each test subcondition. Error bars represent the upper and lower CI from RTs from a Linear Mixed Effects model.

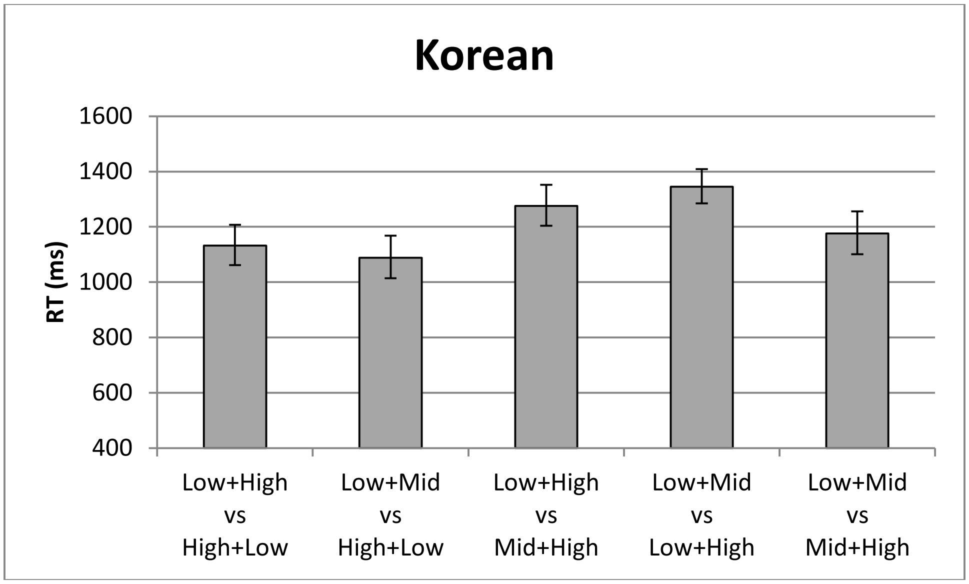


Figure 4.12e: Reaction times (ms) for the Korean on each test subcondition. Error bars represent the upper and lower CI from RTs from a Linear Mixed Effects model.

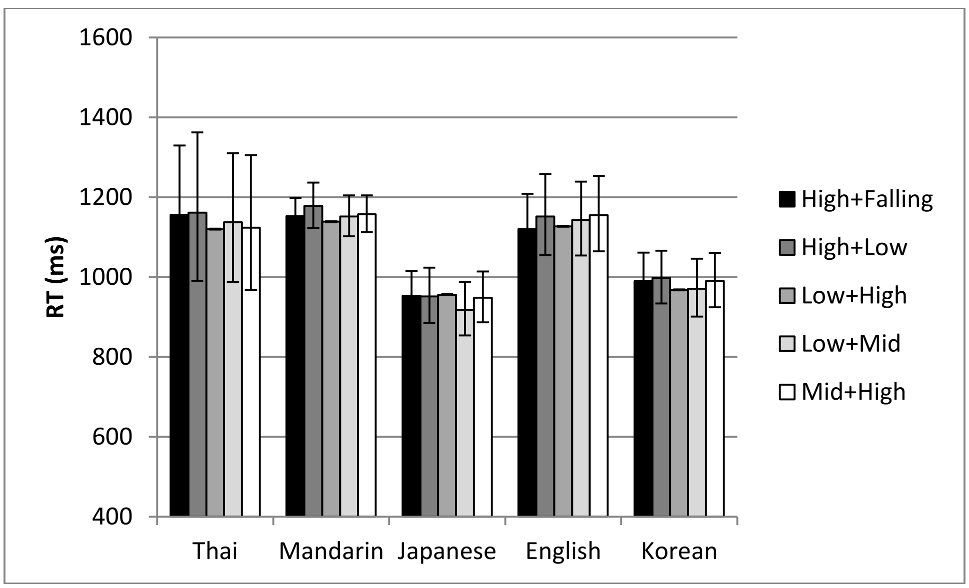


Figure 4.13: RTs (ms) for each group on each control subcondition. Error bars represent the upper and lower CI from RTs from a Linear Mixed Effects model.

A Linear Mixed Effects model was run on the LogRTs with Language (Thai, Mandarin, Japanese, English, Korean) and Subcondition (individual tonal comparisons for both the tone and segment subconditions) as fixed factors, declaring Subjects as a random factor. When looking at the Type III tests of fixed effects, the F-tests showed a main effect of Subcondition (F(9, 7723) =

100.25, $p < 0.001$), a significant effect of Language ($F(4, 104.15) = 8.48$, $p < 0.001$), and a significant interaction between the two factors ($F(36, 7723) = 1.93$, $p = 0.001$).

The interaction shows that the performance of the groups varied on both the Test and Control subconditions. As for reaction times, the Japanese and Korean groups were faster overall while the English and Mandarin groups were slower overall. The Japanese group was significantly faster than the English group on all the test conditions and on three of the five control conditions. The Korean group also was significantly faster than the English group on all the test tone conditions and on three out of the five control conditions. The Japanese group was significantly faster than the Mandarin group on every condition, both test and control (See Table 4.14 below). The Korean group were also significantly faster than the Mandarin group on the Low+Mid vs Mid+High (p = 0.038). Additionally, all the groups showed significant within-group differences on the RTs on all the subconditions whether test or control as seen on Univariate Tests.

Table 4.14: Statistically significant LogRTs on each tonal comparison (i.e., subcondition) for each group

Comparison	Reaction times
Tone overall	Japanese > English (*p* < 0.001) Korean > English (*p* = 0.002) Japanese > Mandarin (*p* = 0.001)
LH-HL	Japanese > English (*p* < 0.001) Korean > English (*p* = 0.016) Japanese > Mandarin (*p* = 0.020)
LM-HL	Japanese > English (*p* < 0.001) Korean > English (*p* = 0.002) Japanese > Mandarin (*p* = 0.025)
LH-MH	Japanese > English (*p* < 0.001) Korean > English (*p* = 0.030) Japanese > Mandarin (*p* < 0.001)
LM-LH	Japanese > English (*p* = 0.001) Korean > English (*p* = 0.034) Japanese > Mandarin (*p* = 0.007)
LM-MH	Japanese > English (*p* = 0.001) Japanese > Mandarin (p = 0.001)
Segment overall	Japanese > English (*p* = 0.002) Korean > English (*p* = 0.036) Japanese > Mandarin (*p* < 0.001) Korean > Mandarin (*p* < 0.001)
HF (control)	Japanese > English (*p* = 0.013) Japanese > Mandarin (*p* < 0.001) Korean > Mandarin (*p* = 0.002)
HL (control)	Japanese > English (*p* = 0.011) Japanese > Mandarin (*p* < 0.001) Korean > Mandarin (*p* =0.001)
LH (control)	Japanese > English (*p* =0.021) Korean > English (*p* = 0.038) Japanese > Mandarin (*p* < 0.001) K > M (*p* = 0.001)
LM (control)	Japanese > English (*p* = 0.001) Korean > English (*p* = 0.036) Japanese > Mandarin (*p* < 0.001) Korean > Mandarin (p = 0,001)
MH (control)	Japanese > English (*p* = 0.003) Korean > English (*p* = 0.46) Japanese > Mandarin (*p* < 0.001) Korean > Mandarin (*p* = 0.001)

Analysis by Korean Subgroup

To examine the accuracy data on each individual condition of the two Korean dialects, a GEE model declared the fixed factors Language (i.e., two Korean dialects) and Subcondition (individual tonal comparisons for both the tone and segment subconditions), with Subjects as a random factor. The Type III tests of fixed effects showed a main effect of Language ($\chi^2(1)=40.20$, $p < 0.001$), a significant effect of Subcondition ($\chi^2(9) = 500.08$, $p < 0.001$), and a marginal interaction between the two factors ($\chi^2(8)=13.38$, $p = 0.10$). There was no significant difference between the two Korean dialects on the accuracy rates of the individual tone comparisons on the ABX disyllabic task. However, both groups did not show significant within-group differences on the accuracy rates between the test subconditions with p-values on each comparison greater than 0.005.

For reaction times (RTs) for each individual comparison between the two Korean dialects, a Linear Mixed Effects model was run on the LogRTs with Language (i.e., Dialects: standard Korean, Kyungsang Korean) and Subcondition (individual tonal comparisons for both the tone and segment subconditions) as fixed factors, with Subjects as random factors. When looking at the Type III tests of fixed effects, the F-tests showed a main effect of Subcondition ($F(9, 1764) = 39.86$, $p < 0.001$), no significant effect of Language ($F(1, 23.05) = 0.62$, $p = 0.438$), and no significant interaction between the two factors ($F(9, 1764) = 1.11$, $p = 0.351$). There was no significance in RTs between the two Korean dialects on the individual tone comparisons. However, both groups did not show significant within-group differences between the RTs on either the test or control subconditions as seen on Univariate tests.

4.3.3 Discussion

In response to the overarching research question as to whether the varying presence of certain features in the L1, specifically lexically-contrastive pitch, aid in the naïve perception of non-native tone, the results for overall accuracy rates from the disyllabic task appear to substantiate this viewpoint. As to whether this subsequently results in a linguistic hierarchy of perceptual ability, again the answer appears to be affirmative as evidenced in the overall accuracy results. However, a linguistic hierarchy is not evidenced to the detailed extent predicted, i.e., there is not a clear-cut statistical difference on performance between *all* five groups tested.

In contrast, performance on individual disyllabic tone comparisons diverges at times from the overall disyllabic tone accuracy results with subtle differences. First, there are two individual tone conditions (e.g., Low+Mid vs Mid+High, Low+Mid vs Mid+High) where all groups patterned similarly with no statistical differences. On the other three individual tone conditions, however, the language groups display a mixture of language-specific variety with statistical differences not necessarily reflective of the performance on overall tone perception. For example, on the Low+High vs Mid+High comparison, the results are “mixed”: all groups are less accurate on this comparison, but the difference in accuracy levels is statistically significant between some groups. In.short, we see that despite differences in overall levels of accuracy, the pattern of performance across the specific comparisons is similar in all groups. However, we observe language-specific differences as well, specifically in the disparity in accuracy levels on a given comparison between language groups. For example, only the Mandarin, Japanese and Korean groups performed at lower levels on the Low+Mid vs Low+High comparison compared to three out of the four remaining comparisons (except the Low+High vs Mid+High). Therefore, to account for both the universal trend and the subtle language-specific tendencies, it is necessary to determine the exact nature of lexically-contrastive pitch in each L1, particularly whether the disyllabic stimuli allow the English group (as well as possibly the other groups) to apply the pitch differences of word stress to Thai tone perception.

The RTs offer a more nuanced picture of tone perception as shaped by the L1. On the overall tone condition a clear trend appears: The L1 Japanese group is faster than the L1 English and L1 Mandarin groups while the L1 Korean group is faster than the L1 English. However, as this trend is also exhibited on the overall segment condition, it appears that condition does not seem to matter and that this is perhaps merely a type of strategy. As for individual tone comparisons, we see the same trend, but with slight variation on certain individual tone comparisons. In a few cases, we see that the L1 Korean group is not faster than the L1 English group and/or is even faster than the L1 Mandarin group. Thus, RTs like accuracy rates on the individual tone comparisons demonstrate both universal and language-specific tendencies.

Several questions concerning the results may provide clues for these tendencies. First, we ask why the overall accuracy rates did not define a more robust difference between the groups. In fact, there is only one statistically significant outcome on overall tone accuracy: The Thai group outperformed the English group. This result puts the Thai group at the top of a proposed perceptual hierarchy and the English group at the bottom end, with the other groups of Mandarin, Japanese and Korean in between. Moreover, the overall RT results do not necessarily clarify the hierarchy either. For example, the one difference between the RTs on the two overall tone and segmental conditions is that in addition to the statistical differences on the overall tone RTs, the Korean group is also faster than the Mandarin group on the control condition. This faster speed of the Korean group over the English group along with the Korean group being unexpectedly more accurate overall is also unexpected. This result, therefore, reinforces the case of the lower saliency of pitch in English word stress vis-à-vis the other correlates (cf., Kochanski et al., 2005), equating the saliency to that of having no lexically-contrastive pitch.

Second, we must consider the original premise for including disyllabic stimuli in the current study: to reflect English word stress and therefore, allow the English group to exploit English word stress in their perception of Thai tones. Yet, this appears not to be the case considering both the accuracy and RT scores of the English group. However, disyllabic stimuli may have helped all the groups. For example, the Japanese may also require disyllabic stimuli as it is reflective of Japanese pitch accent usage. It was surmised that having adjacent tones to reflect

L1 usage of word stress and pitch accent would make perception easier. While this may be the case for the Japanese, the same appears not to be true for the English speakers. Indeed, the possibility that having to hear six tones (i.e., three tokens of disyllabic stimuli) in one trial may actually be more confusing (i.e., greater cognitive load) cannot be discounted.

Third, in order to obtain a faithful description of perceptual ability in any one language group, we must not merely consider performance vis-à-vis other groups on individual tones, but also performance vis-à-vis the other tone comparisons within one language group. For instance, some tones or tonal comparisons are easier to perceive than others (cf., pitch height or pitch direction) in the naïve perception of non-native tone. Performance on each individual tone comparison cumulates to determine overall performance while simultaneously indicating where perceptual "weaknesses" arise. For example, we would intuit that the striking difference between the two tone patterns in Low+High vs High+Low disyllabic tonal comparison alone should help any listener, specifically as the two very different contours are reflective of word stress. Yet, this is precisely where the English and Korean speakers differ statistically from the Thai group. However, on closer examination we see that the English speakers did not do poorly on this tonal comparison (*M* = 80.7% accuracy) compared to their accuracy rates on the other comparisons (i.e., 63%, 81.8%, 80.2%, 81.8%). In the case of the Koreans, their performance was middling (*M* = 85.3%) with a low of 70.1% for the Low+High vs Mid+High comparison to a high of 92% for the Low+Mid vs Mid+High comparison. Thus, to explain this statistical difference between the groups we see that it is rather that the Thai group performed their best on this comparison.

Fourth, performance on individual tone comparisons illuminates language-specific differences in how tones may be perceived and thus, possibly processed. To illustrate, there is only one significance in accuracy on the Low+Mid vs Low+High comparison: the Thai group outperformed the Japanese group. This may be expected as Japanese uses either Low or High pitch in their phonology to differentiate moras in minimal pairs where the relative height of the high mora may phonetically vary to actually be “mid” compared to other high pitched moras. That is, a Mid tone is perceived as a High tone due to downstepping of high pitches in Japanese so that in Japanese there may only be [+High]. Thus, phonologically the Japanese may equate the Mid tone to the High tone and therefore, have difficulty in distinguishing the two and merely note that the second syllable is higher than the first syllable. We perhaps see the same phenomenon of the Japanese possibly equating the Mid tone as an allotone of the High tone on the Low+Mid vs Mid+High comparison, thus hearing what would be the equivalent of Low+High vs High+High. Indeed, the Japanese performed equally as well as the Thai (91.9% and 91.7%, respectively) on this comparison. Yet, this interpretation of the Japanese categorizing the Mid tone as an “allotone” of the High pitch in Japanese seems not to bear out when looking at their performance on the Low+High vs Mid+High comparison. In fact, the Japanese performed at a much lower rate of accuracy on the Low+High vs Mid+High (71.7%) than on the Low+Mid vs Low+High (77.2%). Thus, regarding the Mid tone merely as an allotone of the High tone is not a sufficient explanation. Additionally, we may have to add another factor, namely endpoint and/or register. Examining all three tone comparisons, we see that only the Low+High vs Mid+High comparison ends at the same point of High and contains a disyllabic tone pattern crossing registers, i.e., from Low to High. It should be noted that the Thai group performed poorly on the same Low+High vs Mid+High comparison. This can be attributed to the difficulty the Thai have in hearing the difference between the Low and Mid tones which are phonetically close (i.e., Abramson, 1976) so that they may hear Low+High vs Low+High instead. The phonetic similarity between the Low and Mid tone combined with the fact some of the disyllabic samples cross over registers (low to high) and the fact that the Mid tone may be an allotone for the High tone may account for the difference in performance on these three tone comparisons.

Turning to RTs of individual comparisons we see an overall trend regardless of whether for the tone condition or segmental condition, both overall and for individual tone comparisons: J, K > E, and J > M. There are few deviations from this pattern with the exception of a few cases where the Korean group was faster than the Mandarin group. Both the Japanese and Korean group may strategize speed over accuracy. In comparison, the Mandarin group may take more time as they are attempting to map their L1 tones onto "similar" Thai tones. As for the English group it may be merely that the functional load of the pitch correlate of word stress in English is indeed less salient than that of length, intensity and/or vowel quality. Thus, the mechanisms behind the similar slow reaction times of the English and Mandarin group are different.

In short the perception of Thai tones is shaped by the L1 with accuracy rates commensurate to the functionality of lexically-contrastive pitch in the L1. As such, a simple hierarchy of at least a top and bottom group emerges with possibly a three-tiered hierarchy when looking at individual tonal performance in terms of accuracy rates and RTs. However, individual tonal comparisons paint a slightly different picture revealing both universal and language-specific tendencies. In sum, as for the question motivating the use of disyllabic stimuli in this task, it appears that disyllabic stimuli are more difficult to perceive with seemingly little aid from word stress for the English group. These observations and interpretations thereof will be discussed in more detail in the next section comparing the results from both the monosyllabic and disyllabic ABX tasks.

4.4 Summary and general discussion

The monosyllabic and disyllabic ABX task determined that speakers of a different L1 vary in their naïve perception of a non-native tone system. Results also appear to indicate that the varying degree to which language groups perceive Thai tones corresponds to the varying presence of lexically-contrastive pitch in the L1. That is, in line with the Feature Hypothesis (McAllister et al., 2002), the greater the functionality of lexically-contrastive pitch in the L1, the greater the perceptual ability for Thai tones. Subsequently, the varying functionality of lexically-contrastive pitch in the L1 results in a linguistic hierarchy of perceptual ability among these various L1s. This conclusion agrees with that of the pilot study. Disyllabic stimuli do not appear to aid L1 English speakers in their naïve perception of tone; In fact, disyllabic stimuli appear to result in lower perceptual accuracy rates. Examining the perception of individual tone patterns has revealed that indeed some tones or tonal comparisons are easier or more difficult while demonstrating both universal and language-specific tendencies in the naïve perception of Thai tones. Simultaneously, analyses of individual tone comparisons defy a Feature Hypothesis framework, given the different nature of tone (i.e., large inventory of tone patterns) as compared to the target of the McAllister et al. study, i.e., vowel length.

Overall findings and comparison to the pilot study

On both the ABX monosyllabic or disyllabic task, we see that the Thai group generally performs at significantly higher levels of accuracy over the English group and to a lesser extent over the Japanese and Korean groups (i.e., Thai > Japanese, Korean on only the monosyllabic ABX task), hinting at a hierarchy of performance in the perceptual accuracy of Thai tones. As for reaction times, the Japanese group and to a lesser extent the Korean group perform at faster rates than the English group. The Japanese are also faster than the Mandarin group but to a lesser extent than they are against the English group. Taken together, we may interpret a hierarchy of performance as follows: Thai, Mandarin > Japanese, Korean > English. What runs counter to predictions is the performance of the Korean group over the English group in the overall hierarchy and the fast RTs of the Japanese and Korean groups vis-à-vis the Mandarin and English groups, respectively.

First, for accuracy rates, the Korean group's performance may be due to several reasons as discussed in the previous discussion sections: 1) influence of the Korean pitch-accent Kyungsang dialect, 2) influence of the emerging pitch accent of standard Korean, 3) influence of L2 English on the Korean group, and 4) the low saliency of pitch as a correlate of English word stress. Any of these may alone account for the results, or a combination of two or more of these reasons. In response, better analyses of the results of the current study, i.e., stricter filtering of the background of the Korean participants, may be necessary. The design of the current study might need to be either made to put more cognitive load on the participants to tease apart the groups and/or more finely designed to account for English word stress or dialectal features. Also, testing Korean participants with very low English language exposure and proficiency might be necessary. As for the low saliency of word stress in English, the use of disyllabic stimuli seems not to aid English speakers. Yet, we must still consider the possibility that the English group may have been able to exploit word stress in their perception of Thai tones, but that the other groups were able to do so as well, obscuring the effect in the English group's performance

Observations on the RTs point to several possible conclusions. It was expected that the Japanese group would be faster than the English group due to the functionality of lexically-contrastive pitch. However, we did not expect the Mandarin group to be slower than the Japanese group. Nor was it expected that the Korean group would be faster than the English group. Again, the Mandarin group may be slower due to the larger inventory of the tones in Mandarin while the Japanese may be faster due to the limited inventory of lexically-contrastive pitch in Japanese of merely a high vs low pitched mora. As for the faster rates of the Korean group vis-à-vis the English group, again it may be due to the possible reasons listed in the above paragraph. Thus, looking at RTs while we see the outline of a perceptual hierarchy of Thai tones according to the influence of lexically-contrastive pitch in the L1 in terms of functional load and inventory, we also witness unexpected phenomena that remain somewhat opaque in their cause.

Monosyllabic and disyllabic stimuli

Overall accuracy data for the disyllabic task paints a slightly different picture of Thai tonal perception among the five language groups than the overall accuracy data for the monosyllabic task. On the ABX monosyllabic task the Thai outperformed the Japanese, English and Korean groups but not the Mandarin group. In contrast, on the ABX disyllabic task the Japanese and Mandarin groups have in a sense "swapped positions". Combining the accuracy results of the two ABX tasks seem to place a divide between the Mandarin and Japanese groups along with the Thai group on one side and the English and Korean groups on the other side. This would paint the broad strokes of a perceptual hierarchy as influenced by the L1 based on the results of this task. Comparable performance in accuracy rates on the control condition of segments would bolster the claim that the test tone condition is influenced by the L1 and that the influence is most likely due to the presence/absence of lexically-contrastive pitch and its functional load and inventory. These results are, therefore, broadly in line with those from the pilot study in Chapter 3.

As for reaction times, the Japanese group and to a lesser extent the Korean group perform at faster rates than the English group. The Japanese are also faster than the Mandarin group but to a lesser extent than they are against the English group. These observations point to several possible conclusions. It is expected that the Japanese group would be faster than the English group due to the functionality of lexically-contrastive pitch. However, we do not expect the Mandarin group to be slower than the Japanese group. Nor do we expect the Korean group to be faster than the English group. Again, the Mandarin group may be slower due to the larger inventory of the tones in Mandarin while the Japanese may be faster due to the limited inventory of lexically-contrastive pitch in Japanese of merely a high vs low pitched mora. As for the faster rates of the Korean group vis-à-vis the English group, there may be several possible causes: influence of L2 English, emergence of pitch in Standard Korean, and most likely the low salience of pitch in English word stress.. Thus, while we see the outline of a perceptual hierarchy of Thai tones according to the influence of lexically-contrastive pitch in the L1 in terms of functional load and inventory, we also witness unexpected phenomena that remain somewhat opaque.

Additionally, it must also be kept in mind that while there may not be a difference in the perceptual accuracy level between the Mandarin and Japanese speakers, there is most likely a difference in how they perceive tone. That is, the Mandarin listeners might hear tones while the Japanese may hear just a high pitch vs a low pitch in strings of moras. For example, the Japanese appear to have done better on disyllabic task than monosyllabic task, suggesting disyllabic stimuli may have allowed this group to compare the pitches of two mora against one another for easier perception. Differences in how tone is perceived, particularly in disyllabic stimuli, may also apply to the performance of the Korean group, who are faster than the English group and just as accurate. Separating the two Korean dialect groups again did not show any significant difference between these two groups as was previously suggested by the pilot study analyses. A greater cognitive load on the disyllabic task was predicted for the same participants as compared to the monosyllabic task while also predicting “greater help” for L1 English speakers who may need two syllables to better hear the difference in tones as a reflection of word stress in English (cf., “a record” vs “to record”). However, results from the disyllabic task did not support this hypothesis as the results from the monosyllabic and disyllabic tasks did not differ.

Moreover, the statistically weaker results hinting at a performance hierarchy could be due to several design “flaws”. First, we compare the stimuli in the two tasks of the pilot study and current study. In the pilot study, the four-paradigm trials (i.e., ABB, ABA, BAA, BAB) used different segments for each of the four parts of the paradigm while in the current task the same segments were used for each of the four parts. Thus, in the pilot study AXB task, we have [ba:] HHR, [pu:] RRH, [su:] HRR, and [t^ha:] RHH, for example, while in the current ABX task, we have [no:i] used for all four trials of the four-part paradigm. This may have caused two differences: greater awareness to listen for tonal differences and an inadvertent introduction of a “short training” practice. The fact that the current study used only two different segment strings (i.e., [no:i], [p^huai]) for all the tonal comparisons may have caused the participants to focus on tone whenever the two types of segments were heard rather than having to expend “energy” on determining whether vowels or consonants were different (i.e., lower cognitive load).

Analyses of individual tone comparisons

The individual tone comparisons illuminate Thai tone perception somewhat differently from the overall accuracy rates and RTs for tone perception while simultaneously accounting for these overall tendencies. Among these comparisons, we see both universal trends in line with the overall accuracy rates and RTs, but also language-specific trends primarily for accuracy rates with fewer for the RTs for individual tone comparisons. In fact the RTs for the individual tone comparisons are generally more reflective of RTs for overall tone perception where the RTs trended the same on both the tone and segment conditions. As such, in contrast to the differences arising on the accuracy rates for overall tone perception and individual tone comparisons, the RTs provide less illuminating observations. More importantly, a greater nuanced picture of tone perception is seen when considering the results of the individual tone comparisons. It is here that we see the differences more clearly among the various L1s, obtaining clues as to what may be occurring in the perception of non-native lexically-contrastive pitch as shaped by the native lexically-contrastive pitch system. This perception in turn hinges primarily upon the conceptualization of what constitutes functionality of lexically-contrastive pitch in both the L1 and non-native systems.

Re-interpretating functionality in light of the results of the current study

As noted, the Korean group appears to perform at levels comparable to or surpassing that of the English group in defiance of the predicted hierarchy of accuracy in the naïve perception of Thai tones as shaped by the functionality of lexically-contrastive pitch in the L1. While several possibilities are explored and/or put forth in this section to account for this result, there is one remaining plausible reason. We might redefine the functionality of lexically-contrastive pitch in both Korean and English and possibly Japanese.

Functionality should be reconceptualized, taking into account whether pitch is employed exclusively to mark word stress and whether when working in tandem with other correlates, pitch is the most salient of the correlates or merely marginal (i.e., English). Additionally, we need to examine the difference between languages possessing minimal pairs that differ only by pitch and languages which do not possess minimal pairs that differ only by pitch, but yet feature various pitch patterns for lexical items (i.e., Korean). In short, English may be equivalent to a language without any lexically-contrastive pitch, i.e., Korean, especially as Korean may actually feature pitch, however marginal, as a feature marking prominence in lexical items.

Additionally, in the case of Japanese, we must consider the functional load of lexically-contrastive pitch. It is said that only 14% of minimal pairs in Japanese are distinguished by pitch as opposed to minimal pairs that are distinguished by segmentals (Shibata & Shibata, 1990). In contrast, the functional load of tone in Mandarin is considered to be much higher than that of word stress in English. Furthermore, the functional load of tone in Mandarin is considered as high as the functional load of vowels in Mandarin. However, each tone pattern in Mandarin is considered to carry a varying degree of functional load (Surendan & Levow, 2004). While it is acknowledged that the greater degree of functional load of tone in Mandarin is higher than that of pitch accent in Japanese, what does this actually say about the functional load of pitch accent or tone for that matter? While 14% may seem like a low number, one must ask what the percentage of minimal pairs in English, for example, are distinguished by the [i] and [ɪ] vowels and whether a low number for the functional load of this vowel pair would make these two vowels somehow less salient and therefore, more difficult to perceive. Thus, it might follow that while the functional load of lexically-contrastive pitch used to distinguish minimal pairs in Japanese is low, this does not in fact make the perception of these pitch pattern types any less salient than segmentals even if pitch may seem less stable than vowels or consonants. It is a question to ponder when analyzing the results of this study. In connection, pitch patterns must be considered for the possible number of patterns, similarity between patterns, and/or prevalence of pitch types (e.g., contour, level) influencing possible weighting.

Thus, when defining the functionality of pitch, three additional properties should be taken into account as well: 1) *exclusivity* to signal lexical contrast, 2) *functional load*, and 3) *inventory of pitch patterns*. The last property will be discussed along with the second property in the following section as they both bear upon the forging of the framework needed to analyze the results of the current study.

Interpretation of results: Feature Hypothesis, PAM and PIM

Applying the Feature Hypothesis (McAllister et al., 2002) to overall tone perception appears relatively straightforward: the greater the functionality of the lexically-contrastive pitch in the L1, the greater the salience of pitch in the L1 and therefore, the greater the ability to perceive non-native lexically contrastive pitch. Thus, the interpretation is built on the presence or absence of lexically-contrastive pitch in the L1. However, there are several points which need to be clarified under such a model.

First, lexically-contrastive pitch in the form of tones exhibits several patterns. As such, when considering results on the individual tone comparisons within a Feature Hypothesis framework, the question arises as to what is being measured in the various L1s. Do we measure the presence or absence of each tonal pattern? If so, to what extent can a tone pattern in the L1 match the counterpart tone pattern in the non-native language? Under the Feature Hypothesis reconciling this fact is somewhat difficult. In the pilot study, corollaries or properties as mentioned above could be added to address this situation, e.g., inventory of pitch patterns.

Additionally, we may be able to apply the PAM model (Best, 1995) to remedy this situation. Under PAM, we would be able to address whether L2 tone patterns might be mapped onto L1 tone patterns depending on the degree of similarity between the two. Under PAM we question then whether pitch accent and word stress have categories and whether listeners can access and map target non-native tone patterns onto their L1 pitch patterns. So & Best (2010) have made attempts to do so for Japanese, but there do not appear to be any studies that have tried to seriously do so for English word stress. However, Francis et al. (2008) believe that L1 English speakers should be able map target non-native tone patterns onto the intonational categories of English, e.g., statement and question intonational categories. Several thoughts come to mind in response to such attempts. First, we must test whether all languages may access intonational categories despite being a different domain from that for word stress. Also, this assumes that a Falling tone or High+Low pattern would be quite salient for English speakers and yet, in this study we saw that on these patterns the English group still underperformed other groups. Naturally, we must assume also that the other language speakers may access similar question intonation (albeit with differences) which may neutralize the English group's advantage. Second, there is the question as to whether the Japanese listeners indeed do have LH, HL, and HH lexical categories as is assumed in the So & Best study (2010). Finally, the PAM model (Best, 1995) is built on listeners hearing similarities between non-native and native sounds. Yet, it is not clearly defined what exactly listeners must hear that defines whether or not a non-native sound is similar or not to an L1 sound? This study does not by all means discount the interpretations for non-native tone by the above-mentioned studies. Rather, the current study wishes to explore another possible interpretation based on features to account for what listeners may rely on to determine similarity.

We recall the Phonological Interference Model (PIM) in the Brown study (1998) addressed in the Background section of this study. Referencing this model, we might consider basing interpretation of the results by the participants in this study on the feature of [+ high pitch] for example. Doing so, would allow us to apply the Feature Hypothesis when this feature rather than on a more nebulous category such as tone. This approach I believe would be more in line with the original intent of the researchers of the Feature Hypothesis as they examined vowel length. If tone were then seen as comparable to vowels and consonants, then, it is only natural that we break it down into features to better analyze. Thus, in the PAM model we would still be able to consider tone patterns, but in order to determine what the listeners are hearing to determine the similarity between the L1 and non-native target sound we should examine the features of these sounds. As such, when looking at segmentals, we might look at the extent to which [+round] for vowels or [+nasality] for consonants is employed in one language when interpreting the perception of speakers of this language on non-native sounds employing one of these features more extensively. In this respect, PIM may bridge the two models of the Feature Hypothesis and PAM and provide a greater insight into what constitutes similarity between the L1 and target non-native sound. In fact, PIM and Feature Hypothesis both are similar in that they rely on features, with PIM (Brown, 1998) being put forth earlier than the Feature Hypothesis (McAllister et al., 2002). Thus, I propose combining these two models with PAM resulting in what we might call the PIM-PAM approach.

The results of the Japanese group on their naïve perception of Thai tones may illustrate the possibilities of such a combined approach. We recall that despite having a smaller inventory and functional load for pitch accent vis-à-vis those in a tonal language like Mandarin to the extent of being considered an "impoverished tone language," the Japanese group performed at accuracy levels at times equal to those of the Mandarin group and with RTs more than often faster than those of the Mandarin group. Assuming that Japanese indeed has the three pitch accent categories of LH, HL, and HH (although the HH pattern may still be debatable), we then would have difficulty in determining why when compared to the Mandarin group the Japanese at times are equal in terms of accuracy while surpassing in terms of speed. That is, the Mandarin rising tone, falling and high tones would be the LH, HL and HH pitch accent categories, respectively. Thus, the Mandarin group would have the additional pattern of the dipping-rising tone and possibly the "no-tone" pattern. It is possible that extra tones could cause the Mandarin group to be slower. As to whether this would cause the Mandarin group to be more or less accurate is subject to interpretation. Also, if we look at the individual tone comparisons we see that while the two languages pattern similarly at times (e.g., Low vs Rising comparison differ statistically from the Thai group for both groups), they also diverge (e.g., Low vs Mid for only Mandarin differs statistically; Low vs High for only Japanese differs statistically). Again, this may reflect differences in the inventory of categories or could be attributed to another approach for Japanese pitch accent patterns, namely a string of low and high pitches.

While this current study does not discount again a unitary approach to the analysis of tone perception, we must also consider a compositional approach as well. In fact, a compositional approach combining the Moraic Alignment for Thai tones (Morén & Zsiga, 2006) appears to align perfectly with what the Japanese group may be doing in their perception of Thai tones. Japanese features either a high or low pitch on a mora which interacts with the syllable (e.g, heavy syllables composed of L+H on the first and second mora of a word becoming H+H). Such an interpretation is generally supported by the results on the ABX monosyllabic task where. Yet, there are possibly a few loose ends: there are a few cases difficult to resolve under such an approach, and the fall in pitch can be considered to be the most salient feature of accent in Japanese (Sugiyama, 2012).

Finally, just as we must consider speakers of two dialects, we should consider speakers of two languages. The following chapter examines L1 English speakers who have learned Mandarin Chinese as an L2. This group took part in two experiments: they first participated in the same ABX task with Thai tones described in this chapter, and also took part in a lexical task in their L2 Mandarin. The objective was to determine 1) the possible influence of the acquired L2 Mandarin tones of this group on the naïve perception of Thai tones; and 2) the extent to which L2 learners of Mandarin succeed in encoding tonal information in lexical representations of Mandarin words.

5. L2 Learners' perception and acquisition of tones

5.1 Introduction

This chapter examines the naïve perception of Thai tones by L1 speakers of English who have learned another tonal language as an L2, i.e., Mandarin Chinese. Additionally, this chapter looks at the lexical encoding of Mandarin tones by the same participants. In total four experiments are reported in this chapter: 1) Thai ABX monosyllabic task (Section 5.2), 2) Thai ABX disyllabic task (Section 5.3), 3) Mandarin Chinese Lexical Decision task with Repetition Priming (Section 5.5), and 4) Mandarin Chinese Phonological Pronunciation task (Section 5.6). This L1 English/L2 Mandarin group is compared to a native English group (with no learning experience of Mandarin Chinese) and to a native Mandarin group. These tasks were devised in order to answer the following overarching research question:

1. RQ2: Does the learning of a tone language aid in the naïve perception of the tones of another non-native tone language?

In relation to this first research question, the following related subquestions are also asked:

1.1 RQ2.a: Do L2 learners of a tone language differ in their naïve perception of another tone language as compared to participants with the same L1 who are not learning a tonal L2?

1.2 RQ2.b: Do L2 learners of a tone language perform at levels comparable to L1 speakers of that tone language in a linguistic hierarchy of perceptual ability of another tone language?

1.3 RQ2.c: Do L2 learners of a language perform in a similar manner to L1 speakers of that one language, i.e., Are some tones or tonal comparisons easier to perceive than others (cf., pitch height or pitch direction) in the naïve perception of another tone language? Do the two groups react in the same manner?

The tasks also ask another overarching question:

2. RQ3: Are L2 learners of a tone language able to lexically encode tonal information to constrain lexical access?

The results of the ABX tasks will answer research questions 1 and 3. A further comparison between these three groups on their performance on individual tone comparisons will answer questions 2 and 4, to determine which tones are easier or more difficult for all three groups.

Regarding Research Question 5, the L2 Learner group participated in a Lexical Decision task with Repetition Priming to determine whether Mandarin tonal contrasts are acquired and encoded lexically in participants' lexical representations.

In section 5.6, a pronunciation proficiency task was added to provide a means of measuring pronunciation proficiency across the learners despite differences in course levels, experience, etc. among these learners.

5.2 Experiment 2a: ABX Thai tones, monosyllabic

We now turn to the methodology and results for a monosyllabic ABX task which tested Learners of Mandarin as an L2 on their naïve perception of Thai tones. Their results are compared to that of native speakers of Mandarin and that of native speakers of English who have not learned Mandarin as an L2.

5.2.1 Methodology

5.2.1.1 Participants

Participants were recruited from Mandarin Chinese language courses, both summer and academic-year courses. Students were required to have completed at least second-year Mandarin Chinese at Indiana University. In total, 23 learners were recruited (L2 Mandarin/L1 English=23, females=10). The ages ranged from 18-53 (average age: 24.4). On the whole, participants were either undergraduate or graduate students at Indiana University. For example, approximately half of the Thai participants were professors at a university in Thailand while some of the other language participants were spouses of students and other non-students employed in the local community (See demographic and language background information presented in Appendix H, pp. 305-314).

5.2.1.2 Stimuli and conditions

The stimuli and conditions were the same as in Experiment 1a.

5.2.1.3 Procedure

The procedure was the same as in Experiment 1a.

5.2.2 Results

The data of the learners were coded and screened for outliers in the same way as for the other groups in Experiment 1a. Similarly, RT data were also log-transformed, and the same analysis models were applied for this set of data.

Reaction times (RT) that were 300ms and below and any over 2500ms were removed. In total, only 0.00014% of the RT datapoints were removed (N = 3). The RT data were examined and were found to be skewed. Hence, a log-transformation was applied to the RT (Log-RT) to obtain a normal distribution. Both the Log RTs and the “more-understandable” converted RTs are reported with their lower and upper confidence intervals (CI) in all the following results tables in this chapter.

One participant indicated that they had misunderstood the ABX task and had essentially done an “XAB” task and not an ABX task. Their score was below two SD from the Learner group mean on the over test condition on the ABX monosyllabic task and therefore, their data were excluded. A second outlier was found among the Mandarin speakers who also performed beyond 2SD from the accuracy means on the control condition. The data from these two participants was also removed from the RT analysis. In addition, data for two learners who were born in China and then, later adopted and grew up speaking English (cf., Pierce et al., 2014) were removed. Also, data from three learners who were heritage speakers were removed due to long-term exposure to a tonal language since birth.

After cleaning the data, a Generalized Estimating Equation (GEE) model for a binary response was fitted to the data and run for overall accuracy as the data structure is categorical. A Linear Mixed Effects model was fitted to the data and run on the LogRT. Language (Mandarin, Learners, English) and Condition (test vs control) were declared as fixed factors with Subjects declared as a random factor. Sidak correction for multiple comparisons was used over Bonferroni. Both the Log RTs and the RTs in ms are reported with their lower and upper confidence intervals (CI) in all the following tables.

5.2.2.1 Analysis by condition: Test and control

Since we are mainly interested in observing the effect of acquiring a tonal L2 on non-native tone perception (Thai), this section compares the data of the Learners to the performance of the Mandarin and English speakers from Experiment 1a. For convenience, their data are repeated here. For each group, a mean accuracy score and mean RTs were computed for the overall test and control conditions and each subcondition (i.e., individual tonal comparisons).

Accuracy rates on the test condition were lower than those on the control condition across all groups. The learners' accuracy rates were similar to those of the Mandarin group on the test condition. On the control condition the Mandarin and English groups performed at lower accuracy rates than the Learners.

Table 5.1: Mean accuracy (%) for each group on test and control conditions

	Accuracy	
Language group	Test (SE)	Control (SE)
Mandarin (n=30)	85.0 (1.6)	93.6 (1.0)
Learners (n=18)	86.3 (2.3)	96.1 (0.8)
English (n=23)	79.5 (2.6)	93.5 (1.4)

Note: SE = standard error.

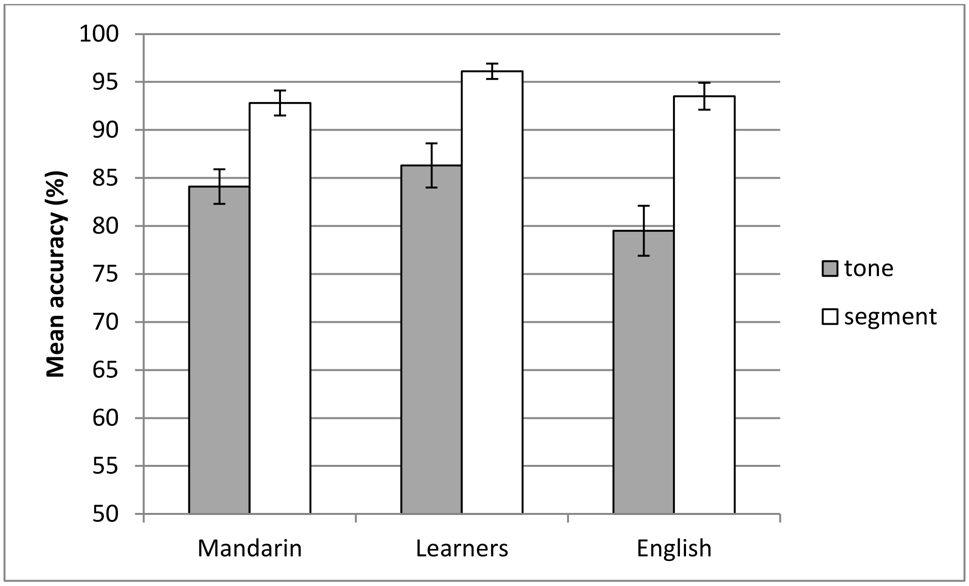


Figure 5.1: Mean accuracy (%) for each group on test and control conditions. Error bars represent +/-1 SE.

The un-aggregated data were used for statistical analysis. Accuracy data are analyzed first with RT analysis following.

Accuracy Data

To examine the accuracy data on the test and control conditions, a GEE model was fitted with the fixed factors Language (Mandarin, Learners, English) and Condition (test vs control), with Subjects as a random factor. The Type III tests of fixed effects showed no effect of Language ($\chi^2(2) = 4.15$, $p = 0.126$), a significant effect of Condition ($\chi^2(1) = 151.31$, $p < 0.001$), and no significant interaction between the two factors ($\chi^2(2) = 4.39$, $p = 0.111$).

Again, interaction shows that the groups did not differ from one another statistically on either the Test or Control condition. That is, there was no between-group difference in accuracy rates for either test, $\chi^2(2) = 4.26$, $p = 0.119$, or control condition, $\chi^2(2) = 4.86$, $p = 0.088$.

However, the main effect of condition demonstrates that there is a difference between the test and control condition for all groups. We see that there are differences in the mean accuracy rates between the Test condition (Mandarin: $M = 85\%$, Learners: $M = 86.3\%$, English: $M = 79.5\%$)

versus those on the Control condition (Mandarin: *M* = 93.6%, Learners: *M* = 96.1%, English: *M* = 93.5%) (See Table 5.1). Univariate tests revealed that performance differed significantly between conditions for all three groups: English ($p < 0.001$), Learners ($p < 0.001$), Mandarin ($p < 0.001$).

RT Data

Mean RTs and LogRTs for each group and condition are presented in Table 5.2. In the test condition, the learner group and Mandarin group performed at fairly equal speeds, and were faster than the English group. On the control condition, both the English group and the Mandarin group were slower than the learner group.

Table 5.2: Mean RTs (ms) for each group on test and control conditions

	RT		Log RT	
Language group	Test (CI)	Control (CI)	Test (CI)	Control (CI)
Mandarin (n=31)	1088 (1040; 1137)	951 (906; 997)	3.04 (3.02; 3.06)	2.98 (2.96; 3.0)
Learners (n=17)	1053 (989; 1120)	855 (790; 926)	3.02 (3.0; 3.05)	2.93 (2.9; 2.97)
English (n=23)	1170 (1108; 1236)	991 (923; 1063)	3.07 (3.05; 3.09)	3.0 (2.97; 3.03)

Note: CI = Confidence interval (lower; upper).

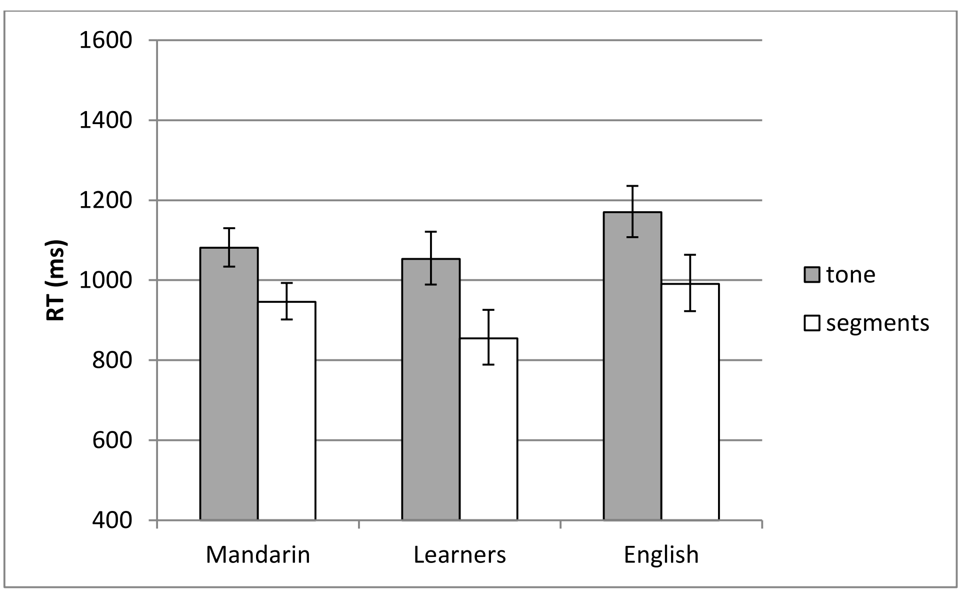


Figure 5.2: Mean RT (ms) for each group on test and control conditions. Error bars represent the upper and lower CI from RTs from a Linear Mixed Effects model.

A Linear Mixed Effects model was run on the LogRT with Language (Mandarin, Learners, English) and Condition (test vs control) as fixed factors, declaring Subjects as a random factor. For reaction times (RTs) when looking at the Type III tests of fixed effects, the F-tests showed a main effect of Condition ($F(1, 10117) = 845.2$, $p < 0.001$), a significant effect of Language ($F(2, 68.01) = 3.99$, $p < 0.023$), and a significant interaction between the two factors ($F(2, 10117) = 13.26$, $p < 0.001$).

There is a main effect of condition between the mean RTs on the Test condition (Mandarin: $M = 1081$ ms, Learners: $M = 1053$ ms, English: $M = 1170$ ms) and the mean RTs on the Control condition (Mandarin: $M = 946$ ms, Learners: $M = 855$ ms, English: $M = 991$ ms). On the test condition, the Learners were significantly faster than the English group ($p = 0.037$), but not faster than the Mandarin ($p = 0.874$). Additionally, the Learners were faster than the English on the control condition ($p = 0.021$) (See Table 5.2). Univariate tests demonstrate that the performance of each group differed significantly between conditions: English ($p < 0.001$), Learners ($p < 0.001$), Mandarin ($p < 0.001$).

The interaction shows that groups differ from one another on either the Test or the Control condition (i.e., between-group difference). Univariate tests show that performance differed significantly between groups on both the test condition ($F(2, 78.82) = 3.45$, $p = 0.037$) and control condition ($F(2, 77.73) = 3.62$, $p = 0.031$). The Learners were faster than the English group on both the tone condition ($p = 0.037$) and segment condition ($p = 0.021$).

5.2.2.2 Analysis by subcondition

We now turn to the comparison of individual tones for both accuracy and RTs. We examine the individual tonal comparisons to determine whether the perception of one tone type may be more difficult than perceiving another tone type. (Table 5.3 presents the individual accuracy means for test and control items in each subcondition for each group while Table 5.5 presents the RTs).

Accuracy Data

Accuracy rates for the test condition among the three groups showed a wider range of scores. In contrast, the accuracy rates for the control condition were much more uniform with higher overall scores than those on the test tone conditions. There are simultaneously similar tendencies among all three groups while at the same time some language-specific variation between the three groups. For example, all three groups performed at almost the lowest level of accuracy on the Low vs Mid and Low vs Rising comparisons. In contrast, on the High vs Falling comparison, the Mandarin and Learner group performed similarly while the English group performed at a much lower level of accuracy (See Table 5.3 and Figures 5.3a-c).

Table 5.3: Mean accuracy for each group on each tonal comparison, and statistically significant differences between groups in accuracy rates

Test subconditions	Mean accuracy (SE)		
	Mandarin	Learners	English
FR	94.2 (1.4)	92.7 (2.1)	90.6 (2.6)
LF	90.4 (2.2)	93.4 (2.4)	86.5 (3.7)
LR	75.0 (3.2)	78.7 (5.5)	64.6 (4.3)
MF	86.3 (2.5)	85.3 (2.8)	83.9 (3.4)
MR	92.5 (1.7)	89.7 (2.4)	84.9 (3.0)
HF	92.1 (2.1)	91.9 (2.5)	80.7 (4.1)
HR	85.0 (2.9)	89.0 (3.1)	77.6 (4.2)
LM	71.7 (3.6)	77.2 (4.8)	72.4 (4.4)
LH	84.6 (2.8)	89.0 (3.1)	79.7 (4.4)
MH	77.9 (3.2)	75.7 (4.2)	74.5 (3.2)
Control subconditions			
F	94.2 (1.4)	97.8 (1.0)	93.5 (1.7)
R	92.3 (1.4)	94.5 (1.6)	93.5 (1.5)
L	93.3 (1.3)	97.1 (0.9)	93.2 (1.6)
M	94.4 (1.0)	94.5 (1.3)	92.7 (2.0)
H	94.0 (1.4)	96.7 (1.2)	94.8 (1.4)

Note: SE=standard error.
T = Thai; M = Mandarin, J = Japanese, E = English, K = Korean; L = low tone, M = mid tone, H = high tone, F = falling tone, R = rising tone. Hence, LH = Low tone vs High tone comparison, HF = High tone vs Falling tone comparison, etc.

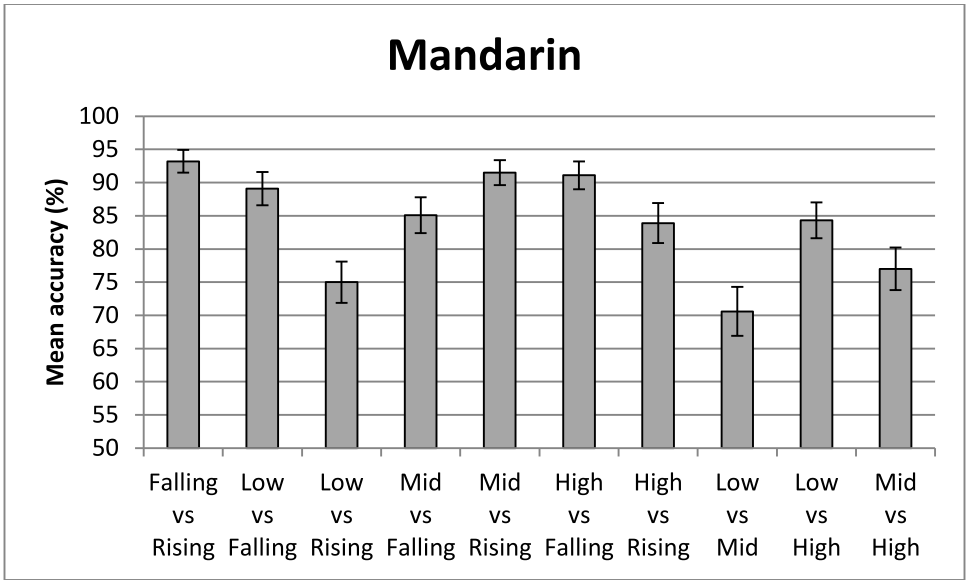


Figure 5.3a: Mean accuracy (%) for the Mandarin group on each test subcondition. Error bars represent +/-1 SE.

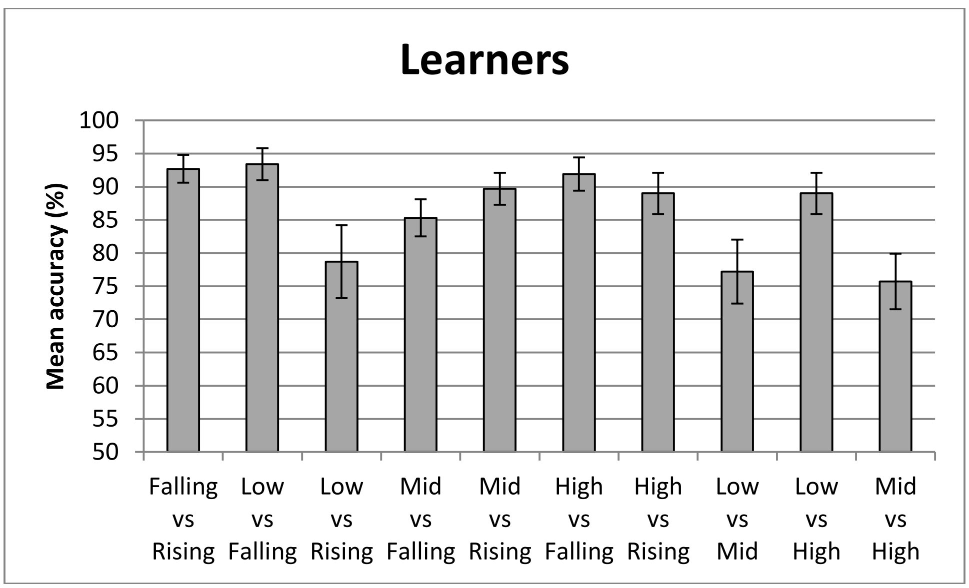


Figure 5.3b: Mean accuracy (%) for the Learner group on each test subcondition. Error bars represent +/-1 SE.

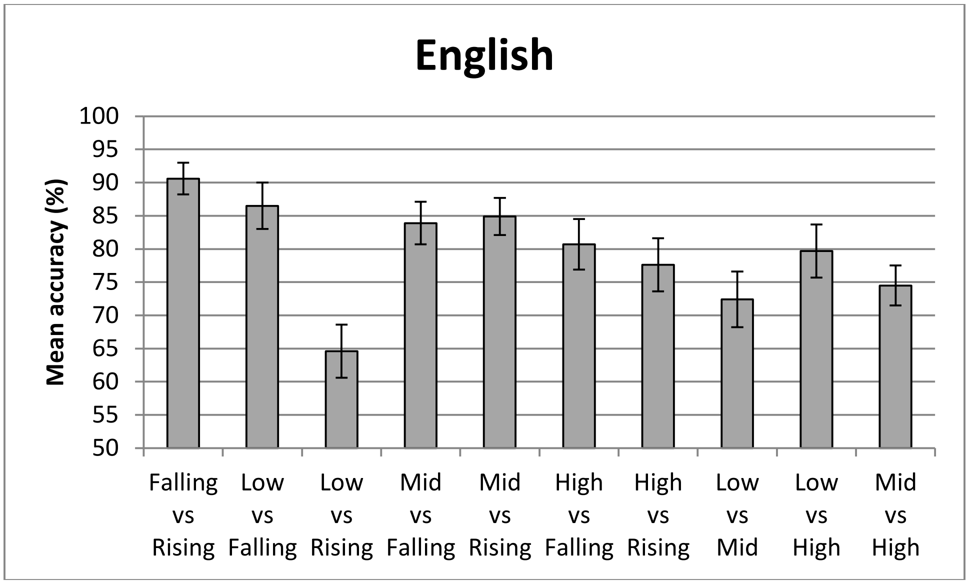


Figure 5.3c: Mean accuracy (%) for the English group on each test subcondition. Error bars represent +/-1 SE.

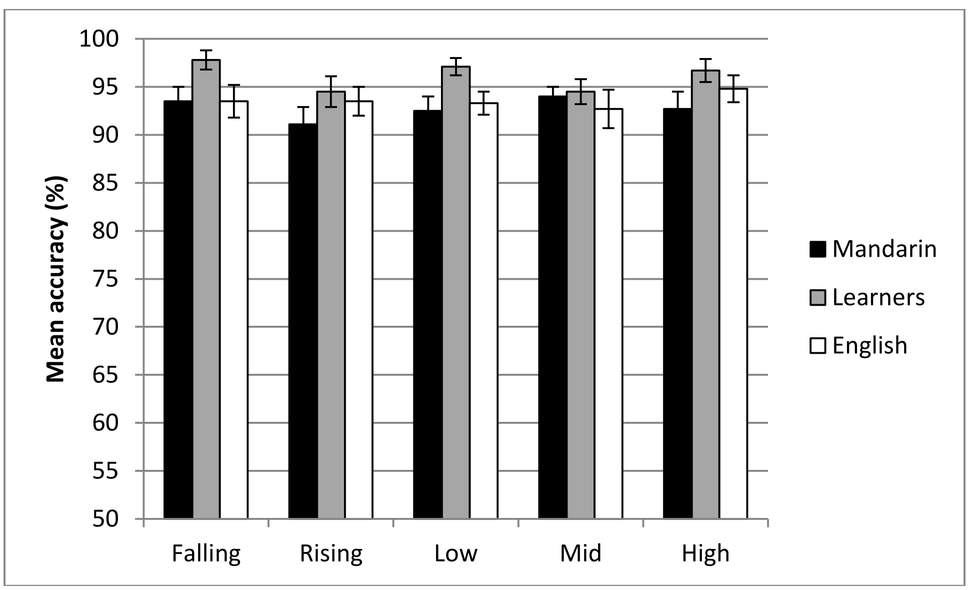


Figure 5.4: Mean accuracy (%) for each group on each control subcondition. Error bars represent +/-1 SE.

To examine the accuracy data for each individual tone comparison, a GEE model was fitted with the fixed factors Language (Mandarin, Learners, English) and Subcondition (individual tonal comparisons for both the tone and segment subconditions), with Subjects as a random factor. The Type III tests of fixed effects showed no effect of Language ($\chi^2(2) = 4.67$, $p = 0.097$), a significant effect of Subcondition ($\chi^2(14) = 488.36$, $p < 0.001$), and a significant interaction between the two factors ($\chi^2(28) = 74.34$, $p = 0.001$).

We see no cases of a statistically significant difference in accuracy on the test or control comparisons between the three groups. However, overall test results show that all three groups showed significant within-group differences on the accuracy rates between subconditions, specifically for two control subconditions: Falling ($p = 0.018$) and Low ($p = 0.013$), thus, accounting for the significant interaction.

RT Data

RT and LogRT data for each group and each subcondition are presented in Table 5.4. RTs in the control condition were lower (all but one below 1000 ms) and more uniform than those on the test condition. The RTs for the test condition ranged from the fastest performance by the Learners on the Low vs Falling comparison (*M* = 949 ms) to the slowest speed by the English group on the Mid vs High comparison (*M* = 1373 ms). Overall, the learners performed faster on the control subconditions than both the Mandarin and English groups. On the test subconditions the learners were again faster than the English group but not necessarily faster than the Mandarin group.

Again, while we see some universal tendencies among the three groups, we also see some language-specific ones as well. For example, all three groups were generally faster on the Falling vs Rising and Low vs Falling comparisons while their slowest scores were generally on the Low vs Mid and Mid vs High comparisons. Yet, it is the latter Mid vs High comparison where the three groups diverge: English (*M* = 1373 ms) were much faster than the Mandarin (*M* = 1191 ms) and the Learners (*M* = 1202 ms). On the other comparisons, the gaps between the RTs are generally not as big (See Table 5.5 and Figures 5.5a-c).

Table 5.4: Mean RTs (ms) and LogRTs for each group on each subcondition

Test subconditions	RT (CI)			Log RT (CI)		
	Mandarin	Learners	English	Mandarin	Learners	English
FR	948 (891; 1009)	960 (890; 1037)	1064 (988; 1146)	2.98 (2.95; 3.0)	2.98 (295; 3.02)	3.03 (3.0; 3.06)
LF	988 (929; 1052)	949 (878; 1025)	1071 (994; 1154)	3.0 (2.97; 3.02)	2.98 (2.94; 3.01)	3.03 (3.0; 3.06)
LR	1170 (1097; 1247)	1206 (1114; 1304)	1303 (1205; 1410)	3.07 (3.04; 3.1)	3.08 (3.05; 3.12)	3.12 (3.08; 3.15)
MF	1091 (1023; 1164)	1018 (929; 1116)	1146 (1064; 1236)	3.04 (3.01; 3.07)	3.01 (2.97; 3.05)	3.06 (3.03; 3.09)
MR	1122 (1052; 1194)	1070 (994; 1151)	1211 (1124; 1305)	3.05 (3.02; 3.08)	3.03 (3.0; 3.06)	3.08 (3.05; 3.12)
HF	1050 (986; 1119)	1020 (943; 1103)	1102 (1022; 1189)	3.02 (2.99; 3.05)	3.01 (2.97; 3.04)	3.04 (3.01; 3.08)
HR	1047 (982; 1114)	1025 (959; 1096)	1139 (1055; 1228)	3.02 (2.99; 3.05)	3.01 (2.98; 3.04)	3.06 (3.02; 3.09)
LM	1202 (1125; 1285)	1156 (1076; 1242)	1255 (1162; 1355)	3.08 (3.05; 3.11)	3.06 (3.03; 3.09)	3.1 (3.07; 3.13)
LH	1099 (1033; 1172)	1020 (953; 1092)	1146 (1063; 1236)	3.04 (3.01; 3.07)	3.01 (2.98; 3.04)	3.06 (3.03; 3.09)
MH	1180 (1107; 1259)	1202 (1118; 1292)	1373 (1272; 1482)	3.07 (3.04; 3.1)	3.08 (3.05; 3.11)	3.14 (3.1; 3.17)
Control subconditions						
F	951 (897; 1005)	853 (783; 929)	985 (920; 1054)	2.98 (2.95; 3.0)	2.93 (2.89; 2.97)	2.99 (2.96; 3.02)
R	940 (887; 995)	856 (795; 921)	999 (933; 1069)	2.97 (2.95; 3.0)	2.93 (2.9; 2.96)	3.0 (2.97; 3.03)
L	938 (885; 993)	870 (788; 960)	963 (900; 1031)	2.97 (2.95; 3.0)	2.94 (2.9; 2.98)	2.98 (2.95; 3.01)
M	953 (900; 1007)	850 (785; 920)	1029 (961; 1101)	2.98 (2.95; 3.0)	2.93 (2.89; 2.96)	3.01 (2.98; 3.04)
H	951 (897; 1005)	850 (784; 921)	979 (915; 1047)	2.98 (2.95; 3.0)	2.93 (2.89; 2.96)	2.99 (2.96; 3.02)

Note: CI = Confidence interval (lower; upper).

T = Thai; M = Mandarin, J = Japanese, E = English, K = Korean ; L = low tone, M = mid tone, H = high tone, F = falling tone, R = rising tone. Hence, LH = Low tone vs High tone comparison, HF = High tone vs Falling tone comparison, etc.

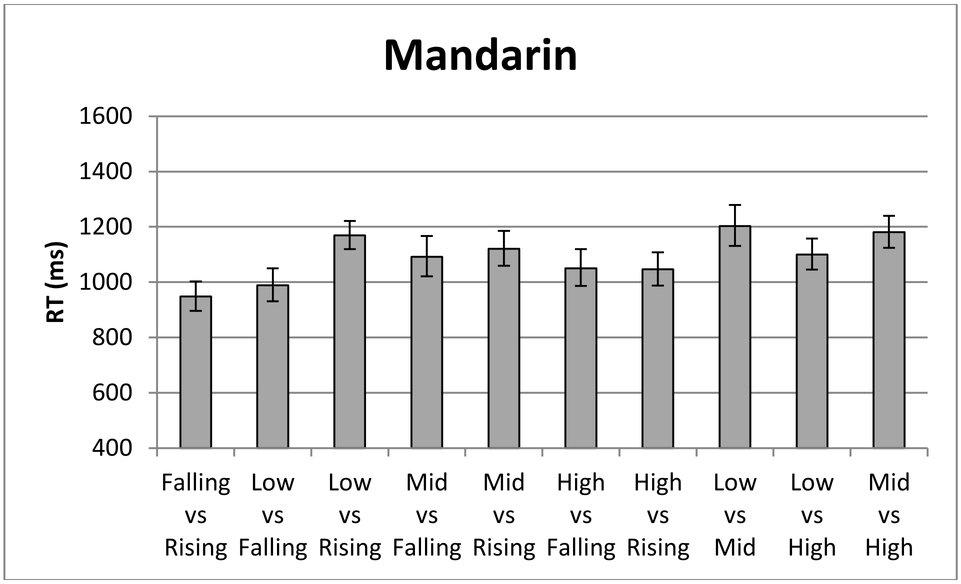


Figure 5.5a: RTs (ms) for the Mandarin group on each test subcondition. Error bars represent the upper and lower CI from RTs from a Linear Mixed Effects model.

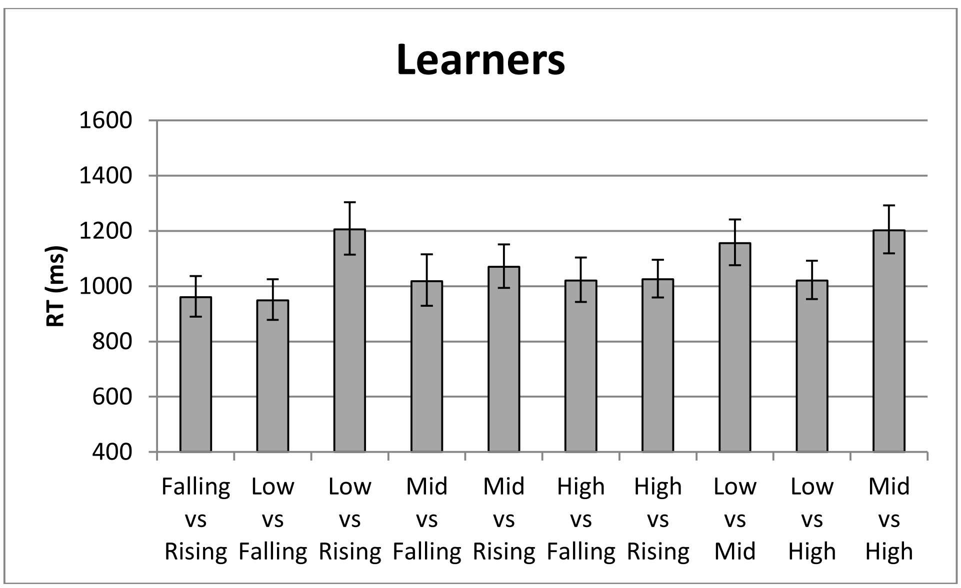


Figure 5.5b: RTs (ms) for the Learners on each test subcondition. Error bars represent the upper and lower CI from RTs from a Linear Mixed Effects model.

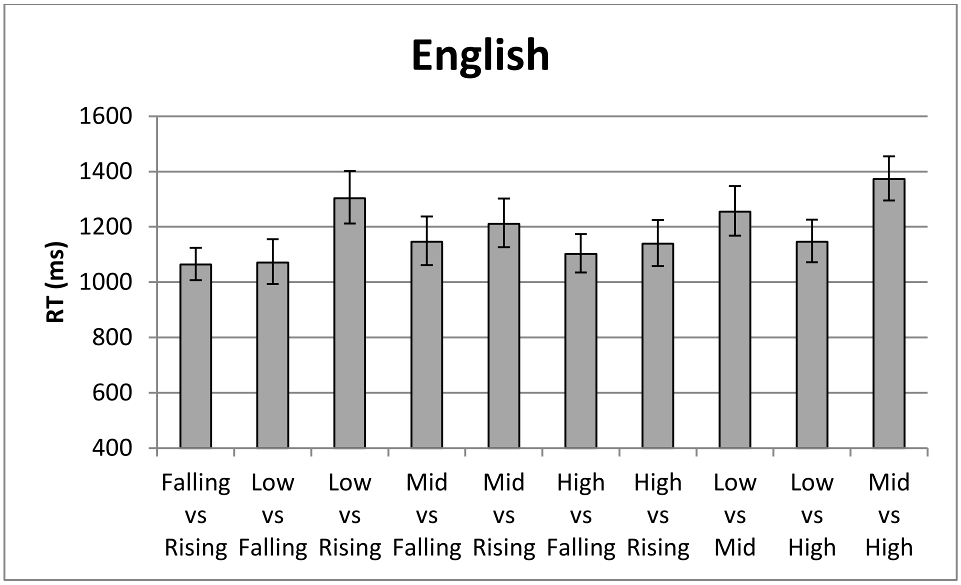


Figure 5.5c: RTs (ms) for the English group on each test subcondition. Error bars represent the upper and lower CI from RTs from a Linear Mixed Effects model.

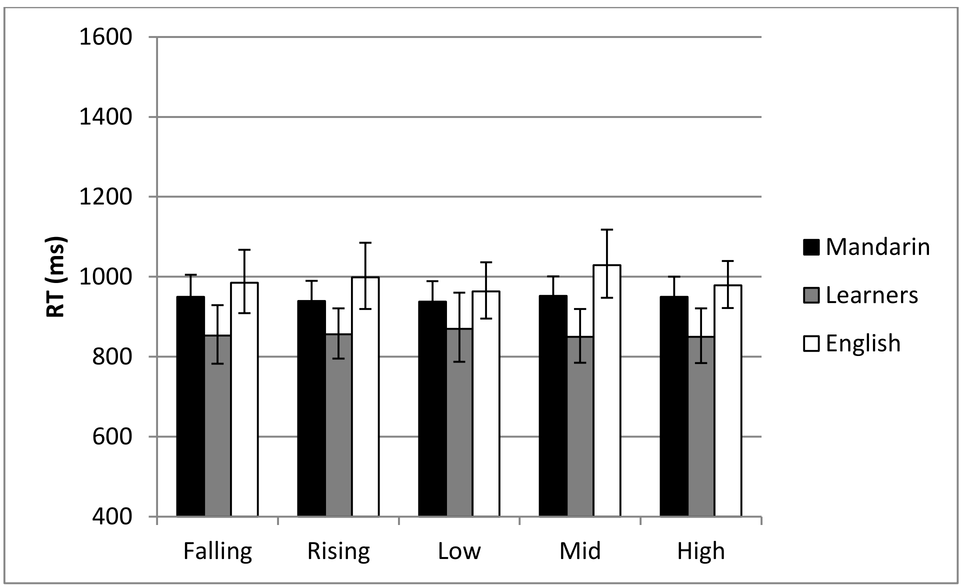


Figure 5.6: RTs (ms) for each group on each control subcondition. Error bars represent the upper and lower CI from RTs from a Linear Mixed Effects model.

A Linear Mixed Effects model was run on the LogRT with Language (Mandarin, Learners, English) and Subcondition (individual tonal comparisons for both the tone and segment subconditions) as fixed factors and Subjects declared as a random factor. When looking at the

Type III tests of fixed effects, the F-tests showed a main effect of Sub-condition ($F(14, 9977) = 86.12$, $p < 0.001$), a significant effect of Language ($F(2, 69.59) = 3.64$, $p < 0.031$), and a significant interaction between the two factors ($F(28, 9977) = 2.0$, $p = 0.001$).

The interaction demonstrates the groups vary on both the Test and Control subconditions (See Table 5.5). The Learners were faster than the English group overall. On the following test subcondition, the Learners were significantly faster than the English: Low vs Rising ($p = 0.037$), Low-High ($p = 0.049$), and Mid vs High ($p = 0.015$). The Mandarin group was faster than the English group on the Falling vs Rising ($p = 0.013$), Low vs Rising ($p = 0.037$) and Mid vs High ($p = 0.001$).

For the control subconditions, the Learners were faster than the English group on the Falling (p = 0.048), Rising ($p = 0.019$). Mid ($p = 0.003$) and High ($p = 0.018$). The Learners were statistically faster than the Mandarin group on the Mid but marginally so ($p = 0.051$) (See Table 5.5). Additionally, Univariate tests demonstrate that all the groups show significant within-group differences between RTs on subconditions, specifically on the RTs of two test subconditions: for Falling vs Rising ($p < 0.042$) and Mid vs High ($p < 0.007$). That is, all the performance of all the groups on these two test subconditions varied from their performances on the other test subconditions.

Table 5.5: Statistically significant difference in mean LogRTs between groups on each tonal comparison (i.e., subcondition)

Comparison	Reaction times
Tone overall	Learners > English (*p* = 0.037)
FR	Mandarin > English (*p* = 0.013)
LF	--
LR	Mandarin > English (*p* = 0.037)
MF	--
MR	--
HF	--
HR	--
LM	--
LH	Learners > English (*p* = 0.049)
MH	Learners > English (*p* = 0.015) Mandarin > English (*p* = 0.001)
Segment overall	Learners > English (*p* = 0.021)
F	Learners > English (*p* = 0.048)
R	Learners > English (*p* = 0.019)
L	--
M	Learners > English (*p* = 0.003) Learners > Mandarin (*p* = 0.051)
H	Learners > English (*p* = 0.018)

Note: F = Falling, R = Rising, L = Low, M = Mid, H = High tone.

5.2.3 Discussion

Accurate results for the overall tone condition and each individual tone subcondition do not statistically distinguish the three groups of L1 Mandarin, Learners, and L1 English from one another. On the other hand, the RTs reveal differences between the three groups on their perception of Thai tones.

The Learners are significantly faster than the English group on overall tone perception. The Learners were also faster than the English group on the segmental condition, thereby essentially neutralizing the effect of condition. Yet, the Mandarin group was not faster than the English group nor slower than the Learners. This may provide an insightful clue as to how the groups differ in their perception of tone despite similar accuracy rates. Thus, these two seemingly contradictory

RT trends are most likely attributable to the same cause: exposure to tone, but with differing results.

These differences in RTs between the Mandarin and Learner groups may be due to processing differences. Again, we can reference the Stagray et al. study (1993) which showed that native listeners are more categorical in their perception of tones than non-native listeners who therefore appear to hear the slight phonetic differences between tones that are categorized as the same tone by native speakers. It is then this fact that may account for the difference in RTs by the Mandarin and Learner groups. In short, while it may seem to counter the Stagray et al study, the fact that the Mandarin group has strong categories for tones to compare the Thai tones to may give them more reason to pause in their response. Thus, the fact that the Mandarin group is neither statistically slower than the English group nor faster than the Learners is in of itself a hint that the three groups process tones differently.

There are both universal trends for RTs on individual tone comparisons. It appears that all groups have difficulty as shown by slower RTs on the Low vs Rising, Low vs Mid and Mid vs High comparisons while all groups exhibit faster RTs on comparisons with a Falling tone. The seemingly universal trend may be attributed to the similarity in both shape and height of the tones despite the phonetic difference of the High tone rendering the Mid and High different in shape. For example, it does not come as a surprise then that the Low vs Mid tone comparison results in slower RTs across all groups. This comparison has been noted in the literature as being difficult even for native Thai speakers due to the similarity of the two tones in shape and their close proximity in pitch height.

Language-specific differences also are present in the RTs for the individual tone comparisons. We see that the Mandarin group is faster than the English group on three of the tone comparisons: Falling vs Rising, Low vs Rising and Mid vs High comparisons. These results may reflect word stress in English and tone patterns in Mandarin. The Low and Rising result is expected as these two tones are similar in height, i.e., lower register, (despite differences in contour) and therefore, both tones may sound similar to unstressed syllables in English. If English considers any presence of tone as being similar to the pitch of word stress, then it is natural that the English group reacts more slowly on the Mid vs High comparison as both may seem to bear stress to an L1 English ear. In the case of the Mandarin group, there is either a Rising or High tone which may be similar to their counterpart Rising and High tones in Mandarin and thus, the performance of the Mandarin group is not unexpected. In sum, the L1 in the form of word stress for the English group and in the form of the tone patterns for the Mandarin group appears to play a large role in perception.

Exposure to Mandarin appears to have aided the Learners in their perception of Thai tones. For example, the Learner group is faster than the English group on the Mid vs High comparisons. We expect the difference on the Mid vs High as the Mandarin group was statistically faster as well. It is specifically the High tone that the Learners may have acquired as the Mid tone does not exist in Mandarin. Moreover, the Learners are also faster than the English group on the Low vs High comparison. However, this is surprising as the Mandarin group did not perform statistically faster on this comparison vis-à-vis the English group. Thus, while we witness an overall positive effect of learning Mandarin, it appears to be limited in some instances. In short, this positive effect appears to range from easier processing as indicated by faster RTs on easy comparisons (e.g., Falling vs Rising), but also to performance affected by the presence or absence of counterpart target tones in the tonal inventory of Mandarin (e.g., high performance on High vs Falling, poor performance on Low vs Mid).

In sum, exposure to Mandarin tones through learning Mandarin as an L2 has an effect on the the perception of Thai tones. However, this exposure affects only the perception of some tones and not others. We also observe universal tendencies more than language-specific tendencies. Additionally, we detect a seemingly inexplicable tendency for Learners to be faster across the board; this phenomenon may actually be highlighted by the fact that the Mandarin group is slower due to processing affected by the categorization of tones, providing a clue as to a difference between the Mandarin and Learner group. Thus, we must consider not only accuracy rates but also those of RTs when interpreting results. In the next section, we turn to see whether this performance on the perception of Thai monosyllabic stimuli holds on the perception of Thai disyllabic stimuli where we expect a positive effect of English word stress.

5.3 Experiment 2b: ABX Thai tones, disyllabic

We turn now to the methodology and results of the disyllabic ABX task. The disyllabic ABX task may put a greater cognitive load onto the same participants than the monosyllabic ABX task might. This might possibly tease out more statistically significant results. It may also conversely provide “greater help” to the English group which may need two syllables to better hear the difference in tones reflective of word stress in English.

5.3.1 Methodology

5.3.1.1 Participants

The participants were the same as in Experiment 2a (Disyllabic).

5.3.1.2 Stimuli

The stimuli were the same as in Experiment 1b (Disyllabic).

5.3.1.3 Procedure

The procedure was the same as in Experiment 1b.

5.3.2 Results

The data of the learners were coded and screened for outliers in the same way as for the other groups in Experiment 1a. Similarly, RT data were also log-transformed, and the same analysis models were applied for this set of data.

Reaction times (RT) that were 300ms and below and any over 2500ms were removed. In total, only 0.00009% of the RT datapoints were removed (N = 1). The RT data were examined and were found to be skewed. Hence, a log-transformation was applied to the RT (Log-RT) to obtain a normal distribution.

The same five learners were excluded in this task as in experiment 2a. The one participant who indicated that they had misunderstood the ABX monosyllabic task and had essentially done an "XAB" task and not an ABX task, participated in the ABX disyllabic task. Their score was below two SD from the Learner group mean on the overall test condition on the ABX disyllabic task and therefore, their data was excluded. In addition, two Mandarin speakers were removed as well since they also performed beyond two SD from the mean accuracy on the control condition. In addition, data for two learners who were born in China and then, later adopted and grew up speaking English (cf., Pierce et al., 2014) was removed. Also, three learners who were heritage speakers were removed due to long-term exposure to a tonal language since birth.

After cleaning the data, a Generalized Estimating Equation (GEE) model for a binary response was fitted to the data and run for overall accuracy as the data structure is categorical. A Linear Mixed Effects model was fitted to the data and run on the LogRT. Language (Mandarin, Learners, English) and Condition (test vs control) were declared as fixed factors with Subjects declared as a random factor. Sidak correction for multiple comparisons was used over Bonferroni. Both the Log RTs and the RTs in ms are reported with their lower and upper confidence intervals (CI) in all the following tables.

5.3.2.1 Analysis by condition: Test and control

Since the focus of this task is on the acquisitional robustness of a tonal L2 (i.e., Mandarin) on non-native tone perception (i.e., Thai), the data of the Learners are compared to the performance of the Mandarin and English groups from Experiment 1b. The data of these two groups from Experiment 1b are repeated here. The table below present the mean accuracy score and mean RTs for each of the three groups on each condition and subcondition (i.e., individual tonal comparisons).

Accuracy Data

Accuracy rates for the test condition were lower than those for the control condition across the three groups as seen in Table 5.6. On the test condition, the Learners performed the most accurately (*M* = 87.5%) followed by the Mandarin group (*M* = 85.9%) and then, the English group (*M* = 77.5%). The scores on the control condition were higher and also generally closer together than those for the test condition with the Learners performing the most accurately (*M* = 96.7%) with the English (*M* = 96.6%) and the Mandarin group somewhat lower (*M* = 97.2%) at almost the same accuracy rate.

Table 5.6: Mean accuracy (%) for each group on test and control conditions

	Accuracy	
Language group	Test (SE)	Control (SE)
Mandarin (n=29)	85.9 (2.0)	96.7 (0.7)
Learners (n=18)	87.5 (1.8)	97.2 (0.7)
English (n=23)	77.5 (3.1)	96.6 (0.7)

Note: SE = standard error.

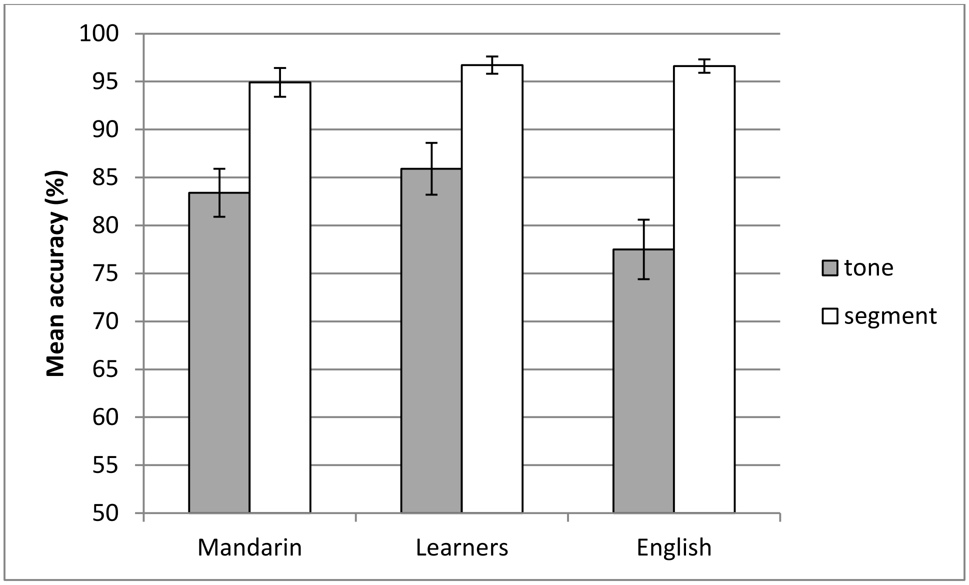


Figure 5.7: Mean accuracy (%) for each group on test and control. Error bars represent +/-1 SE.

The un-aggregated data were used for statistical analysis. Accuracy data are analyzed first with RT analysis following.

A GEE model was fitted with the fixed factors Language (Mandarin, Learners, English) and Condition (test vs control), with Subjects as a random factor. The Type III tests of fixed effects showed no main effect of Language ($\chi^2(2) = 3.52$, $p = 1.72$), a significant effect of Condition ($\chi^2(1) = 210.86$, $p < 0.001$), and no interaction between the two factors ($\chi^2(2) = 4.84$, $p = 0.089$).

A significant effect of condition demonstrates that the groups differed on the test and control conditions. We see the following results in Table 5.6: Mandarin: $M = 85.9\%$, Learners: $M = 87.5\%$, English: $M = 77.5\%$ versus Mandarin: 96.7%, Learners: 97.2%, English: 96.6%, respectively (See Table 5.8). However, overall test results revealed that performance differed significantly between the groups on the test condition ($\chi^2(2) = 7.88$ $p < 0.019$) but not on the control conditions ($\chi^2(2) = 4.31$ $p < 0.806$). We see one case only of a difference between groups on the test condition: the Learners were more accurate overall than the English group ($p = 0.016$) (See Table 5.8).

RT Data

RT and LogRT are presented in Table 5.7. RTs were slower on the test condition (*M* = 1294 for Mandarin, *M* = 1265ms for Learners, *M* = 1368 ms for English) than the RTs on the control condition by approximately 100-200ms (*M* = 1168ms for Mandarin, *M* = 1026 ms for Learners, *M* = 1140 ms for English).

Table 5.7: Mean RT (ms) and LogRT for each group on test and control conditions

	RT		Log RT	
Language group	Test (CI)	Control (CI)	Test (CI)	Control (CI)
Mandarin (n=31)	1294 (1227; 1363)	1168 (1109; 1231)	3.11 (3.09; 3.14)	3.07 (3.05; 3.14)
Learners (n=18)	1265 (1179; 1357)	1026 (957; 1100)	3.10 (3.07; 3.13)	3.01 (2.98; 3.04)
English (n=23)	1368 (1287; 1454)	1140 (1073; 1210)	3.14 (3.11; 3.16)	3.06 (3.03; 3.08)

Note: CI = Confidence interval (lower; upper).

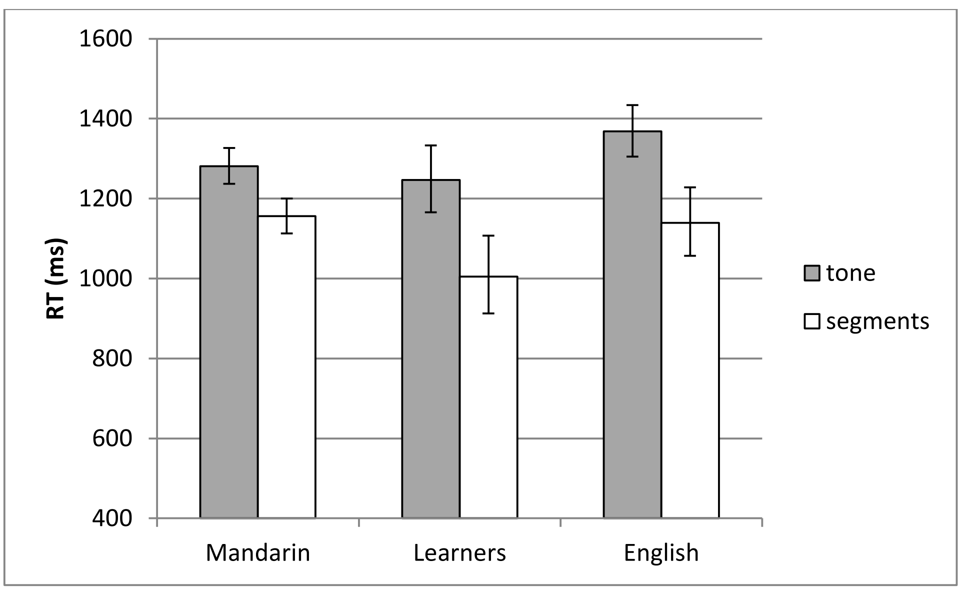


Figure 5.8: RTs (ms) for each group on test and control conditions. Error bars represent the upper and lower CI from RTs from a Linear Mixed Effects Model.

A Linear Mixed Effects model was run on the LogRT with Language (Mandarin, Learners, English) and Condition (test vs control) as fixed factors, declaring Subjects as a random factor. For reaction times (RTs) when looking at the Type III tests of fixed effects, the F-tests showed a main effect of Condition ($F(1, 4974) = 550.67$, $p < 0.001$), a significant effect of Language ($F(2, 67.09) = 3.43$, $p = 0.038$), and a significant interaction between the two factors ($F(2, 4974) = 24.06$, $p < 0.001$).

There is a main effect of condition between the mean RTs on the Test condition (Mandarin: M = 1294 ms, Learners: M = 1265 ms, English: M = 1368 ms) versus the Control condition (Mandarin: M = 1168 ms, Learners: M = 1026 ms, English: M = 1140ms) (See Table 5.7). Univariate tests revealed that performance did not differ significantly between groups on the test condition ($F(3, 72.53) = 2.54$, $p < 0.086$) but did vary on the control condition ($F(2, 72.48) = 6.15$, $p < 0.003$). In Table 5.11, we see that the Learners were faster than the Mandarin group on every control subcondition and faster than the English group on the two control subconditions carrying the Low+Mid and Mid+High disyllabic tone patterns.

5.3.2.2 Analysis by subcondition

The next section looks at the comparison of individual tones for both accuracy and RTs. Individual tonal comparisons may illuminate whether the perception of one tone type may be more difficult than another tone type. (Table 5.8 presents the individual accuracy means for the test and control items in each subcondition for each group while Table 5.10 presents the RTs).

Accuracy Data

Accuracy rates for the control condition ranged from a low on the control condition using the High+Falling by the English group (M = 94.8%); the high score was on the Mid+High control condition by the Learners (M = 98.5%). In contrast, accuracy rates for the test condition ranged from a low on the Low+High vs Mid+High by the English group (M = 63%) to a high score on the Low+High vs High+Low by the Learners (M = 95.6%).

Again, we see both universal tendencies among the three groups as well as a few language-specific tendencies. All three groups performed the least accurately on the Low+High vs

Mid+High comparison. The Mandarin and Learner groups performed well on the Low+High vs High+Low and Low+Mid vs High+Low comparisons ranging 92.2-95.6%, but relatively less accurately on the Low+Mid vs Low+High comparison at 85.3% and 87.5%, respectively . In contrast, the English group performed on four out of five comparisons at the same level of accuracy ranging from 80-82%, excluding their lowest accuracy rate of *M* = 63 % on the Low+High vs Mid+High comparison.

Table 5.8: Mean accuracy (%) for each group on each subcondition, and statistically significant differences between groups in accuracy rates

Test subconditions	Mean accuracy (SE)			Comparison	Accuracy
	Mandarin	Learners	English	Tone overall	Learners > English (*p* = 0.16)
LH-HL	92.2 (2.4)	95.6 (1.8)	80.7 (3.5)	LH-HL	Learners > English (*p* = 0.01) Mandarin > English (*p* = 0.21)
LM-HL	95.3 (1.5)	91.2 (2.5)	81.8 (3.5)	LM-HL	Mandarin > English (*p* = 0.001)
LH-MH	66.4 (4.3)	72.1 (4.2)	63.0 (5.1)	LH-MH	--
LM-LH	85.3 (3.3)	87.5 (2.5)	80.2 (3.5)	LM-LH	--
LM-MH	90.1 (2.1)	91.2 (2.3)	81.8 (3.6)	LM-MH	--
Control subconditions				Segment overall	--
HF	97.4 (1.1)	95.6 (1.4)	94.8 (1.8)	HF	--
HL	97.0 (1.2)	97.1 (1.3)	96.9 (1.1)	HL	--
LH	97.4 (0.9)	97.8 (1.2)	97.9 (1.0)	LH	--
LM	96.6 (1.2)	97.1 (1.7)	95.8 (1.4)	LM	--
MH	95.3 (1.3)	98.5 (1.0)	97.4 (1.0)	MH	--

Note: SE=standard error.
T = Thai; M = Mandarin, J = Japanese, E = English, K = Korean; L = low tone, M = mid tone, H = high tone, F = falling tone, R = rising tone. Hence, LM = Low tone + Mid tone disyllabic word, LH = Low tone + High tone disyllabic word, MH = Mid tone + High tone disyllabic word, HL = High tone + Low tone disyllabic word, HF = High tone + Falling tone disyllabic word.

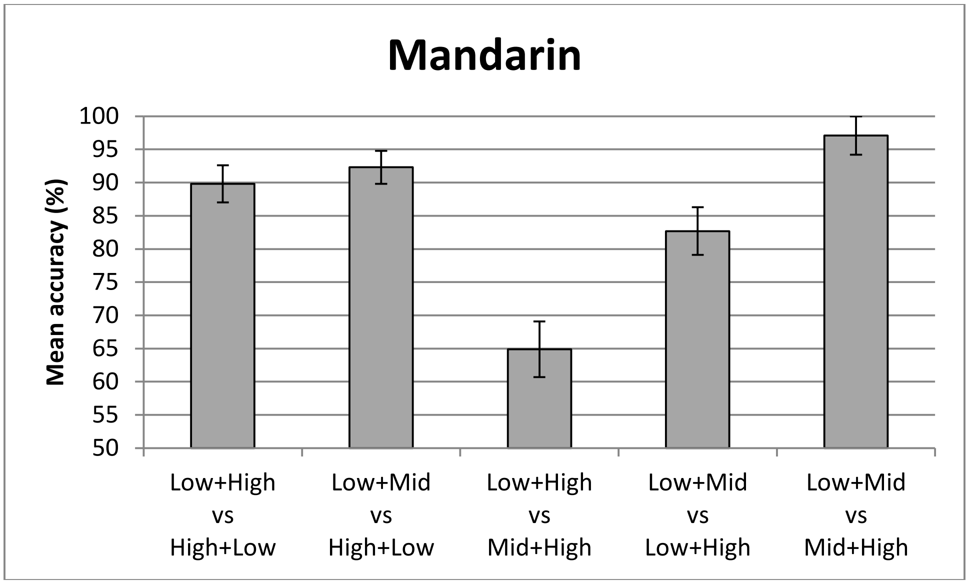


Figure 5.9a: Mean accuracy (%) for the Mandarin group on each test subcondition. Error bars represent +/-1 SE.

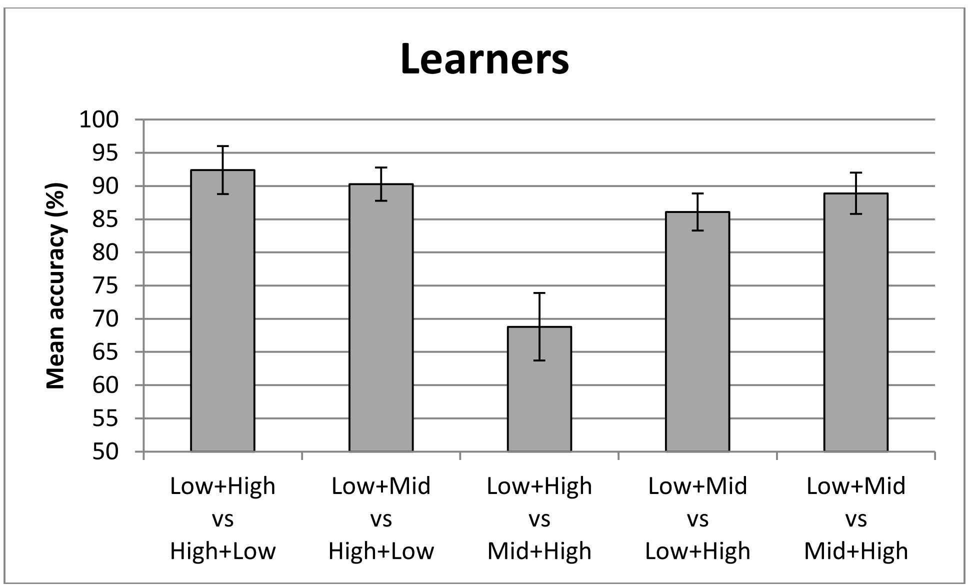


Figure 5.9c: Mean accuracy (%) for the Learner group on each test subcondition. Error bars represent +/-1 SE.

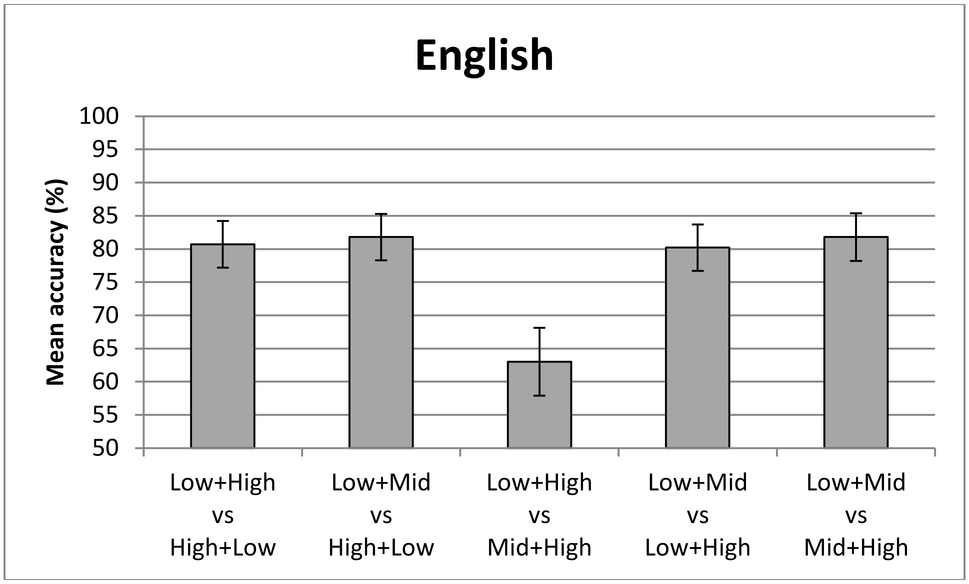


Figure 5.9c: Mean accuracy (%) for the English group on each test subcondition. Error bars represent +/-1 SE.

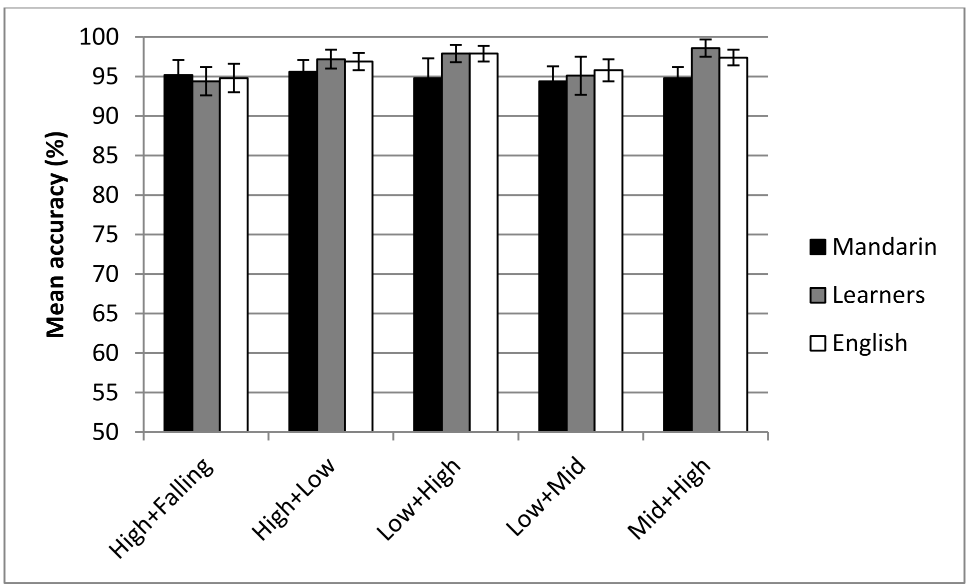


Figure 5.10: Mean accuracy (%) for each group on each control subcondition. Error bars represent +/-1 SE.

A GEE model was fitted with the fixed factors Language (Mandarin, Learners, English) and Subcondition (individual tonal comparisons for both the tone and segment conditions), with Subjects as a random factor. As for individual tonal comparisons, the Type III tests of fixed effects showed no main effect of Language ($\chi^2(2) = 4.35$, $p = 0.114$), a significant effect of Subcondition

($\chi^2(9) = 362.91$, $p < 0.001$), and a significant interaction between the two factors ($\chi^2(18) = 59.45$, $p < 0.001$).

The interaction is likely driven by the differences in performance between groups on the Test condition (See Table 5.12). The Mandarin group was more accurate than the English group on the Low+Mid vs High+Low comparison ($p = 0.040$). Additionally, overall test results demonstrate that all three groups show significant within-group differences between accuracy rates on subconditions, specifically on the accuracy of two test subconditions: Low+High vs High+Low ($p = 0.001$) and Low+Mid vs High+Low ($p = 0.001$).

RT Data

As for RTs on the control subconditions, they ranged from the fastest time on the Low+Mid pattern comparison by the Learners (*M* = 999 ms) to the slowest speed on subcondition using the High+Low pattern the by the Mandarin group (*M* = 1191 ms). In general, the Learners were the fastest overall while both the Mandarin and English groups were somewhat slower. The RTs for the test condition ranged from the fastest speed on the Low+High vs High+Low comparison by the Learners (*M* = 1190 ms) to the slowest speed on the Low+Mid vs Low+High by the English group (*M* = 1484 ms). Again, the Learners were generally faster than the other groups, but less so on the test condition than they were on the control condition. Unlike the trend on the control condition, the English group was slower than the Mandarin group on the test condition with some differences in RTs between the two groups in excess of 100 ms.

Table 5.9: Mean RT (ms) and LogRT for each group on each subcondition

Test subconditions	RT (CI)			Log RT (CI)		
	Mandarin	Learners	English	Mandarin	Learners	English
LH-HL	1197 (1127; 1271)	1190 (1099; 1288)	1301 (1214; 1396)	3.08 (3.05; 3.1)	3.08 (3.04; 3.11)	3.11 (3.08; 3.15)
LM-HL	1197 (1127; 1274)	1192 (1101; 1291)	1297 (1210; 1391)	3.08 (3.05; 3.11)	3.08 (3.04; 3.11)	3.11 (3.08; 3.14)
LH-MH	1384 (1297; 1476)	1313 (1209; 1427)	1459 (1356; 1569)	3.14 (3.11; 3.17)	3.12 (3.08; 3.15)	3.16 (3.13; 3.2)
LM-LH	1454 (1365; 1542)	1423 (1314; 1541)	1484 (1384; 1592)	3.16 (3.14; 3.19)	3.15 (3.12; 3.19)	3.17 (3.14; 3.2)
LM-MH	1309 (1231; 1390)	1240 (1146; 1343)	1332 (1243; 1429)	3.12 (3.09; 3.14)	3.09 (3.06; 3.13)	3.13 (3.09; 3.16)
Control subconditions						
HF	1165 (1097; 1236)	1018 (941; 1102)	1121 (1047; 1200)	3.07 (3.04; 3.09)	3.01 (2.97; 3.04)	3.05 (3.02; 3.08)
HL	1191 (1122; 1266)	1059 (979; 1145)	1152 (1076; 1233)	3.08 (3.05; 3.1)	3.03 (2.99; 3.06)	3.06 (3.03; 3.09)
LH	1151 (1085; 1221)	1030 (952; 1113)	1127 (1053; 1206)	3.06 (3.04; 3.09)	3.01 (2.98; 3.05)	3.05 (3.02; 3 .08)
LM	1168 (1099; 1239)	999 (923; 1080)	1143 (1068; 1223)	3.07 (3.04; 3.09)	3.0 (2.97; 3.03)	3.06 (3.03; 3.09)
MH	1162 (1094; 1234)	1027 (950; 1111)	1155 (1079; 1236)	3.07 (3.04; 3.09)	3.01 (2.98; 3.05)	3.06 (3.03; 3.09)

Note: CI = Confidence Interval.
T = Thai; M = Mandarin, J = Japanese, E = English, K = Korean ; L = low tone, M = mid tone, H = high tone, F = falling tone, R = rising tone. Hence, LM = Low tone + Mid tone disyllabic word, LH = Low tone + High tone disyllabic word, MH = Mid tone + High tone disyllabic word, HL = High tone + Low tone disyllabic word, HF = High tone + Falling tone disyllabic word.

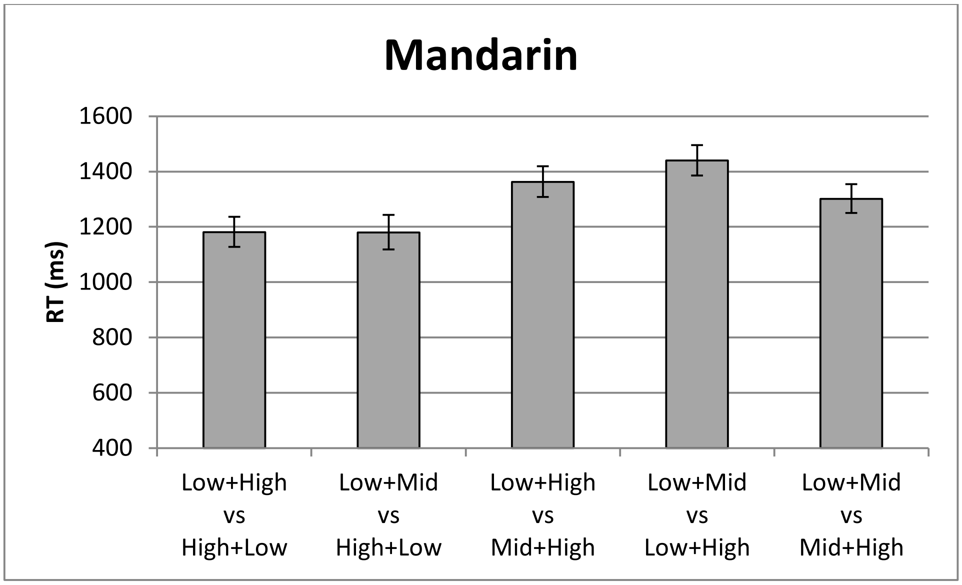


Figure 5.11a: RT (ms) for the Mandarin group on each test subcondition. Error bars represent the upper and lower CI from RTs from a Linear Mixed Effects model.

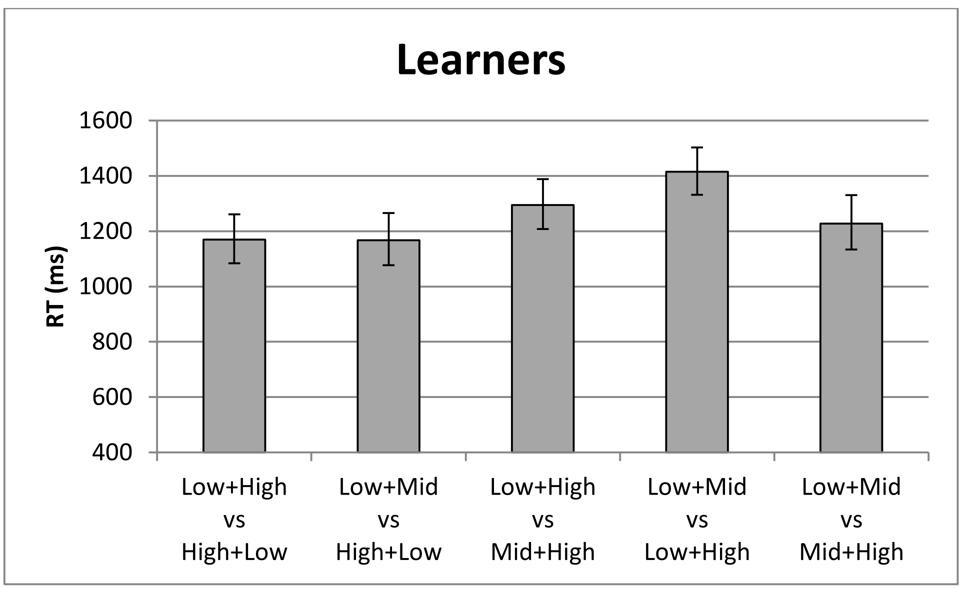


Figure 5.11b: RT (ms) for the Learner group on each test subcondition. Error bars represent the upper and lower CI from RTs from a Linear Mixed Effects model.

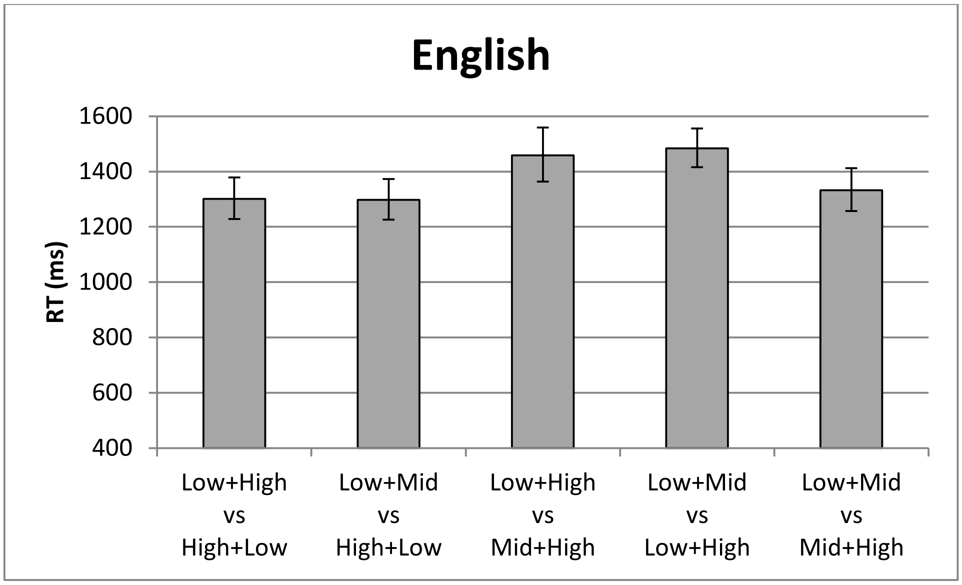


Figure 5.11c: RT (ms) for the English group on each test subcondition. Error bars represent the upper and lower CI from RTs from a Linear Mixed Effects model.

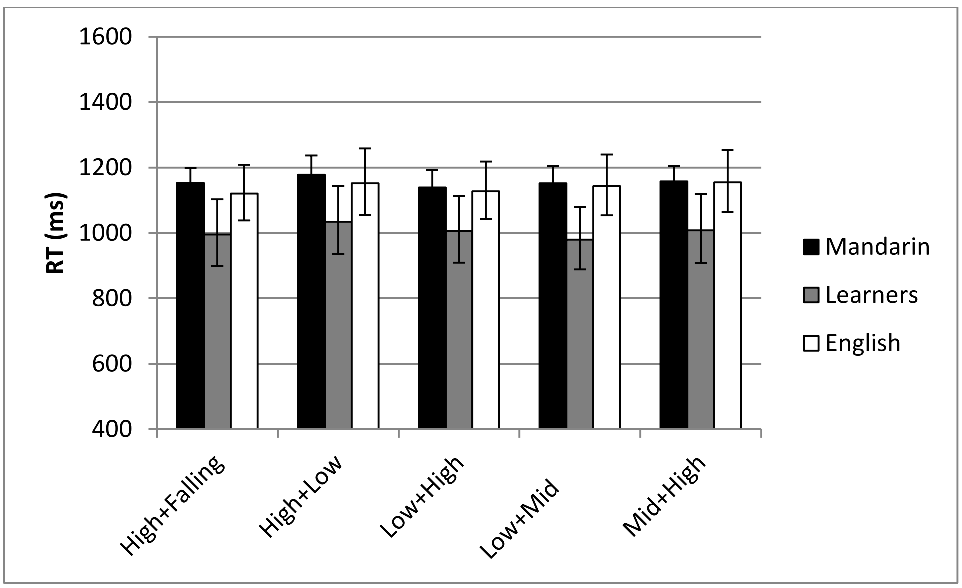


Figure 5.12: RT (ms) for each group on each control subcondition. Error bars represent the upper and lower CI from RTs from a Linear Mixed Effects model.

A Linear Mixed Effects model was run on LogRT with Language (Mandarin, Learners, English) and Subcondition (individual tonal comparisons for both the tone and segment conditions) as fixed factors, declaring Subjects as a random factor. When looking at the Type III tests of fixed effects, the F-tests showed a main effect of Condition ($F(9, 4949) = 84.40$, $p < 0.001$), a significant effect of Language ($F(2, 67.13) = 3.456$, $p = 0.037$), and a significant interaction between the two factors ($F(18, 4949) = 3.34$, $p < 0.001$).

The interaction demonstrates the performance of the groups vary only on the Control subcondition (See Table 5.14). The Learners were faster than the Mandarin group on all the Control subconditions and faster than the English on two of the control subconditions, namely the condition carrying the disyllabic Low+Mid pattern and that carrying the disyllabic Mid+High pattern (See Table 5.11). Additionally, Univariate tests demonstrate that all three groups patterned alike in their RTs showing a significant difference between the Low+High vs High+Low pattern ($p < 0.048$) vis-à-vis their performance on the other test subconditions. All three groups performed alike on the Low+Mid vs High+Low RT pattern ($p = 0.053$) compared to their performance on the other test subconditions.

Table 5.10: Statistically significant differences between groups in LogRTs on each tonal comparison

Comparison	Reaction times
Tone overall	--
LH-HL	--
LM-HL	--
LH-MH	--
LM-LH	--
LM-MH	--
Segment overall	Learners > Mandarin ($p = 0.024$)
HF	Learners > Mandarin ($p = 0.009$)
HL	Learners > Mandarin ($p = 0.024$)
LH	Learners > Mandarin ($p = 0.034$)
LM	Learners > English ($p = 0.008$) Learners > Mandarin ($p = 0.003$)
MH	Learners > English ($p = 0.022$) Learners > Mandarin ($p = 0.015$)

Note: F = Falling, R = Rising, L = Low, M = Mid, H = High tone.

5.3.3 Discussion

In this section, the three groups of L1 Mandarin, Learners and L1 English were tested in terms of accuracy rates and RTs on their perception of Thai tones using disyllabic stimuli which were introduced as they may reflect English word stress. There are few differences among the three groups in their accuracy rates. In contrast, there are more differences in reaction times.

The English group appears not to be able to access English word stress. For example, we see that the Mandarin group was statistically more accurate on the Low+High vs High+Low and Low+Mid vs High+Low comparisons. This is unexpected as this comparison reflects English word stress. In fact, exposure to Mandarin may conversely aid in the perception of the Low+High vs High+Low comparison. The Learners were statistically more accurate than the English group on the Low+High vs High+Low comparison. Hence, it may be that the pitch correlate of word stress is not sufficiently salient and that exposure to Mandarin may cause sensitivity to pitch height to become more salient. In short, the Mandarin group and Learner group may map these patterns onto the Falling and Rising tones in Mandarin while the English group appears not to be able to access or at least apply English word stress patterns.

Additionally, as noted the Mandarin group is more accurate than the English group on the Low+Mid vs High+Low comparison. This may be because despite the fact that a Mid tone does not exist in Mandarin, the Mandarin group is attuned to direction as a distinguishing feature of the inventory of Mandarin tones, e.g., three contour tones vs one level tone in Mandarin. In contrast, the Learners appear not to have not picked up as great a sensitivity to pitch direction evidenced by their statistically similar performance on the Low+Mid vs High+Low comparison as compared to the English group.

Nevertheless, universal tendencies seem more prevalent than language-specific tendencies. For example, the accuracy levels of the English group were relatively flat (i.e., comparable to one another). To a lesser extent this was the case for the other groups except for the Low+High vs

Mid+High comparison where all three groups performed poorly. Therefore, the English group merely did not perform as accurately overall as the two other groups with exposure to Mandarin. Moreover, we see that a combination of characteristics aid in perception. First, having two differences (e.g., height and contour) between disyllabic stimuli helps in perception, e.g., Low+Mid vs Mid+High. Also, comparing tones in clearly different registers helps, e.g., High+Low vs Low+High. The most difficult comparison occurs when there is only one difference (e.g., height). For example, when comparing only the Low and Mid tones, e.g., Low+High vs Mid+High, all three groups show the lowest accuracy rates. Thus, the difficulty in perceiving the individual tone comparisons appears to occur when all listeners must gauge the difference between the Low and Mid tones. Again, this is not the case when comparing the Mid and High tones, e.g., Low+Mid vs Low+High, especially as the phonetic shape of the High tone differs from that of the Mid tone while the phonetic shape of the Mid tone is similar to that of the Low tone even though all three tones vary in pitch height.

Moreover, we must further qualify these perceptual differences by two more points of interpretation. First, the presence of a Mid or High tone may be equal to word stress in English whereas a Low tone equals no stress. Also, the Mid tone may be associated with the High tone as an allotone, making it difficult to differentiate the two. Second, direction may favor listeners exposed to Mandarin due to the large number of contour tones in the tonal inventory of Mandarin so that the differences between Low and Mid vs Mid and High may be problematic. These two points allow for a focus on the categorization of the features of height and contour, thereby accounting for the similarities between the L1 and non-native target tones within a PAM (Best, 1995) approach.

As for RT results, none are significantly different on the overall tone condition or individual tone conditions between the three groups. However, on the overall segmental condition and each of the individual segmental conditions, the Learner group is faster than the Mandarin group. The Learners are also faster than the English group on the Low+Mid and Mid+High segmental tonal comparisons.

One has to question whether the faster RTs are a strategic effect (i.e., lower accuracy with faster RT vs higher accuracy vs slower RT) or a reflection of processing difficulty in the case of the English group where lower accuracy rates accompany slower RTs. This question is asked in light of the fact that the Learners also reacted more quickly than the English group on one of the control segment conditions. Thus, exposure to tone may merely make such learners more comfortable or familiar with tone and aid them in their perception (i.e., reduce anxiety levels). In contrast, the English group may feel intimidated by what most native English speakers perceive to be a difficult language due to the daunting "alien" tones. Also, the Learners also reacted marginally more quickly than the Mandarin group on several of the control segment conditions (stimuli using Low+Mid, Mid+High, High+Falling). We would expect the Learners to be slower the Mandarin group, but the difference may lie in the degree of robustness in encoding tone, i.e., native vs non-native (cf., Stagray and Downs, 1993). That is, the Mandarin group may be slowed down by the categorization of tones in Mandarin when considering the Thai tones as this group most likely accesses stronger robust representations of Mandarin tones.

Additionally, the reaction times only further underline the positive influence of Mandarin exposure on the naïve perception of Thai tones as both the Learners and Mandarin group reacted more quickly than the English group. For example, we see slower RTs with lower accuracy rates on the Low+High vs Mid+High, corroborating the difficulty in differentiating the Low and Mid tones. Thus, again we see that looking at both accuracy rates and RTs is necessary to see a more nuanced picture of tone perception.

In sum, even though it would appear that exposure to Mandarin has a positive influence on the perception of Thai tones, there are also universal difficulties regardless of the linguistic background of the participants. Additionally, exposure to Mandarin aids in hearing disyllabic patterns reflective of English word stress while exposure to English does not.

5.4 Summary and general discussion

In these two sections (5.2 ABX Thai tones, monosyllabic and 5.3 ABX Thai tones, disyllabic), we see that exposure to one tone language (i.e., Mandarin) influences the naïve perception of another tone language (i.e., Thai). That is, the Learners behave differently from the English group which has not learned Mandarin, exhibiting faster RTs on both monosyllabic and disyllabic stimuli and greater accuracy on the disyllabic stimuli.

Therefore, unexpectedly the disyllabic stimuli aid participants with exposure to Mandarin tones in their perception of Thai tones. In contrast, the disyllabic stimuli do not aid the English group. The disyllabic stimuli were expected to aid the English group in their naïve perception of Thai tones as the disyllabic stimuli were thought to better reflect English word stress. Indeed, on the Low+Mid-High+Low comparison the Mandarin group is more accurate than the English group while on the Low+High-High+Low comparison both the Mandarin group and Learner group are statistically more accurate than the English group. Combined these results hint that exposure to Mandarin tones, namely most likely the Rising and Falling tones, may aid in perception of disyllabic tones reflective of the rising and falling stress patterns of English word stress, e.g., Low+High, High+Low. Conversely, the English group appears to be unable to access English word stress in their perception of these two disyllabic stimuli. Thus, we see that exposure to

Mandarin has a positive effect on the naïve perception of Thai tones as witnessed in the diverging performance between the Learner group and English group. The question, however, remains as to the basis for this positive effect, i.e., lower anxiety levels and/or robust encoding of tone, and the exact nature of this effect.

In sum, exposure to a tone language shapes the naïve perception of another tone language. Indeed, the Learners perform at levels comparable to the Mandarin group on a hierarchy on the perception of Thai tones. However, Learners do not pattern in the exact same manner as the Mandarin group. In addition to a few differences in accuracy rates and RTs on individual tone comparisons, the Learner group is faster overall than the Mandarin group, regardless of condition (e.g., tones or segments), including both monosyllabic or disyllabic stimuli.

Thus, while exposure to a tonal system appears to be positive without a doubt, the very nature of this exposure remains somewhat obscure. There are several plausible scenarios to account for these differences between the Mandarin and Learner groups. For instance, we must consider the interaction between the L1 and L2 (Weinreich, 1953; Flege et al., 2003) and therefore, expect to have differences in the perception of Thai tones reflective of the varying levels of strength lexically-contrastive pitch of English and Mandarin against one another (i.e., depending on which is the L1 and L2) (cf., Feature Hypothesis, McAllister et al., 2002). Most likely, however, the robustness of the encoding of L2 tones and their subsequent scope of categorization (i.e., features and/or “tonal space”) might be the strongest candidate as to why the Learners and Mandarin group differ to some extent on the naïve perception of Thai tones. This avenue of inquiry will be explored in the next section describing a Lexical Decision task with Repetition Priming using Mandarin tones involving Learners and the Mandarin group. We now move on to the next experiment in this study to determine the acquisition of Mandarin tones by the Learner group as compared to the Mandarin group.

5.5 Experiment 3: Lexical Decision with Repetition Priming

A Lexical Decision task with Repetition Priming has two components. The participant decides whether a stimulus is a real word or not in the target language, in this case Mandarin. If the participant cannot determine if the words presented are real or not, then, their data are excluded from analyses as their proficiency level in the language would be deemed too low, as long as the items selected are common words in the target language. As discussed in the literature review section (2. Background), in a lexical priming task, there are two main conditions: *repetition* and *minimal-pair*. In the condition *repetition*, an item is presented twice in the list of stimuli whereby the second one appears within a designated time range from the first (See Methodology below). For words (not for non-words), native speakers as well as learners are expected to react more quickly to the second time the same word is heard (i.e., repetition priming) when making the decision if it is a real word or not.

In the other condition (minimal pair), two items that form a minimal pair are presented. However, in the case of Mandarin, a minimal pair that varies by tone only (e.g., 'mother' [ma] with tone 1 and 'horse' [ma] with tone 3) is not normally expected to prime native listeners or highly proficient learners as in the case of repetition, because tonal information is accessed if it is encoded. On the other hand, if a learner does not have a robust lexical representation of the two different tones for this minimal pair, they may react more quickly on hearing the second word of the pair as they would consider them to be "the same word". Priming can only occur for minimal pairs (i.e., 'mother' [ma] with tone 1 vs 'horse' [ma] with tone 3), if the learners have not acquired a robust representation of tone in Mandarin (i.e., if tone is not encoded in Mandarin).

The research questions for this experiment concerning the Learners are as follows:

1. Is the presence of an additional language system in the form of a second language in bilingual speakers robust enough to affect naive perception of non-native tone?

In relation to this research question, the following related subquestions are also asked:

2. Do L2 learners of Mandarin differ in their naïve perception of Thai tone as compared to non-learners of Mandarin from the same L1 background?
3. Do L2 Mandarin learners perform at levels comparable to L1 speakers of Mandarin in a linguistic hierarchy of perceptual ability of Thai tone?
4. Do L2 Mandarin learners perform in a similar manner to L1 speakers of Mandarin, i.e., Are some tones or tonal comparisons easier to perceive than others (cf., pitch height or pitch direction) in the naïve perception of Thai tone in the same manner for both groups?

Accordingly, if the results from the ABX monosyllabic and disyllabic tasks answer these questions, the Lexical Decision task with Repetition Priming should bolster the following claims:

1) Learners have indeed acquired Mandarin tones and it is therefore most likely that these acquired Mandarin tones have influenced their performance on their naïve perception of Thai (and not necessarily "comfort" with tones) (i.e., question 2 above).

2) If indeed Learners have encoded Mandarin tones with a robustness comparable to that of the Mandarin group (as evidenced by the results on the Lexical Decision task with Repetition Priming), then the Learners should react similarly to the Mandarin group on all tasks. However, the Learners might still react differently than the Mandarin group due to the degree of robustness of the acquired Mandarin tones and their native English. The question is again one of familiarity or encoding which motivates the Lexical Decision task with Repetition Priming in the current study (i.e., questions 3 and 4 above).

5.5.1 Methodology

5.5.1.1 Participants

Two groups participated in this experiment: a native speaker group as a control (the same native speakers in Experiments 1a and 1b) and the same learners as in Experiments 2a and 2b. In total, 54 participants were recruited. On the whole, participants were either undergraduate or graduate students at Indiana University. For example, approximately half of the Thai participants were professors at a university in Thailand while some of the other language participants were spouses of students and other non-students employed in the local community (See demographic and language background information presented in Appendix H, pp. 305-314).

Each participant was recruited either through flyers distributed on campus or through English language support courses, Mandarin as a foreign language courses or Second Language Studies courses at Indiana University or by word of mouth (i.e., participants would contact their friends or classmates). All participants were paid $20 for participation in the two ABX tasks and the Lexical Decision task. All procedures were approved by the Indiana University Institutional Review Board.

5.5.1.2 Stimuli and conditions

Four types of experimental stimuli were created, in addition to distractors: Real word pairs differing by tone, real word pairs differing by segmentals, non-word pairs differing by tone and non-word pairs differing in either consonants or vowels. For each type, eight pairs of stimuli were selected. Table 4.13 presents a sample of these four experimental stimuli types

Table 5.11: Examples of real word and non-word test and control stimuli

	Test items (tones)			Control items (segments)		
	IPA	Tone	Gloss, character, *pinyin*	IPA	Tone	Gloss, character, *pinyin*
Real words	[tʰaŋ]	1	'soup' 湯, *tang*	[dʒja]	1	'house, home' 家, *jia*
	[tʰaŋ]	2	'candy' 糖, *tang*	[dʒjɛ]	1	'to pick up (the phone) etc.' 接, *jie*
	[tʰjɛn]	1	'heaven, sky' 天, *tian*	[tʃʰəŋ]	2	'to become' 成, *cheng*
	[tʰjɛn]	2	'to be sweet' 甜, *tian*	[tʃʰaŋ]	2	'to be long' 長, *chang*
	[xə]	1	'to drink' 喝, *he*	[ʃau]	3	'to be few' 少, *shao*
	[xə]	2	'and' 和 or 'river'河, *he*	[dʒau]	3	'to search for' 找, *zhao*
Non-words[11]	[hiŋ]	1	*hing*	[sə]	1	*se*
	[hiŋ]	2	*hing*	[so:]	1	*sou*
	[nue]	1	*nui*	[siɛn]	2	*sian*
	[nue]	2	*nui*	[sən]	2	*sen*
	[gi:]	1	*gii*	[bua]	3	*bua*
	[gi:]	2	*gii*	[buɛ]	3	*bue*

*See Appendix F (pp. 296-300) for a full list of the test and control stimuli used in the Lexical Decision task.

For the test pairs (differing by tone), eight minimal pairs of words and eight minimal pairs of non-words differing only in tone 1 and 2 were chosen (See Table X). Mandarin tones 1 and 2

[11] While participants noted possible similarities to words in Taiwanese and other Mandarin dialects, most were not specific about the details due to a lack of recall after participating in the tasks. However, after examination of the stimuli by a citizen of Taiwan, it was found that there were two words among the training session non-words which are close to words in Taiwanese (which is a Chinese language mutually unintelligible with Mandarin Chinese). Among the distracters there were four non-words similar to Taiwanese words and an additional two non-words (e.g., [lən], [tən]) in the pronunciation of Taiwanese Mandarin (which is a dialect of Mandarin Chinese spoken in Taiwan) which are considered to be words ([ləŋ] 'to be cold', [təŋ] 'to wait' in standard Mandarin on Mainland China (i.e., PRC). More importantly, among the non-words serving as counterparts to the control segmental items, there were six words considered to be words in Taiwanese. Finally, among the non-words serving as counterparts to the test tone items, there were five words considered to be words in Taiwanese. Of these five words, there was one minimal pair (i.e., [juɔ] with tone 1 and 2). Despite the presence of these Taiwanese words, all participants performed above 70% on determining whether stimuli were Mandarin words or not. Presence of items similar to Taiwanese words in the training helped to remind participants that they were determining whether the stimuli were Mandarin Chinese. Also, it should be noted that many people, especially younger people, in Taiwan do not speak Taiwanese and therefore, there is question as to how robust priming may be among the participants from Taiwan. See Appendix F (pp. 292-296) for list of stimuli that are close to Taiwanese or Taiwanese-Mandarin words.

were chosen because of their comparable lengths, unlike tone 3 which is longer and tone 4 which is shorter than both tones 1 and 2. Also, tones 1 and 2 are spoken in a different register: upper register for tone 1 and lower register for tone two. This clear difference between the two tones was important to facilitate a good perceptual discrimination of the stimuli. Pairs of non-words (carrying tone 1 and 2) were constructed by creating non-words that sounded like Mandarin words but were actually not words in standard Mandarin (i.e., phonotactically possible in Mandarin).

For the control pairs (differing by segment), eight minimal pairs of words and of non-words carrying the same tones were chosen. All four tones were used here. Four pairs differed in the vowel while four pairs differed in the beginning consonant. Non-words were constructed in the same way as for the tone non-words. A roughly equal number of pairs with each tone were chosen. There were an additional 64 distracters. These consisted of 32 real words and 32 non-words (15 monosyllabic real words, 17 disyllabic real words, 16 monosyllabic non-words and 16 disyllabic non-words). Since the experimental items were repeated, it was important to also repeat a part of the distracters. Half of this list of 64 was repeated taking every odd item. In total (including the repetition) there were 96 distracters, and 64 experimental (test and control) items. In total, 160 items were presented in the lexical decision task.

Table 5.12: Examples of real word and non-word distracters

Distracters	IPA	Tone(s)	Gloss, character, *pinyin*
Real words	[ʃan]	1	‘mountain’ 山, *shan*
	[gau]	1	‘to be high’ 高, *gao*
	[ʃin]	1	‘to be new’ 新, *xin*
	[tʰoː]	2	‘head’ 頭, *tou*
	[nan]	2	‘to be difficult’ 難, *nan*
Non-words[1]	[fə]	1	*fe*
	[mʷun]	1	*mun*
	[dʒʷun]	1	*zhun*
	[thjo]	1	*tiu*
	[fjɛ]	2	*fie*

*See Appendix F (pp. 296-300) for a full list of the distracters used in the Lexical Decision task.

Exposure to the four tones was balanced throughout the entire task. The number of times each tone was heard, whether in a monosyllabic or disyllabic word, was as follows: tone 1 = 56, tone 2 = 55, tone 3 = 53 and tone 4 = 44. Of the 160 stimuli, 48 were disyllabic or 30% of the 160.

The frequency of the chosen words was checked by consulting the Chinese Proficiency Test (汉语水平考试 Hànyǔ Shuǐpíng Kǎoshì). All the chosen words were on the either the beginning or elementary word lists. Non-words were checked by a native speaker of Mandarin from Xi'an, China to ensure that the non-words seemed "Mandarin-like" and did not violate Mandarin phonotactics. Care was taken to avoid creating words that resembled English whether in the word or non-word case.

Each pair of items was then inserted into a list for presentation to participants. To counterbalance the pairings across the two conditions (repetition or minimal-pair), four lists are needed (repetition: AA, BB; minimal-pair: AB, BA). Each minimal pair was presented in the list in one of four orders: AB, AA, BA, or BB. That is, one member of the minimal pair was presented as the first word (i.e., A) and then, followed by the other member (i.e., B) or by another token of itself again (i.e., A), resulting in the AB, AA orders. This was reversed and done for the other member of the minimal pair (i.e., BA, BB orders). Each pairing of a given experimental stimulus is only presented in one list, and all types of pairings are counterbalanced across the lists[12]. The second item of a pair was manually inserted into each list at a random distance of 8 to 20 items after the first item of the pair. Random distances were selected using Random.org at

[12] For instance, the pair ma1-ma2 (i.e., [ma] tone 1 "mother" vs [ma] tone 2 "hemp"] appears four times, once per list. In List 1 (aa), ma1 will be followed by itself ma1 (note that we use a difference acoustic token for the second occurrence). In List 2 (ab), ma1 will be paired with its tonal counterpart ma2. Lists 3 and 4 are the mirror images. List 3 (bb) will contain ma2 and ma2 (same), and List 4 (ba) will present ma2 followed by ma1.

http://www.random.org/ such that the corresponding counterpart of a targeted pair appeared within 8-20 items after the first item of the pair was presented in the list.

Each list consisted of 160 items with 64 test and control items (16 test real word tone items, 16 control real word segment items, 16 target non-word test items and 16 control non-word segment items). Each list was divided into four blocks to allow three breaks in the task for the participants to rest. Additionally, in order for each participant to use his or her dominant hand to answer if an item was a real word, a version of the program for each list was created for left- or right-handed participants. See Appendix K for the four lists (pp. 330-346).

The presence of repetition priming will be examined by looking at the real word minimal pairs in each group. While priming is expected for all groups for the words in the 'repetition' condition for both tone and segments, we expect Mandarin participants to not experience repetition priming in the minimal pair condition, for either tone or segment. For the L2 learners, if a tonal contrast is not encoded, "false priming" may occur. For example, the listener perceives [ma] tone 1 'mother' and [ma] tone 3 'horse' to be the same word and so, reacts more quickly upon hearing the second word even though it is actually a different word (i.e., mismatch). Such repetition priming is not expected for the Learner group on the control (segment) minimal pair condition.

5.5.1.3 Speaker for Mandarin stimuli Elicitation procedure

A 33-year old female speaker from Xi'an, China, recorded the Mandarin stimuli by repeating the first member of a minimal pair three times in isolation and then, the second member of the minimal pair three times. For the target stimuli she said the first member with tone 2 and then, the second member with tone 1, repeating this pattern for the counterpart non-words.

Two different tokens of each stimulus were needed, resulting in an "a" token and "b" token (both the second or third recording of the stimuli of the test and control stimuli were used in the task). The stimuli were recorded using a Shure SM7B microphone. A Shure M267 4-channel mixer was used to sum all the signals or in this case amplify signals from a dynamic microphone. A Motu 828 (audio interface) converter was used to convert analog to digital signals. A G5 Mac using Peak LE 5.2 digital audio workstation (i.e., a computer plus Peak LE audio software) was used to record and make edits on the stimuli. All files were recorded at 48kHz/24bit and left unprocessed. All the stimuli were spliced into individual wav. files and checked using Praat (Boersma & Weenink, 2015; Boersma, 2001) to verify that the tones were labelled correctly. Stimuli were normalized at 13HZ for all Mandarin stimuli using Audition software.

5.5.1.4 Procedure

The four lists were used evenly among the Mandarin group participants (n=31) so that approximately 7-8 participants each from the Mandarin native speaker group while 5-6 participants each from the Learner group (n=23) were tested on each list. The ordering prevented any possible bias due to ordering and/or stimuli.

Each participant sat in front of the computer, wearing high-quality headphones. Participants were instructed to press keys on the computer keyboard as quickly as possible to indicate whether the stimulus presented was a real word of Mandarin or not. They listened to the auditory stimuli at a self-selected comfortable volume level. At the beginning of the Lexical Decision task, they took a mini training session consisting of 12 items. This training was merely a short version of the actual Lexical Decision task providing feedback such as “CORRECT” or “INCORRECT”. Each button for the YES (real word) or NO (non-word) answer was clearly marked on the left and right alt key respectively on the computer keyboard. This allowed the participant to become accustomed to the task. The researcher sat near the participant as they took the training session. After finishing the training session, each participant was asked whether they understood what was expected of them and were allowed to ask questions. They then proceeded to the actual Lexical Decision task. When the participants finished the Lexical Decision task, they were debriefed, i.e., asked if they had any comments or questions about the task. At the end of participation, the participants were thanked and given $10 for participation in the Lexical Decision task. In the case of the Learners, they were told that raw scores for their tasks would be sent to them at a later time. A time-out of 2500ms was set for participants to respond to each trial (i.e., push the button on the computer to answer whether the X stimulus was a real word or not). The entire procedure generally required between 45-75 minutes in total. Appendix K (pp. 330-347) presents the DMDX (Forster & Forster, 2003) scripts for this task.

5.5.2 Results

As was the case of Experiments 1 and 2, the data from the language background questionnaire allowed for the screening of participants with non-target profiles. The same five participants (heritage speakers and those with exposure to a tone language) were excluded in this task.

For each participant, a mean accuracy score and a mean RT for the lexical decision response were computed for each condition (for words and for non-words). No participant was excluded on the basis of low performance on the distractor items, and no item was excluded on the basis of low performance by the Mandarin native speaker participants. No outliers among either the Mandarin or Learner group were found who scored less than 70% on the real word vs non-word accuracy response. Statistical analyses were run for overall accuracy and RTs for each condition (repeated and minimal-pair), for words and non-words. The RT data were found to be skewed. Hence, a log-transformation was applied to the RT (Log-RT) to obtain a normal distribution. Only the “more-understandable” converted RTs are reported with their lower and upper confidence intervals (CI) in all the following results tables and descriptions in this chapter. Table 5.13 presents the RTs for words while Table 5.15 presents the RT for non-words. Sidak was used over the more conservative Bonferroni to estimate significance.

Table 5.13 presents RTs to words by both groups while Table 5.14 presents LogRTs. RTs for both groups are generally somewhat higher when hearing the first member of a tone minimal pair than for hearing the first member of a segment minimal pair. RTs are generally quite similar between both groups. Figure 5.13 shows the difference in RT between the first and the second occurrence in graphical form (i.e., RT2-RT1). Bars above zero indicate that the RT of the second member of a pair was faster (smaller) than to the first.

Table 5.13: Mean RT (ms) for word pairs in each condition and for each occurrence, for both groups

Language Group	Occurrence	Tone		Segment	
		Minimal Pair	Repetition	Minimal Pair	Repetition
Learners	first	1164 (87.7; 94.8)	1117 (84.1; 91.0)	1140 (85.9; 92.9)	1094 (82.4; 91.8)
	second	1143 (86.1; 93.1)	1019 (76.7; 82.9)	1109 (83.5; 90.3)	1023 (77.1; 83.3)
	difference (2nd – 1st)	-21	-98	-31	-71
Mandarin	first	1143 (68.9; 70.5)	1122 (65.2; 69.2)	1067 (62.0; 65.8)	1072 (62.3; 66.1)
	second	1145 (66.6; 70.7)	1033 (60.0; 63.7)	1052 (61.1; 64.9)	1064 (61.8; 65.7)
	difference (2nd – 1st)	-.6	-89	-15	-7.4

Note: CI = confidence interval (lower; upper

Table 5.14: Mean LogRT (ms) for word pairs in each condition and for each occurrence, for both groups

Language Group	Occurrence	Tone		Segment	
		Minimal Pair	Repetition	Minimal Pair	Repetition
Learners	first	3.067 (0.034; 0.034)	3.05 (0.034; 0.034)	3.06 (0.034; 0.034)	3.04 (0.035; 0.035)
	second	3.06 (0.034; 0.034)	3.01 (0.034; 0.034)	3.05 (0.034; 0.034)	3.01 (0.034; 0.034)
	difference (2nd – 1st)	-0.007	-0.04	-0.011	-0.03
Mandarin	first	3.06 (0.027; 0.026)	3.05 (0.026; 0.026)	3.03 (0.026; 0.026)	3.03 (0.026; 0.026)
	second	3.06 (0.026; 0.026)	3.01 (0.026; 0.026)	3.02 (0.026; 0.026)	3.03 (0.026; 0.026)
	difference (2nd – 1st)	-0.0	-0.04	-0.01	-0.0

Note: CI = confidence interval (lower; upper

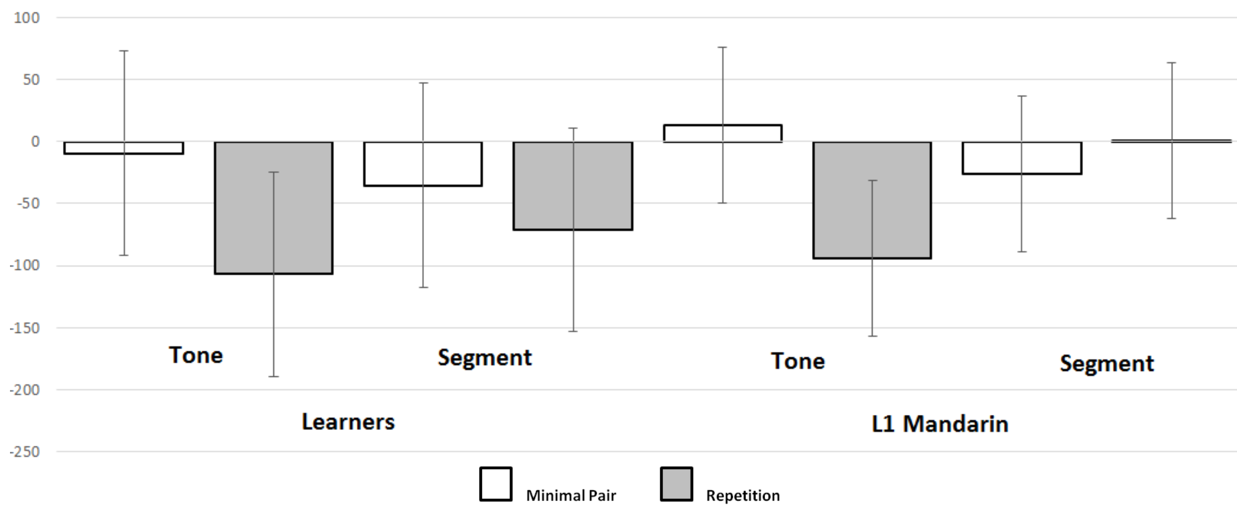


Figure 5.13: Difference in mean RT (in ms) to words for each condition and each group (RT2-RT1)

Table 5.15 presents RTs to non-words by both groups while Table 5.16 presents LogRTs. The Mandarin group reacted at fairly similar rates whether for the same or minimal pair words on both the tone and segment condition. The Learner group, however, reacted faster upon hearing the second member of the same and minimal pair on the tone condition. This group also reacted somewhat faster upon hearing the second member of a same pair on the segment condition, but exhibited approximately the same reaction time rate upon hearing the second member of the minimal pairs on the segment condition.

Table 5.15: Mean RT (ms) for non-word pairs in each condition and for each occurrence, for both groups

Language Group	Occurrence	Tone		Segment	
		Minimal Pair	Repetition	Minimal Pair	Repetition
Learners	first	1186 (101.8; 108.4)	1233 (103.3; 112.8)	1167 (97.8; 103.8)	1178 (98.7; 104.7)
	second	1125 (91.8; 100.0)	1156 (94.4; 105.7)	1167 (97.8; 106.7)	1148 (93.8; 105.0)
	difference (2nd – 1st)	-61.164	-76.993	0	-29.452
Mandarin	first	1135 (73.3; 75.6)	1104 (71.3; 73.5)	1094 (70.7; 75.5)	1072 (69.2; 71.4)
	second	1135 (70.9; 78.4)	1122 (72.5; 74.7)	1089 (68.0; 72.5)	1091 (70.5; 75.4)
	difference (2nd – 1st)	0	17.939	-5.026	19.921

Note: CI = confidence interval (lower; upper).

Table 5.16: Mean LogRT (ms) for non-word pairs in each condition and for each occurrence, for both groups

Language Group	Occurrence	Tone		Segment	
		Minimal Pair	Repetition	Minimal Pair	Repetition
Learners	first	3.07 (0.039; 0.038)	3.09 (0.038; 0.038)	3.07 (0.038; 0.037)	3.07 (0.038; 0.037)
	second	3.05 (0.037; 0.037)	3.06 (0.037; 0.038)	3.07 (0.038; 0.038)	3.06 (0.037; 0.038)
	difference (2nd – 1st)	-0.02	-0.03	0.0	-0.01
Mandarin	first	3.06 (0.029; 0.028)	3.04 (0.029; 0.028)	3.04 (0.029; 0.029)	3.03 (0.029; 0.028)
	second	3.06 (0.028; 0.029)	3.05 (0.029; 0.028)	3.04 (0.028; 0.028)	3.04 (0.029; 0.029)
	difference (2nd – 1st)	0.0	0.01	0.0	-0.01

Note: CI = confidence interval (lower; upper).

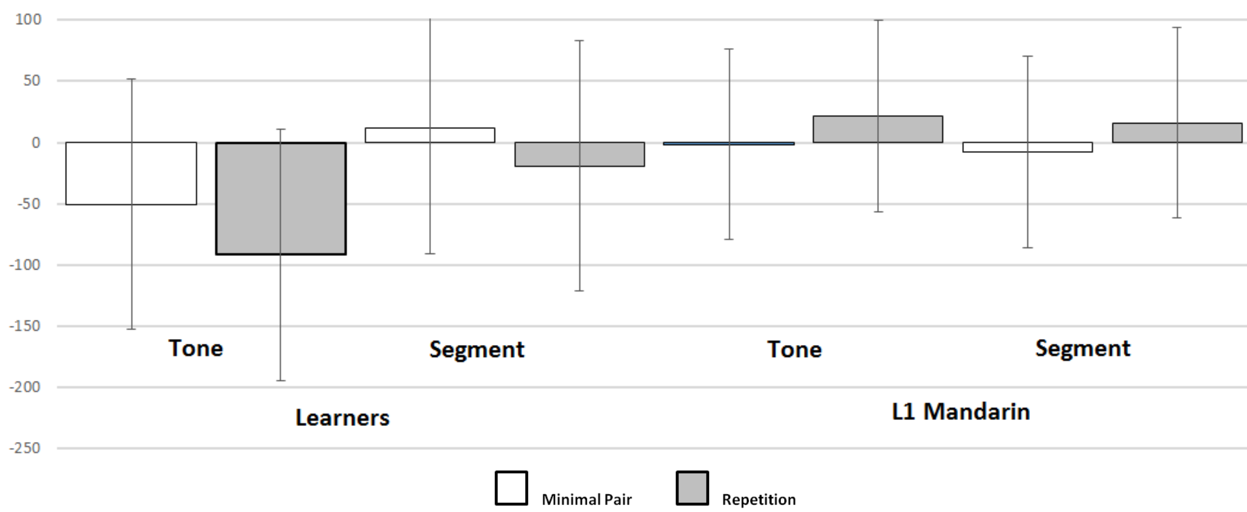


Figure 5.14: Difference in mean RT (in ms) to words for each condition and each group (RT2-RT1)

For results on real words, a Linear Mixed Effects model was run on the LogRT, declaring Condition (tone vs segments) and Language (Mandarin, Learners) as fixed factors and Subjects as a random factor. When looking at the Type III tests of fixed effects, the F-tests show a main effect of Condition ($F(1, 1444.212) = 7.108$, $p = 0.008$). There is no significant effect of Language ($F(1, 46.81) = 0.102$, $p = 0.751$). We also see a main effect of Minimal pair_same ($F(1, 1444) = 18.22$, $p < 0.001$) where Minimal pair_Same compares the RTs on minimal pairs vs the RTs on the same word pairs.

Also, a marginal interaction was evident between Language and Minimalpair_same ($F(1, 1444)=3.63$, $p = 0.57$) and Condition and Minimalpair_same ($F(1, 1444)=3.74$, $p = 0.53$). An analysis for non-words was also run.

Table 5.17: Type III Tests of Fixed Effects for the overall reaction times on the Lexical Decision task with Repetition Priming

Type III Tests of Fixed Effects[a, b]				
Source	Numerator df	Denominator df	F	Sig.
Intercept	1	46.812	120601.650	.000
Language	1	46.812	.102	.751
condition	1	1444.212	7.108	.008
minpair_same	1	1444.190	18.219	.000
Language * condition	1	1444.212	1.255	.263
Language * minpair_same	1	1444.190	3.633	.057
condition * minpair_same	1	1444.266	3.744	.053
Language * condition * minpair_same	1	1444.266	1.259	.262
a. lexical = Word				
b. Dependent Variable: RT_log.				

The RTs of both the Learners and Mandarin group differ on the repetition of the same tone condition but not on the repetition of the same segmental condition. To be exact, for the repetition of the same tone words, the learners show a significant difference in the first and second reaction times ($p = 0.011$). The Mandarin group also display a significant difference on the same tone words ($p = 0.004$). For the repetition of the same segment words, the Learners show a marginal difference between the first and second reaction times ($p = 0.090$). However, the Mandarin groups does not show any statistical difference between the first and second reactions times for the same segment words ($p = 0.979$). There were no statistical differences for either the Mandarin group or the Learners on the minimal pair words, whether differing by tone or segmentals as determined by a two-tailed test.

5.5.3 Discussion

The results on the Lexical Decision task allowed us to examine the encoding of Mandarin tone by both the Learners and the Mandarin group. We expected the Mandarin group to be primed (i.e., react faster) on the second occurrence of the same word, and not to be primed on the occurrence of the second member of a minimal pair differing by either tone or segment. As for the

Learners, we predicted that if they had acquired a fairly robust representation of Mandarin tone, they too would prime on the second occurrence of the same word but not when presented with the second member of a minimal pair. In the case of a lack of representation or weak representation of Mandarin tone, we predicted that the Learners might exhibit a "false prime" on the occurrence of the second member of minimal pair. That is, they would consider the minimal pairs differing by tone to be the same word.

Priming generally occurs among the Mandarin speakers and Learners on both tones and segmentals in the condition "repetition," i.e., where the same word is repeated twice. As such, the instrument to test priming is overall valid. More importantly, it seems that Learners have been able to encode the tones of Mandarin Chinese. The one unexpected result is that repetition priming on segmentals does not occur among the native speakers of Mandarin (condition "repetition"). There are several possibilities as to why this is so. One would be that the task is not valid, meaning that the task is not extracting an accurate phonological representation of the target stimuli, i.e., task effect. However, the task seems to be valid if looking at the other conditions and the Learners. Another possibility concerns the stimuli. For instance, there may be greater discrepancies between Mandarin dialects in terms of segmentals versus tones. That is, certain segmentals may vary to a greater extent than the mere four tones, e.g., nasals such as [n] and [ŋ] as in "to be cold" [lən] in Taiwanese Mandarin vs [ləŋ] in standard Mainland Mandarin Chinese. Furthermore, in addition to knowledge of other Mandarin dialects, participants may be familiar with other "Chinese languages" (e.g., Taiwanese, Cantonese), e.g., [sua], on the training session was often said to be a word in a Mandarin dialect or Chinese language like Taiwanese.

Also, looking at the non-words, as expected no difference in RT between the first and the second occurrence for the control items (segmentals) is visible for the Mandarin group. That is, we expect them to react with the same speed on the non-words regardless of whether a non-word stimulus is the first or second occurrence. What is unexpected are the faster RTs on the second hearing of tone non-words, for both “repetition” and “minimal pair” conditions. This observation might call into question the faster reaction times upon hearing the second member of the “same pair” of the real words varying by tone except for the fact that the times for the second member of the minimal pair of the real words varying by tone are the same (i.e., lack of a difference in RT between the 1^{st} and 2^{nd} occurrences) for the Learner group.

Thus, we see that generally priming occurs for tone in both the case of the Mandarin and Learner group, indicating that exposure to Mandarin results in a fairly robust encoding of the phonological representation of tone. Also, at least in the case of the Learners, priming on segmentals is corroborated by their performance on the segmental condition.

In sum, the Learners of Mandarin appear to have a robust encoding of the phonetic and tonal representation of Mandarin words. This result alone bodes well for learners and teachers of Mandarin as second language.

5.6 Experiment 4: Pronunciation assessment – Phonological Proficiency Task

A pronunciation elicitation task was included to provide a means to make comparisons across a variety of Learners who exhibited a wide range of learning/exposure backgrounds. This task assessed participant's proficiency in terms of pronunciation rather than by either the level of the Mandarin Chinese course they were taking or had taken or a placement exam of some type. This was considered ideal as pronunciation and global language proficiency do not necessarily reflect one other (e.g., 'native-like' pronunciation but limited vocabulary and poor grammar or 'extremely non-native-like' pronunciation with high-level grammar and a rich vocabulary).

5.6.1 Methodology

5.6.1.1 Participants

The participants were the same as in section 5.5. Experiment 3, namely the Learners (See 5.5.1.1).

5.6.1.2 Stimuli

L2 Mandarin learners were given a sheet with questions and a reading passage. Learners chose from either a script with the questions and reading passage written in simplified characters or a script written in traditional characters. Simple questions and a reading passage were used to elicit both spontaneous speech and a reading sample was thought to give a more balanced look at each participant's pronunciation. Questions were simple in order to elicit 'felicitous' samples of pronunciation without applying too great of a cognitive load on the speaker. The paragraph for reading purposes was written in *pinyin* (i.e., phonetic alphabet). The questions were also written in *pinyin* with an English translation while the reading passage was written in Mandarin characters and *pinyin* only without an English translation. *Pinyin*, or the Roman alphabet used by the People's Republic of China to teach the pronunciation of Chinese characters to their people and to learners of Mandarin, was offered to avoid conflating possible pronunciation issues with reading issues, particularly of recalling the pronunciation of Chinese characters. Nevertheless, it is still possible that pronunciation of English spelling conventions could interfere with the pronunciation of Mandarin *pinyin*. Tones were marked with diacritics used in textbooks but tone sandhi was not marked (e.g., tone 3 + tone 3 was not marked as tone 2 + tone 3) as it is usually not marked for native speakers when learning Mandarin as children nor in texts for learners of Mandarin as an L2. The questions and reading passage were checked by a native speaker of Mandarin from the PRC (i.e., Mainland China) to ensure authenticity (See Appendix M, pp. 351-352 for the questions and reading passage).

5.6.1.3 Procedure

Participants were asked to first introduce themselves, answer three questions, count from one to 20, and then, read the passage. Participants' answers and reading were recorded in a soundproof booth using Praat (Boersma & Weenink, 2015; Boersma, 2001). The participant then signed for the remuneration fee, was thanked, and told that they would be sent their results on the tasks after the results of all the participants had been collected and analyzed.

Three samples were extracted from the recording of each participant:

1) An approximately 7-second excerpt of the answer to "Why do you study Mandarin?"

2) A section of the participant counting from 1-10

3) The participant's reading of the final two sentences from the reading passage: *Míngnián, wǒ yào qù Táiwān xué Zhōngwén. Bìyè yǐhòu, wǒ yào zài Zhōngguó zuò shēngyi* (Next year I want to go to Taiwan to study Mandarin. After I graduate, I want to do business in China.).

5.6.1.4 Praat script

The samples were later presented to three native speakers who rated them from 1 to 7 on a Likert scale with 1 being poor and 7 being native-like. Three evaluators were enlisted to rate the pronunciation of the L2 Mandarin learners. One was a male PhD student in the Linguistics Department at Indiana University from Taiwan with a specialty in phonetics and L2 phonology. The second was a female PhD student in Second Language Studies from Xi'an, PRC with a specialty in pragmatics. The third was a female PhD student in the Linguistics department from Taiwan who was specializing in phonetics and psycholinguistics.

Three samples from each of the 23 learners were presented to the three raters for a total of 69 samples.

Any long pauses beyond 400ms were cut down to 400ms to prevent the raters from conflating fluency with pronunciation. The samples were randomized, and there was no training practice. Raters were told to rate pronunciation in terms of vowels, consonants and tones, but to try to ignore fluency, pausing, grammar, any odd meanings, possible poor splicing, etc. The rating screen for the Phonological Proficiency task with the question, replay button and "re-rate" button looked as follows (Figure 5.17):

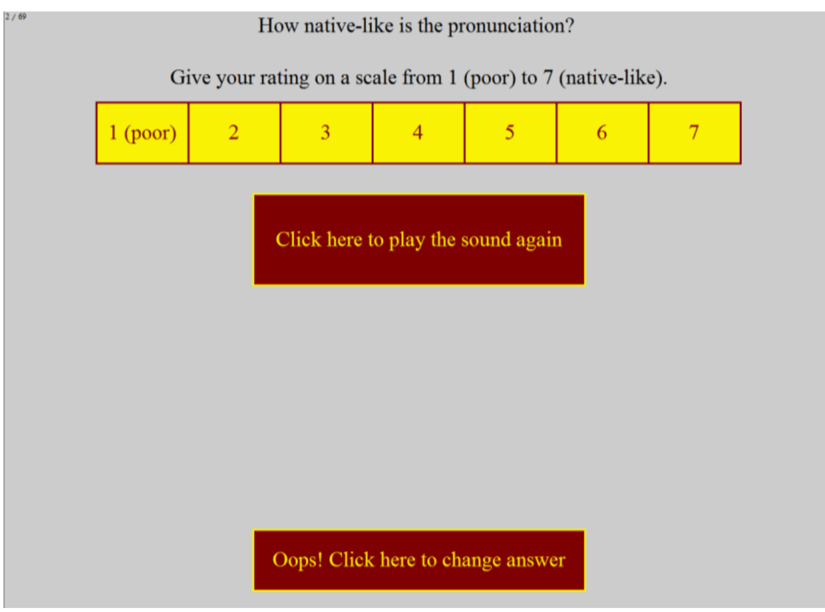


Figure 5.17: Screen shot of the Pronunciation Proficiency task

5.6.2 Results

The participants who were removed from the analysis for the ABX monosyllabic, ABX disyllabic and Lexical Decision task were retained for the proficiency rating task although their results were not evaluated. Again, as was done in the analysis for the Thai ABX tasks and Mandarin Lexical Decision task, two learners who were born in China and then, later adopted and grew up speaking English were removed (cf., Pierce et al., 2014) along with the three learners who were heritage speakers with long-term exposure to a tonal language since birth.

As was the case of Experiments 1 and 2, the data from the language background questionnaire allowed for the screening of participants with non-target profiles. The same five participants (heritage speakers and those with exposure to a tone language) were excluded in this task.

Scores are presented for three types of samples: counting from 1-10 (numbers), answering the question as to why the participant studies Mandarin (question), and short excerpt from the participant reading a passage (reading). Three scores for each sample along with an average of the three are presented in Table 5.18 below. The average overall scores for the three types of samples vary. The totals with averages are as follows: Number (76.3, 4.2), Question (68.3, 3.8), and Reading (68, 3.78). At first glance, it appears that the pronunciation of numbers is higher in scores, but at closer examination of scores, we see that the evaluator one scored the participants on both the question and reading samples with much lower scores, dragging the overall scores for these two sample types down (See Table 5.18 below). Otherwise, the overall scores appear to be relatively similar for each sample type.

There appears to be a slight divergence of scores between the learners who have not studied in a Mandarin-speaking country for at least six months and the learners who have lived and studied in a Mandarin-speaking country. The average scores are for learners who have not studied abroad is 3.45 (or a total of 31.01 for all nine participants) while the average score learners who have studied abroad is 4.43 (or a total of 39.87 for all nine participants). While there appears to be overall higher scores for the participants who have studied or lived in a Mandarin-speaking country, there are a few lower scores among this group as well. In contrast, the scores of the participants who have not studied or lived in a Mandarin-speaking country are in general lower, but there are also a few relatively higher scores. An independent-samples t-test was conducted to compare the two groups. However, there was no significant difference in the scores between the less advanced learner group (M = 3.45, SD = 0.84) and the more advanced learner group (M = 4.43, SD = 1.56); $t(7)=1.47$, $p=0.184$. Moreover, five out of the nine overall scores of the learners who have studied or lived in a Mandarin-speaking country are higher than the lowest overall score of the heritage speaker/adoptee group. Among the group of learners who have lived and studied in a Mandarin-speaking country, there are relatively low scores as well despite their experience in a Mandarin-speaking country (e.g., 2.33 and 2). The highest overall score of 6.33 was obtained by one learner with experience in a Mandarin-speaking country. Simultaneously, the lowest overall score of 2 can also be found among the learners who have lived or studied overseas.

Table 5.18: Phonological Proficiency Task scores for each participant by type, evaluator and total average

	Numbers			Question			Reading			Total Average Score (higher = more fluent)
Evaluator	1	2	3	1	2	3	1	2	3	
Participant										
L119*	1	3	1	2	2	3	1	4	1	2
L064	3	4	1	1	3	3	1	3	1	2.22
L106*	4	3	3	1	3	3	1	2	1	2.33
L005	3	3	1	2	2	4	1	4	2	2.44
L123	4	2	6	1	4	1	3	3	3	3
L080	2	3	3	2	5	4	3	4	4	3.33
L039*	5	3	3	1	4	2	4	4	4	3.33
L025	4	3	2	5	3	4	3	4	2	3.34
L120	3	3	4	2	5	6	2	4	3	3.56
L118	5	3	3	2	4	2	4	5	7	3.89
L032*	6	3	4	3	4	6	3	5	5	4.33
L122	3	7	4	3	4	4	4	6	6	4.56
L087	4	6	7	4	3	4	4	5	5	4.67
L100*	6	3	7	4	4	7	4	4	5	4.89
L035*	6	6	7	4	5	6	4	4	5	5.22
L135*	5	6	7	5	6	7	3	5	5	5.44
L109*	7	7	7	6	6	6	4	6	5	6
L054*	7	6	7	5	6	7	5	7	7	6.33
Overall average	4.3	4.1	4.3	2.9	4.1	4.4	3	4.4	3.9	3.94
Scores of participants excluded from the study										
L083	5	5	7	4	5	4	4	5	3	4.67
L098	6	6	7	5	6	5	5	5	4	5.44
L115	7	7	7	6	6	7	7	7	7	6.78
L117	6	5	5	4	4	6	4	5	3	4.67
L121	6	7	7	6	7	6	5	6	7	6.33

Note: L083 and L117 are adoptees while L098, L115 and L121 are heritage speakers of Mandarin or have exposure to a “Chinese language” such as Cantonese or Taiwanese.
* Asterisk indicates learners who have lived or studied in a Mandarin-speaking country.

5.6.3 Discussion

It seems that exposure to Mandarin by living or studying in a Mandarin-speaking country results in a higher evaluation of a learner’s pronunciation by native speakers. However, statistical tests do not support a difference between groups varying by exposure (i.e., long-term in-country exposure of six or more months versus no or short-term in-country exposure) experience in a

Mandarin-speaking country. Thus, perhaps more stringent guidelines for evaluation of samples or more stringent methods for obtaining samples is required to tease apart results of the groups. Also, overall scores do not indicate necessarily what the evaluators were adopting as criteria (cf., Munro and Derwing, 1995) as most likely evidenced by the discrepancy in overall scores between evaluator 1 and the other two evaluators on the question and reading samples. That is, even if the evaluators are truly only concentrating on pronunciation, we do not know whether segmentals played a greater or lesser role than suprasegmentals. Plus, we do know what aspect of suprasegmentals are being evaluated, e.g., fluency, bias toward perhaps one tone being mispronounced, etc. Again, a speak-aloud protocol used during evaluation of tones would shed more light on what exactly the evaluators were considering.

As for sample type (i.e., reading vs spontaneous speech, question vs "rote" speech, numbers), there does not seem to be any consistent variation in scores. Thus, it seems that pronunciation is consistent across sample types despite possible differences in cognitive load, e.g., lowest for reciting numbers with possibly highest for spontaneous answers to questions and/or reading.

5.7 General discussion (all experiments in this chapter)

In this chapter, we focused on the perception of tone by English learners of Mandarin as an L2. The focus was on an interrelated two-pronged approach: second-language acquisition/lexical encoding of Mandarin tones and the naïve perception of Thai tones. In turn, this group was compared to two groups: an English counterpart group who had never learned Mandarin Chinese or any other tone language and a Mandarin group who to varying degrees spoke English as a second language.

Research question three asked whether L2 learners of a tone language are able to lexically encode tonal information to constrain lexical access. The repetition priming task determined that the L2 Mandarin learners have encoded Mandarin tones in their lexical representations for words. Their production of Mandarin tones (pronunciation) exhibits varying degrees of proficiency whereas their perception of Mandarin tones is on the whole more consistently robust across the group.

Moreover, this robustness of the encoding of Mandarin tones is attested to by the answer to research question two asking whether the learning of a tone language aids in the naïve perception of the tones of another non-native tone language. Specifically, one subquestion to research question two asks whether L2 learners of a tone language differ in their naïve perception of another tone language as compared to participants with the same L1 who are not learning a tonal L2. Indeed, we can answer definitively that on naïve perception of Thai tones the performance differs between Learners and a group of English speakers who have not learned Mandarin Chinese. That is, we witness the superior performance of the learner group on the naïve perception of Thai tones as compared to the lower performance of the counterpart English "monolingual" group.

However, performance by the Learner group did not pattern consistently with that of the Mandarin group, particularly in terms of RTs on the naïve perception of Thai tones. Thus, we answer the third subquestion to research question two. Namely, we address whether L2 learners of a language perform in a similar manner to L1 speakers of that one language. That is, are some tones or tonal comparisons easier to perceive than others (cf., pitch height or pitch direction) in the naïve perception of another tone language? And do the two groups react in the same manner? The answer appears to be that some tones are easier for those with exposure to Mandarin and that both groups, i.e., Learners and Mandarin group, react in a similar manner in terms of accuracy levels but less so in RT speeds. Thus, this observation also answers another subquestion to research question two. To be exact, it answers whether L2 learners of a tone language perform at levels comparable to L1 speakers of that tone language in a linguistic hierarchy of perceptual ability of another tone language. The answer is that overall they generally do, but as noted there are differences in RTs as well as discrepancies in accuracy rates on some individual tone comparisons. For example, on RTs the Learner group was significantly faster than the English group on many of the individual tone comparisons while the Mandarin group was not.

As suggested, this trend in RTs highlighting differences between the Learners and Mandarin group suggests differences between the two groups despite overall comparable accuracy performances. These RT differences may be attributed to several possible causes. The Learner group may have weaker and/or a different or smaller scope of categories for Mandarin tones as compared to that of the Mandarin group which conversely allows them to react faster, perhaps due to being able to hear more subtle differences (cf., Stagray & Downs, 1993). Additionally, while their L1 English is expected to be an influential factor (cf., Interaction Hypothesis, Flege, Yeni-Komshian & Liu, 1999; Weinreich, 1953), the slower RTs of the native English "monolingual" group on the naïve perception of Thai tones leaves little room to speculate that there is influence from their native English on the faster RTs in the case of the Learner group.

Additionally, we see universal tendencies across the three groups of English, Mandarin and Learners. For example, scores by speakers of a non-tonal language may be lower, but they oftentimes reflect the same trend of the other groups, merely at lower accuracy rates. That is, we see similar patterns in individual tonal comparisons relative to the other individual tonal comparisons within a group. Oftentimes, all groups exhibit their lowest score on the same individual tonal comparison as well as their highest score on the same individual tonal comparison. However, we also group-specific tendencies most likely due to the functionality of lexically-contrastive pitch in the L1. These language- or group-specific tendencies emerge in individual tone comparisons and thereby, illuminate the actual mechanisms behind true perception obscured by accuracy rates for overall tone perception.

Also, the addition of disyllabic stimuli to the disyllabic ABX task did not render tone perception easier for the English group. This group performed similarly on the monosyllabic ABX task. Thus, we do not see the influence of word stress in English. English speakers, therefore, seem not to be able to access word stress in their perception of Thai tones.

Thus, we see that exposure to a tone language through learning does boost the naïve perceptual accuracy of non-native tones by a learner. Indeed, the perceptual ability of a learner differs from that of a monolingual counterpart. Furthermore, it patterns overall similarly to that of a native speaker of Mandarin, but does not necessarily entirely reflect that of an L1 speaker of that tone language.

6. Conclusions and future directions

Explained in broad strokes the original overarching objectives of the current study were twofold. First, the study attempts to answer whether or not lexically-contrastive pitch in the L1 has an effect on the naïve perception of non-native tones (i.e., Thai) with varying degrees of effect depending on this functionality of lexically-contrastive pitch in the L1. Second, after determining the baseline effect from lexically-contrastive pitch in the L1, the study further attempts to explore the effect of having learned a tonal language (i.e., Mandarin) as an L2 on the naïve perception of non-native tones of another unknown tonal language (i.e., Thai). Each of these objectives was further qualified by questions concerning the role of individual tone types and type of stimuli (i.e., monosyllabic, disyllabic).

I first discuss whether each of the research questions of this study was answered or not. Then, I move onto the implications of these answers. The first overarching research question of the three is as follows:

1. RQ1: Does the varying presence of certain features in the L1, specifically lexically-contrastive pitch, aid in the naïve perception of non-native tone, thus resulting in a linguistic hierarchy of perceptual ability?

Two subquestions to this first overarching question are:

1.1 RQ1.a: Are some tones or tonal comparisons easier to perceive than others in the naïve perception of non-native tone?

1.2 RQ1.b: Are tones presented in disyllabic stimuli easier or more difficult to perceive compared to monosyllabic stimuli?

Results from this study show that the naïve perception of non-native tone does indeed vary according to the L1 of the listener. It is highly likely that lexically-contrastive pitch in the L1 shapes this perception given the fact that the influence of intonation is the same across languages and "blocked" by the domain of a word for lexically-contrastive pitch and phrase for intonation. In short, the greater the functionality of the lexically-contrastive pitch in the L1, the greater the salience of pitch in the L1 and therefore, superior perception of non-native features of lexically contrastive pitch.

Thus, a linguistic hierarchy of some nature does emerge according to the functionality of lexically-contrastive pitch in the L1. The results of the accuracy rates on the combined tasks of the study supports a hierarchy of perception as follows from high to low accuracy: Thai > Mandarin > Japanese > English, Korean. Therefore, the hierarchy is not necessarily clearly defined as was the case for the pilot study. Also, the reaction times contradict the predicted hierarchy in some respects with Japanese being faster than the L1 Mandarin group, which is higher up the hierarchy of accuracy due to the greater functionality of tone in Mandarin. Additionally, the L1 Korean group was faster than the L1 English group, another unexpected "leap frog" in the hierarchy. Yet, these tendencies span across the segmental items as well, indicating an overall tendency not specific to the tone items, and thereby, suggesting other processes. That is, the L1 Mandarin group may be slowed down by the very fact that Mandarin has tones which require more processing of the Thai tones when attempting to map Thai tones as complete units (cf., unitary model) onto Mandarin tones (cf., categorization, Stagray and Downs, 1993). In contrast, the Japanese may be perceiving the Thai tones in a dimoraic series of [+High] and [-High] pitch units. Thus, for Japanese Thai tones are "ideally" perceived as each stressed syllable in Thai must be bimoraic (Bennet, 1994 cited in Zsiga and Nitisaroj, 2007) and Thai tones may be "decomposed" into high or low pitches aligned with one and/or both moras in a syllable (Moraic Alignment Hypothesis, Morén and Zsiga, 2006).

The final question is whether the L1 Korean group is faster or the L1 English group is slower. We might hazard that the L1 Korean group is faster as they show a tendency to be faster than the L1 Mandarin group as well. The speed may be attributable to several factors. First, we cannot entirely discount the possible influence of lexically-contrastive pitch in Korean either through the emergence of such pitch in Seoul Korean among younger Koreans born after 1965 (Silva, 2006) for Seoul Korean speakers or possibly through the Kyungsang dialect for dialect speakers. Furthermore, this pitch usage may be similar to that in Japanese and thus, account for the speed of both groups.

Nor can we discount the possible influence of English as an L2 by the participants. In contrast, the L1 English group appears to not be able to access the pitch correlate of word stress, suggesting that pitch is indeed a negligible feature of word stress. However, the functional load of pitch to mark stress in both English and Korean must be considered. First of all, there are few minimal pairs in English differing by word stress. Also, in English word stress pitch may not be as salient as the other correlates. In standard Korean, on the other hand, there appears to be no minimal pairs differing by pitch. However, the Korean language features pitch as a part of a word's phonetic representation which must be remembered as its placement within a word varies, even if pitch does not contrast minimal pairs. As such, an answer to the speed of the L1 Korean group on the perception of Thai tones may lie in the use of pitch in Korean words.

Additionally, there may be non-linguistics reasons as to why the hierarchy of perception in the current study is not necessarily clearly defined as was in the case of the pilot study. This may be chalked up to differences in the testing instrument. As mentioned previously, the stimuli in the two tasks of the pilot study and current study varied. In the pilot study, the four-paradigm trials (i.e., ABB, ABA, BAA, BAB) used different segments for each of the four parts of the paradigm while in the current task the same string of segments was used for each of the four parts. For example, in the pilot study AXB task, we have [ba:] HHR, [pu:] RRH, [su:] HRR, and [t^ha:] RHH whereas in the current ABX task, we have one string of segments such as [no:i] used for all four trials of the four-part paradigm. This may have allowed participants to be trained on these items or at least be cued to focusing on tone for these trials. Thus, the cognitive load of the traditional presentation used in the current study may be lower than that of the "unique" presentation of the pilot study and thereby, prevent the extraction of a faithful representation of the perception of Thai tones by each of the groups.

Nevertheless, there is an overall hierarchy which is determined by a combination of language-specific and universal tendencies as evidenced in individual tonal comparisons. These individual tone comparisons provide a more nuanced insight into the perception of Thai tones and seem to come to bear upon the theoretical framework of perception as will be discussed further in this chapter.

Universal tendencies exist in the sense that certain tonal comparisons appear to be easier or harder across all language groups although the relative accuracy scores may be higher or lower between groups. Language-specific tendencies exist as some tonal comparisons appear to be easier or harder for only some of the language groups. Thus, the hierarchy of performance is formed by these discrepancies in performance between each language group on these individual tonal comparisons. It is then the performance on each individual tonal comparison that precisely illuminates exactly how lexically-contrastive pitch in the L1 shapes the naïve perception of the non-native tone system of Thai. Simultaneously, these results also greatly shape the very framework needed to interpret these results.

Theoretical framework

The pilot study derived specific predictions based on the functionality of a phonetic dimension (Feature Hypothesis, McAllister et al., 2002) for the naïve perception of tone, establishing a baseline for tone perception focused on the functional use of linguistic pitch in four language types. In order to model the data, a more narrowly defined *Pitch Functionality Hypothesis* was put forth: The degree to which pitch differentiates lexical items in the L1 (i.e., lexically-contrastive functionality) shapes the naïve (= non-learner) perception of a non-native lexically-contrastive pitch system, as in this case, of a non-native tone system. Unlike the Feature Hypothesis, the *Pitch Functionality Hypothesis* takes into account several constraints specific to lexically-contrastive pitch, going beyond defining functionality merely as "greater usage" or "degree of usage." This approach is in response to the fact that vowel length the focus of the Feature Hypothesis study by McAllister et al. (2002) differs from lexically-contrastive pitch in that lexically-contrastive pitch is a complex phenomenon encompassing a large typological variety of pitch patterns, analogous to vowel/consonant inventories. In short, the *Pitch Functionality Hypothesis* works within the Feature Hypothesis but allows for a more streamlined, comprehensive approach to guide future study of the phenomenon of L2 perception of lexically-contrastive pitch across typologically-diverse usages of lexically-contrastive pitch. Additionally, it must be noted that the McAllister et al. study looked at L2 learners while the pilot study for this current study looked at naïve listeners. This difference, however, does not affect the formulation of the *Pitch Functionality Hypothesis* as will be fleshed out below.

Unlike the Feature Hypothesis, the *Pitch Functionality Hypothesis* requires several other constraints due to the pitch pattern inventory. The first constraint concerning the definition of pitch functionality in the L1 takes into consideration the prosodic domain of pitch contrasts. The findings of the pilot study suggest that the specific prosodic domain in which pitch differentiates lexical items also constrains performance: domain overlap yields most accurate performance. For instance, Mandarin uses pitch to signal lexical contrast at the syllable domain, which is also the case for Thai, whereas in Japanese, the prosodic domain of pitch contrasts is rather the prosodic word. When domains do not overlap, it appears more difficult to map L1-pitch usage to the non-native pitch contrasts. Similarly, if intonational categories which typically require a phrasal domain cannot be mapped or associated with tonal contrasts in a smaller prosodic domain such as the syllable, then in the case of English, would the pitch correlate of word stress aid in tone perception? This question motivated the introduction of disyllabic stimuli for the current study.

Additionally, three other constraints are taken into account when defining pitch functionality: 1) *exclusivity* to signal lexical contrast, 2) *functional load*, and 3) *inventory of pitch patterns*. *Exclusivity* refers to whether lexically-contrastive pitch is used solely to differentiate words in the L1. For example, in Japanese this appears to be the case while word stress in English includes other correlates such as vowel length, spectral quality and intensity in addition to pitch. *Functional load* refers to the extent and/or number of minimal pairs differentiated in the L1. That is, tone languages such as Mandarin require pitch to distinguish a far larger number of lexical items in comparison to pitch accent (Pierrehumbert & Beckman, 1988) and word stress languages. Finally, *inventory of pitch patterns* refers to the number and type of patterns possible. For example, Mandarin has four tones, one level and three contour tones, while Thai has five tones which are phonologically described as three level and two contour tones but may have three contour tones as the High tone phonetically patterns after the Rising tone in shape. Also, the composition of the tone inventory in a language implies a possible bias toward both pitch height and/or direction. Furthermore, it must be noted that differences exist between similar tone types across languages, i.e., the rising tones in Thai and Mandarin vary in shape.

In short, motivated by the above interpretation of the results in the pilot study, the current study also adopted the Feature Hypothesis (McAllister et al., 2002) in the form of the *Pitch Functionality Hypothesis* with constraints as the primary framework for interpreting results. Also, in response to the possibility of tone-to-tone mapping for Mandarin tones and possibly Japanese pitch accent patterns, the current study keeps in mind the Perceptual Assimilation Model (Best, 1995), which was conceived to interpret the naïve perception of non-native sounds, and its subsequent follow-up PAM-2 model (Best & Tyler, 2007), which is a modified PAM model to account for the interpretation of L2 sounds acquired by learners.

For example, within the *Pitch Functionality Hypothesis*, tone-to-tone mapping may be applied where there is domain overlap, and where pitch categories are robust as seems to be the case for L1 tone language speakers. In this case, a model like PAM-L2 (Best & Tyler, 2007) might be applied to make straightforward predictions for cross-linguistic tone-to-tone perception although such an approach needs to be more clearly elaborated upon and subsequently tested. Current approaches are examining this possibility (Hao, 2012, or So & Best, 2010). However, the question remains as to the extent to which lexically-contrastive pitch categories can be defined for non-tonal language speakers (e.g., Stagray & Downs, 1993). A clearer case for the existence of lexically-contrastive pitch categories could be made for pitch-accent languages such as Swedish or Japanese as such languages are defined as a subclass of tone languages (Yip, 2002: 2) and therefore, allow the application of a tone-to-tone mapping approach as well.

Additionally, based upon the results of the ABX tasks in the current study, the role of phonological features in the L1 on the perception of L2 sounds (Phonological Interference Model, Brown, 1998) was deemed as the optimal means for uniting the *Pitch Functionality Hypothesis* and PAM models into a more comprehensive model. The Feature Hypothesis as indicated by its very name looks at the functionality (i.e., referred to as ‘prominence’ by McAllister et al.) of features in the L1 to determine the degree to which a non-native sound can be perceived and produced by L2 Learners. In contrast, PAM (Best, 1995) often refers to the degree of similarity between L1 and L2 sounds in determining the degree to which a non-native sound can be perceived, but originally did not provide a means for determining what similarity constitutes. However, if sounds are broken down into features as Brown (1998) has done, then, it appears to be possible to more clearly ascertain why certain sounds are easier or more difficult to perceive and hence, produce.

Here, we should again note the differences on several accounts between the Feature Hypothesis (McAllister et al., 2002), and by extension the *Pitch Functionality Hypothesis*, and the PAM model (Best, 1995). First of all, the Feature Hypothesis (McAllister et al., 2002) looks at the absence or presence of a lexically-contrastive feature and then, the extent of the presence of this feature in the L1. In comparison, the PAM model (Best, 1995) also looks at the absence or presence of a lexically-contrastive feature and then, the similarity between the L1 feature/category and its counterpart L2 feature/category. The differences lie in the initial focus of both models. The Feature Hypothesis looked at the suprasegmental of vowel length while PAM looked at segmentals although attempts are being made to apply the model to tones (cf., So & Best, 2010). Also, the Feature Hypothesis looked at learners and PAM looked at naïve listeners. PAM has since been modified as PAM-L2 (Tyler & Best, 2007) to account for learners while the pilot study (Schaefer & Darcy, 2014) and the current study has looked at naïve listeners within the framework of the Feature Hypothesis. Thus, combined the PAM (Best, 1995) and PAM-2 (Tyler & Best, 2007) attempt to account for both the perception of segmentals and suprasegmentals by naïve listeners and learners, In contrast, the Feature Hypothesis has not been applied to the perception of segmentals while naïve perception might be extrapolated from learner data. The Feature Hypothesis appears to be difficult to apply to segmental categories. Nevertheless, the Feature Hypothesis might be applied to the degree of a certain feature used for segmentals, e.g., the rounding of vowels.

Another difference concerns the absence of the target feature. In the absence of a segment, the PAM model predicts a high degree of perception if the segment is perceived as language (e.g., clicks) while the Feature Hypothesis predicts a very low degree of perception. If "similar" sounds are present in the L1, the PAM model then moves onto determining the difficulty of perception according to the degree of similarity. Yet, the PAM model does not determine what accounts for the degree of similarity other than to note the result of perceptual similarity. That is, the PAM model describes the difficulty of L2 sounds by how the listener maps the two sounds onto one, two or no L1 sound. This is where precisely the two models might possibly overlap again. That is, both models would look at the feature that is in the L1 that facilitates or deters perception. Then, we could channel the Feature Hypothesis in terms of the degree to which a certain feature is used in the L1. Now, the PAM model also looks at categories which are perhaps merely a streamlined bundle of features (e.g., place/manner of articulation, voicing for consonants) that compose a sound whether a vowel, consonant, or tone pattern. Thus, the notion of categories and features are kept in mind as a possibility to combine the two models and to account for the naïve perception of tones.

At this juncture is where Brown's Phonological Interference Model (1998) might possibly provide the solution to unite the Feature Hypothesis (McAllister et al., 2002)/Pitch Functionality Hypothesis and PAM models (Best, 1995). In her analysis, Brown posits that the use of the feature [coronal] to distinguish two phonemes in Mandarin Chinese (i.e., alveolar /s/ and retroflex /ʂ/)

allows L1 speakers of Mandarin to perceive the difference between the /l/ and /r/ in English as the feature needed to distinguish these two approximants is the feature [coronal]. In opposition, the phonology of Japanese does not "activate the feature [coronal] in its Universal Grammar toolbox" as the feature [coronal] is not exploited in the Japanese language to distinguish two phonemes from one another. Accordingly, L1 speakers of Japanese have difficulty in perceiving the difference between the English /l/ and /r/ as Japanese phonology lacks this "lexically-contrastive [coronal]". As such, just as the Feature Hypothesis relies on features to determine the acquisition (and by extension naïve perception) of a non-native sound based upon the L1 phonological inventory of features rather than phonemes, i.e., bundle of features (e.g., coronal vs /t/). PAM (Best, 1995) might avail itself of the same means in determining the degree of similarity between an L1 phoneme to an L2 target phoneme. In this manner, melding Brown's phonological feature geometry with PAM, we might conjecture that since a Japanese flap is [+approximant] as are the English [l] and [r], this most likely accounts for why this phoneme is considered similar and used by the Japanese for both [l] and [r] in L2 English or "single category" under the PAM model. Additionally, the lack of lexically-contrastive [coronal] in Japanese accounts for why the [l] and [r] cannot be distinguished. Thus, Brown's Phonological Interference Model (PIM, 1998) based on features is merely qualified by the functionality or prominence (the term used in the McAllister et al., 2002) of a feature. Combining PIM (Brown, 1998) and PAM (Best, 1995) with the Feature Hypothesis (McAllister et al., 2002) perhaps subsumed under Brown's PIM, we would create a comprehensive model to explain naïve and L2 perception of both non-native segmentals and suprasegmentals, namely the PIM-PAM model

In the case of the perception of tone, this same approach could also be applied if tones are broken down into features from a bundle of features which equals a phoneme or toneme. To define height, we might have [+high pitch] and possibly [+low pitch] to determine mid-level tones (unless these are indeed allotones). We might also need features to define direction while adding more such as [+creaky voice] for other tones of various languages. If we adopt a compositional analysis of tones based on the Moraic Alignment Hypothesis of Morén and Zsiga (2006) in their analysis of Thai, we can view contour tones as a mora-based sequence of [+high pitch] and [-high pitch]. Thus, in addition to features for segmentals of vowels and consonants, we might add features for suprasegmentals such as tone. In this respect, just as we streamline a bundle of segmental features into a phoneme, we can streamline a bundle of suprasegmental features into tonemes.

Such a model would neatly account for the Japanese listener's high rate of accuracy and reaction times over other groups. All of the Thai tones except for the Mid tone may not be accounted for if we consider that the Japanese phonology features [+high pitch], but the Mid tone in Thai may be perceived again an "allotone" for High pitch in Japanese as phonologically high-pitched moras in Japanese can be realized phonetically at lower heights. Results on the naïve perception of Thai tones by L1 speakers of Japanese substantiate such an approach on several accounts. However, the functional load of pitch accent to differentiate Japanese minimal pairs is a relatively low 14% (Shibata & Shibata, 1990) with some figures ranging to lower or higher approximations. However, just as there may be a low number of minimal pairs distinguished by [l] and [r] in English, this does not minimalize the effect of being able to hear these two phonemes. Additionally, we cannot discount the fact that pitch is used throughout the Japanese language and not merely to distinguish minimal pairs. To illustrate, in addition to marking regional dialects, pitch contributes to overall intelligibility when working in tandem with segmentation and thus, has a "functional intelligibility load. Finally, it must be remembered that pitch is the only correlate to mark accent in Japanese, making L1 speakers of Japanese highly sensitive to pitch placement.

Thus, due to sensitivity to pitch in their L1, the Japanese might map the tones onto Japanese [+high pitch] or [+low pitch]. Additionally, we might keep in mind that the Japanese may take the syllabic stimuli they hear and "convert" them into two moras bearing several possible sequences of high and low pitches. This approach would be very plausible for the perception of Thai tones as it is in line with the compositional model for Thai tones as analyzed by (Morén & Zsiga, 2006). Thus, even for the monosyllabic stimuli the Japanese might be able to map the High, Rising and Falling tones of Thai. Also, in the case of the Low and Mid tones, in a PAM model they may be considered either one category or two categories with varying degree of being either equally or better fit. For example, comparing a Mid tone with a Rising tone could possibly be two categories with each tone being easily mapped: the Rising tone corresponding to a LH bimoraic string for Japanese and the Mid tone corresponding to a HH moraic string if Mid tone is considered an allotone of the High tone. If the Mid tone, however, is viewed as a string of LL bimoraic string, then, we have the case of possibly "uncategorizable" as we might for the Low tone itself.

Also, for the Japanese in many cases it could well be merely a judgment of whether a tone is in the low or high register. Thus, an ABX trial composed of a Rising and High tone would be quite simple to differentiate correctly while a trial composed of Rising and Low tone would not, which is precisely what the accuracy rates show. Thus, we see it is simply a decision made on pitch height. This is corroborated by the fact that for the Japanese differentiating a High tone against a Low tone was not problematic while doing so when comparing a Mid and Low tone and Mid and High tone was.

In the case of English, we might expect something similar to the case of Japanese, as word stress in English in syllables of two or more could merely be the combination of high and low pitches to create a rise and fall. However, unlike Japanese, English speakers do not refer to the domain of the mora. And perhaps more decidedly is the fact that English stress is defined by several correlates of which length, intensity and vowel quality may be more salient (cf., Kochanski, Grabe, Coleman & Rosner, 2005). The "non-exclusive" nature of pitch in English word stress may be sufficient to lower the perceptual accuracy of Thai tone by L1 English speakers vis-à-vis the other groups including the L1 Korean group. Results from the current study seem to indicate that the application of the pitch correlate to the naïve perception of pitch in non-native tone is difficult.

Moreover, results show that disyllabic stimuli do not seem to aid L1 English listeners in their naïve perception of tones. Disyllabic stimuli may have aided speakers of other languages. Disyllabic stimuli may have thus revealed language-specific trends. For example, as mentioned the Low+High vs High+Low comparison seems to show that neither English word stress nor intonation play a strong role in the perception of Thai tones. For word stress, it is likely that pitch is not sufficiently salient while for intonation it is likely that the domain differs and so, does not come into play.

As for native speakers of Mandarin, there is the possibility of tone-to-tone mapping. When examining the perception of a non-native L2 tone inventory by L1 speakers of another tone language, we can look at their naïve perception of the non-native L2 tones by determining which L2 tones they map onto which L1 tones with reference to the degree of categorization of the L1 tone (cf., Stagray & Downs, 1993). Now, we can see entire tone patterns being mapped. Results from this study show that the L1 Mandarin group had difficulty when comparing the Low and Rising tones and Low and Mid tones. This is expected as both the Thai Low and Mid tones do not have "exact" counterparts in Mandarin. Also, these tones are phonetically similar to one another in either register or shape. Here we can see where the tone patterns in the L1 influence non-native tone perception. The other comparisons involving the High and Falling tones generally did not present problems for the native Mandarin speakers. Mandarin has a Rising tone and so, this was not expected to be a problem. For example, the Rising vs High tone comparison did not cause any problems. We now examine whether learners of Mandarin pattern like the Mandarin group, their English counterpart group or both or neither.

Learners of tone

The second overarching research questions concerns the learning of tones of a second language and the degree to which exposure to an L2 tone system (i.e., Mandarin) shapes the naïve perception of a non-native tone system that has not been learned (i.e., Thai). This second research question is simply as follows:

2. RQ2: Does the learning of a tone language aid in the naïve perception of the tones of another non-native tone language?

This second question is further qualified by several sub-questions which served as the motivation behind further tasks in this study

2.1 RQ2.a: Do L2 learners of a tone language differ in their naïve perception of another tone language as compared to participants with the same L1 who are not learning a tonal L2?

2.2 RQ2.b: Do L2 learners of a tone language perform at levels comparable to L1 speakers of that tone language in a linguistic hierarchy of perceptual ability of another tone language?

2.3 RQ2.c: Do L2 learners of a language perform in a similar manner to L1 speakers of that one language, i.e., Are some tones or tonal comparisons easier to perceive than others (cf., pitch height or pitch direction) in the naïve perception of another tone language? Do the two groups react in the same manner?

Finally, the third overarching research question concerns the encoding of tones of a second language.

3. RQ3: Are L2 learners of a tone language able to lexically encode tonal information to constrain lexical access?

Results show that exposure to a tone language through learning results in the acquisition of tones. More importantly, it seems that not only can tone be learned but it can be robustly encoded as is discussed in this section. Indeed, the presence of an additional language system in the form of a second language in bilingual speakers is robust enough to affect naïve perception of non-native tone. That is, we see that the Learners differ in their naïve perception of Thai tone as compared to non-learners of Mandarin from the same L1 language, namely English. This is evidenced not only in the higher accuracy levels but also in the faster reaction times of the Learner group. Moreover, it would seem that the Learners perform at levels comparable to native speakers of Mandarin in a linguistic hierarchy of perceptual ability of Thai tone where the two groups fall at similar positions on the hierarchy. Lastly, we see that the Learners appear to generally perform in a similar manner to native speakers of Mandarin. The two groups generally find the same tonal comparisons easier to perceive than others although not with a complete overlap. Again, differences between the Learners and Mandarin group appear primarily in the RTs. In short, these results are encouraging for both learners and teachers as well.

To determine exactly what is occurring, we need to make the following comparisons between the three language groups of Mandarin, English, and Learners:

1. Where does the Mandarin group differ from the English group?
2. Where does the Learner group differ from the Mandarin group?
3. Where does the Learner group differ from the English group?

4. Where does the Learner group differ from both the Mandarin and English group?

We see that there were no differences in accuracy among the three groups on the monosyllabic stimuli, but differences emerged on the RTs highlighting where the groups diverge in their perception. In contrast, on the disyllabic stimuli, we see differences in accuracy and RTs among the three groups. Overall, it is generally either the Learners or the Mandarin group or both these groups performing either more accurately or faster than the English group.

There are two possible explanations for these results: 1) Learners develop a "level of comfort" with tones and/or 2) robust encoding of Mandarin tones occurs but not to the level of native speakers. This first possibility is reinforced by the fact that the Learners are even faster on segmental comparisons while the English group is slower. The second possibility concerns categorization.

Exposure to Mandarin has enabled the Learners to build a more nuanced tone space beyond a lower and upper register reflective of English word stress, allowing this group to categorize tones more finely, i.e., where the Mid tone lies. Yet, as non-native speakers of Mandarin the Learners may not have as strong categories for tone allowing them to hear subtle differences more easily. The Mandarin group, on the other hand, has a more "entrenched" tonal space which makes mapping Thai tones which differ from their L1 Mandarin tones more difficult as there are more mapping possibilities. That is, as predicted under a PAM (Best, 1995) approach, a low tone may be close to either a rising tone or dipping tone in Mandarin, slowing the Mandarin group in making a decision. Also, the tone categories are more "set" so that it is harder to hear the differences in tones that may be within one category in Mandarin. For example, the Thai Low tone or Mid tone may seem to be the same as the Mandarin Rising tone.

Furthermore, a robust representation of Mandarin tones is evident in the naïve perception of Thai tones. For example, encoding of the High tone is highly likely the reason both the Mandarin group and Learner group were faster than the English group on the Mid vs High tone comparison on the Monosyllabic ABX task. This observation is further reinforced by the Learner group's performance on the Low vs High tone over the English group. Additionally, we see that on the High+Low vs Low+High comparison, both the Mandarin group and Learners were more accurate than the English group. This result alone seems to indicate not only that the disyllabic stimuli do not help the English group, but also that it is Mandarin that does. Thus, again we see that pitch does not seem to be a salient correlate of word stress in English while the robust representations of Mandarin tones does. Moreover, this last result seems to indicate that intonation is not accessed either as these patterns reflect the intonation pattern of statements and questions in English. However, there is the possibility that the less salient correlate of pitch used in word stress might be "activated" to its full potential through exposure to Mandarin. Or that exposure to Mandarin strengthens the Learners ability to categorize pitch in either a lower or upper register.

However, Learners may still not have as robust a representation as the native Mandarin speakers. For example, only the Mandarin group was more accurate on the Low+Mid vs High+Low comparison on the disyllabic task. The Mandarin group are perhaps able to map the Mid tone somewhere in their L1 inventory. Also, the Learners do not consider a Mid tone as an allotone of a High tone. That is, Learner do not consider the mere presence of pitch as "stressed" and absence as "unstressed" reflective of English word stress. This again may negate the effect of the pitch correlate of English word stress.

Learners of a tonal language as an L2 are able to lexically encode tonal information to constrain lexical access. Learners have encoded Mandarin tones as attested to by the results of the priming component of the Lexical Decision task with Repetition Priming. This is further reinforced by the greater accuracy of the Learners on the naïve perception of Thai tones as compared to the counterpart English group. In general, the Learners demonstrate consistent levels of accuracy on the perception ABX tasks. This observation is noted despite the expectation that speaking English as an L1 would have a powerful influence (cf., Interaction Hypothesis, Flege, Yeni-Komshian & Liu, 1999; Weinreich, 1953), lowering accuracy rates and slowing down RTs as witnessed among the English group in their perception of Thai tones. Now, the Learners may have weaker representations of tone and/or different or limited categories for Mandarin tones in comparison to the Mandarin group (e.g., categorization and perception of Thai tones by Learners and Mandarin group, cf., Stagray & Downs, 1993). This may be seen in the varying phonological proficiency scores where production of Mandarin tones appears non-native-like, even differing among the Learners. We can speculate that this is most likely not merely a production problem but a perceptual one as well. In short, Learners are able to encode tonal information but they may not naturally do so exactly in the manner as the native Mandarin speakers.

Limitations of the study

The current study exhibits the following limitations:

1) How to interpret the data within the current theoretical framework still has room for more thought and analyses. The concept of feature remains unclear as does the concept of similarity between sounds.
2) The ABX task is a broad assessment of perception and a good first step, but a categorization task such as listening to tones and then, grouping them according to ones the listener feels are similar may provide a more in-depth idea of what listeners are doing when they hear tones rather than only a "two-choice" answer ABX task. Also, on this note, adding a neither answer could be enlightening.

3) Stimuli are limited to Thai, but adding tones that are similar and occur within the same register may provide a more fine-grained interpretation of tone perception
4) A better analysis of the L1s may be necessary to assess the role and prominence of lexically-contrastive pitch. This includes possible dialects and a more refined examination of intonational categories and the conceptualization of categories for pitch accent in Japanese and stress accent in English.
5) The logistics of recruiting participants is somewhat complicated. This includes not only limitations in number and therefore, uniform background, but also L2 English influence. Specifically, the numbers of Japanese and Thai speakers are low at Indiana University. Also, many L1 Mandarin and L1 Korean speakers have varied exposure to English, including many with several years and early exposure in country. Additionally, L1 Mandarin speakers display variation in their dialect of Mandarin.

Future steps

Future directions may include changes to the tasks or the implementation of entirely different tasks in order to answer any lingering questions from the current study.

A categorization task may show even more clearly which tones are perceived to be similar. Additionally, a speak-out-loud protocol might be added to the categorization task to better understand what participants are thinking and therefore, what they hear. Also, different tasks or stimuli applying more cognitive load on the listener may tease out results to see if this would result in greater differences among the groups reflective of the varying typological usage of lexically-contrastive pitch.

To better understand how tone-to-tone mapping under a more PAM-centered approach, adding participants who speak Vietnamese or Cantonese both of which have different tones from both Thai and Mandarin may illuminate perceptual similarity between tones by native speakers. For example, in the Vietnamese tonal inventory there are mid falling and mid rising tones which are further distinguished by phonation while Cantonese also features a mid-rising tone. Thus, adding such tone patterns may help to understand the concept of tonal space, particularly the space between the low and high registers. Under such an approach, we must also keep in mind constraints concerning the inventory of tones in terms of numbers and types of tones, the functional load of tone in a language, and whether pitch is used exclusively as length or phonation may also come into play to cue listeners to the tone. We may add synthetic tones that would better determine what listeners are perceptive to. For example, a high tone that is flat or a wider range of level tones to see where the boundaries of the categorization of a flat tone form high to mid to low occurs (cf., VOT categorization; for tones: Gandour, 1983). We might also create non-words using English intonation patterns (e.g., Well. Well! Well?) and test whether English speakers can access intonation in order to perceive the differences between these stimuli. Moreover, features for tones must be more carefully thought through when choosing tones for testing referencing the case of the perception of English [l] and [r] by L1 speakers of Japanese and Mandarin which differ in that Mandarin also possesses the lexically-contrastive feature of [+coronal].

To further understand how lexically-contrastive pitch of different sound systems influences the naïve perception of non-native tones, looking at two other types of speakers may further shed light on the naïve perception of non-native tones. Looking at Swedish listeners' perception in contrast to the Japanese listeners' perception may verify the results that L1 speakers of pitch accent perform high on the hierarchy of tone perception although Japanese and Swedish vary in their use of lexically-contrastive pitch. Adding Spanish may also bolster the claim that speakers of word-stress languages would outperform speakers of languages which do not feature lexically-contrastive pitch. This is because unlike English Spanish marks stress with only pitch which would seemingly make pitch the most salient feature of Spanish word stress and therefore, be accessed perhaps more readily in the naïve perception of non-native tones.

Also, testing dialect speakers may be informative. For example, testing dialect speakers of Japanese may demonstrate how exactly lexically-contrastive pitch in Japanese influences the naïve perception of non-native tone. For example, the Kansai dialect features words that have only two low pitches unlike standard Japanese while some dialects in Japan do not use pitch to differentiate words. In addition, testing Korean speakers of standard Korean and the pitch-accent Kyungsang dialect in Korean may eliminate any possible influence of L2 English. Moreover, a closer look at a non-pitch-accent standard and pitch-accent dialect as in the case of Korean as compared to a pitch-accent standard and non-pitch-accent dialect as in the case of Japan would also immensely tell us more about the formation and subsequent effect and interaction of a phonological system of a bidialectal person and serve as a suprasegmental counterpart to the studies on the perception of segmentals by bilingual Spanish/Catalan speakers.

Additionally, examining learners of Mandarin with an L1 different from English could be enlightening as well. Looking at L2 Mandarin learners with L1 Japanese or L1 Korean would determine if learners of languages with varying degrees of functionality of lexically-contrastive pitch achieve the same levels of accuracy as the L2 Mandarin/L1 English group of this study and react in the same manner to the individual tonal comparisons. We would determine if like the McAllister et al study (2002) an L2 Mandarin/L1 Japanese group would outperform the L2 Mandarin/L1 English group and if an L2 Mandarin/L1 Korean group would underperform the L2 Mandarin/L1 English group. This would then be a closer replication of the original McAllister et al. study but on a different suprasegmental. Possibly, adding an L2 Mandarin/L1 Swedish group may also bolster the claims that speaking a pitch-accent as an L1 aids in the learning of a tonal language. Differences in performance between such groups would also clarify how exactly learners differ in their encoding of Mandarin tone as compared to native speakers.

The current study provides a baseline for the naïve perception of Thai tones and learner perception of Mandarin tones. The study shows that lexically-contrastive pitch in the L1 shapes the naïve perception of a non-native tone system to varying degrees determined by the functionality of lexically-contrastive pitch in the L1. Also, the study shows that learners of a tone language can robustly encode the tones of a language learned as an L2 albeit not exactly in the manner of native speakers. Furthermore, the L2 tones learned do undeniably shape the naïve perception of the non-native tones of another tone language that has not been learned. Combined these conclusions demonstrate that learners do not necessarily start their learning of a tone language empty-handed and can exploit lexically-contrastive pitch in the L1 to learn non-native tones. This is indeed heartening news for learners and teachers of tone languages. However, more work remains to be done in order to fully understand how both learners and naïve listeners of a tone or non-tone language as an L1 perceive non-native tones.

References

Abramson, A. S. (1975). The tones of Central Thai: Some perceptual experiments. In J. G. Harris & J. R. Chamberlain (Eds.), *Studies in Thai linguistics in honor of William J. Gedney* (pp. 1-16). Bangkok: Central Institute of English Language.

Abramson, A. S. (1976). Thai tones as a reference system. In T. W. Gething, Jimmy G. Harris, & Pranee. Kullavanijaya (Eds.), *Tai linguistics in honor of Fang-Kuei Li* (pp. 1-12). Bangkok: Chulalongkorn University Press.

Abramson, A. S. (1978). Static and dynamic acoustic cues in distinctive tones. *Language and Speech*, *21*, 319-325.

Abramson, A. S. (1979). The coarticulation of tones: an acoustic study of Thai. In T. Thongkum, V. Panupong, P. Kullavanijaya & M. R. Kalaya Tingsabadh (Eds.), *Studies in Tai and Mon-Khmer Phonetics and Phonology in Honor of Eugénie J. A. Henderson* (pp. 1-9). Bangkok: Chulalongkorn University Press.

Beckman, M. E. (1986). *Stress and non-stress accent.* Dordrecht, Holland: Foris.

Beckman, M., Hirschberg, J. & Shattuck-Hufnagel, S. (2005). The original ToBI system and the evolution of the ToBI framework. In S-A. Jun (Ed.) *Prosodic Typology: The Phonology of Intonation and Phrasing* (pp. 9-54). Oxford University Press.

Bennett, J.F. *(*1994*).* Iambicity in Thai*, Studies in the linguistic sciences, 24*(*1* /*2*), 39-57.

Best, C. (1995). A direct realist view of cross-language speech perception. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 171-204). Baltimore, MD: York Press.

Best, C. & Tyler, M . (2007). Nonnative and second-language speech perception: Commonalities and complementarities. In O. S. Bohn & M. Munro (Eds.), *Second-language speech learning: The role of language experience in speech perception and production. A festschrift in honor of James E. Flege* (pp. 13-34). Amsterdam: John Benjamins.

Bialystok, E. & Hakuta, K. (1999). Confounded age: Linguistic and cognitive factors in age differences for second language acquisition. In D. Birdsong (Ed.), *Second language acquisition and the critical period hypothesis* (pp 161-181). Mahwah, NJ: Lawrence Erlbaum Associates.

Binh, N. N. (2003). *Elementary Vietnamese: Revised edition*. Tokyo: Tuttle Publishing

Boersma, P. & Weenink, D. (2015). Praat: doing phonetics by computer [Computer program]. Version 5.4.08, retrieved 20 April 2015 from http://www.praat.org/

Boersma, Paul (2001). Praat, a system for doing phonetics by computer. *Glot International 5*(9/10), 341-345.

Bohn, O.-S. (1995). Cross-language speech perception in adults: First language transfer doesn't tell it all. In W. Strange (Ed.), *Speech perception and linguistic experience. Issues in cross-language research* (pp. 273-304). Baltimore: York Press.

Braun, B. & Johnson, E. K. (2011). Question or tone 2? How language experience and linguistic function guide pitch processing. *Journal of Phonetics*, *39*, 585-594.

Brown, C. (1998). The role of the L1 grammar in the L2 acquisition of segmental structure. *Second Language Research*, *14*(2), 136-193.

Burnham, D., & Mattock, K. (2007). The perception of tones and phones. In O.-S Bohn & M. J. Munro (Eds), *Language Experience in Second Language Speech Learning: In honor of James Emil Flege* (pp. 259-280). Amsterdam: John Benjamins.

Burnham, D., & Francis, E. (1997). The role of linguistic experience in the perception of Thai tones. In A. S. Abramson (Ed.), *Southeast Asian linguistic studies in honour of Vichin Panupong* (pp. 29-47). Bangkok: Chulalongkorn University Press.

Burnham, D., Francis, E., Webster, D., Luksaneeyanawin, S., Attapaiboon, C., Lacerda, F., & Keller, P. (1996). Perception of lexical tone across languages: Evidence for a linguistic mode of processing. In H. T. Bunnell & W. Idsardi (Eds.), *Proceedings of the Fourth International Conference on Spoken Language Processing 1* (pp. 2514-2517). Wilmington, DE: Applied Science and Engineering Laboratories.

Burnham, D., Francis, E., Webster, D., Luksaneeyanawin, S., Lacerda, F., & Attapaiboon, C. (1996). Facilitation or attenuation in the development of speech mode processing? Tone perception over linguistic contexts. In P. McCormack & A. Russell (Eds.), *Proceedings of the Sixth Australian International Conference on Speech Science and Technology* (pp. 587-592). Canberra: Australian Speech Science and Technology Association.

Burnham, D., Kirkwood, K., Luksaneeyanawin, S., & Pansottee, S. (1992). Perception of Central Thai tones and segments by Thai and Australian adults. In *Pan-Asiatic Linguistics: Proceedings of the Third International Symposium on Language and Linguistics* (pp. 546-560). Bangkok: Chulalongkorn University Press.

Chang, Y.-C., Hallé, P., Best, C. T. & Abramson, A. (2008). Do non-native language listeners perceive Mandarin tone continua categorically? In *Proceedings of the 8th Phonetic Conference of China and the International Symposium on Phonetic Frontiers.* Beijing, China.

Chao, Y. R. (1948). *Mandarin primer*. Cambridge, MA: Harvard University Press.

Chiao, W.-H., Kabak, B. & Braun, B. (2011). When more is less: Non-native perception of level tone contrasts. In *Psycholinguistic Representation of Tone Conference* (pp. 42-45). Hong Kong, China.

Cooper, N., Cutler, A., & Wales, R. (2002). Constraints of lexical stress on lexical access in English: Evidence from native and non-native listeners. *Language and Speech, 45(3),* 207-228.

Cutler, A. & Otake, T. (1999). Pitch accent in spoken word recognition in Japanese. *The Journal of the Acoustical Society of America, 105*(3), 1877-1888.

Cutler, A., Dahan, D. & Van Donselaar, W. A. (1997). Prosody in the comprehension of spoken language: A literature review. *Language and Speech, 40*(2), 141-202.

Cui, R. & van Heuven, V. J. (2011). Mutual intelligibility of English vowels by Chinese dialect speakers. In W.-S. Lee, & E. Zee (Eds.), *Proceedings of the 17th International Congress of Phonetic Sciences 2011 [ICPhS XVII]* (pp. 544-547). Hong Kong: Department of Chinese, Translation and Linguistics, City University of Hong Kong.

Darcy, I., Dekydtspotter, L., Sprouse, R. A., Glover, J., Kaden, C., McGuire, M. and Scott, J. H. G. (2012). Direct Mapping of Acoustics to Phonology: On the lexical encoding of front round vowels in L1 English-L2 French acquisition. *Second Language Research 28,* 5-40.

Darcy, I., Daidone, D., & Kojima, C. (2013). Asymmetric lexical access and fuzzy lexical representations in second language learners. *The Mental Lexicon, 8*(3), 372-420.

Deutsch, D., Henthorn, T., Marvin, E., & Xu, H-S. (2006). Absolute pitch among American and Chinese conservatory students: Prevalence differences, and evidence for a speech-related critical period. *Journal of the Acoustical Society of America, 119*(2), 719-722.

Duanmu, S. (2004). Tone and non-tone languages: An alternative to language typology and parameters. *Language and Linguistics 5*(4), 891-924.

Dupoux, E., Sebastián-Gallés, N., Navarrete, E. & Peperkcamp, S. (2008). Persistent stress 'deafness': The case of French learners of Spanish. *Cognition, 106*(2), 682-706.
Dupoux, E., Peperkamp, S., & Sebastián, N. (2001). A robust method to study stress 'deafness'. *Journal of the Acoustical Society of America, 110*, 1606-1618.
Dupoux, E., Kakehi, K., Hirose, Y., Pallier, C., & Mehler, J. (1999). Epenthetic vowels in Japanese: A perceptual illusion? *Journal of Experimental Psychology: Human Perception and Performance, 25*(6), 1568-1578.
Dupoux, E., Pallier, C., Sebastián, N., & Mehler, J. (1997). A destressing "deafness" in French? *Journal of Memory and Language, 36*, 406-421.
Eliasson, S. (1997). Tone in second language acquisition. In R. Hickey & S. Puppel (Eds.), *Language history and linguistic modeling: Vol. 2* (pp. 1273–1289). *Berlin: Mouton de Gruyter.*
Feldhausen, I., Pešková, A., Kireva, E. & Gabriel, C. (2011). Categorical perception of Porteño nuclear accents. In W.-S. Lee, & E. Zee (Eds.), *Proceedings of the 17th International Congress of Phonetic Sciences 2011 [ICPhS XVII]* (pp. 116-119). Hong Kong: Department of Chinese, Translation and Linguistics, City University of Hong Kong.
Flege, J. E. (1993). Production and perception of a novel, second-language phonetic contrast. *Journal of Acoustic Society of America 93*(3), 1589-1608.
Flege, J. E. (1995). Second language speech learning: Theory, findings, and problems. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 233-277). Baltimore, MD: York Press.
Flege, J. E. (1999). Age of learning and second language speech. In D. Birdsong (Ed.), Second *Language Acquisition and the Critical Period Hypothesis* (pp. 101-131). Mahwah, NJ: Lawrence Erlbaum Publishers.
Flege, J. E., Schirru, C. & MacKay, I. R. A. (2003). Interaction between the native and second language phonetic subsystems. *Speech Communication 40*, 467-491.
Flege, J. E., Yeni-Komshian, G. H., & Liu, S. (1999). Age constraints on second-language acquisition. *Journal of Memory and Language, 41*, 78-104.
Flege, J. E., Frieda, E. M, & Nozawa, T. (1997). Amount of native-language (L1) use affects the pronunciation of an L2. *Journal of Phonetics 25*, 169-186.
Fok Chan, Y. Y. (1974). *A perceptual study of tones in Cantonese*. Hong Kong: Centre of Asian Studies, University of Hong Kong.
Fon, J., & Chiang, W.-Y. (1999). What does Chao have to Say about Tones? A Case Study of Taiwan Mandarin. *Journal of Chinese Linguistics*, *27*(1), 13-37.
Forster, K. I. & Forster, J. C. (2003). DMDX: A Windows Display Program with Millisecond Accuracy. *Behavior Research Methods, Instruments, & Computers 35*(1), 116-124.
Francis, A. L., Ciocca, V., Ma, L., & Fenn, K. M. (2008). Perceptual learning of Cantonese lexical tones by tonal and non-tonal language speakers. *Journal of Phonetics*, *36*(2), 268-294.
Fry, D. B. (1958). Experiments in the perception of stress. *Language and Speech, 1*, 126-52.
Fujisaki, H., & Sugito, M. (1977). Onsei no butsuriteki seishitsu. [The physical quality of sound]. In S. Ono and T. Shibata (Eds.), *Iwanami Koza: Nihongo 5, On'in* (pp. 63-106). Tokyo: Iwanami Shoten.
Braun, B., Galts, T. & Kabak, B. (2014). Lexical encoding of L2 tones: the role of L1 stress, pitch accent and intonation. *Second Language Research 30*(3), 323-350.

Gandour, J. (1983). Tone perception in Far Eastern languages. *Journal of Phonetics, 11*, 49-175.
Gandour, J. (1984). Tone dissimilarity judgments by Chinese listeners. *Journal of Chinese Linguistics, 12* (2), 235-261.
Gandour, J., Ponglorpisit, S. & Dardarananda, R. (1992). Tonal disturbances in Thai after brain damage, *Journal of Neurolinguistics 7*(1-2), 133-145.
Gandour, J., Petty S. H., Dardarananda, R. (1988). Perception and production of tone in aphasia**.** *Brain and language 35*(2), 201-40.
Gandour, J. & Harshman, R. (1978). Crosslanguage differences in tone perception: a multidimensional scaling investigation. *Language and Speech, 21*, 1-33.
Gerrits, E., & Schouten, M. E. H. (2004). Categorical perception depends on the discrimination task. *Perception and Psychophysics 66*(3), 363-376.
Goto, H. (1971). Auditory perception by normal Japanese adults of the sounds "l" and "r". *Neuropsychologia 9*, 317–323.
Guion, S. G. & Pederson, E. (2007). Investigating the role of attention in phonetic learning. In O.-S. Bohn & M. Munro (Eds.), *Language experience in second language speech learning* (pp. 57-77). Amsterdam: John Benjamins.
Gussenhoven, C. (2004) *The Phonology of Tone and Intonation*. Sound files. Cambridge: Cambridge University.
Hallé, P. A., Chang, Y-C. & Best, C.T. (2004). Identification and discrimination of Mandarin Chinese tones by Mandarin Chinese vs French listeners. *Journal of Phonetics, 32*, 395-421.
Hao, Y.-C. (2012). Second language acquisition of Mandarin Chinese tones by tonal and non-tonal language speakers. *Journal of Phonetics, 40,* 269-279.
Hao, Y.-C. (2008). Intonational interference on the perception and production of Mandarin Chinese tones by English-speaking second language learners. *Journal of the Acoustical Society of America, 124*(4), 2595.
Harada, T. (2006). The acquisition of single and geminate stops by English-speaking children in a Japanese immersion program. *Studies in Second Language Acquisition 28*(4), 601-632.
Haraguchi, S. (1999). Accent. In N. Tsujimura (Ed.), *The Handbook of Japanese Linguistics (pp.* 1-30)*. Malden, MA: Blackwell Publishers*.
Hayes-Harb, R. & Masuda, K. (2008). Development of the ability to lexically encode novel second language phonemic contrasts. *Second Language Research*, *24*, 5-33.
Hed, A. (2014). Perception and production of Swedish word accents by Somali L1 speakers. *In M. Heldman (Ed.), Proceedings from FONETIK 2014* (pp. 105-110). Stockholm: Stockholm University.
Hirst, D. and Di Cristo, A. (1998). A survey of intonation systems. In D. Hirst & A. Di Cristo (Eds.), *Intonation systems* (pp. 1-45). Cambridge: Cambridge University Press.
Holbrook, A. & Lu H-T. (1969). A study of intelligibility in whispered Chinese. *Speech Monographs*, 36, 464-466.
Hung, T-H. & Lee, C-Y. (2008). Processing linguistic and musical pitch by English-speaking musicians and non-musicians. In M. K. M. Chan & H. Kang (Eds.), *Proceedings of the 20th North American Conference of Chinese Linguistics (NACCL-20) 2008. Volume 1* (pp. 139-145). Columbus Ohio: The Ohio State University.
Hyman, L. M. (2009). How (not) to do phonological typology: the case of pitch-accent. *Language Sciences, 31*(2-3), 213-238.
Jiang, N. (2012). *Conducting reaction time research in second language studies*. New York:

Routledge.
Jun, J., Kim, J., Lee, H., & Jun, S-A. (2006). The prosodic structure and pitch accent of Northern Kyungsang Korean. *Journal of East Asian Linguistics, 15*(4), 289-317.
Kadyamusuma, M. R., De Bleser, R., Mayer, J., & Vasishth, S. (2009). Production and Perception of Shona lexical tones in Left and Right hemisphere damaged patients. *Proceedings of the Second Clinical linguistics conference*, Madrid, Spain.
Khouw E. & Ciocca V. (2007). Perceptual correlates to Cantonese tones. *Journal of Phonetics, 35(1),* 104-117.
Kim, J.-S. (2011). Perception of Lexical Pitch Accent by Kyungsang and Cholla Korean Listeners. In W.-S. Lee, & E. Zee (Eds.), *Proceedings of the 17th International Congress of Phonetic Sciences 2011 [ICPhS XVII]* (pp. 1070-1073). Hong Kong: Department of Chinese, Translation and Linguistics, City University of Hong Kong.
Kim, J.-S., & de Jong, K.J. (2007). Perception and Production in the Pitch Accent System of Korean. In J. Trouvain and W. J. Barry (Eds.), Proceedings of the 16th International Congress of Phonetic Sciences 2007 [ICPhS XVI] (pp. 1273-1277). Dudweiler: Pirrot.
Kim-Renaud, Y-K. (2009). *Korean: An essential grammar*. New York: Routledge.
Kindaichi, H. (1947). Tookyoogo ni okeru “hana” to “hana” no kubetsu – tookyoo akusento shinnidankan kyoochooron [The difference between “hana”(flower) and “hana” (nose) in the Tokyo dialect – New theory emphasizing the two-levels of the Tokyo accent]. *Kokugo* [*The Japanese language*], Summer.
Kirby, J. P. (2011b). Vietnamese (Hanoi Vietnamese). *Journal of the International Phonetic Association 41*(3), 381-392.
Kirby, J. P. (2010). Dialect experience in Vietnamese tone perception. *Journal of the Acoustical Society of America. 127*(6), 3749-3757.
Kochanski, G., Grabe, E., Coleman, J., & Rosner, B. (2005). Loudness predicts prominence: fundamental frequency lends little. *The Journal of the Acoustical Society of America, 118*(2), 1038-1054.
Kondo, M. (2007). Acoustic realization of lexical accent and its effects on phrase intonation in English speakers’ Japanese. In J. Trouvain and W. J. Barry (Eds.), Proceedings of the 16th International Congress of Phonetic Sciences 2007 [ICPhS XVI] (pp. 1649-1652). Dudweiler: Pirrot. Saarbrücken, Germany: Saarland University.
Kondo, M. (2009). Is acquisition of L2 phonemes difficult? Production of English stress by Japanese. In M. Bowles, T. Ionin, S. Montrul & A. Tremblay (Eds.), *Proceedings of the 10th Generative Approaches to Second Language Acquisition Conference (GASLA 2009)* (pp. 105-112). Somerville, MA: Cascadilla Proceedings Project.
Kratochvil, P. (1987). The case of the third tone. In Ma (Ed.), *Wang Li memorial volumes* (pp. 253–276). Hong Kong: The Chinese Language Society of Hong Kong.
Kubozono, H. (1999). Mora and syllable. In N. Tsujimura (Ed.), *The handbook of Japanese linguistics* (pp. 31-61). Malden, MA: Blackwell Publishers
Kuhl, P., Conboy, B., Coffey-Corina, S., Padden, D., Rivera-Gaxiola, M., & Nelson, T. (2008). Phonetic learning as a pathway to language: New data and native language magnet theory expanded (NLM-e). *Philosophical Transactions B*, 363, 979-1000.
Kuhl, P. K. & Iverson, P. (1995). Linguistic experience and the ‘perceptual magnet effect.’ In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 121-154). Baltimore, MD: York Press.
Ladefoged, P. & Johnson, K. (2010). *A course in phonetics (6th ed.).* Boston: Thomson

Wadsworth.
Lado, R. (1957). *Across Cultures*. Ann Arbor, MI: University of Michigan Press.
Lai Yuk Wah, E. (2002). *Prosody and prosodic transfer in foreign language acquisition: Cantonese and Japanese*. München: LINCOM Europa.
Lai, Y. & Zhang, J. (2008). Mandarin Lexical Tone Recognition: The Gating Paradigm. *Kansas Working Papers in Linguistics, Vol. 30*, 183-194.
Liang, J., and van Heuven, V. J. (2007). Chinese tone and intonation perceived by L1 and L2 Listeners. In C. Gussenhoven and T. Riad (Eds.), *Tones and Tunes, Volume 2: Experimental studies in word and sentence prosody* (pp. 27-61). Berlin: Mouton de Gruyter.
Llama, R., Cardoso, W., & Collins, L. (2010). The Influence of Language Distance and Language Status on the Acquisition of L3 Phonology. *International Journal of Multilingualism 7(1)*, 39-57.
Leather, J. (1983). Speaker normalization in perception of lexical tone. *Journal of Phonetics*, *11*, 373-382.
Leather, J. (1990). Perceptual and productive learning of Chinese lexical tone by Dutch and English speakers. In J. Leather & A. James (Eds.), *New Sounds 90: Proceedings of the Amsterdam Symposium on the Acquisition of Second-Language Speech* (pp. 72-89). Amsterdam: University of Amsterdam.
Lee, C-Y. (2009). Identifying isolated, multispeaker Mandarin tones from brief acoustic input: A perceptual and acoustic study. *Journal of the Acoustical Society of America*, *125*(2), 1125-1137.
Lee, C.-Y., Tao, L., & Bond, Z. S. (2008). Identification of acoustically modified mandarin tones by native listeners. *Journal of Phonetics*, *36*(4), 537-563.
Lee, C.-Y. (2007). Does horse activate mother? Processing lexical tone in form priming. *Language and Speech 50* (1), 101-123.
Lee, I. & Ramsey, S. R. (2000). The Korean Language. Albany, New York: State University of New York Press.
Lee, L. & Nusbaum, H. C. (1993). Processing interactions between segmental and suprasegmental information in native speakers of English and Mandarin Chinese. *Perception & Psychophysics*, *53*, 157-165.
Lenneberg, E. 1967. *Biological Foundations of Language.* New York: Wiley & Sons.
Li, B. & Shuai, L. (2011). Effects of native language on perception of level and falling tones. In W.-S. Lee, & E. Zee (Eds.), *Proceedings of the 17th International Congress of Phonetic Sciences 2011 [ICPhS XVII]* (pp. 1202-1205). Hong Kong: Department of Chinese, Translation and Linguistics, City University of Hong Kong.
Li, C. N.& Thompson, S. A. (1977). The Acquisition of Tone in Mandarin-Speaking Children, *Journal of Child Language 4*(2), 185-199.
Lin, Y.-H. (2007). *The sounds of Chinese*. Cambridge: Cambridge University Press.
Liu, S. & Samuel, A.G. (2004). Perception of Mandarin lexical tones when F0 information is neutralized. *Language & Speech, 47*, 109-138.
Llisterri, J., & Poch-Olivé, D. (1987). Phonetic interference in bilingual's learning of a third language. In *Proceedings of the 11th International Congress of Phonetic Sciences* 1987 [ICPhs XI] (pp. 134-147). Austria, Tallinn, Estonia: Academy of Sciences of the Estonian SSR.
Martin, S. E. (1992). *A reference grammar of Korean.* Tokyo: Tuttle Publishing.

McAllister, R., Flege, J. E., & Piske, T. (2002). The influence of L1 on the acquisition of Swedish quantity by native speakers of Spanish, English and Estonian. *Journal of Phonetics, 30*, 229-258.

McCawley, J. D. (1978). What is a tone language? V. Fromkin (Ed.), *Tone: a linguistic survey (pp.* 113-131). New York: Academic Press,

McGinnis, S. (1997). Tonal distinction errors by beginning Chinese language students: A comparative study of American English and Japanese native speakers. In S. McGinnis (Ed.), *Chinese pedagogy: An emerging field* (pp. 81-91). Columbus, OH: Ohio State University Foreign Language Publications.

Mixdorff, H., Luksaneeyanawin, S., Fujisaki, H., & Charnavit, P. (2002). Perception of tone and vowel quality in Thai. *Paper presented at the 7th International Conference on Spoken Language Processing,* Denver Colorado, September, 2002.

Miyawaki, K., Strange, W., Verbrugge, R.R., Liberman, A.M., Jenkins, J.J., & Fujimura, O. (1975). An effect of linguistic experience: the discrimination of [r] and [l] by native speakers of Japanese and English. *Perception and Psychophysics 18* (5), 331-340.

Moen, I. (1993). Functional lateralization of the perception of Norwegian word tones – Evidence from a dichotic listening experiment. *Brain and Language, 44*, 400–413.

Morén, B. & Zsiga, E.C. (2006). The lexical and post-lexical phonology of Thai tones. *Natural Language and Linguistic Theory 24* (1), 113-178.

Nguyen, H. T. & Macken, M. A. (2008). Factors affecting the production of Vietnamese tones: A study of American learners. *Studies in Second Language Acquisition, 30*, 49-77.

Ota, M., Hartsuiker, R.J., & Haywood, S.L. (2009). The KEY to the ROCK: Near-homophony in nonnative visual word recognition. *Cognition,* 111, 263-69.

Otaka, H. (2008). *Phonetics and phonology of moras, feet, and geminate consonants in Japanese*. Lanham, MD: University Press of America, Inc.

Otake, T. & Cutler, A. (1999). Perception of suprasegmental structure in a nonnative dialect. *Journal of Phonetics, 27*(3), 229-253.

Otake, T. and Higuchi, M. (2004). A role of pitch accent in spoken-word recognition in accentless Japanese dialect: Evidence from Fukushima listeners. In *Proceedings of Speech Prosody 2004* (pp. 41-44). Nara, Japan.

Oyama, S. (1976). A sensitive period for the acquisition of a nonnative phonological system. *Journal of Psycholinguistic Research, 5*, 261-285.

Oyama, S. (1979). The concept of the sensitive period in developmental studies. *Merrill-Palmer Quarterly, 25(2),* 83-102.

Pallier, C., Dehaene, S., Poline, J.-B., LeBihan, D., Argenti, A.-M., Dupoux, E., & Mehler, J. (2003). Brain imaging of language plasticity in adopted adults: Can a second language replace the first? *Cerebral Cortex 13*(2), 155-161.

Pallier, C., Colomé, A., & Sebastian-Gallés, N. (2001). The influence of native-language phonology on lexical access: Concrete exemplar-based vs. abstract lexical entries. *Psychological Science, 12*(6), 445-449.

Pallier, C., Bosch, L., & Sebastian-Gallés, N. (1997). A limit on behavioral plasticity in speech perception. *Cognition, 64,* B9-B17.

Pan, H-H. (2004). Focus and Taiwanese unchecked tones. In C. Lee, M. Gordon & D. Büring (Eds.), *Topic and focus: A cross-linguistic perspective* (pp. 197-216). Dordrecht, Netherlands: Kluwer Academic Publishers.

Patel, A. D., Xu, Y. & Wang. B. (2010). The role of F0 variation in the intelligibility of

Mandarin sentences. In M. Hasegawa-Johnson (Ed.), *Proceedings of Speech Prosody 2010* (100890, pp. 1-4). Chicago, IL.

Peng, S-H. (1997). Production and perception of Taiwanese tones in different tonal and prosodic contexts. *Journal of Phonetics 25*, 371-400.

Pierce, L. J., Klein, D., Chen, J.-K., Delcenserie, A., & Genesee, F. (2014). Mapping the unconscious maintenance of a lost first language. *Proceedings of the National Academy of Sciences of the United States of America (PNAS) 111*(48), 17314-17319.

Pierrehumbert, J., & Beckman, M. (1988). *Japanese tone structure.* Cambridge, MA: MIT Press.

Polivanov, E. D. (1931). La perception des sons d'une langue étrangère. *Travaux du Cercle Linguistique de Prague 4,* 79–96.

Poss, N., Hung, T-H, & Will, U. (2008). The effects of tonal information on lexical activation in Mandarin. In M. K. M. Chan & H. Kang (Eds.), *Proceedings of the 20th North American Conference on Chinese Linguistics* (NACCL-20) (Volume 1, pp. 205-211). Columbus, OH: The Ohio State University.

Purcell, E & Suter, R (1980). Predictors of pronunciation accuracy: a re-examination. *Language Learning 30* (2), 271-87

Repp, B. H. & Lin, H.-B. (1990). Integration of segmental and tonal information in speech perception: A cross-linguistic study. *Journal of Phonetics, 18*, 481-495.

Sakuma, K. (1929). *Nihon onseigaku* [Japanese phonetics]. Tokyo: Kazama Shobo.

Schaefer, V. & Darcy, I. (2014). Lexical function of pitch in the first language shapes cross-linguistic perception of Thai tones. *Laboratory Phonology 5*(4), 489-522.

Schaefer, V. & Darcy, I. (2014). Pitch Prominence matters: Perception of Thai tones by Seoul Korean and Kyungsang Korean speakers. In Cardoso, W., & Trofimovich, P. (Eds.), *Concordia Working Papers in Applied Linguistics, 5*, 597-611.

Schaefer, V. & Darcy, I. (2014). Linguistic Prominence of Pitch within the Native Language Determines Accuracy of Tone Processing. In R. T. Miller, K. I. Martin, C. M. Eddington, A. Henery, N. Marcos Miguel, A. M. Tseng, A. Tuninetti, & D. Walter (Eds.), *Selected Proceedings of the 2012 Second Language Research Forum (*pp. 1-14). Somerville, MA: Cascadilla Proceedings Project.

Schwartz, B. & Sprouse, R. (1996). L2 cognitive states and the full transfer/full access model. *Second Language Research*, *12*(1), 40-72.

Sebastián-Gallés, N. (2004). Cross-language speech perception. In D. B. Pisoni and R. E. Remez (Eds.), *The handbook of speech perception* (pp. 546-566). Oxford: Black Publishing.

Sereno, J. A. & Wang, Y. (2007). Behavioral and cortical effects of learning a second language: The acquisition of tone. In O.-S.Bohn & M. J. Munro (Eds.), *Language experience in second language speech learning: In honor of James Emil Flege* (pp. 241-258). Amsterdam: John Benjamins.

Sheldon, A., & Strange, W. (1982). The acquisition of /r/ and /l/ by Japanese learners of English: Evidence that speech production can precede speech perception. *Applied Psycholinguistics*, *3,* 243-261.

Shen, X., Lin, M., and Yan, J. (1993). F0 turning point as an F0 cue to tonal contrast: A case study of Mandarin tones 2 and 3. *Journal of the Acoustical Society of America 93*, 2241-2243.

Silva, D. J. (2006). Acoustic evidence for the emergence of tonal contrast in contemporary Korean. *Phonology, 23*, 287-308.

Sluijter, A. M. C. & van Heuven V. J. (1996). Acoustic correlates of linguistic stress and accent in Dutch and American English. In H. T. Bunnell & W. Idsardi (Eds.), *Proceedings of the Fourth International Conference on Spoken Language Processing [ICSLP 96]* (pp. 630-633). Wilmington, Delaware: University of Delaware, I. DuPont Institute.

So, C. K. & Best, C. K. (2010). Cross-language perception of non-native tonal contrasts: effects of native phonological and phonetic influences. *Language and Speech, 53*, 273-293.

So, C. K. (2010). Categorizing Mandarin tones in Japanese pitch-accent categories: The role of phonetic properties. *Interspeech 2010 Second language studies: Acquisition, learning, education and techonology* (pp. O1-3). Waseda University, Tokyo, Japan.

So, C. K. (2006). Perception of non-native tonal contrasts: Effects of native phonological and phonetic influences. In P. Warren, & C. I. Watson (Eds.), *Proceedings of the 11th Australian International Conference on Speech Science & Technology* (pp. 438-443). Auckland, New Zealand: University of Auckland.

Song, J. J. (2005). *The Korean language: structure, use and context*. London: Routledge.

Soto-Faraco, S., Sebastián-Gallés, N.' & Cutler, A. (2001). Segmental and suprasegmental mismatch in lexical access. *Journal of Memory and Language, 45*(3), 412-432.

Stagray, J. R., and Downs, D. (1993). Differential sensitivity for frequency among speakers of a tone and a nontone language. *Journal of Chinese Linguistics, 21*(1), 143-163.

Strange, W. (2011). Automatic selective perception (ASP) of first and second language speech: A working model. *Journal of phonetics 39*(4), 456-466.

Sugito, M. (1972). Ososagari-koo: Dootai-sokutei ni yoru nihongo akusento no kenkyu [Delayed pitch fall: An acoustic study of Japanese accent]. *Shoin Joshi Daigaku Ronshuu 10*. Reprinted in M. Tokugawa (Ed.), *Akusento [Accent]* (pp. 201-229). Tokyo: Yuseido, 1980.

Sugiyama, Y. (2012). *The production and perception of Japanese pitch accent*. Cambridge: Cambridge Scholars Publishing.

Sugiyama, Y. (2008). *The nature of Japanese pitch accent: An experimental study*. Unpublished doctoral dissertation, State University of New York at Buffalo.

Sukekawa, Y., Choi, H., Maekawa, K., & Sato, S. (1995). Perception of pitch accent by Korean learners of Japanese and its implications. *Denshi Joho Tsushin Gakkai gijutsu kenkyu hokoku: shingaku giho, 95*(41)/19950518, 61-66.

Sukegawa, Y., Choi, H., Maekawa, K., & Sato, S. (1995). Perception of pitch accent by Korean learners of Japanese and its implications. *The Institute of Electronics, Information and Communication Engineers (IEICE) technical report: Speech 95*(41), 61-66.

Sumner, M., & Samuel, A. G. (2009). The role of experience in the processing of cross-dialectal variation. *Journal of Memory and Language*, *60*, 487-501.

Sun, S. H. (1998). *The development of a lexical tone phonology in American adult learners of standard Mandarin Chinese.* Honolulu, HI: University of Hawai'i Press.

Torgerson Jr., R. C. (2005). *A comparison of Beijing and Taiwan Mandarin tone register: An acoustic analysis of three native speech styles*. Unpublished master's thesis, Brigham Young University.

Tronnier, M. & Zetterholm, E. (2013). Tendencies of Swedish word accent production by L2 learners of with tonal and non-tonal L1. In E.L. Asu & P. Lippus (Eds.), *Proceedings of Nordic Prosody XI (pp.* 391-400). Tartu, Estonia.

Tronnier, M. & Zetterholm, E. (2014).Appropriate tone accent production in L2 Swedish by L1 speakers of Somali. In Cardoso, W., & Trofimovich, P. (Eds.), *Concordia Working Papers in Applied Linguistics, 5*, 722-736.

Trubetzkoy, N. (1969). *Principles of phonology* (C. A. M. Baltaxe, Trans.). Berkeley, CA: University of California Press. (Original work published 1939)

Tuc, H.-D. (2003). *Vietnamese-English bilingualism: patterns of code-switching.* New York: RoutledgeCurson.

Utsugi, A., Koizumi, M., and Mazuka, R. (2010). The perception of non-native lexical pitch accent by speakers of 'accentless' Japanese dialects. In M. Hasegawa-Johnson (Ed.), *Proceedings of Speech Prosody 2010* (100090, pp. 1-4). Chicago, IL

van Dommelen, W. A. & Husby, O. (2009). Native and non-native identification of Norwegian word tones. In M. Vainio, R. Aulanko, & O. Aaltonen (Eds.), *Nordic Prosody: Proceedings of the Xth Conference* (pp. 237-248). Helsinki: Peter Lang.

van Dommelen, W., & Husby, O. (2009). Perception of Norwegian word tones by Chinese and German listeners. In M.A. Watkins, A.S. Rauber, & B.O. Babtista (Eds.), *Recent Research in Second Language Phonetics/Phonology: Perception and* Production (pp. 308-321). Cambridge Scholars Publishing.

Van Lancker, D. (1980). Cerebral lateralization of pitch cues in the linguistic signal. *Papers in Linguistics: International Journal of Human Communication, 13,* 201-277.

Van Lanker, D., & Fromkin, V. A., (1973). Hemispheric specialization for pitch and 'tone': Evidence from Thai. *Journal of Phonetics 1*, 101-109

Vance, T. J. (1995). Final accent vs. no accent: utterance-final neutralization in Tokyo Japanese. *Journal of Phonetics, 23*, 487-499.

Vance, T. J. (1987). *An introduction to Japanese phonology*. Albany, N.Y.: State University of New York Press.

Vũ, T. P. (1981). *The acoustic and perceptual nature of tone in Vietnamese* (Unpublished doctoral dissertation) Australia National University, Canberra.

Vu, T.T., Nguyen, D.T., Luong, M.C., & Hosom, J.P. (2005). Vietnamese large vocabulary speech recognition. In *Proceedings of InterSpeech* (pp. 1689-1692). Lisbon, Portugal.

Wang, W. S-Y. (1976). Language change. In S.R. Hamad, H. D. Steklis , J. Lancaster (Eds.), *Origins and Evolution of Language and Speech. Annals of the New York Academy of Sciences, Vol. 280* (pp. 61-72). New York: The New York Academy of Sciences.

Wang, Y., Behne, D. M., Jongman, A. & Sereno, J. A. (2004). The role of linguistic experience in the hemispheric processing of lexical tone. *Applied Linguistics 25*, 449-466.

Wang, Y., Jongman, A., & Sereno, J. A. (2001). Dichotic perception of Mandarin tones by Chinese and American listeners. *Brain and Language*, *78*, 332–348.

Wang, Y., Jongman, A, and Sereno, J.A. (2006). L2 acquisition and processing of Mandarin Chinese tones. in P. Li, E. Bates, L.H. Tan, and O. Tseng (Eds.), *The Handbook of East Asian Psycholinguistics* (pp.250-256). Cambridge: Cambridge University Press .

Wang, Y., Sereno, J. A., Jongman, A., & Hirsch, J. (2003). fMRI evidence for cortical modification during learning of Mandarin lexical tone. *Journal of Cognitive Neuroscience 15*, 7, 1019-1027.

Wang, Y. Spence, J.A, Jongman, A. & Sereno, J. A. (1999). Training American listeners to perceive Mandarin tones. *Journal of the Acoustical Society of America, 106*(6), 3649-3658.

Wayland, R. P. & Guion, S. G. (2003). Perceptual discrimination of Thai tones by naïve and experienced learners of Thai. *Applied Psycholinguistics 24*, 113-129.

Wayland, R. P. & Guion, S. G. (2004). Training English and Chinese listeners to perceive Thai tones: A preliminary report. *Language Learning, 54*, 681-712.
Wayland, R. P., Herrera, E. & Kaan, E. (2010). Effects of musical experience and training on contour pitch perception. *Journal of Phonetics 38* (4), 654-662.
Weber, A. & Cutler, A. (2004). Lexical competition in non-native spoken-word recognition. *Journal of Memory and Language*, 50, 1–25.
Weinreich, U. (1953). *Languages in Contact.* The Hague: Mouton Publishers.
Werker, J. F., & Tees, R. C. (1984). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development, 7*, 49-63.
Whalen, D. H. & Xu ,Y. (1992). Information for Mandarin tones in the amplitude contour and in brief segments. *Phonetica*, *49*, 25-47.
White, C. M. (1981). Tonal pronunciation errors and interference from English intonation. *Journal of Chinese Language Teachers Association 16*(2), 27-56
Wood, C. C. (1974). Parallel processing of auditory and phonetic information in speech discrimination. *Perception & Psychophysics, 15*, 501-508.
Wrembel, M. (2010). L2-accented speech in L3 production. *International Journal of Multilingualism 7(1)*, 75-90.
Wu, E-C. & Fon, J. (2010). The effect of Min proficiency on the realization of Mandarin tones in Mandarin-Min bilinguals. In M. Hasegawa-Johnson (Ed.), *Proceedings of Speech Prosody 2010* (100420, pp. 1-4). Chicago, IL.
Yeh, H-N., Chan, H-C., & Cheng, Y-S. (2004). Language use in Taiwan: Language proficiency and domain analysis. *Journal of Taiwan Normal University: Humanities & Social Sciences*, *49*(1), 75-108.
Yip, M. (2001). Tone in East Asian Languages. In J. Goldsmith (Ed.) *The Handbook of Phonological Theory* (pp. 475-494). Cambridge, MA: Blackwell Publishers Ltd.
Yip, M. (2002). *Tone*. Cambridge: Cambridge University Press.
Yip, M., Leung, P.-Y., & Chen, H.-C. (1998). Phonological similarity effects in Cantonese spoken-word processing. *Proceedings of ICSLP'98*, *5*, 2139 – 2142. Sydney, Australia.
Yoshida, Y. Z., & Zamma, H. (2001). The Accent System of the Kyoto Dialect of Japanese: A Study on Phrasal Patterns and Paradigms. In J. van Weijer and T. Nishihara (Eds.), *Issues in Japanese Phonology and Morphology* (pp. 215-244). Berlin: De Gruyter Mouton.
Zsiga, E. C. (2007). Modeling diachronic change in the Thai tonal space. *University of Pennsylvania Working Papers in Linguistics*, *13*(1), 395-408.
Zsiga, E. & Nitisaroj, R. (2007). Tone features, tone perception, and peak alignment in Thai. *Language and Speech*, *50* (3), 343-383.

Appendix

A. Stimuli used in Pilot Study

Segments	High tone	Mid tone	Low tone	Falling tone	Rising tone
[ba:]	x	bar	shoulder	crazy	X
[tʃa:]	slow	tea	X	X	X
[di:]	X	good	X	X	X
[hu:]	shrunken	X	X	X	ear
[dʒa:]	X	X	leader	bright	particle to call s.o.'s attention
[kʰa:i]	X	to spit out	net; limit	camp	to sell
[ma:]	horse	to come	X	X	dog
[ma:i]	wood	X	X	window	to indicate
[mi:]	X	to have	egg noodles	X	bear
[na:]	younger maternal uncle/aunt	rice field	X	face; season	thick
[pʰi:]	X	X	X	older brother/ sister	ghost
[pʰu:]	X	land; ground	X	male; person	X
[ru:]	to know	hole	X	X	X
[su:]	X	X	to arrive at	to fight	X
[tʰa:]	to dare	to smear on	X	if; posture	X
[wa:]	X	Thai measure of length	X	to say	X

X indicates a nonce word with the segments of that row and the tone of that column.

B. Language background questionnaire used in Pilot Study

Questionnaire

[L2 language perception task]

Participant number/pseudonym ____

Native language(s) ____

Home country ____________________ Gender M F Age __________

Languages studied (Please state the language/number of years studied/when you were first exposed to the language/how you studied the language):

1.

2.

3.

4.

5.

Overseas experience (Please state where you have lived for a considerable amount of time outside of your home country. Please do not include short trips.)

C. Consent form for Pilot Study

Study #08-12931

INDIANA UNIVERSITY – BLOOMINGTON

INFORMED CONSENT STATEMENT

[L2 language perception task]

You are invited to participate in a research study. The purpose of this study is to test your perception of sounds in a foreign language, specifically Thai and/or Japanese.

INFORMATION

You will be asked to fill out a questionnaire describing your language background (i.e., native language and other languages you have learned). You will then listen to two tasks, consisting of sets of words in a foreign language, and determine if the words you hear are similar or not. This should require a maximum of 45~60 minutes (5 minutes for the questionnaire and 20 minutes for each task).

Approximately 9-45 people will be participating in the research.

Your answers will be recorded in writing on an answer sheet.

RISKS

There should be no risks.

BENEFITS

This study hopes to shed light on the perception of sounds in a foreign language and to apply the results to improving the teaching of language.

CONFIDENTIALITY

You will not be asked to record your name anywhere, but only information about your language background and other information. You will be given a code number/pseudonym, and this will be used in any possible publications or conference presentations.

CONTACT

If you have questions at any time about the study or the procedures, (or you experience adverse effects as a result of participating in this study) you may contact the researcher, Vance Schaefer, at 809 E. Hunter Ave. Apt. 16, Bloomington, IN 47401, (812) 339-0621, and vkschaef@indiana.edu.

If you feel you have not been treated according to the descriptions in this form, or your rights as a participant in research have been violated during the course of this project, you may contact the office for the Indiana University Bloomington Human Subjects Committee, Carmichael Center L03, 530 E. Kirkwood Ave., Bloomington, IN 47408, 812/855-3067, by e-mail at iub_hsc@indiana.edu.

PARTICIPATION

Your participation in this study is voluntary; you may refuse to participate without penalty. If you decide to participate, you may withdraw from the study at any time without penalty and without loss of benefits to which you are otherwise entitled. If you withdraw from the study before data collection is completed your data will be returned to you or destroyed.

CONSENT

I have read this form and received a copy of it. I have had all my questions answered to my satisfaction. I agree to take part in this study.

Subject's signature____ Date ___________________

Consent form date: March 11, 2008

D. Participant Background in Pilot Study

L1 Thai participants

code	age	hometown	overall scores (error rate %)				RTs				language	status	time abroad
THM1 (1)	25	Bangkok	12.5	4	1	1	1062	982	815	834	E, F	Educ	US - 3
THM2 (20)	32	Lop Buri	10.4	3	0	2	1179	1033	900	714	E	IEP	US- 1

L1 Japanese speakers

code	age	home	overall scores (error rate %)				RTs				language	status	time abroad
JPF1 (2)	50	Osaka, Nagoya	25	5	2	5	1914	1491	1796	1594	E, F	TESOL	US – 1.6
JPF2 (3)	30	Tokyo	44	6	6	9	1293	1396	1497	1175	E, F	Ling	US – 2.4
JPF3 (4)	35	Shizuoka	19	5	0	4	1108	844	979	936	E, A	Educ	US – 15, UK - 0.10, Egypt – 0.2
JPF4 (6)	21	Hokkaido	38	5	3	10	1238	1330	1396	1120	E, G	IEP, undergrad	US – 0.6
JPF5 (7)	34	Tochigi	25	6	0	6	1554	1150	1570	1346	E, F	Ling	US – 8 years, 4 mo.
JPF6 (10)	25	Ibaraki	27	4	3	6	1786	1345	1551	1544	E, F	Educ	US - 3
JPF7 (12)	43	Aichi	33	5	2	9	1498	1359	1328	1328	E, F	Japanese	US - 9
JPF8 (17)	35	Okayama	25	5	3	4	1167	1045	1241	962	E, S	Educ	US – 10, Aus – 0.4
JPF9 (26)	34	Chiba	27	7	0	6	1752.	1393	1629	1577	E, G, SWE	Linguistics?	US - 9
JPF10 (42)	33	Tokyo	21	4	0	6	1992	1196	1677	1311	E, K, M, S	Music	US – 8.6, UK - 4
JPF11 (44)	25	Yamaguchi	21	4	0	6	1550	1286	1475	1169	E, F, G	African …	US – 1.6
JPM1 (22)	45	Gifu/Tokyo?	25	5	2	5	1191	1078	1299	1391	E, F	Ling	US – 3.6

*Shaded area indicates this subject was cut from the final analysis as she varied from the general target group.

L1 English speakers

code	age	home	overall scores (error rate %)				RTs				language	status	time abroad
EF1 (5)	50		42	4	5	11	1618	1563	1401	1252	J, S, K, HCE	TESOL - done	Japan - 15
EF2 (13)	27		31	4	5	6	1273	1556	1388	1090	G, R	TESOL - done	Russia – 1.6
EF3 (18)	23		29	3	5	6	1200.	1420	1167	1081	S, Q	TESOL, Educ	S. America – 0.6 (?)/x3?
EF4 (25)	30		29	5	4	5	1297.84	1168	1187	975	S, F, G	French	France – 1.6~1.9
EF5 (31)	35		40	4	3	12	1753	1837	1664	1226	F, R, S, J	SLS	France /Belgium – 0.6
EF6 (33)	34		48	5	4	14	1791	1482	1613	1121	S, M, F, Q	SLS	Spanish-speaking – 2.6+, China 0.3-0.6/3x
EF7 (39)	45		42	7	5	8	1714	1632	1591	1361	P, S, Dan	TESOL – done	Denmark - 1
EF8 (45)	35		38	6	5	7	1637	1491	1564	1223	S, F, H	Spanish/SLS	S. America – 3.9, France – 0.4
EF9 (46)	25		23	4	1	6	1680	1143	1600	1136	S, F, Wolof	TESOL	Spanish – 0.8~0.9
EF10 (47)	25		21	4	1	5	1639	1387	1368	1402	A, F, G	TESOL	Morrocco – 1, French – 0.2-0.3
EM1 (28)	25		19	3	1	5	1185	1174	1059	1099	V, J	English	Japan – 0.9-1
EM2 (32)	37		50	3	9	12	1380	1458	1664	1687	S,P	Educ	Brazil – 2,

													Spanish – 1, French - ?
EM3 (40)	28		35	4	3	10	1185	1174	1059	1610	G	German	Germany – 2.9

*Shaded area indicates these subjects were cut from the final analysis as they had exposure to either a tone language or Japanese.

L1 Korean speakers

code	age	home	overall scores (error rate %)				RTs				language	status	time abroad
KR1 (8)	35	Seoul	42	6	4	10	1154	1171	1150	1020	E, G	Visiting scholar	US – 0.6, Germany – 0.1
KR2 (9)	30	Seoul	56	8	5	14	1800	1931	1634	1522	E, J	Business?	US – 1.6
KR3 (14)	29	Busan	31	5	2	8	1380	1516	1237	1161	E, G, F, J	SLS	US - 5
KR4 (15)	32	Seoul	29	3	1	10	1525	1452	1510	1211	G, J	SLS	US – 4.6, Canada – 0.4
KR5 (19)	30	Jinju	25	5	0	7	1421	1343	1414	1177	E, F	Ling	US – 3.6
KRF6 (21)	27	Seoul	35	6	1	10	1732	1397	1305	1178	E, F, J, M	SLS	US - 4
KRF7 (41)	35	Seoul	54	7	6	13	2239	2317	2211	1496	E, J, R, etc.	Slavic	US – 6.6, Japan – 1, Russia, 1
KRM1 (34)	27	Busan	21	3	3	4	1433	1527	1464	1413	E, F, G, J	undergrad	US – 5, Japan – 1.6
KRM2 (37)	47	Cholla	27	6	2	5	1340	1080	1173	1109	E, F, G, J	Ling	US - 5
KRM3 (43)	30	Seoul	50	7	5	12	1572	1435	1422	1096	E, J	Ling	US - 5

*shading indicates use of pitch accent in that dialect

L1 Mandarin speakers

code	age	home	Scores				RTs				language	status	time abroad
MNF1 (11)	26	ROC	19	4	1	4	747	794	806	870	E, S	Ling	US - 4
MNF2 (16)	31	ROC	21	4	1	5	1182	1177	1180	1162	T, E, J, S	SLS	US – 4.6
MNF3 (23)	24	ROC	21	6	0	4	1006	869	1096	972	T, Hakka, E, J	TESOL, Educ	US - 2
MNF4 (24)	29	ROC	21	4	1	5	1434	1086	1198	1329	T, E, F	Ling	US - 4
MNF5 (27)	24	PRC, Shandong	14.6	3	0	4	1361	1300	1265	1271	E, F, J	Chinese	US – 1.9
MNF6 (36)	27	ROC	8	1	0	3	1059	660	1025	1007	T, E, F, G, J, S, SWA	SLS	US – 4.6, Japan, - 0.2, UK – 0.3
MNM1 (29)	26	PRC, Hebei	31	7	3	5	1881	1338	1862	1243	E, J, F, Dan	SPEA	US – 2.6, Denmark - 1
MNM2 (30)	27	PRC, Shandong	21	6	0	4	1812	1395	1578	1372	E	Chem	US – 3.6
MNM3 (35)	30	ROC	12.5	3	2	1	1335	1193	1210	1156	T, E, R, J, G	Ling	US – 5.6, + 0.2
MNM4 (38)	27	PRC, Shanghai	10	3	0	2	1366	1289	1370	1163	E. G	Chem	US – 2.6

E. DMDX Script used in AXB task in Pilot Study

ABX program in the DMDX format:

```
<ep> <azk> <NumberOfItems 112> <Scramble 112> <ContinuousRun> <Delay 118>
<Timeout 3000> <id "Keyboard"> <mr +Space> <MapNegativeResponse "+Left Ctrl">
<MapPositiveResponse "+Right Ctrl"> <vm 1280,800,800,32,60> <eop>

$
0 <line -6> "You will hear a sequence of THREE sounds.", <line -4> "Please determine if
the SECOND sound is similar to the FIRST or THIRD sound.",<line -2> "Press the LEFT Control
Key, if the SECOND sound is similar to the FIRST sound.",<line 0> " Press the RIGHT Control
Key, if the SECOND sound is similar to the THIRD sound.",<line 2> "Please react as quickly as
possible.", <line 4> "A plus sign (+) appears at the start of each group of three
words.",<line 6> "You will first do a practice test.",<line 8> "Press SPACEBAR to start.";
999 <ms% 2000> / ;
$

-1001 <cfb "CORRECT"> <wfb "INCORRECT"> <tlfb "HURRY UP!"> <ms% 500> "+"/ <wav 2>
"fltwaa2"/ <ms% 500> / <wav 2> "mltwaa2"/ <ms% 500> / <wav 2> "fmtwaa2" * /;
\
+1002 <ms% 500> "+"/<wav 2> "fltsuu2"/ <ms% 500> / <wav 2> "mhtsuu2"/ <ms% 500> /
<wav 2> "fhtsuu2" * /;

-1003 <ms% 500> "+"/<wav 2> "fftmii2"/ <ms% 500> / <wav 2> "mftmii2"/ <ms%
500> / <wav 2> "frtmii2" * /;

+1004 <ms% 500> "+"/<wav 2> "frtruu2"/ <ms% 500> / <wav 2> "mftruu2"/ <ms%
500> / <wav 2> "fftruu2" * /;

-1005 <ms% 500> "+"/<wav 2> "ffthuu2"/ <ms% 500> / <wav 2> "mfthuu2"/ <ms%
500> / <wav 2> "fmthuu2" * /;

+1006 <ms% 500> "+"/<wav 2> "fhtkai2"/ <ms% 500> / <wav 2> "mftkai2"/ <ms%
500> / <wav 2> "fftkai2" * /;

-1007 <ms% 500> "+"/<wav 2> "fltmaa2"/ <ms% 500> / <wav 2> "mltmaa2"/ <ms%
500> / <wav 2> "frtmaa2" * /;

+1008 <ms% 500> "+"/<wav 2> "fmtbpuu2"/ <ms% 500> / <wav 2> "mrtbpuu2"/ <ms%
500> / <wav 2> "frtbpuu2" * /;

-2001 <ms% 500> "+"/<wav 2> "fdiskau2"/ <ms% 500> / <wav 2> "mdiskau2"/ <ms%
500> / <wav 2> "fdisgau2" * /;

-2002 <ms% 500> "+"/<wav 2> "fdisngae2"/ <ms% 500> / <wav 2> "mdisngae2"/
<ms% 500> / <wav 2> "fdisyae2" * /;

-2003 <ms% 500> "+"/<wav 2> "fdisnguu2"/ <ms% 500> / <wav 2> "mdisnguu2"/
<ms% 500> / <wav 2> "fdisngoo2" * /;

-2004 <ms% 500> "+"/<wav 2> "fdispuu2"/ <ms% 500> / <wav 2> "mdispuu2"/ <ms%
500> / <wav 2> "fdispoo2" * /;
```

```
+2005 <ms% 500> "+"/<wav 2> "fdischua2"/ <ms% 500> / <wav 2> "mdistua2"/ <ms%
500> / <wav 2> "fdistua2" * /;

+2006 <ms% 500> "+"/<wav 2> "fdisjuu2"/ <ms% 500> / <wav 2> "mdischuu2"/ <ms%
500> / <wav 2> "fdischuu2" * /;

+2007 <ms% 500> "+"/<wav 2> "fdisnii2"/ <ms% 500> / <wav 2> "mdisnee2"/ <ms%
500> / <wav 2> "fdisnee2" * /;

+2008 <ms% 500> "+"/<wav 2> "fdisgii2"/ <ms% 500> / <wav 2> "mdisgee2"/ <ms%
500> / <wav 2> "fdisgee2" * /;

$
0 <line 0> ”Practice is finished. Take a break.”,<line 1> ”Press SPACEBAR
when ready for the test”;

999 <nfb> <ms% 2000> / ;
$

\

-100001 <ms% 500> "+"/<wav 2> "fltbaa2"/ <ms% 500> / <wav 2> "mltbaa2"/ <ms% 500> /
<wav 2> "fmtbaa2" * /;

-100002 <ms% 500> "+"/<wav 2> "fltbpuu2"/ <ms% 500> / <wav 2> "mltbpuu2"/ <ms% 500> /
<wav 2> "fhtbpuu2" * /;

+100003 <ms% 500> "+"/<wav 2> "fmtchaa2"/ <ms% 500> / <wav 2> "mltchaa2"/
<ms% 500> / <wav 2> "fltchaa2" * /;

+100004 <ms% 500> "+"/<wav 2> "fhtdii2"/ <ms% 500> / <wav 2> "mltdii2"/ <ms%
500> / <wav 2> "fltdii2" * /;

-100005 <ms% 500> "+"/<wav 2> "fmthuu2"/ <ms% 500> / <wav 2> "mmthuu2"/ <ms%
500> / <wav 2> "fhthuu2" * /;

-100006 <ms% 500> "+"/<wav 2> "fhtkai2"/ <ms% 500> / <wav 2> "mhtkai2"/ <ms%
500> / <wav 2> "fmtkai2" * /;

+100007 <ms% 500> "+"/<wav 2> "fltmaa2"/ <ms% 500> / <wav 2> "mmtmaa2"/ <ms%
500> / <wav 2> "fmtmaa2" * /;

+100008 <ms% 500> "+"/<wav 2> "fltmai2"/ <ms% 500> / <wav 2> "mhtmai2"/ <ms%
500> / <wav 2> "fhtmai2" * /;

-100009 <ms% 500> "+"/<wav 2> "fmtmii2"/ <ms% 500> / <wav 2> "mmtmii2"/ <ms% 500> /
<wav 2> "fltmii2" * /;

-100010 <ms% 500> "+"/<wav 2> "fhtnaa2"/ <ms% 500> / <wav 2> "mhtnaa2"/ <ms%
500> / <wav 2> "fltnaa2" * /;

+100011 <ms% 500> "+"/<wav 2> "fmtpii2"/ <ms% 500> / <wav 2> "mhtpii2"/ <ms% 500> /
 <wav 2> "fhtpii2" * /;

+100012 <ms% 500> "+"/<wav 2> "fhtruu2"/ <ms% 500> / <wav 2> "mmtruu2"/ <ms%
500> / <wav 2> "fmtruu2" * /;
```

```
-100013 <ms% 500> "+"/<wav 2> "frtsuu2"/ <ms% 500> / <wav 2> "mrtsuu2"/ <ms% 500> /
 <wav 2> "fftsuu2" * /;

-100014 <ms% 500> "+"/<wav 2> "ffttaa2"/ <ms% 500> / <wav 2> "mfttaa2"/ <ms%
500> / <wav 2> "frttaa2" * /;

+100015 <ms% 500> "+"/<wav 2> "frtwaa2"/ <ms% 500> / <wav 2> "mftwaa2"/ <ms% 500> /
 <wav 2> "fftwaa2" * /;

+100016 <ms% 500> "+"/<wav 2> "fftyaa2"/ <ms% 500> / <wav 2> "mrtyaa2"/ <ms%
500> / <wav 2> "frtyaa2" * /;

-100017 <ms% 500> "+"/<wav 2> "frtbaa2"/ <ms% 500> / <wav 2> "mrtbaa2"/ <ms% 500> /
<wav 2> "fftbaa2" * /;

-100018 <ms% 500> "+"/<wav 2> "fftbpuu2"/ <ms% 500> / <wav 2> "mftbpuu2"/
<ms% 500> / <wav 2> "frtbpuu2" * /;

+100019 <ms% 500> "+"/<wav 2> "frtchaa2"/ <ms% 500> / <wav 2> "mftchaa2"/ <ms% 500> /
<wav 2> "fftchaa2" * /;

+100020 <ms% 500> "+"/<wav 2> "fftdii2"/ <ms% 500> / <wav 2> "mrtdii2"/ <ms%
500> / <wav 2> "frtdii2" * /;

-100021 <ms% 500> "+"/<wav 2> "frthuu2"/ <ms% 500> / <wav 2> "mrthuu2"/ <ms% 500> /
 <wav 2> "ffthuu2" * /;

-100022 <ms% 500> "+"/<wav 2> "fftkai2"/ <ms% 500> / <wav 2> "mftkai2"/ <ms%
500> / <wav 2> "frtkai2" * /;

+100023 <ms% 500> "+"/<wav 2> "frtmaa2"/ <ms% 500> / <wav 2> "mftmaa2"/ <ms% 500> /
<wav 2> "fftmaa2" * /;

+100024 <ms% 500> "+"/<wav 2> "fftmai2"/ <ms% 500> / <wav 2> "mrtmai2"/ <ms%
500> / <wav 2> "frtmai2" * /;

-100025 <ms% 500> "+"/<wav 2> "fltmii2"/ <ms% 500> / <wav 2> "mltmii2"/ <ms% 500> /
<wav 2> "frtmii2" * /;

-100026 <ms% 500> "+"/<wav 2> "frtnaa2"/ <ms% 500> / <wav 2> "mrtnaa2"/ <ms%
500> / <wav 2> "fltnaa2" * /;

-100027 <ms% 500> "+"/<wav 2> "fltpii2"/ <ms% 500> / <wav 2> "mltpii2"/ <ms% 500> /
<wav 2> "fftpii2" * /;

-100028 <ms% 500> "+"/<wav 2> "fftruu2"/ <ms% 500> / <wav 2> "mftruu2"/ <ms%
500> / <wav 2> "fltruu2" * /;

-100029 <ms% 500> "+"/<wav 2> "fmtsuu2"/ <ms% 500> / <wav 2> "mmtsuu2"/ <ms% 500> /
<wav 2> "frtsuu2" * /;

-100030 <ms% 500> "+"/<wav 2> "frttaa2"/ <ms% 500> / <wav 2> "mrttaa2"/ <ms%
500> / <wav 2> "fmttaa2" * /;

-100031 <ms% 500> "+"/<wav 2> "fmtwaa2"/ <ms% 500> / <wav 2> "mmtwaa2"/ <ms% 500> /
<wav 2> "fftwaa2" * /;
```

```
-100032 <ms% 500> "+"/<wav 2> "fftyaa2"/ <ms% 500> / <wav 2> "mftyaa2"/ <ms%
500> / <wav 2> "fmtyaa2" * /;

$
0 <line 0> ”Take a break.”,<line 1> ”Press SPACEBAR when ready to continue”;
$

-100033 <ms% 500> "+"/<wav 2> "fhtbaa2"/ <ms% 500> / <wav 2> "mhtbaa2"/ <ms% 500> /
<wav 2> "frtbaa2" * /;

-100034 <ms% 500> "+"/<wav 2> "frtbpuu2"/ <ms% 500> / <wav 2> "mrtbpuu2"/
<ms% 500> / <wav 2> "fhtbpuu2" * /;

-100035 <ms% 500> "+"/<wav 2> "fhtchaa2"/ <ms% 500> / <wav 2> "mhtchaa2"/ <ms% 500> /
<wav 2> "fftchaa2" * /;

-100036 <ms% 500> "+"/<wav 2> "fftdii2"/ <ms% 500> / <wav 2> "mftdii2"/ <ms%
500> / <wav 2> "fhtdii2" * /;

+100037 <ms% 500> "+"/<wav 2> "flthuu2"/ <ms% 500> / <wav 2> "mrthuu2"/ <ms% 500> /
<wav 2> "frthuu2" * /;

+100038 <ms% 500> "+"/<wav 2> "frtkai2"/ <ms% 500> / <wav 2> "mltkai2"/ <ms%
500> / <wav 2> "fltkai2" * /;

+100039 <ms% 500> "+"/<wav 2> "fltmaa2"/ <ms% 500> / <wav 2> "mftmaa2"/ <ms% 500> /
 <wav 2> "fftmaa2" * /;

+100040 <ms% 500> "+"/<wav 2> "fftmai2"/ <ms% 500> / <wav 2> "mltmai2"/ <ms%
500> / <wav 2> "fltmai2" * /;

+100041 <ms% 500> "+"/<wav 2> "fmtmii2"/ <ms% 500> / <wav 2> "mrtmii2"/ <ms% 500> /
<wav 2> "frtmii2" * /;

+100042 <ms% 500> "+"/<wav 2> "frtnaa2"/ <ms% 500> / <wav 2> "mmtnaa2"/ <ms%
500> / <wav 2> "fmtnaa2" * /;

+100043 <ms% 500> "+"/<wav 2> "fmtpii2"/ <ms% 500> / <wav 2> "mftpii2"/ <ms% 500> /
<wav 2> "fftpii2" * /;

+100044 <ms% 500> "+"/<wav 2> "fftruu2"/ <ms% 500> / <wav 2> "mmtruu2"/ <ms%
500> / <wav 2> "fmtruu2" * /;

+100045 <ms% 500> "+"/<wav 2> "fhtsuu2"/ <ms% 500> / <wav 2> "mrtsuu2"/ <ms% 500> /
<wav 2> "frtsuu2" * /;

+100046 <ms% 500> "+"/<wav 2> "frttaa2"/ <ms% 500> / <wav 2> "mhttaa2"/ <ms%
500> / <wav 2> "fhttaa2" * /;

+100047 <ms% 500> "+"/<wav 2> "fhtwaa2"/ <ms% 500> / <wav 2> "mftwaa2"/ <ms% 500> /
<wav 2> "fftwaa2" * /;

+100048 <ms% 500> "+"/<wav 2> "fftyaa2"/ <ms% 500> / <wav 2> "mhtyaa2"/ <ms%
500> / <wav 2> "fhtyaa2" * /;

-200001 <ms% 500> "+"/<wav 2> "fdisboo2"/ <ms% 500> / <wav 2> "mdisboo2"/
<ms% 500> / <wav 2> "fdispoo2" * /;
```

```
-200002 <ms% 500> "+"/<wav 2> "fdisbpoo2"/ <ms% 500> / <wav 2> "mdisbpoo2"/
<ms% 500> / <wav 2> "fdisboo2" * /;

-200003 <ms% 500> "+"/<wav 2> "fdisbpoo2"/ <ms% 500> / <wav 2> "mdisbpoo2"/
<ms% 500> / <wav 2> "fdispoo2" * /;

-200004 <ms% 500> "+"/<wav 2> "fdisbpau2"/ <ms% 500> / <wav 2> "mdisbpau2"/
<ms% 500> / <wav 2> "fdisbpoo2" * /;

-200005 <ms% 500> "+"/<wav 2> "fdispau2"/ <ms% 500> / <wav 2> "mdispau2"/
<ms% 500> / <wav 2> "fdispoo2" * /;

-200006 <ms% 500> "+"/<wav 2> "fdischua2"/ <ms% 500> / <wav 2> "mdischua2"/
<ms% 500> / <wav 2> "fdischuu2" * /;

-200007 <ms% 500> "+"/<wav 2> "fdisjea2"/ <ms% 500> / <wav 2> "mdisjea2"/
<ms% 500> / <wav 2> "fdisjua2" * /;

-200008 <ms% 500> "+"/<wav 2> "fdisjea2"/ <ms% 500> / <wav 2> "mdisjea2"/
<ms% 500> / <wav 2> "fdisjuu2" * /;

-200009 <ms% 500> "+"/<wav 2> "fdisngoo2"/ <ms% 500> / <wav 2> "mdisngoo2"/
<ms% 500> / <wav 2> "fdisngau2" * /;

-200010 <ms% 500> "+"/<wav 2> "fdisnguu2"/ <ms% 500> / <wav 2> "mdisnguu2"/
<ms% 500> / <wav 2> "fdisngau2" * /;

-200011 <ms% 500> "+"/<wav 2> "fdisngaa2"/ <ms% 500> / <wav 2> "mdisngaa2"/
<ms% 500> / <wav 2> "fdisnaa2" * /;

-200012 <ms% 500> "+"/<wav 2> "fdisduu2"/ <ms% 500> / <wav 2> "mdisduu2"/
<ms% 500> / <wav 2> "fdisdua2" * /;

-200013 <ms% 500> "+"/<wav 2> "fdisdua2"/ <ms% 500> / <wav 2> "mdisdua2"/
<ms% 500> / <wav 2> "fdistua2" * /;

-200014 <ms% 500> "+"/<wav 2> "fdisdtoo2"/ <ms% 500> / <wav 2> "mdisdtoo2"/
<ms% 500> / <wav 2> "fdistoo2" * /;

-200015 <ms% 500> "+"/<wav 2> "fdisgee2"/ <ms% 500> / <wav 2> "mdisgee2"/
<ms% 500> / <wav 2> "fdisgae2" * /;

-200016 <ms% 500> "+"/<wav 2> "fdiskee2"/ <ms% 500> / <wav 2> "mdiskee2"/
<ms% 500> / <wav 2> "fdiskii2" * /;

$
0 <line 0> ”Take another break.”,<line 1> ”Press SPACEBAR when ready to
continue”;
$

-200017 <ms% 500> "+"/<wav 2> "fdischae2"/ <ms% 500> / <wav 2> "mdischae2"/
<ms% 500> / <wav 2> "fdischua2" * /;

-200018 <ms% 500> "+"/<wav 2> "fdispea2"/ <ms% 500> / <wav 2> "mdispea2"/
<ms% 500> / <wav 2> "fdispua2" * /;
```

```
-200019 <ms% 500> "+"/<wav 2> "fdisjii2"/ <ms% 500> / <wav 2> "mdisjii2"/
<ms% 500> / <wav 2> "fdischii2" * /;

-200020 <ms% 500> "+"/<wav 2> "fdisjee2"/ <ms% 500> / <wav 2> "mdisjee2"/
<ms% 500> / <wav 2> "fdisjae2" * /;

-200021 <ms% 500> "+"/<wav 2> "fdisngae2"/ <ms% 500> / <wav 2> "mdisngae2"/
<ms% 500> / <wav 2> "fdisnae2" * /;

-200022 <ms% 500> "+"/<wav 2> "fdisnee2"/ <ms% 500> / <wav 2> "mdisnee2"/
<ms% 500> / <wav 2> "fdisngee2" * /;

-200023 <ms% 500> "+"/<wav 2> "fdisdae2"/ <ms% 500> / <wav 2> "mdisdae2"/
<ms% 500> / <wav 2> "fdisdtae2" * /;

-200024 <ms% 500> "+"/<wav 2> "fdisdee2"/ <ms% 500> / <wav 2> "mdisdee2"/
<ms% 500> / <wav 2> "fdisdtee2" * /;

+200025 <ms% 500> "+"/<wav 2> "fdisdtae2"/ <ms% 500> / <wav 2> "mdistae2"/
<ms% 500> / <wav 2> "fdistae2" * /;

+200026 <ms% 500> "+"/<wav 2> "fdisgoo2"/ <ms% 500> / <wav 2> "mdisgau2"/
<ms% 500> / <wav 2> "fdisgau2" * /;

+200027 <ms% 500> "+"/<wav 2> "fdiskoo2"/ <ms% 500> / <wav 2> "mdiskau2"/
<ms% 500> / <wav 2> "fdiskau2" * /;

+200028 <ms% 500> "+"/<wav 2> "fdisgea2"/ <ms% 500> / <wav 2> "mdisgae2"/
<ms% 500> / <wav 2> "fdisgae2" * /;

+200029 <ms% 500> "+"/<wav 2> "fdisbplaa2"/ <ms% 500> / <wav 2> "mdisbpaa2"/
<ms% 500> / <wav 2> "fdisbpaa2" * /;

+200030 <ms% 500> "+"/<wav 2> "fdispra2"/ <ms% 500> / <wav 2> "mdisbplaa2"/
<ms% 500> / <wav 2> "fdisbplaa2" * /;

+200031 <ms% 500> "+"/<wav 2> "fdisrap2"/ <ms% 500> / <wav 2> "mdisrak2"/
<ms% 500> / <wav 2> "fdisrak2" * /;

+200032 <ms% 500> "+"/<wav 2> "fdisrat2"/ <ms% 500> / <wav 2> "mdisrak2"/
<ms% 500> / <wav 2> "fdisrak2" * /;

+200033 <ms% 500> "+"/<wav 2> "fdisrat2"/ <ms% 500> / <wav 2> "mdisrap2"/
<ms% 500> / <wav 2> "fdisrap2" * /;

+200034 <ms% 500> "+"/<wav 2> "fdisgaeo2"/ <ms% 500> / <wav 2> "mdisgaao2"/
<ms% 500> / <wav 2> "fdisgaao2" * /;

+200035 <ms% 500> "+"/<wav 2> "fdisglap2"/ <ms% 500> / <wav 2> "mdiskrap2"/
<ms% 500> / <wav 2> "fdiskrap2" * /;

+200036 <ms% 500> "+"/<wav 2> "fdisyae2"/ <ms% 500> / <wav 2> "mdisyee2"/
<ms% 500> / <wav 2> "fdisyee2" * /;

+200037 <ms% 500> "+"/<wav 2> "fdisyua2"/ <ms% 500> / <wav 2> "mdisyea2"/
<ms% 500> / <wav 2> "fdisyea2" * /;
```

```
+200038 <ms% 500> "+"/<wav 2> "fdisyuu2"/ <ms% 500> / <wav 2> "mdisyuu2"/
<ms% 500> / <wav 2> "fdisyua2" * /;

+200039 <ms% 500> "+"/<wav 2> "fdisgea2"/ <ms% 500> / <wav 2> "mdisgee2"/
<ms% 500> / <wav 2> "fdisgee2" * /;

+200040 <ms% 500> "+"/<wav 2> "fdispuu2"/ <ms% 500> / <wav 2> "mdispau2"/
<ms% 500> / <wav 2> "fdispau2" * /;

-200041 <ms% 500> "+"/<wav 2> "fdisngae2"/ <ms% 500> / <wav 2> "mdisngaa2"/
<ms% 500> / <wav 2> "fdisngaa2" * /;

+200042 <ms% 500> "+"/<wav 2> "fdisdee2"/ <ms% 500> / <wav 2> "mdisdua2"/
<ms% 500> / <wav 2> "fdisdua2" * /;

+200043 <ms% 500> "+"/<wav 2> "fdisjae2"/ <ms% 500> / <wav 2> "mdisjua2"/
<ms% 500> / <wav 2> "fdisjua2" * /;

+200044 <ms% 500> "+"/<wav 2> "fdispua2"/ <ms% 500> / <wav 2> "mdispau2"/
<ms% 500> / <wav 2> "fdispau2" * /;

+200045 <ms% 500> "+"/<wav 2> "fdisbea2"/ <ms% 500> / <wav 2> "mdisbpau2"/
<ms% 500> / <wav 2> "fdisbpau2" * /;

+200046 <ms% 500> "+"/<wav 2> "fdisyae2"/ <ms% 500> / <wav 2> "mdisyua2"/
<ms% 500> / <wav 2> "fdisyua2" * /;

+200047 <ms% 500> "+"/<wav 2> "fdisgae2"/ <ms% 500> / <wav 2> "mdisgea2"/
<ms% 500> / <wav 2> "fdisgea2" * /;

+200048 <ms% 500> "+"/<wav 2> "fdistae2"/ <ms% 500> / <wav 2> "mdistua2"/
<ms% 500> / <wav 2> "fdistua2" * /;

$
0 <line -1> "The End...thank you for participating", <line 1> "Press Esc to
Save the data";
$
```

F. Stimuli used for the ABX Monosyllabic, ABX Disyllabic and Lexical Decision with Repetition Priming task in the current dissertation study

ABX monosyllabic stimuli, conditions and number of trials

Test (tone) (80 trials**)		Control (segmental) (80 trials**)		Training (segmental only) (10 trials**)	
Tone types	segments	Tone types	Segment pairs	Tone types	Segment pairs
Falling-Rising Low-Falling Low-Rising Mid-Falling Mid-Rising High-Falling High-Rising Low-Mid Low-High Mid-High	[no:i] โนย* [p^huai] พวย	Falling Rising Low Mid High	[bɛ:o] เบว - [tɛ:o] เตว [wia] เวีย - [t^hia] เทีย [ua] อัวะ - [ia] เอีย [ria] เรีย - [rɯ:i] รืย	Falling Rising Low Mid High	[wuːi] วูย - [p^huːi] พูย [dua] ดัว - [ŋiːn] งีน

*Spellings in the Thai script may not reflect the actual tones. A complex combination of spelling conventions determines the tone of any given syllable in Thai. First, a vowel within a syllable may be marked above by one of four diacritics denoting tone: ่, ้, ๊, ๋. These diacritics are referred to as one (ไม้เอก), two (ไม้โท), three (ไม้ตรี) and four (ไม้จัตวา), respectively. A syllable may also not be marked by any of the four diacritics. This is, however, where the similarity to the markings of Mandarin tones in *pinyin* ends. The tones of syllables in Thai are further determined in conjunction with three other factors. The first factor is the class of the consonant (i.e., low, mid, high *which do not correlate to the tones with the same designations*) preceding the vowel. The consonant does not necessarily represent phonetically different consonants. For example, there are several letters that represent [k^h] such as ข ('high' consonant) and ค ('low' consonant), e.g., ไข่ [k^hai] low tone 'egg' ควาย [k^hwaai] mid tone 'water buffalo'. The second factor is whether the syllable is "alive" or "dead" (i.e., long vowel/open syllable or short vowel/closed syllable, respectively). The third factor is whether one of two silent letters (ห [h] or อ [ɔ]) precedes the initial consonant of the syllable. This third factor is somewhat analogous to English where adding a final 'e' to the word 'fin' changes the pronunciation of the 'i' as in the case of 'fine.' While students are taught the rules in school, anecdotally many Thai appear to spell and read by rote memorization. Again, this is analogous to English where generally spelling is memorized despite being taught a few rules like "'i' after 'e' except after 'c'" etc.

** The tone test condition = 10 tone types x 2 segments types x 4 orders (AAB, ABB, BBA, BAA) = 80 trials; the control test condition = 5 tone types x 4 segment pairs x 4 orders (AAB, ABB, BBA, BAA) = 80 trials; training = 5 tone types x 2 segment pairs = 10 trials

ABX disyllabic stimuli, conditions and number of trials

Test (=tone) (40 trials**)		Control (=segmentals) (40 trials**)		Training (=segmentals only) (10 trials**)	
Tone types	Segments	Tone types	Segment pairs	Tone types	Segment pairs
Low+High-High+Low	[dua.p^{h}uui] ตัวพูย*	High+Falling	[p^{h}uːi.wuːi] พูยวูย - [ruːi.wuːi] รูยวูย	High+Falling	[noːbɛːo] โนเบว - [luai.p^{h}uai] ลูยพวย
Low+Mid-High+Low	[kiːŋ.kæ] กึงแก	High+Low	[teo.fiːŋ] เตวฟีง - [teo.foːŋ] เตวโฟง	High+Low	[roːbɛːo] โรเบว - [luːi.p^{h}uai] ลวยพวย
Low+High-Mid+High		Low+High		Low+High	
Low+Mid-Low+High		Low+Mid		Low+Mid	
Low+Mid-Mid+High		Mid+High		Mid+High	

*Spellings in the Thai script may not reflect the actual tones.
** The tone test condition = 5 tone types x 2 segments x 4 orders (AAB, ABB, BBA, BAA) = 40 trials; the control test condition = 5 tone types x 2 segments x 4 orders (AAB, ABB, BBA, BAA) = 40 trials; training = 5 tone types x 2 segments = 10 trials

Lexical Decision stimuli

Test items (tones) Real words			Test Nonword counterparts			Control items (segments) Real words			Control Nonword counterparts		
IPA	Tone	Gloss, character, *pinyin*	IPA	Tone	*pinyin*	IPA	Tone	Gloss, character, *pinyin*	IPA	Tone	*pinyin*
[t^{h}aŋ]	1	'soup' 湯, *tang*	[hiŋ]	1	*hing*	[t͡ɕja]	1	'house, home' 家, *jia*	[sɕ][1]	1	*se*
[t^{h}aŋ]	2	'candy' 糖, *tang*	[hiŋ]	2	*hing*	[t͡ɕjɛ]	1	'to pick up (the phone) etc.' 接, *jie*	[səʊ̯][1]	1	*sou*
[t^{h}jɛn]	1	'heaven, sky' 天, *tian*	[nweɪ]	1	*nui*	[t͡ʂ^{h}əŋ]	2	'to become' 成, *cheng*	[sjɛn][1]	2	*sian*
[t^{h}jɛn]	2	'to be sweet' 甜, *tian*	[nweɪ]	2	*nui*	[t͡ʂ^{h}aŋ]	2	'to be long' 長, *chang*	[sən]	2	*sen*
[xɕ]	1	'to drink' 喝, *he*	[kiː][1]	1	*gii*	[ʂɑʊ̯]	3	'to be few' 少, *shao*	[pua]	3	*bua*
[xɕ]	2	'and' 和 or 'river'河, *he*	[kiː]	2	*gii*	[t͡ʂɑʊ̯]	3	'to search for' 找, *zhao*	[puɛ]	3	*bue*
[t^{h}iŋ]	1	'to listen to' 聽, *ting*	[juɔ][1]	1	*yuo*	[ɕɑʊ̯]	3	'good' 好, *hao*	[kjɛn]	3	*gian*

[tʰiŋ]	2	'to stop' 停, *ting*	[juɔ][1]	2	*yuo*	[pɑʊ̯]	3	'to be full' 飽, *bao*	[kʰjɛn][1]	3	*kian*
[t͡ʂʰuan]	1	'to wear' 穿, *chuan*	[fjɛn]	1	*fian*	[kweɪ]	4	'expensive' 貴, *gui*	[tʰjəʊ̯]	4	*tiu*
[t͡ʂʰuan]	2	'boat' 船, *chuan*	[fjɛn]	2	*fian*	[kwɔ]	4	'to pass, cross' 過, *guo*	[sjəʊ̯]	4	*siu*
[t͡ʂʰa]	1	'to insert' 插 or 'fork' 叉, *cha*	[pʰəʊ̯]	1	*pou*	[fan]	4	'meal' 飯, *fan*	[faɪ]	4	*fai*
[t͡ʂʰa]	2	'tea' 茶, *cha*	[pʰəʊ̯][1]	2	*pou*	[pan]	4	'half' 半, *ban*	[jaɪ]	4	*yai*
[t͡ɕʰjɛn]	1	'1,000' 千, *qian*	[tʰua]	1	*tua*	[ta]	4	'to be big' 大, *da*	[nja]	1	*nia*
[t͡ɕʰjɛn]	2	'money' 錢, *qian*	[tʰua][1]	2	*tua*	[ti]	4	'land' 地, *di*	[nən][1]	1	*nen*
[t͡ɕʰi]	1	'seven' 七, *qi*	[mjəʊ̯]	1	*miu*	[t͡saɪ]	4	'to be, exist' 在, *zai*	[t͡sʰəʊ̯][1]	1	*cou*
[t͡ɕʰi]	2	'to ride (a horse, motorbike)' 騎, qi	[mjəʊ̯]	2	*miu*	[maɪ]	4	'to sell' 賣, *mai*	[fəʊ̯]	1	*fou*

[1] Indicates a nonword that is close to a real word in Taiwanese.
[2] Indicates a nonword that is close to the pronunciation of a real word in Taiwanese-Mandarin.

Distracters					
Real words	Tone(s)	Gloss, character, *pinyin*	Non-words	Tone(s)	*pinyin*
[ʂan]*	1	'mountain' 山, *shan*	[fɤ]*	1	*fe*
[kɑʊ̯]	1	'to be high' 高, *gao*	[mun]	1	*mun*
[ɕin]*	1	'to be new' 新, *xin*	[t͡ʂun]*	1	*zhun*
[tʰəʊ̯]	2	'head' 頭, *tou*	[tʰjəʊ̯]*	1	*tiu*
[nan]*	2	'to be difficult' 難, *nan*	[fjɛ]*	2	*fie*
[ɕiaŋ]	3	'to think' 想, *xiang*	[sjɛn][1]	2	*sian*
[ɕœ]*	3	'snow' 雪, *xue*	[jweɪ]	2	*yui*
[kəʊ̯]	3	'dog' 狗, *gou*	[t͡ʂweɪ]*[1]	2	*zhui*
[ʂəʊ̯]*	3	'hand'	[kjɑʊ̯]	2	*giao*

		手, *shou*			
[t͡ʂɨ]	3	‘paper’ 紙, *zhi*	[pwan]*	3	*buan*
[ʂweɪ]*	3	‘water’ 水, *shui*	[lən][2]	3	*len*
[təŋ]	3	‘to wait’ 等, *deng*	[tən]*[2]	3	*den*
[t͡səʊ̯]*	3	‘to go/walk’ 走, *zou*	[nweɪ]	3	*nui*
[xua]	4	‘conversation’ 話, *hua*	[pəʊ̯n]*	3	*bon*
[tɑʊ̯]*	4	‘to arrive’ 到, *dao*	[fi]	3	*fi*
[fan.kwan]	4	‘restaurant’ 飯館, *fanguan*	[faɪ]*	3	*fai*
[t͡ʂəʊ̯ŋ.wən]*	1-2	‘Chinese language’ 中文, *zhongwen*	[fjɛn.fi]	1-1	*fianfi*
[ʂəŋ.xwɔ]	1-2	‘life’ 生活, *shenghuo*	[fjɛn.kʰeɪ]*	1-3	*fiankei*
[ʂən.tʰi]*	1-3	‘body’ 身體, *shenti*	[tʰən.pʰɤ]	2-2	*tenpe*
[ʂaŋ.tjɛn]*	1-4	‘store’ 商店, *shangdian*	[nweɪ.kjɛ]*	2-3	*nuigie*
[jin.jœ]*	1-4	‘music’ 音樂, *yinyue*	[ton.teɪ]	2-4	*dondei*
[ʂɨ.t͡ɕjɛn]	2-1	‘time’ 時間, *shijian*	[kʰweɪ.fɤ]*	3-2	*kuife*
[t͡ɕjɛ.xun]*	2-1	‘to get married’ 結婚, *jiehun*	[pɤ.tən]	3-4	*beden*
[ɕœ.ɕjɑʊ̯]	2-4	‘school’ 學校, *xuexiao*	[kʰeɪ.kʰweɪ]*	3-4	*keikui*
[lɑʊ̯.ʂɨ]*	3-1	‘teacher’ 老師, *laoshi*	[fjɛ.fjɛ]	4-1	*fiefie*
[xwɔ.t͡ʂʰɤ]	3-1	‘train’ 火車, *huoche*	[tʰəʊ̯n.lən]	4-2	*tonlen*
[kan.t͡ɕœ]*	3-2	‘feeling’ 感覺, *ganjue*	[kʰeɪ.nun]*	4-2	*keinun*
[jɛn.t͡ɕiŋ]	3-4	‘glasses’ 眼鏡, *yanjing*	[lən.pɤ]*[1]	4-2	*lenbe*
[wən.tʰi]*	4-2	‘question/problem’ 問題, *wenti*	[səʊ̯n.ɻeɪ][1]	4-3	*sonrei*
[tjɛn.nɑʊ̯]	4-3	‘computer’	[mun.pɤ]	4-3	*munbe*

		電腦, *diannao*			
[t͡swɔ.jɛ]	4-4	‘homework’ 作業, *zuoye*	[fun.mun]*	4-3	*funmun*
[ʂɤ.xweɪ]	4-4	‘society’ 社會, *shehui*	[pʰɤ.fjɛ]	4-3	*pefie*

* Asterisk indicates that this distracter item was inserted twice. That is, the above table contains 64 different distracters of which 32 were repeated for a total of 96 distracters in the task.
[1] Indicates a nonword that is close to a real word in Taiwanese.
[2] Indicates a nonword that is close to the pronunciation of a real word in Taiwanese-Mandarin.

Training items					
Real words	Tone(s)	Gloss, character, *pinyin*	Nonwords	Tone(s)	*pinyin*
[t͡ʂʰɨ]	1	‘to eat’ 吃, *chi*	[sua][1]	1	*sua*
[t͡ʂu]	1	‘pig’ 猪, *zhu*	[pɤ]	1	*be*
[tweɪ]	4	‘to be correct’ 對, *dui*	[fweɪ]	4	*fui*
[ʂwɔ]	1	‘to say’ 說, *shuo*	[lua][1]	4	*lua*
[tjɛn. xua]	4-4	‘telephone’電話, *dianhua*	[fjɑʊ̯]	4	*fiao*
[ʂɨ.t͡ɕjɛ]	4-4	‘world’ 世界, *shijie*	[kjɛ.tən]	3-1	*kieden*

[1] Indicates a nonword that is close to a real word in Taiwanese.
[2] Indicates a nonword that is close to the pronunciation of a real word in Taiwanese-Mandarin.

G. Language background questionnaire used in Current Study

Subject's ID	______________
Nb - Group	______________

Language Background Questionnaire

This questionnaire is designed to learn about your language history. Information collected will be stored in a secured laboratory. No identifying information will be made available unless you specifically give permission in writing to do so. All answers are strictly confidential. Any information you provide will not be distributed to outside parties. Thank you!

1. **Date of experiment:_______________ [session _______]**
2. **Sex : M F**
3. **Date of birth** (month/day/year): ....
4. Are you **left or right-handed**? L R Both
5. **Place of birth**: a. City: .... b. State/Province & Country: ....
6. **Your native language(s) – please be specific (e.g. British English)?**

7. Father's native language: ....

8. Mother's native language: ....

9. As a child, what languages were spoken in your home and by whom (e.g., by parents, guardians, grandparents, or relatives)? For example, these can be languages that you frequently heard but didn't understand or speak, languages that you understood but did not speak, or languages that you both understood and spoke. Please indicate whether you spoke and/or understood any of these languages.

...

...

10. As a child, what languages were spoken outside your home (e.g., kindergarten or school)? These can be languages that you frequently heard but didn't speak, languages that you understood but did not speak, or languages that you both understood and spoke. Please indicate whether you spoke and/or understood any of these languages.

...

...

11. Do you speak or understand a dialect of your native language, e.g., *Kyungsang Korean, Fukushima or Kumamoto or Osaka Japanese*, *Taiwanese Mandarin, etc*. Please indicate how well you speak or understand this dialect.

....

....

12. Number of **years of formal education** in your native country _________ and/or in the U.S. ___________.

If you had formal education in other countries, which countries are they and how many years?

Country: _____________ Years: ____________ Country: _____________ Years: __________

Country: _____________ Years: ____________ Country: _____________ Years: __________

If you need more space, please use the free space below.

What do (or did) you study in college? ....

What is (or was/were) your occupation(s)? ....

Are you a Graduate/Undergraduate Student at IU? **G** **U** [in my __________ year of study]

13. If you know **any languages other than your native language**, list the language(s) and indicate how you learned these languages and at what age you first started to learn them. Please include English if it is not your native language. Also estimate your ability to speak spontaneously, understand, read and write each language on a scale from "1" (i.e., your ability is very poor) to "11" (i.e., your ability is very good), by writing the number in the box.

Language or Dialect:					
How did you learn these languages ? *(more than one response possible, place an "x" or a number in the boxes below:)*					
At home					
At school					
At a language-school					
At the university					
Living in the country					
Through professional or private contacts					
Other (please specify)					
Estimate your ability to **speak** this language spontaneously *(none* 1 2 3 4 5 6 7 8 9 10 11 *perfectly)*					
Estimate your ability to **understand** spoken language *(none* 1 2 3 4 5 6 7 8 9 10 11 *perfectly)*					
Estimate your ability to **read** this language *(none* 1 2 3 4 5 6 7 8 9 10 11 *perfectly)*					
Estimate your ability to **write** this language *(none* 1 2 3 4 5 6 7 8 9 10 11 *perfectly)*					
Estimate how strong **your accent** is in this language *(very strong* 1 2 3 4 5 6 7 8 9 10 11 *no accent)*					
First *<u>exposure</u>* to this language (spoken): (age)					
First *<u>use</u>* of this language (spoken) : (age)					

If you need more space, please use the free space above or below.

14. Have you ever had any kind of a speech or hearing disorder? Yes No

If "Yes", please explain: ....
...

15. Please list the places you have lived more than 6 months in chronological order:

- **Where I have lived (earliest to most recent): From (year) to (year):**

1 ....

2 ....

3 ....

4 ....

If you need more space, please use the free space below.

16. Do you take part in any musical activities? **Yes No**
(for example, do you play an instrument, sing in a choir, etc.)
……………………………………………………………………………

17. Have you taken any of the following classes at Indiana University?

T101 Pronunciation Yes No

T101 Speaking Fluency Development Yes No

T101 Academic Listening Yes No

Any linguistics/SLS/phonetics classes Yes No Which : ……………………………………………..

H. Participant Background in Current Study

L1 Thai participants (n=9)

code	age	gender	place of birth	languages/ dialects	overseas experience	occupation	studies	musical ability	right-/ left-handed
T043	39	F	Bangkok	Isaan dialect, E, J		Banquet supervisor	Communication Arts/Tourism Management	No	R
T044	29	F	Phichit	E		Teacher	Sports science	No	R
T045	23	F	Samutprakarn	E, C		Student	Law	No	R
T078	41	F	Bangkok	E		University lecturer	Nursing Science, Exercise Physiology	No	R
T079	41	F	Saraburi	E		University lecturer	Nursing Science, Counseling Psychology	Yes	R
T082	26	M	Bangkok	Isaan/Southern/Northern dialects, E	Oxford, UK	University lecturer	Law	No	R
T113	25	F	Khuankhanoon	Southern dialect, E	NJ, CA	Student	Mathematics education	Yes, guitar, drums	R
T124	26	F	Bangkok (Rayong)	E		Student	Law	No	R
T126	31	F	Ubon Ratchatani	Isaan dialect, E, V, C, F		Student (Lecturer previously)	Education (Folklore)	No	R

*Shaded area indicates this subject was cut from the final analysis as she varied from the general target group.

L1 English participants (n=25)

code	age	gender	place of birth	languages/ dialects	overseas experience (6 months+)	occupation	studies	musical ability	right-/ left-handed
E006	25	M	Fort Wayne, IN	SP		English teacher	Linguistics, Spanish	No	R

E013	34	M	Franklin, IN	SP, G	Seoul, Korea; Baguio, Philippines; Jochiwon, Chungcheng, Korea	machinist	General Studies	Yes	R
E021	48	F	San Francisco, CA	SP, F	Paris, France	ESL instructor	Psychology	No	L
E023	29	F	Monterey, CA	F, Italian	Stuttgart, Germany	Student	French, Linguistics	Yes	R
E028	29	F	Minneapolis, MN	F		Student, Teacher, Assessment	Second language studies	Yes	R
E029	32	F	Minneapolis, MN	F		Artist, Floral design	Theater Arts, Dance	Yes	R
E031	27	M	Salem, MA	G, J, SW, Czech		ESL lecturer	German (BA), TESOL & Applied Linguistics (MA)	Yes	R
E037	29	M	San Juan, Puerto Rico	SP	Vigo, Spain	ESL instructor	Spanish, English (BA), TESOL & Applied Linguistics (MA)	No	R
E053	25	F	Overland Park, KS	SP	Queretaro, Mexico	English teacher	Linguistics	No	R
E056	33	M	Hartford, CT	Russian, Ukrainian	Ukraine	Teacher	Political Science, Russian, TESOL & Applied Linguistics	Yes	R
E057	41	M	Fargo, ND	SP, Polish, F, Turkish	Ireland, Poland, Argentina, Colombia	Graduate student	History, Linguistics	Yes	R
EO60	23	M	Laramie, WY	SP, Modern Hebrew, Arabic, Portuguese		Student	Hispanic Linguistics, Management, Hebrew, Economics/Political Science (BA), Second language studies (MA)	No	R
E061	24	F	Fort Wayne, IN	SP		Developmental therapist	Speech-Language pathology	No	R
E062	25	F	Dayton, OH	SP, F	Santiago, Chile	Student	Linguistics, Spanish	Yes	R
E073	23	F	West Lafayette, IN	F, Arabic (MSA, Jordanian)		ESL teacher	TESOL, Arabic	Yes	R

E074	19	F	Columbus, IN	F, G		Student, Sales associate	Journalism, International studies, Linguistics	No	R
E081	24	F	Richmond, VA	G, SP		Student, ESL teacher	Speech Pathology & Audiology, Linguistics	No	L
E131	28	M	Concord, NH	SP		ESL instructor	Cross-Cultural studies	Yes	R
E132	43	F	Camden, NJ	G, SP, F, SW		Teacher	German, Spanish	No	R
E133	25	M	Plymouth, MA	Italian	Milan, Italy	Student	Romance languages and literatures	No	L
E134	50	F	San Jose, CA	Danish, SP, F	Copenhagen, Denmark, Ouagadougou, Burkina Faso	ESL lecturer	Art	No	R
E136	18	F	Columbus, OH	F		Student	Business	No	L
E137	29	M	Hammond, IN	SP, Mandarin Chinese	Madrid/La Rioja/Basque country/Barcelona, Spain, Shanghai, PRC	English teacher	Biology, Spanish	No	R
K138	44	F	Alexandria, VA	G	US, Australia	University administrator	Biology	No	R
E139	49	F	Grand Forks, ND	F		Academic advisor	Education		L

*Shaded area indicates this subject was cut from the final analysis as she varied from the general target group.

L1 Japanese participants (n=23)

code	age	gender	place of birth	languages/ dialects	overseas experience	occupation	studies	musical ability	right-/ left-handed
J002	36	F	Tokyo	Kansai dialect, E, F	US	Language instructor (English, Japanese)	English, Linguistics	Yes	R
J009	33	M	Kasugai, Aichi	Nagoya dialect, Osaka dialect, E	US	Student	Finance, Linguistics	No	R
J016	31	F	Togane, Chiba	Kansai dialect, E	US, Canada	Student	Linguistics	No	R
J017	23	F	Hokota, Ibaraki	Dialects, E	US	Student	English education	No	R

J018	21	F	Date, Hokkaido	Hokkaido dialect, E, F	US	Student	English language	Yes	R
J026	27	F	Hamamatsu, Shizuoka	Dialects, E	US	Part-time job	Education	Yes	R
J027	Not given	F	Fukuoka	Dialect, E, K	US	Student	Business (accounting)	Yes	R
J034	37	M	Okazaki, Aichi	Mikawa dialect, all dialects, E	US	Consultant	Political Science	No	R
J040	29	F	Syunan-city, Yamaguchi	Osaka dialect, E	US	Graduate student	African-American & African Diaspora Studies	Yes	R
J076	33	M	Kasumigaura, Ibaraki	Dialects understand, E, G	US	Bank	Economics	No	R
J077	23	M	Tokyo	Dialects understand, E	US	Student	Economics	No	R
J085	28	M	Tahara, Aichi	Osaka/Nagoya/Mikawa dialects, E	US	Accountant	Accounting	No	R
J097	21	F	Sapporo, Hokkaido	Hokkaido dialect, E, P, C	US	Student	English, Business	No	R
J099	30	M	Osaka	Osaka dialect, E	US	Engineer, Researcher	Business	No	R
J101	29	M	Shizuoka	Osaka/ Hiroshima dialects, E, SW	US	Kindergarten teacher, cram school teacher	Swahili, African culture	Yes	R
J102	32	M	Chita, Aichi	Nagoya dialect, E	US	Office worker	Sociology	Yes	R
J104	33	M	Saitama	Kyushu/Hokkaido/ Osaka/Nagoya dialects, E	US	Tax, Accounting	Business administration	No	R
J105	31	F	To-on City, Ehime	Ehime dialect, E, M	US, PRC	Student	English/American literature	Yes	R
J107	21	M	Izumi, Kagoshima	Kagoshima/Osaka dialects, E	US	Student	English/American history	No	R
J116	18	M	Ojiya, Niigata	Dialects, Tagalog, E	US, Philippines	Student	IEP	Yes	R
J127	33	F	Yokohama, Kanagawa	Dialects, E	US	Human resources, translation	Education, city planning, nonprofit management	No	R

J129	27	M	Yokohama, Kanagawa	E, KM	US, Uganda	Student	International studies	No	R
J130	26	F	Tokyo	Osaka/other dialect, E	US	Student	Educational technology	No	R

*Shaded area indicates this subject was cut from the final analysis as she varied from the general target group.

L1 Korean participants (n=28)

code	age	gender	place of birth	languages/ dialects	overseas experience	occupation	studies	musical ability	right-/ left-handed
K001	26	M	Seoul	Dialects, E,C, J	US	student	International business management	Yes	R
K008	28	F	Seoul	E, J	US, Australia	Middle school English teacher	Computer Science Education	Yes	R
K014	23	M	Seoul	E, C	US	Student	Economics	Yes	R
K019	37	F	Seoul	J, G	US, Canada (when 3 yrs old)	Graduate student	German, English	Yes	R
K020	34	F	Incheon	E, F, J	US	Graduate student	French, Literature, English	No	R
K022	27	F	Seoul	E	US	Graduate student	Education (Higher education, student affairs)	No	R
K038	22	M	Buchon	Geranamdo, E	US	student	Economics and Public Policies	Yes	R
K042	30	F	Seoul	J	US	Business owner (trader)	Japanese language and literature	Yes	R
K048	39	F	Seoul	E	US, Canada	Official	Atmospheric science	No	R
K050	31	F	North Namwon, North Cholla	E	US	Teacher	Education, Social sciences	Yes	R
K084	32	F	Seoul	Kyungsang dialect, E, J, M	US, Australia	Government official	Geology	No	R
K086	26	M	Kyungsangnamd o	Kyungsang dialect, E, J	US, (US airbase in Korea)	Student	Computer science	Yes	R
K088	39	M	Busan, Kyungsang	E, G	US	"salaryman"	Economics	No	R

K089	32	F	Seoul	E, F	US	Marketing employee	Communication	No answer	R
K091	32	F	Seoul	E, F	US, Canada, Hong Kong	Graduate student	Business administration	Yes	R
K092	32	F	Daegu	Daegu dialect, J	US	English language instructor	English language and literature	No	R
K093	39	F	Seoul	Kyungsang dialect, E	US	Student	Economics	Yes	R
K094	41	F	Seoul	Kyungsang dialect, E	US	Stock (investment) company	International trade	Yes	R
K095	32	F	Seoul	E, C	US	Public officer	Chinese Culture	No	R
K096	54	F	Yangpyeong, Gyeongki	Only Korean	US	Store owner	--	No	R
K103	34	F	Busan	E	US, New Zealand	Graduate student	English literature & language	No	R
K108	24	M	Anyang, Kyeong-gido	Jeju/Kyungsang dialects, E	US, Australia	Student	Sports	No	Both
K110	26	M	Daejeon	Chungchang dialect, E, C	US, PRC	Student	International trade IEP	Yes	R
K111	20	M	Gwang-ju	Junla dialect, E	US	Student	Industrial engineering	No	L
K112	21	M	Seoul	E	US	Student	Business, computer engineering	No	R
K114	22	M	Busan, Kyungsang	Kyungsang dialect, E	US	Student	IEP??	No	Both
K125	20	F	Tokyo	E	US, Canada	Student	Physics	Yes	L
K128	27	F	Seoul	E, S, J	US, Canada	Student	Studio art (sciences for a few years)	No	R

*Shaded area indicates this subject was cut from the final analysis as she varied from the general target group.

L1 Mandarin participants (n=31)

code	age	gender	place of birth	languages/ dialects	overseas experience	occupation	studies	musical ability	right-/ left-handed

M003	36	F	Kaohsiung, ROC	Taiwanese Mandarin, E, TW, Hakka, Finnish, SW	US, Finland, Sweden	Student, teacher, interpreter	Journalism, Business administration	Yes	R
M004	36	M	Chia-Yi, ROC	Taiwanese Mandarin, E, F, TW, J	US	Student, teacher	English literature	Yes	R
M007	24	F	Changchuan, Jilin, PRC	Taiwanese Mandarin, E, F	US	Student	English translation	No	R
M010	19	F	Yantai, Shandong, PRC	Yantai dialect, E	US	Student	Business	Yes	L, Both
M011	19	F	Ningbo, PRC	Ningbo dialect, E	US	Student	Business	Yes	R
M012	19	F	Zhejiang, PRC	Dongyang dialect, E	US	Student	Biology	No	R
M015	21	F	Handan, Hebei, PRC	Dialects, E	US	Student	Psychology	Yes	R
M024	24	M	Taitung, ROC	Taiwanese Mandarin, E, TW, I	US	Student	Music	Yes	R
M030	21	F	Xiangtan, Hunan, PRC	Hunan dialect, E, J	US	Student	TESOL	No	R
M033	18	M	Changde, Hunan, PRC	Changde dialect, E, J	US	Student	Finance	Yes	R
M036	24	F	Maanshan, Anhui, PRC	E, G	US	English teacher	English	Yes	R
M041	22	F	Shanghai, PRC	Shanghaiese, E	US	Student	Finance	Yes	R
M046	19	F	Chengdu, Sichuan, PRC	Sichuan dialect, E	US	Student	Accounting, Finance	No	R
M047	24	F	Dalian, Liaoning, PRC	Dalian dialect, E, K	US	Student	Mathematics	Yes	R
M049	19	F	Shenzhen, Guangdon, PRC	Inner Mongolian dialect, E	US	Student	Accounting, Finance	Yes	R
M051	22	F	Fuyang, Zhejiang, PRC	Fuyang dialect, E	US	Student	Accounting, Finance	Yes	R
M052	20	F	Shanghai, PRC	Shanghaiese, E, J	US	Student	Accounting	No	R

M055	30	F	Kaohsiung, ROC	TW, Hakka, E, J	US	Teacher	Political Science, Foreign languages & literatures, teacher education	No	R
M058	31	F	Chia-Yi, ROC	Taiwanese Mandarin, E, TW	US	Researcher	Public Health	No	R
M059	37	M	Taipei, ROC	Taiwanese Mandarin, E, TW, J	US	Server	Electrical engineering	No	R
M063	36	F	Taipei, ROC	TaiwaneseMandarin, E, TW, J	US	Product manager, Interaction designer	Industrial design	No	R
M065	28	F	Beijing, PRC	Beijing dialect, Shanghaiese, E, G	US	Banker	English	No	Both, R dominant
M066	33	F	Dezhou, Shandong, PRC	Shandong dialect, Shanghaiese, E	US	Product management	Computer science technology/Communications	No	R
M067	32	M	Hsinchu, ROC	Taiwanese Mandarin, E, TW, J	US	Student	Chemistry	Yes	R
M068	32	F	Taipei, ROC	Taiwanese Mandarin, E, TW, J	US	Student	Biology	Yes	R
M069	28	F	Chengdu, Sichuan, PRC	Sichuan Mandarin, E	US	Civil servant	Law	No	R
M070	25	F	Zunyi, Guizhou, PRC	Guizhou Mandarin, E	US	Legal English editor, administrative assistant	Law	No	R
M071	22	F	Beijing, PRC	Beijing Mandarin, Jiangsu Mandarin, Cantonese, E, F, J	US	Student	Law	Yes	R
M072	23	F	Changsha, Hunan, PRC	Xiang Mandarin, E	US	Student	Law	Yes	R
M075	19	F	Taipei, ROC	TW, E, G	US	Student	Special education/pre-occupational therapy	No	R
M090	34	M	Taipei, ROC	Taiwanese Mandarin, E, TW, J, G, R	US	Graduate student	Russian, Linguistics	No	R

*Shaded area indicates this subject was cut from the final analysis as she varied from the general target group.

L2 Mandarin Learners/L1 English participants (n=23)

code	age	gender	place of birth	languages	overseas experience	occupation	studies	musical ability	right-/ left-handed
L005	21	F	Fort Wayne, IN	SP, R, F, G, M		Student	Psychology, Mandarin Chinese	Yes	R
L025	21	M	Elmhurst, IL	SP, GK, M		Student	Supply chain, International business, Mandarin Chinese	Yes	R
L032	28	M	Arlington, VA	SP, Polish, M, Malay	PRC	Graduate student, teacher	International relations, Mandarin Chinese	Yes	R
L035	26	M	Roue, GA	SP, M, G, F	ROC	Graduate student	East Asian studies	Yes	R
L039	26	M	Dayton, OH	SP, M	PRC	Graduate student, English teacher	Linguistics	Yes	R
L054	27	F	Silver Spring, MD	M	ROC	Graduate student, Chinese Mandarin instructor	Mandarin Chinese	Yes	L (write and use fork with left, other right)
L064	53	M	Rochester, NY	M, F, G	France	Computer technician	Physics, Astronomy	Yes	R
L080	21	M	Scottsdale, AZ	M		Student	Mandarin Chinese	Yes	R
L083	19	F	Nanning, Guangxi, PRC	F, M	PRC	Student	Mandarin Chinese, International studies, Sociology	Yes	R
L087	34	M	DC	SP, Latin, G, F, M		Graduate student, Software developer, music teacher	Computer science (BS), Music theory (MA)	Yes	R
L098	18	F	Lafayette, LA	SP, M		Student	Business	No	R
L100	27	M	Madison, WI	M	ROC, PRC	Graduate student, Editor, Teaching assistant	International studies, Asian studies	No	L
L106	32	F	Eden Prairie, MN	M, Polish		Academic Advisor, Chinese flagship coordinator	Music, Asian studies (BA), Library science, Ethnomusicology (MA)	Yes	R

L109	44	M	Akron, OH	M, Latin	PRC	Higher Education Administration	English literature, Music (BA), Mandarin Chinese (graduate)	Yes	R
L115	20	F	Nanjing, Jiangsu, PRC	E, K	PRC	Student	East Asian Studies, Mandarin Chinese	Yes	R
L117	19	F	Xiamen, Fujian, PRC	SP, M		Student, retail sales associate, service writer assistant	East Asian studies	Yes	R
L118	21	M	Mountain View, CA	SP, M		Student	Finance, Mandarin Chinese	Yes	L
L119	23	F	Winston Salem, NC	M	PRC	Graduate student, Researcher	Sociology	Yes	R
L120	18	M	Chicago, IL	M		Student	Business	Yes	L (play sports with right)
L121	18	F	Chicago, IL	SP, M	PRC	Student	Biology, Chemistry, Mandarin Chinese	Yes	R
L122	20	F	Summit, NJ	SP,M		Student	Linguistics, Mandarin Chinese	Yes	R
L123	21	M	Bloomington, IN	M		Student	English, Lingustics, Mandarin Chinese	No	R (write with left, throw with left)
L135	19	M	Cincinnati, OH	S, M	PRC	Student	Business, Mandarin Chinese	Yes	R

*Shaded area indicates this subject was cut from the final analysis as she varied from the general target group.

**E = English, SP = Spanish, F = French, G = German, M = Mandarin, J = Japanese, K = Korean, TW = Taiwanese, SW = Swedish, V = Vietnamese, C = Chinese, R = right, L = left, ROC = Republic of China (Taiwan), PRC = People's Republic of China (Mainland China), US = United States, IEP = Intensive English Program (pre-matriculated university students)

I. DMDX Script used in monosyllabic AXB task in Current Study

```
<ep> <azk> <NumberOfItems 160> <scramble 40> <ContinuousRun> <Delay 158>
<FrameDuration 250> <Timeout 2500> <id "Keyboard"> <mr +Space>
<MapNegativeResponse "+Right Alt"> <MapPositiveResponse "+Left Alt"> <vm
1280,1024,1024,32,60> <eop>


$
0 <line -8> "ABX Instructions",<line -6> "You will hear 3 pseudo words in
a row.", <line -4> "You have to  decide whether", <line -2> "the third one is
similar to the first or the second word.",<line 0> "Press the LEFT Alt A button if the
third one is the same as the first word",<line 2> "Press the RIGHT Alt B button if the
third one is the same as second word",<line 4> "Answer as quickly as
possible.",<line 8> "Press SPACEBAR to start with a short practice.";

999 <ms% 2000> "+"  /;
$

+100000001<cfb " CORRECT"><wfb " INCORRECT"><tlfb "Too
slow"><ms% 500> "+" / <wav 2> "1MCwuiM" / <ms% 500> / <wav 2> "1MCpuiM" / <ms% 500> /
<wav 2> "2MCwuiM" * /;

-200000002<cfb " CORRECT"><wfb " INCORRECT"><tlfb "Too
slow"><ms% 500> "+" / <wav 2> "1MCwuiR" / <ms% 500> / <wav 2> "1MCpuiR" / <ms% 500> /
<wav 2> "2MCpuiR" * /;

-200000003<cfb " CORRECT"><wfb " INCORRECT"><tlfb "Too
slow"><ms% 500> "+" / <wav 2> "1MCpuiF" / <ms% 500> / <wav 2> "1MCwuiF" / <ms% 500> /
<wav 2> "2MCwuiF" * /;


+100000004<cfb " CORRECT"><wfb " INCORRECT"><tlfb "Too
slow"><ms% 500> "+" / <wav 2> "1MCpuiL" / <ms% 500> / <wav 2> "1MCwuiL" / <ms% 500> /
<wav 2> "2MCpuiL" *
/;

+100000005<cfb " CORRECT"><wfb " INCORRECT"><tlfb "Too
slow"><ms% 500> "+" / <wav 2> "1MCwuiH" / <ms% 500> / <wav 2> "1MCpuiH" / <ms% 500> /
<wav 2> "2MCwuiH" * /;

+100000006<cfb " CORRECT "><wfb " INCORRECT "><tlfb "Too
slow"><ms% 500> "+" / <wav 2> "1MTduaM" / <ms% 500> / <wav 2> "1MVnginM" / <ms% 500> /
<wav 2> "2MTduaM" * /;

-200000007<cfb " CORRECT "><wfb " INCORRECT "><tlfb "Too
slow"><ms% 500> "+" / <wav 2> "1MTduaR" / <ms% 500> / <wav 2> "1MVnginR" / <ms% 500> /
<wav 2> "2MVnginR" * /;

-200000008<cfb " CORRECT "><wfb " INCORRECT "><tlfb "Too
```

```
slow"><ms% 500> "+" / <wav 2> "1MVnginF" / <ms% 500> / <wav 2> "1MTduaF" / <ms% 500> /
<wav 2> "2MTduaF" * /;

+100000009<cfb " CORRECT "><wfb " INCORRECT "><tlfb "Too
slow"><ms% 500> "+" / <wav 2> "1MVnginL" / <ms% 500> / <wav 2> "1MTduaL" / <ms% 500> /
<wav 2> "2MVnginL" *
/;

-200000010<cfb " CORRECT "><wfb " INCORRECT "><tlfb "Too
slow"><ms% 500> "+" / <wav 2> "1MVnginH" / <ms% 500> / <wav 2> "1MTduaH" / <ms% 500> /
<wav 2> "2MTduaH" * /;

$
0 <line -3> "Ready for the real experiment? The feedback is now turned off",
<line -1> "Respond as fast as you can without making mistakes", <line 1> "If you
make a mistake, don't worry and keep going!", <line 4>"...press SPACEBAR when
ready";

999 <nfb> <ms% 2000>  /;
$

\
+111512111 <ms% 350> "+" / <wav 2> "1MTnoiR" / <ms% 500> / <wav 2>
"1MTnoiF" / <ms% 500> / <wav 2> "2MTnoiR" * /;
+111521212 <ms% 350> "+" / <wav 2> "1MTnoiF" / <ms% 500> / <wav 2>
"1MTnoiH" / <ms% 500> / <wav 2> "2MTnoiF" * /;
-211521113 <ms% 350> "+" / <wav 2> "1MTnoiR" / <ms% 500> / <wav 2>
"1MTnoiH" / <ms% 500> / <wav 2> "2MTnoiH" * /;
-211512214 <ms% 350> "+" / <wav 2> "1MTnoiL" / <ms% 500> / <wav 2>
"1MTnoiF" / <ms% 500> / <wav 2> "2MTnoiF" * /;
+111512115 <ms% 350> "+" / <wav 2> "1MTnoiL" / <ms% 500> / <wav 2>
"1MTnoiH" / <ms% 500> / <wav 2> "2MTnoiL" * /;
+111521216 <ms% 350> "+" / <wav 2> "1MTnoiM" / <ms% 500> / <wav 2>
"1MTnoiL" / <ms% 500> / <wav 2> "2MTnoiM" * /;
-211521117 <ms% 350> "+" / <wav 2> "1MTnoiR" / <ms% 500> / <wav 2>
"1MTnoiL" / <ms% 500> / <wav 2> "2MTnoiL" * /;
-211512218 <ms% 350> "+" / <wav 2> "1MTnoiM" / <ms% 500> / <wav 2>
"1MTnoiF" / <ms% 500> / <wav 2> "2MTnoiF" * /;
+111512119 <ms% 350> "+" / <wav 2> "1MTnoiM" / <ms% 500> / <wav 2>
"1MTnoiH" / <ms% 500> / <wav 2> "2MTnoiM" * /;
+111521220 <ms% 350> "+" / <wav 2> "1MTnoiR" / <ms% 500> / <wav 2>
"1MTnoiM" / <ms% 500> / <wav 2> "2MTnoiR" * /;
+121521211 <ms% 350> "+" / <wav 2> "1MTpuaiF" / <ms% 500> / <wav 2>
"1MTpuaiR" / <ms% 500> / <wav 2> "2MTpuaiF" * /;
-221521112 <ms% 350> "+" / <wav 2> "1MTpuaiF" / <ms% 500> / <wav 2>
"1MTpuaiH" / <ms% 500> / <wav 2> "2MTpuaiH" * /;
-221512213 <ms% 350> "+" / <wav 2> "1MTpuaiH" / <ms% 500> / <wav 2>
"1MTpuaiR" / <ms% 500> / <wav 2> "2MTpuaiR" * /;
+121512114 <ms% 350> "+" / <wav 2> "1MTpuaiL" / <ms% 500> / <wav 2>
"1MTpuaiF" / <ms% 500> / <wav 2> "2MTpuaiL" * /;
+121521215 <ms% 350> "+" / <wav 2> "1MTpuaiH" / <ms% 500> / <wav 2>
"1MTpuaiL" / <ms% 500> / <wav 2> "2MTpuaiH" * /;
-221521116 <ms% 350> "+" / <wav 2> "1MTpuaiM" / <ms% 500> / <wav 2>
"1MTpuaiL" / <ms% 500> / <wav 2> "2MTpuaiL" * /;
-221512217 <ms% 350> "+" / <wav 2> "1MTpuaiL" / <ms% 500> / <wav 2>
```

```
"1MTpuaiR" / <ms% 500> / <wav 2> "2MTpuaiR" * /;
-221521118 <ms% 350> "+" / <wav 2> "1MTpuaiF" / <ms% 500> / <wav 2>
"1MTpuaiM" / <ms% 500> / <wav 2> "2MTpuaiM" * /;
-221512219 <ms% 350> "+" / <wav 2> "1MTpuaiM" / <ms% 500> / <wav 2>
"1MTpuaiH" / <ms% 500> / <wav 2> "2MTpuaiH" * /;
+121512120 <ms% 350> "+" / <wav 2> "1MTpuaiM" / <ms% 500> / <wav 2>
"1MTpuaiR" / <ms% 500> / <wav 2> "2MTpuaiM" * /;
+151812135 <ms% 350> "+" / <wav 2> "1MCbpeoF" / <ms% 500> / <wav 2>
"1MCdteoF" / <ms% 500> / <wav 2> "2MCbpeoF" * /;
+151821236 <ms% 350> "+" / <wav 2> "1MCdteoH" / <ms% 500> / <wav 2>
"1MCbpeoH" / <ms% 500> / <wav 2> "2MCdteoH" * /;
-251821137 <ms% 350> "+" / <wav 2> "1MCdteoL" / <ms% 500> / <wav 2>
"1MCbpeoL" / <ms% 500> / <wav 2> "2MCbpeoL" * /;
-251812238 <ms% 350> "+" / <wav 2> "1MCbpeoM" / <ms% 500> / <wav 2>
"1MCdteoM" / <ms% 500> / <wav 2> "2MCdteoM" * /;
+151812139 <ms% 350> "+" / <wav 2> "1MCbpeoR" / <ms% 500> / <wav 2>
"1MCdteoR" / <ms% 500> / <wav 2> "2MCbpeoR" * /;
+161821235 <ms% 350> "+" / <wav 2> "1MCtiaF" / <ms% 500> / <wav 2>
"1MCwiaF" / <ms% 500> / <wav 2> "2MCtiaF" * /;
-261821136 <ms% 350> "+" / <wav 2> "1MCtiaH" / <ms% 500> / <wav 2>
"1MCwiaH" / <ms% 500> / <wav 2> "2MCwiaH" * /;
-261812237 <ms% 350> "+" / <wav 2> "1MCwiaL" / <ms% 500> / <wav 2>
"1MCtiaL" / <ms% 500> / <wav 2> "2MCtiaL" * /;
+161812138 <ms% 350> "+" / <wav 2> "1MCwiaM" / <ms% 500> / <wav 2>
"1MCtiaM" / <ms% 500> / <wav 2> "2MCwiaM" * /;
-261821139 <ms% 350> "+" / <wav 2> "1MCtiaR" / <ms% 500> / <wav 2>
"1MCwiaR" / <ms% 500> / <wav 2> "2MCwiaR" * /;
-281812235 <ms% 350> "+" / <wav 2> "1MVuaF" / <ms% 500> / <wav 2>
"1MViaF" / <ms% 500> / <wav 2> "2MViaF" * /;
+181812136 <ms% 350> "+" / <wav 2> "1MVuaH" / <ms% 500> / <wav 2>
"1MViaH" / <ms% 500> / <wav 2> "2MVuaH" * /;
+181821237 <ms% 350> "+" / <wav 2> "1MViaL" / <ms% 500> / <wav 2>
"1MVuaL" / <ms% 500> / <wav 2> "2MViaL" * /;
-281821138 <ms% 350> "+" / <wav 2> "1MViaM" / <ms% 500> / <wav 2>
"1MVuaM" / <ms% 500> / <wav 2> "2MVuaM" * /;
-281812239 <ms% 350> "+" / <wav 2> "1MVuaR" / <ms% 500> / <wav 2>
"1MViaR" / <ms% 500> / <wav 2> "2MViaR" * /;
+191821235 <ms% 350> "+" / <wav 2> "1MVriaF" / <ms% 500> / <wav 2>
"1MVreuiF" / <ms% 500> / <wav 2> "2MVriaF" * /;
-291821136 <ms% 350> "+" / <wav 2> "1MVriaH" / <ms% 500> / <wav 2>
"1MVreuiH" / <ms% 500> / <wav 2> "2MVreuiH" * /;
-291812237 <ms% 350> "+" / <wav 2> "1MVreuiL" / <ms% 500> / <wav 2>
"1MVriaL" / <ms% 500> / <wav 2> "2MVriaL" * /;
+191812138 <ms% 350> "+" / <wav 2> "1MVreuiM" / <ms% 500> / <wav 2>
"1MVriaM" / <ms% 500> / <wav 2> "2MVreuiM" * /;
+191821239 <ms% 350> "+" / <wav 2> "1MVriaR" / <ms% 500> / <wav 2>
"1MVreuiR" / <ms% 500> / <wav 2> "2MVriaR" * /;
$
0 <line -2> "Take a break. Press SPACEBAR when ready.";
$
-212512211 <ms% 350> "+" / <wav 2> "1MTnoiR" / <ms% 500> / <wav 2>
"1MTnoiF" / <ms% 500> / <wav 2> "2MTnoiF" * /;
+112512112 <ms% 350> "+" / <wav 2> "1MTnoiH" / <ms% 500> / <wav 2>
"1MTnoiF" / <ms% 500> / <wav 2> "2MTnoiH" * /;
+112521213 <ms% 350> "+" / <wav 2> "1MTnoiR" / <ms% 500> / <wav 2>
"1MTnoiH" / <ms% 500> / <wav 2> "2MTnoiR" * /;
-212521114 <ms% 350> "+" / <wav 2> "1MTnoiF" / <ms% 500> / <wav 2>
```

```
"1MTnoiL" / <ms% 500> / <wav 2> "2MTnoiL" * /;
-212512215 <ms% 350> "+" / <wav 2> "1MTnoiL" / <ms% 500> / <wav 2>
"1MTnoiH" / <ms% 500> / <wav 2> "2MTnoiH" * /;
+112512116 <ms% 350> "+" / <wav 2> "1MTnoiL" / <ms% 500> / <wav 2>
"1MTnoiM" / <ms% 500> / <wav 2> "2MTnoiL" * /;
+112521217 <ms% 350> "+" / <wav 2> "1MTnoiR" / <ms% 500> / <wav 2>
"1MTnoiL" / <ms% 500> / <wav 2> "2MTnoiR" * /;
-212521118 <ms% 350> "+" / <wav 2> "1MTnoiF" / <ms% 500> / <wav 2>
"1MTnoiM" / <ms% 500> / <wav 2> "2MTnoiM" * /;
-212512219 <ms% 350> "+" / <wav 2> "1MTnoiM" / <ms% 500> / <wav 2>
"1MTnoiH" / <ms% 500> / <wav 2> "2MTnoiH" * /;
+112512120 <ms% 350> "+" / <wav 2> "1MTnoiM" / <ms% 500> / <wav 2>
"1MTnoiR" / <ms% 500> / <wav 2> "2MTnoiM" * /;
+122512111 <ms% 350> "+" / <wav 2> "1MTpuaiR" / <ms% 500> / <wav 2>
"1MTpuaiF" / <ms% 500> / <wav 2> "2MTpuaiR" * /;
+122521212 <ms% 350> "+" / <wav 2> "1MTpuaiF" / <ms% 500> / <wav 2>
"1MTpuaiH" / <ms% 500> / <wav 2> "2MTpuaiF" * /;
-222521113 <ms% 350> "+" / <wav 2> "1MTpuaiR" / <ms% 500> / <wav 2>
"1MTpuaiH" / <ms% 500> / <wav 2> "2MTpuaiH" * /;
-222512214 <ms% 350> "+" / <wav 2> "1MTpuaiL" / <ms% 500> / <wav 2>
"1MTpuaiF" / <ms% 500> / <wav 2> "2MTpuaiF" * /;
+122512115 <ms% 350> "+" / <wav 2> "1MTpuaiL" / <ms% 500> / <wav 2>
"1MTpuaiH" / <ms% 500> / <wav 2> "2MTpuaiL" * /;
+122521216 <ms% 350> "+" / <wav 2> "1MTpuaiM" / <ms% 500> / <wav 2>
"1MTpuaiL" / <ms% 500> / <wav 2> "2MTpuaiM" * /;
-222521117 <ms% 350> "+" / <wav 2> "1MTpuaiR" / <ms% 500> / <wav 2>
"1MTpuaiL" / <ms% 500> / <wav 2> "2MTpuaiL" * /;
+122521218 <ms% 350> "+" / <wav 2> "1MTpuaiF" / <ms% 500> / <wav 2>
"1MTpuaiM" / <ms% 500> / <wav 2> "2MTpuaiF" * /;
-222521119 <ms% 350> "+" / <wav 2> "1MTpuaiH" / <ms% 500> / <wav 2>
"1MTpuaiM" / <ms% 500> / <wav 2> "2MTpuaiM" * /;
-222512220 <ms% 350> "+" / <wav 2> "1MTpuaiM" / <ms% 500> / <wav 2>
"1MTpuaiR" / <ms% 500> / <wav 2> "2MTpuaiR" * /;
-252812235 <ms% 350> "+" / <wav 2> "1MCbpeoF" / <ms% 500> / <wav 2>
"1MCdteoF" / <ms% 500> / <wav 2> "2MCdteoF" * /;
+152812136 <ms% 350> "+" / <wav 2> "1MCbpeoH" / <ms% 500> / <wav 2>
"1MCdteoH" / <ms% 500> / <wav 2> "2MCbpeoH" * /;
+152821237 <ms% 350> "+" / <wav 2> "1MCdteoL" / <ms% 500> / <wav 2>
"1MCbpeoL" / <ms% 500> / <wav 2> "2MCdteoL" * /;
-252821138 <ms% 350> "+" / <wav 2> "1MCdteoM" / <ms% 500> / <wav 2>
"1MCbpeoM" / <ms% 500> / <wav 2> "2MCbpeoM" * /;
-252812239 <ms% 350> "+" / <wav 2> "1MCbpeoR" / <ms% 500> / <wav 2>
"1MCdteoR" / <ms% 500> / <wav 2> "2MCdteoR" * /;
+162812135 <ms% 350> "+" / <wav 2> "1MCwiaF" / <ms% 500> / <wav 2>
"1MCtiaF" / <ms% 500> / <wav 2> "2MCwiaF" * /;
+162821236 <ms% 350> "+" / <wav 2> "1MCtiaH" / <ms% 500> / <wav 2>
"1MCwiaH" / <ms% 500> / <wav 2> "2MCtiaH" * /;
-262821137 <ms% 350> "+" / <wav 2> "1MCtiaL" / <ms% 500> / <wav 2>
"1MCwiaL" / <ms% 500> / <wav 2> "2MCwiaL" * /;
-262812238 <ms% 350> "+" / <wav 2> "1MCwiaM" / <ms% 500> / <wav 2>
"1MCtiaM" / <ms% 500> / <wav 2> "2MCtiaM" * /;
+162821239 <ms% 350> "+" / <wav 2> "1MCtiaR" / <ms% 500> / <wav 2>
"1MCwiaR" / <ms% 500> / <wav 2> "2MCtiaR" * /;
-282821135 <ms% 350> "+" / <wav 2> "1MViaF" / <ms% 500> / <wav 2>
"1MVuaF" / <ms% 500> / <wav 2> "2MVuaF" * /;
-282812236 <ms% 350> "+" / <wav 2> "1MVuaH" / <ms% 500> / <wav 2>
"1MViaH" / <ms% 500> / <wav 2> "2MViaH" * /;
```

```
+182812137 <ms% 350> "+" / <wav 2> "1MVuaL" / <ms% 500> / <wav 2>
"1MViaL" / <ms% 500> / <wav 2> "2MVuaL" * /;
+182821238 <ms% 350> "+" / <wav 2> "1MViaM" / <ms% 500> / <wav 2>
"1MVuaM" / <ms% 500> / <wav 2> "2MViaM" * /;
-282821139 <ms% 350> "+" / <wav 2> "1MViaR" / <ms% 500> / <wav 2>
"1MVuaR" / <ms% 500> / <wav 2> "2MVuaR" * /;
+192812135 <ms% 350> "+" / <wav 2> "1MVreuiF" / <ms% 500> / <wav 2>
"1MVriaF" / <ms% 500> / <wav 2> "2MVreuiF" * /;
+192821236 <ms% 350> "+" / <wav 2> "1MVriaH" / <ms% 500> / <wav 2>
"1MVreuiH" / <ms% 500> / <wav 2> "2MVriaH" * /;
-292821137 <ms% 350> "+" / <wav 2> "1MVriaL" / <ms% 500> / <wav 2>
"1MVreuiL" / <ms% 500> / <wav 2> "2MVreuiL" * /;
-292812238 <ms% 350> "+" / <wav 2> "1MVreuiM" / <ms% 500> / <wav 2>
"1MVriaM" / <ms% 500> / <wav 2> "2MVriaM" * /;
+192812139 <ms% 350> "+" / <wav 2> "1MVreuiR" / <ms% 500> / <wav 2>
"1MVriaR" / <ms% 500> / <wav 2> "2MVreuiR" * /;
$
0 <line -2> "Take a break. Press SPACEBAR when ready.";
$
-213521111 <ms% 350> "+" / <wav 2> "1MTnoiF" / <ms% 500> / <wav 2>
"1MTnoiR" / <ms% 500> / <wav 2> "2MTnoiR" * /;
-213512212 <ms% 350> "+" / <wav 2> "1MTnoiH" / <ms% 500> / <wav 2>
"1MTnoiF" / <ms% 500> / <wav 2> "2MTnoiF" * /;
+113512113 <ms% 350> "+" / <wav 2> "1MTnoiH" / <ms% 500> / <wav 2>
"1MTnoiR" / <ms% 500> / <wav 2> "2MTnoiH" * /;
+113521214 <ms% 350> "+" / <wav 2> "1MTnoiF" / <ms% 500> / <wav 2>
"1MTnoiL" / <ms% 500> / <wav 2> "2MTnoiF" * /;
-213521115 <ms% 350> "+" / <wav 2> "1MTnoiH" / <ms% 500> / <wav 2>
"1MTnoiL" / <ms% 500> / <wav 2> "2MTnoiL" * /;
-213512216 <ms% 350> "+" / <wav 2> "1MTnoiL" / <ms% 500> / <wav 2>
"1MTnoiM" / <ms% 500> / <wav 2> "2MTnoiM" * /;
+113512117 <ms% 350> "+" / <wav 2> "1MTnoiL" / <ms% 500> / <wav 2>
"1MTnoiR" / <ms% 500> / <wav 2> "2MTnoiL" * /;
+113521218 <ms% 350> "+" / <wav 2> "1MTnoiF" / <ms% 500> / <wav 2>
"1MTnoiM" / <ms% 500> / <wav 2> "2MTnoiF" * /;
-213521119 <ms% 350> "+" / <wav 2> "1MTnoiH" / <ms% 500> / <wav 2>
"1MTnoiM" / <ms% 500> / <wav 2> "2MTnoiM" * /;
-213512220 <ms% 350> "+" / <wav 2> "1MTnoiM" / <ms% 500> / <wav 2>
"1MTnoiR" / <ms% 500> / <wav 2> "2MTnoiR" * /;
-223512211 <ms% 350> "+" / <wav 2> "1MTpuaiR" / <ms% 500> / <wav 2>
"1MTpuaiF" / <ms% 500> / <wav 2> "2MTpuaiF" * /;
+123512112 <ms% 350> "+" / <wav 2> "1MTpuaiH" / <ms% 500> / <wav 2>
"1MTpuaiF" / <ms% 500> / <wav 2> "2MTpuaiH" * /;
+123521213 <ms% 350> "+" / <wav 2> "1MTpuaiR" / <ms% 500> / <wav 2>
"1MTpuaiH" / <ms% 500> / <wav 2> "2MTpuaiR" * /;
-223521114 <ms% 350> "+" / <wav 2> "1MTpuaiF" / <ms% 500> / <wav 2>
"1MTpuaiL" / <ms% 500> / <wav 2> "2MTpuaiL" * /;
-223512215 <ms% 350> "+" / <wav 2> "1MTpuaiL" / <ms% 500> / <wav 2>
"1MTpuaiH" / <ms% 500> / <wav 2> "2MTpuaiH" * /;
+123512116 <ms% 350> "+" / <wav 2> "1MTpuaiL" / <ms% 500> / <wav 2>
"1MTpuaiM" / <ms% 500> / <wav 2> "2MTpuaiL" * /;
+123521217 <ms% 350> "+" / <wav 2> "1MTpuaiR" / <ms% 500> / <wav 2>
"1MTpuaiL" / <ms% 500> / <wav 2> "2MTpuaiR" * /;
+123512118 <ms% 350> "+" / <wav 2> "1MTpuaiM" / <ms% 500> / <wav 2>
"1MTpuaiF" / <ms% 500> / <wav 2> "2MTpuaiM" * /;
+123521219 <ms% 350> "+" / <wav 2> "1MTpuaiH" / <ms% 500> / <wav 2>
"1MTpuaiM" / <ms% 500> / <wav 2> "2MTpuaiH" * /;
```

```
-223521120 <ms% 350> "+" / <wav 2> "1MTpuaiR" / <ms% 500> / <wav 2>
"1MTpuaiM" / <ms% 500> / <wav 2> "2MTpuaiM" * /;
-253821135 <ms% 350> "+" / <wav 2> "1MCdteoF" / <ms% 500> / <wav 2>
"1MCbpeoF" / <ms% 500> / <wav 2> "2MCbpeoF" * /;
-253812236 <ms% 350> "+" / <wav 2> "1MCbpeoH" / <ms% 500> / <wav 2>
"1MCdteoH" / <ms% 500> / <wav 2> "2MCdteoH" * /;
+153812137 <ms% 350> "+" / <wav 2> "1MCbpeoL" / <ms% 500> / <wav 2>
"1MCdteoL" / <ms% 500> / <wav 2> "2MCbpeoL" * /;
+153821238 <ms% 350> "+" / <wav 2> "1MCdteoM" / <ms% 500> / <wav 2>
"1MCbpeoM" / <ms% 500> / <wav 2> "2MCdteoM" * /;
-253821139 <ms% 350> "+" / <wav 2> "1MCdteoR" / <ms% 500> / <wav 2>
"1MCbpeoR" / <ms% 500> / <wav 2> "2MCbpeoR" * /;
-263812235 <ms% 350> "+" / <wav 2> "1MCwiaF" / <ms% 500> / <wav 2>
"1MCtiaF" / <ms% 500> / <wav 2> "2MCtiaF" * /;
+163812136 <ms% 350> "+" / <wav 2> "1MCwiaH" / <ms% 500> / <wav 2>
"1MCtiaH" / <ms% 500> / <wav 2> "2MCwiaH" * /;
+163821237 <ms% 350> "+" / <wav 2> "1MCtiaL" / <ms% 500> / <wav 2>
"1MCwiaL" / <ms% 500> / <wav 2> "2MCtiaL" * /;
-263821138 <ms% 350> "+" / <wav 2> "1MCtiaM" / <ms% 500> / <wav 2>
"1MCwiaM" / <ms% 500> / <wav 2> "2MCwiaM" * /;
+163812139 <ms% 350> "+" / <wav 2> "1MCwiaR" / <ms% 500> / <wav 2>
"1MCtiaR" / <ms% 500> / <wav 2> "2MCwiaR" * /;
+183821235 <ms% 350> "+" / <wav 2> "1MViaF" / <ms% 500> / <wav 2>
"1MVuaF" / <ms% 500> / <wav 2> "2MViaF" * /;
-283821136 <ms% 350> "+" / <wav 2> "1MViaH" / <ms% 500> / <wav 2>
"1MVuaH" / <ms% 500> / <wav 2> "2MVuaH" * /;
-283812237 <ms% 350> "+" / <wav 2> "1MVuaL" / <ms% 500> / <wav 2>
"1MViaL" / <ms% 500> / <wav 2> "2MViaL" * /;
+183812138 <ms% 350> "+" / <wav 2> "1MVuaM" / <ms% 500> / <wav 2>
"1MViaM" / <ms% 500> / <wav 2> "2MVuaM" * /;
+183821239 <ms% 350> "+" / <wav 2> "1MViaR" / <ms% 500> / <wav 2>
"1MVuaR" / <ms% 500> / <wav 2> "2MViaR" * /;
-293812235 <ms% 350> "+" / <wav 2> "1MVreuiF" / <ms% 500> / <wav 2>
"1MVriaF" / <ms% 500> / <wav 2> "2MVriaF" * /;
+193812136 <ms% 350> "+" / <wav 2> "1MVreuiH" / <ms% 500> / <wav 2>
"1MVriaH" / <ms% 500> / <wav 2> "2MVreuiH" * /;
+193821237 <ms% 350> "+" / <wav 2> "1MVriaL" / <ms% 500> / <wav 2>
"1MVreuiL" / <ms% 500> / <wav 2> "2MVriaL" * /;
-293821138 <ms% 350> "+" / <wav 2> "1MVriaM" / <ms% 500> / <wav 2>
"1MVreuiM" / <ms% 500> / <wav 2> "2MVreuiM" * /;
-293812239 <ms% 350> "+" / <wav 2> "1MVreuiR" / <ms% 500> / <wav 2>
"1MVriaR" / <ms% 500> / <wav 2> "2MVriaR" * /;
$
0 <line -2> "Take a break. Press SPACEBAR when ready.";
$
+114521211 <ms% 350> "+" / <wav 2> "1MTnoiF" / <ms% 500> / <wav 2>
"1MTnoiR" / <ms% 500> / <wav 2> "2MTnoiF" * /;
-214521112 <ms% 350> "+" / <wav 2> "1MTnoiF" / <ms% 500> / <wav 2>
"1MTnoiH" / <ms% 500> / <wav 2> "2MTnoiH" * /;
-214512213 <ms% 350> "+" / <wav 2> "1MTnoiH" / <ms% 500> / <wav 2>
"1MTnoiR" / <ms% 500> / <wav 2> "2MTnoiR" * /;
+114512114 <ms% 350> "+" / <wav 2> "1MTnoiL" / <ms% 500> / <wav 2>
"1MTnoiF" / <ms% 500> / <wav 2> "2MTnoiL" * /;
+114521215 <ms% 350> "+" / <wav 2> "1MTnoiH" / <ms% 500> / <wav 2>
"1MTnoiL" / <ms% 500> / <wav 2> "2MTnoiH" * /;
-214521116 <ms% 350> "+" / <wav 2> "1MTnoiM" / <ms% 500> / <wav 2>
"1MTnoiL" / <ms% 500> / <wav 2> "2MTnoiL" * /;
```

```
-214512217 <ms% 350> "+" / <wav 2> "1MTnoiL" / <ms% 500> / <wav 2>
"1MTnoiR" / <ms% 500> / <wav 2> "2MTnoiR" * /;
+114512118 <ms% 350> "+" / <wav 2> "1MTnoiM" / <ms% 500> / <wav 2>
"1MTnoiF" / <ms% 500> / <wav 2> "2MTnoiM" * /;
+114521219 <ms% 350> "+" / <wav 2> "1MTnoiH" / <ms% 500> / <wav 2>
"1MTnoiM" / <ms% 500> / <wav 2> "2MTnoiH" * /;
-214521120 <ms% 350> "+" / <wav 2> "1MTnoiR" / <ms% 500> / <wav 2>
"1MTnoiM" / <ms% 500> / <wav 2> "2MTnoiM" * /;
-224521111 <ms% 350> "+" / <wav 2> "1MTpuaiF" / <ms% 500> / <wav 2>
"1MTpuaiR" / <ms% 500> / <wav 2> "2MTpuaiR" * /;
-224512212 <ms% 350> "+" / <wav 2> "1MTpuaiH" / <ms% 500> / <wav 2>
"1MTpuaiF" / <ms% 500> / <wav 2> "2MTpuaiF" * /;
+124512113 <ms% 350> "+" / <wav 2> "1MTpuaiH" / <ms% 500> / <wav 2>
"1MTpuaiR" / <ms% 500> / <wav 2> "2MTpuaiH" * /;
+124521214 <ms% 350> "+" / <wav 2> "1MTpuaiF" / <ms% 500> / <wav 2>
"1MTpuaiL" / <ms% 500> / <wav 2> "2MTpuaiF" * /;
-224521115 <ms% 350> "+" / <wav 2> "1MTpuaiH" / <ms% 500> / <wav 2>
"1MTpuaiL" / <ms% 500> / <wav 2> "2MTpuaiL" * /;
-224512216 <ms% 350> "+" / <wav 2> "1MTpuaiL" / <ms% 500> / <wav 2>
"1MTpuaiM" / <ms% 500> / <wav 2> "2MTpuaiM" * /;
+124512117 <ms% 350> "+" / <wav 2> "1MTpuaiL" / <ms% 500> / <wav 2>
"1MTpuaiR" / <ms% 500> / <wav 2> "2MTpuaiL" * /;
-224512218 <ms% 350> "+" / <wav 2> "1MTpuaiM" / <ms% 500> / <wav 2>
"1MTpuaiF" / <ms% 500> / <wav 2> "2MTpuaiF" * /;
+124512119 <ms% 350> "+" / <wav 2> "1MTpuaiM" / <ms% 500> / <wav 2>
"1MTpuaiH" / <ms% 500> / <wav 2> "2MTpuaiM" * /;
+124521220 <ms% 350> "+" / <wav 2> "1MTpuaiR" / <ms% 500> / <wav 2>
"1MTpuaiM" / <ms% 500> / <wav 2> "2MTpuaiR" * /;
+154821235 <ms% 350> "+" / <wav 2> "1MCdteoF" / <ms% 500> / <wav 2>
"1MCbpeoF" / <ms% 500> / <wav 2> "2MCdteoF" * /;
-254821136 <ms% 350> "+" / <wav 2> "1MCdteoH" / <ms% 500> / <wav 2>
"1MCbpeoH" / <ms% 500> / <wav 2> "2MCbpeoH" * /;
-254812237 <ms% 350> "+" / <wav 2> "1MCbpeoL" / <ms% 500> / <wav 2>
"1MCdteoL" / <ms% 500> / <wav 2> "2MCdteoL" * /;
+154812138 <ms% 350> "+" / <wav 2> "1MCbpeoM" / <ms% 500> / <wav 2>
"1MCdteoM" / <ms% 500> / <wav 2> "2MCbpeoM" * /;
+154821239 <ms% 350> "+" / <wav 2> "1MCdteoR" / <ms% 500> / <wav 2>
"1MCbpeoR" / <ms% 500> / <wav 2> "2MCdteoR" * /;
-264821135 <ms% 350> "+" / <wav 2> "1MCtiaF" / <ms% 500> / <wav 2>
"1MCwiaF" / <ms% 500> / <wav 2> "2MCwiaF" * /;
-264812236 <ms% 350> "+" / <wav 2> "1MCwiaH" / <ms% 500> / <wav 2>
"1MCtiaH" / <ms% 500> / <wav 2> "2MCtiaH" * /;
+164812137 <ms% 350> "+" / <wav 2> "1MCwiaL" / <ms% 500> / <wav 2>
"1MCtiaL" / <ms% 500> / <wav 2> "2MCwiaL" * /;
+164821238 <ms% 350> "+" / <wav 2> "1MCtiaM" / <ms% 500> / <wav 2>
"1MCwiaM" / <ms% 500> / <wav 2> "2MCtiaM" * /;
-264812239 <ms% 350> "+" / <wav 2> "1MCwiaR" / <ms% 500> / <wav 2>
"1MCtiaR" / <ms% 500> / <wav 2> "2MCtiaR" * /;
+184812135 <ms% 350> "+" / <wav 2> "1MVuaF" / <ms% 500> / <wav 2>
"1MViaF" / <ms% 500> / <wav 2> "2MVuaF" * /;
+184821236 <ms% 350> "+" / <wav 2> "1MViaH" / <ms% 500> / <wav 2>
"1MVuaH" / <ms% 500> / <wav 2> "2MViaH" * /;
-284821137 <ms% 350> "+" / <wav 2> "1MViaL" / <ms% 500> / <wav 2>
"1MVuaL" / <ms% 500> / <wav 2> "2MVuaL" * /;
-284812238 <ms% 350> "+" / <wav 2> "1MVuaM" / <ms% 500> / <wav 2>
"1MViaM" / <ms% 500> / <wav 2> "2MViaM" * /;
+184812139 <ms% 350> "+" / <wav 2> "1MVuaR" / <ms% 500> / <wav 2>
```

```
"1MViaR" / <ms% 500> / <wav 2> "2MVuaR" * /;
-294821135 <ms% 350> "+" / <wav 2> "1MVriaF" / <ms% 500> / <wav 2>
"1MVreuiF" / <ms% 500> / <wav 2> "2MVreuiF" * /;
-294812236 <ms% 350> "+" / <wav 2> "1MVreuiH" / <ms% 500> / <wav 2>
"1MVriaH" / <ms% 500> / <wav 2> "2MVriaH" * /;
+194812137 <ms% 350> "+" / <wav 2> "1MVreuiL" / <ms% 500> / <wav 2>
"1MVriaL" / <ms% 500> / <wav 2> "2MVreuiL" * /;
+194821238 <ms% 350> "+" / <wav 2> "1MVriaM" / <ms% 500> / <wav 2>
"1MVreuiM" / <ms% 500> / <wav 2> "2MVriaM" * /;
-294821139 <ms% 350> "+" / <wav 2> "1MVriaR" / <ms% 500> / <wav 2>
"1MVreuiR" / <ms% 500> / <wav 2> "2MVreuiR" * /;
\
0 <line -1> "The End...thank you very much for participating", <line 1> "Please call the
experimenter.", <line 2> "[Experimenter: Press Esc to Save the data]";
```

J. DMDX Script used in disyllabic AXB task in Current Study

```
<ep> <azk> <NumberOfItems 80> <scramble 40> <ContinuousRun> <Delay 158>
<FrameDuration 250> <Timeout 2500> <id "Keyboard"> <mr +Space>
<MapNegativeResponse "+Right Alt"> <MapPositiveResponse "+Left Alt">
<vm 1280,1024,1024,32,60> <eop>
```

```
$
0 <line -8> ”ABX Instructions”,<line -6> “You will hear 3 pseudo words in
a row.”, <line -4> “You have to  decide whether”, <line -2> “the third one is
similar to the first or the second word.”,<line 0> “Press the LEFT Alt A button if the
third one is the same as the first word”,<line 2> “Press the RIGHT Alt B button if the
third one is the same as second word”,<line 4> “Answer as quickly as
possible.”,<line 8> “Press SPACEBAR to start with a short practice.”;

999 <ms% 2000> "+"  /;
$
```

```
+100000001 <cfb " CORRECT "> <wfb " INCORRECT "> <tlfb "Too
slow"> <ms% 500> "+" / <wav 2> "1DVauLM" / <ms% 500> / <wav 2> "1DCnbLM" / <ms% 500> /
<wav 2> "2DVauLM" * /;

-200000002 <cfb " CORRECT "> <wfb " INCORRECT "> <tlfb "Too
slow"> <ms% 500> "+" / <wav 2> "1DVauLH" / <ms% 500> / <wav 2> "1DCnbLH" / <ms% 500> /
<wav 2> "2DCnbLH" *
/;

-200000003 <cfb " CORRECT "> <wfb " INCORRECT "> <tlfb "Too
slow"> <ms% 500> "+" / <wav 2> "1DCnbMH" / <ms% 500> / <wav 2> "1DVauMH" / <ms% 500> /
<wav 2> "2DVauMH" * /;

+100000004 <cfb " CORRECT "> <wfb " INCORRECT "> <tlfb "Too
slow"> <ms% 500> "+" / <wav 2> "1DCnbHF" / <ms% 500> / <wav 2> "1DVauHF" / <ms% 500> /
<wav 2> "2DCnbHF" * /;

-200000005 <cfb " CORRECT "> <wfb " INCORRECT "> <tlfb "Too
slow"> <ms% 500> "+" / <wav 2> "1DVauHL" / <ms% 500> / <wav 2> "1DCnbHL" / <ms% 500> /
<wav 2> "2DCnbHL" * /;

-200000006 <cfb " CORRECT "> <wfb " INCORRECT "> <tlfb "Too
slow"> <ms% 500> "+" / <wav 2> "1DCrbLM" / <ms% 500> / <wav 2> "1DVuuLM" / <ms% 500> /
<wav 2> "2DVuuLM" * /;

+100000007 <cfb " CORRECT "> <wfb " INCORRECT "> <tlfb "Too
slow"> <ms% 500> "+" / <wav 2> "1DCrbLH" / <ms% 500> / <wav 2> "1DVuuLH" / <ms% 500> /
<wav 2> "2DCrbLH" * /;

+100000008 <cfb " CORRECT "> <wfb " INCORRECT "> <tlfb "Too
slow"> <ms% 500> "+" / <wav 2> "1DVuuMH" / <ms% 500> / <wav 2> "1DCrbMH" / <ms% 500> /
```

```
<wav 2> "2DVuuMH" *
/;

-200000009 <cfb " CORRECT "> <wfb " INCORRECT "> <tlfb "Too
slow"> <ms% 500> "+" / <wav 2> "1DVuuHF" / <ms% 500> / <wav 2> "1DCrbHF" / <ms% 500> /
<wav 2> "2DCrbHF" * /;

+100000010 <cfb " CORRECT "> <wfb " INCORRECT "> <tlfb "Too
slow"> <ms% 500> "+" / <wav 2> "1DCrbHL" / <ms% 500> / <wav 2> "1DVuuHL" / <ms% 500> /
<wav 2> "2DCrbHL" * /;

$
0 <line -4> "Ready for the real experiment?", <line -2> "The
feedback is now turned off  but otherwise it will be almost the same.", <line 0>
"Respond as fast as you can without making mistakes.", <line 2> "If you make a
mistake, don't worry and keep going! Good luck ! ";

999 <nfb> <ms% 2000> /;
$

\
+111512121 <ms% 350> "+" / <wav 2> "1DTdpLH" / <ms% 500> / <wav 2>
"1DTdpHL" / <ms% 500> / <wav 2> "2DTdpLH" * /;

+111521222 <ms% 350> "+" / <wav 2> "1DTdpLH" / <ms% 500> / <wav 2>
"1DTdpMH" / <ms% 500> / <wav 2> "2DTdpLH" * /;

-211521123 <ms% 350> "+" / <wav 2> "1DTdpHL" / <ms% 500> / <wav 2>
"1DTdpLM" / <ms% 500> / <wav 2> "2DTdpLM" * /;

-211512224 <ms% 350> "+" / <wav 2> "1DTdpLM" / <ms% 500> / <wav 2>
"1DTdpLH" / <ms% 500> / <wav 2> "2DTdpLH" * /;

+111512125 <ms% 350> "+" / <wav 2> "1DTdpLM" / <ms% 500> / <wav 2>
"1DTdpMH" / <ms% 500> / <wav 2> "2DTdpLM" * /;

-231521121 <ms% 350> "+" / <wav 2> "1DTkkHL" / <ms% 500> / <wav 2>
"1DTkkLH" / <ms% 500> / <wav 2> "2DTkkLH" * /;

-231512222 <ms% 350> "+" / <wav 2> "1DTkkMH" / <ms% 500> / <wav 2>
"1DTkkLH" / <ms% 500> / <wav 2> "2DTkkLH" * /;

+131512123 <ms% 350> "+" / <wav 2> "1DTkkLM" / <ms% 500> / <wav 2>
"1DTkkHL" / <ms% 500> / <wav 2> "2DTkkLM" * /;

+131521224 <ms% 350> "+" / <wav 2> "1DTkkLH" / <ms% 500> / <wav 2>
"1DTkkLM" / <ms% 500> / <wav 2> "2DTkkLH" * /;

-231521125 <ms% 350> "+" / <wav 2> "1DTkkMH" / <ms% 500> / <wav 2>
"1DTkkLM" / <ms% 500> / <wav 2> "2DTkkLM" * /;

+151812121 <ms% 350> "+" / <wav 2> "1DCpwLH" / <ms% 500> / <wav 2>
"1DCrwHL" / <ms% 500> / <wav 2> "2DCpwLH" * /;
```

```
+151821222 <ms% 350> "+" / <wav 2> "1DCrwLH" / <ms% 500> / <wav 2>
"1DCpwMH" / <ms% 500> / <wav 2> "2DCrwLH" * /;

-251821123 <ms% 350> "+" / <wav 2> "1DCrwHL" / <ms% 500> / <wav 2>
"1DCpwLM" / <ms% 500> / <wav 2> "2DCpwLM" * /;

-251812224 <ms% 350> "+" / <wav 2> "1DCpwLM" / <ms% 500> / <wav 2>
"1DCrwLH" / <ms% 500> / <wav 2> "2DCrwLH" * /;

+151812125 <ms% 350> "+" / <wav 2> "1DCpwLM" / <ms% 500> / <wav 2>
"1DCrwMH" / <ms% 500> / <wav 2> "2DCpwLM" * /;

-281812226 <ms% 350> "+" / <wav 2> "1DVeiHF" / <ms% 500> / <wav 2>
"1DVeoHF" / <ms% 500> / <wav 2> "2DVeoHF" * /;

+181812127 <ms% 350> "+" / <wav 2> "1DVeiHL" / <ms% 500> / <wav 2>
"1DVeoHL" / <ms% 500> / <wav 2> "2DVeiHL" * /;

+181821228 <ms% 350> "+" / <wav 2> "1DVeoLH" / <ms% 500> / <wav 2>
"1DVeiLH" / <ms% 500> / <wav 2> "2DVeoLH" * /;

-281821129 <ms% 350> "+" / <wav 2> "1DVeoLM" / <ms% 500> / <wav 2>
"1DVeiLM" / <ms% 500> / <wav 2> "2DVeiLM" * /;

-281812230 <ms% 350> "+" / <wav 2> "1DVeiMH" / <ms% 500> / <wav 2>
"1DVeoMH" / <ms% 500> / <wav 2> "2DVeoMH" * /;

-212512221 <ms% 350> "+" / <wav 2> "1DTdpLH" / <ms% 500> / <wav 2>
"1DTdpHL" / <ms% 500> / <wav 2> "2DTdpHL" * /;

+112512122 <ms% 350> "+" / <wav 2> "1DTdpMH" / <ms% 500> / <wav 2>
"1DTdpLH" / <ms% 500> / <wav 2> "2DTdpMH" * /;

+112521223 <ms% 350> "+" / <wav 2> "1DTdpHL" / <ms% 500> / <wav 2>
"1DTdpLM" / <ms% 500> / <wav 2> "2DTdpHL" * /;

-212521124 <ms% 350> "+" / <wav 2> "1DTdpLH" / <ms% 500> / <wav 2>
"1DTdpLM" / <ms% 500> / <wav 2> "2DTdpLM" * /;

-212512225 <ms% 350> "+" / <wav 2> "1DTdpLM" / <ms% 500> / <wav 2>
"1DTdpMH" / <ms% 500> / <wav 2> "2DTdpMH" * /;

+132521221 <ms% 350> "+" / <wav 2> "1DTkkHL" / <ms% 500> / <wav 2>
"1DTkkLH" / <ms% 500> / <wav 2> "2DTkkHL" * /;

-232521122 <ms% 350> "+" / <wav 2> "1DTkkLH" / <ms% 500> / <wav 2>
"1DTkkMH" / <ms% 500> / <wav 2> "2DTkkMH" * /;

-232512223 <ms% 350> "+" / <wav 2> "1DTkkLM" / <ms% 500> / <wav 2>
"1DTkkHL" / <ms% 500> / <wav 2> "2DTkkHL" * /;

+132512124 <ms% 350> "+" / <wav 2> "1DTkkLM" / <ms% 500> / <wav 2>
"1DTkkLH" / <ms% 500> / <wav 2> "2DTkkLM" * /;
```

```
+132521225 <ms% 350> "+" / <wav 2> "1DTkkMH" / <ms% 500> / <wav 2>
"1DTkkLM" / <ms% 500> / <wav 2> "2DTkkMH" * /;

-252812221 <ms% 350> "+" / <wav 2> "1DCpwLH" / <ms% 500> / <wav 2>
"1DCrwHL" / <ms% 500> / <wav 2> "2DCrwHL" * /;

+152812122 <ms% 350> "+" / <wav 2> "1DCpwMH" / <ms% 500> / <wav 2>
"1DCrwLH" / <ms% 500> / <wav 2> "2DCpwMH" * /;

+152821223 <ms% 350> "+" / <wav 2> "1DCrwHL" / <ms% 500> / <wav 2>
"1DCpwLM" / <ms% 500> / <wav 2> "2DCrwHL" * /;

-252821124 <ms% 350> "+" / <wav 2> "1DCrwLH" / <ms% 500> / <wav 2>
"1DCpwLM" / <ms% 500> / <wav 2> "2DCpwLM" * /;

-252812225 <ms% 350> "+" / <wav 2> "1DCpwLM" / <ms% 500> / <wav 2>
"1DCrwMH" / <ms% 500> / <wav 2> "2DCrwMH" * /;

-282821126 <ms% 350> "+" / <wav 2> "1DVeoHF" / <ms% 500> / <wav 2>
"1DVeiHF" / <ms% 500> / <wav 2> "2DVeiHF" * /;

-282812227 <ms% 350> "+" / <wav 2> "1DVeiHL" / <ms% 500> / <wav 2>
"1DVeoHL" / <ms% 500> / <wav 2> "2DVeoHL" * /;

+182812128 <ms% 350> "+" / <wav 2> "1DVeiLH" / <ms% 500> / <wav 2>
"1DVeoLH" / <ms% 500> / <wav 2> "2DVeiLH" * /;

+182821229 <ms% 350> "+" / <wav 2> "1DVeoLM" / <ms% 500> / <wav 2>
"1DVeiLM" / <ms% 500> / <wav 2> "2DVeoLM" * /;

-282821130 <ms% 350> "+" / <wav 2> "1DVeoMH" / <ms% 500> / <wav 2>
"1DVeiMH" / <ms% 500> / <wav 2> "2DVeiMH" * /;

$
0 <line -2> "Take a break. Press SPACEBAR when ready.";
$

-213521121 <ms% 350> "+" / <wav 2> "1DTdpHL" / <ms% 500> / <wav 2>
"1DTdpLH" / <ms% 500> / <wav 2> "2DTdpLH" * /;

-213512222 <ms% 350> "+" / <wav 2> "1DTdpMH" / <ms% 500> / <wav 2>
"1DTdpLH" / <ms% 500> / <wav 2> "2DTdpLH" * /;

+113512123 <ms% 350> "+" / <wav 2> "1DTdpLM" / <ms% 500> / <wav 2>
"1DTdpHL" / <ms% 500> / <wav 2> "2DTdpLM" * /;

+113521224 <ms% 350> "+" / <wav 2> "1DTdpLH" / <ms% 500> / <wav 2>
"1DTdpLM" / <ms% 500> / <wav 2> "2DTdpLH" * /;

-213521125 <ms% 350> "+" / <wav 2> "1DTdpMH" / <ms% 500> / <wav 2>
"1DTdpLM" / <ms% 500> / <wav 2> "2DTdpLM" * /;

+133512121 <ms% 350> "+" / <wav 2> "1DTkkLH" / <ms% 500> / <wav 2>
"1DTkkHL" / <ms% 500> / <wav 2> "2DTkkLH" * /;
```

```
+133521222 <ms% 350> "+" / <wav 2> "1DTkkLH" / <ms% 500> / <wav 2>
"1DTkkMH" / <ms% 500> / <wav 2> "2DTkkLH" * /;

-233521123 <ms% 350> "+" / <wav 2> "1DTkkHL" / <ms% 500> / <wav 2>
"1DTkkLM" / <ms% 500> / <wav 2> "2DTkkLM" * /;

-233512224 <ms% 350> "+" / <wav 2> "1DTkkLM" / <ms% 500> / <wav 2>
"1DTkkLH" / <ms% 500> / <wav 2> "2DTkkLH" * /;

+133512125 <ms% 350> "+" / <wav 2> "1DTkkLM" / <ms% 500> / <wav 2>
"1DTkkMH" / <ms% 500> / <wav 2> "2DTkkLM" * /;

-253821121 <ms% 350> "+" / <wav 2> "1DCrwHL" / <ms% 500> / <wav 2>
"1DCpwLH" / <ms% 500> / <wav 2> "2DCpwLH" * /;

-253812222 <ms% 350> "+" / <wav 2> "1DCpwMH" / <ms% 500> / <wav 2>
"1DCrwLH" / <ms% 500> / <wav 2> "2DCrwLH" * /;

+153812123 <ms% 350> "+" / <wav 2> "1DCpwLM" / <ms% 500> / <wav 2>
"1DCrwHL" / <ms% 500> / <wav 2> "2DCpwLM" * /;

+153821224 <ms% 350> "+" / <wav 2> "1DCrwLH" / <ms% 500> / <wav 2>
"1DCpwLM" / <ms% 500> / <wav 2> "2DCrwLH" * /;

-253821125 <ms% 350> "+" / <wav 2> "1DCrwMH" / <ms% 500> / <wav 2>
"1DCpwLM" / <ms% 500> / <wav 2> "2DCpwLM" * /;

+183821226 <ms% 350> "+" / <wav 2> "1DVeoHF" / <ms% 500> / <wav 2>
"1DVeiHF" / <ms% 500> / <wav 2> "2DVeoHF" * /;

-283821127 <ms% 350> "+" / <wav 2> "1DVeoHL" / <ms% 500> / <wav 2>
"1DVeiHL" / <ms% 500> / <wav 2> "2DVeiHL" * /;

-283812228 <ms% 350> "+" / <wav 2> "1DVeiLH" / <ms% 500> / <wav 2>
"1DVeoLH" / <ms% 500> / <wav 2> "2DVeoLH" * /;

+183812129 <ms% 350> "+" / <wav 2> "1DVeiLM" / <ms% 500> / <wav 2>
"1DVeoLM" / <ms% 500> / <wav 2> "2DVeiLM" * /;

+183821230 <ms% 350> "+" / <wav 2> "1DVeoMH" / <ms% 500> / <wav 2>
"1DVeiMH" / <ms% 500> / <wav 2> "2DVeoMH" * /;

+114521221 <ms% 350> "+" / <wav 2> "1DTdpHL" / <ms% 500> / <wav 2>
"1DTdpLH" / <ms% 500> / <wav 2> "2DTdpHL" * /;

-214521122 <ms% 350> "+" / <wav 2> "1DTdpLH" / <ms% 500> / <wav 2>
"1DTdpMH" / <ms% 500> / <wav 2> "2DTdpMH" * /;

-214512223 <ms% 350> "+" / <wav 2> "1DTdpLM" / <ms% 500> / <wav 2>
"1DTdpHL" / <ms% 500> / <wav 2> "2DTdpHL" * /;

+114512124 <ms% 350> "+" / <wav 2> "1DTdpLM" / <ms% 500> / <wav 2>
"1DTdpLH" / <ms% 500> / <wav 2> "2DTdpLM" * /;
```

```
+114521225 <ms% 350> "+" / <wav 2> "1DTdpMH" / <ms% 500> / <wav 2>
"1DTdpLM" / <ms% 500> / <wav 2> "2DTdpMH" * /;

-234512221 <ms% 350> "+" / <wav 2> "1DTkkLH" / <ms% 500> / <wav 2>
"1DTkkHL" / <ms% 500> / <wav 2> "2DTkkHL" * /;

+134512122 <ms% 350> "+" / <wav 2> "1DTkkMH" / <ms% 500> / <wav 2>
"1DTkkLH" / <ms% 500> / <wav 2> "2DTkkMH" * /;

+134521223 <ms% 350> "+" / <wav 2> "1DTkkHL" / <ms% 500> / <wav 2>
"1DTkkLM" / <ms% 500> / <wav 2> "2DTkkHL" * /;

-234521124 <ms% 350> "+" / <wav 2> "1DTkkLH" / <ms% 500> / <wav 2>
"1DTkkLM" / <ms% 500> / <wav 2> "2DTkkLM" * /;

-234512225 <ms% 350> "+" / <wav 2> "1DTkkLM" / <ms% 500> / <wav 2>
"1DTkkMH" / <ms% 500> / <wav 2> "2DTkkMH" * /;

+154821221 <ms% 350> "+" / <wav 2> "1DCrwHL" / <ms% 500> / <wav 2>
"1DCpwLH" / <ms% 500> / <wav 2> "2DCrwHL" * /;

-254821122 <ms% 350> "+" / <wav 2> "1DCrwLH" / <ms% 500> / <wav 2>
"1DCpwMH" / <ms% 500> / <wav 2> "2DCpwMH" * /;

-254812223 <ms% 350> "+" / <wav 2> "1DCpwLM" / <ms% 500> / <wav 2>
"1DCrwHL" / <ms% 500> / <wav 2> "2DCrwHL" * /;

+154812124 <ms% 350> "+" / <wav 2> "1DCpwLM" / <ms% 500> / <wav 2>
"1DCrwLH" / <ms% 500> / <wav 2> "2DCpwLM" * /;

+154821225 <ms% 350> "+" / <wav 2> "1DCrwMH" / <ms% 500> / <wav 2>
"1DCpwLM" / <ms% 500> / <wav 2> "2DCrwMH" * /;

+184812126 <ms% 350> "+" / <wav 2> "1DVeiHF" / <ms% 500> / <wav 2>
"1DVeoHF" / <ms% 500> / <wav 2> "2DVeiHF" * /;

+184821227 <ms% 350> "+" / <wav 2> "1DVeoHL" / <ms% 500> / <wav 2>
"1DVeiHL" / <ms% 500> / <wav 2> "2DVeoHL" * /;

-284821128 <ms% 350> "+" / <wav 2> "1DVeoLH" / <ms% 500> / <wav 2>
"1DVeiLH" / <ms% 500> / <wav 2> "2DVeiLH" * /;

-284812229 <ms% 350> "+" / <wav 2> "1DVeiLM" / <ms% 500> / <wav 2>
"1DVeoLM" / <ms% 500> / <wav 2> "2DVeoLM" * /;

+184812130 <ms% 350> "+" / <wav 2> "1DVeiMH" / <ms% 500> / <wav 2>
"1DVeoMH" / <ms% 500> / <wav 2> "2DVeiMH" * /;

\

0 <line -1> "The End...thank you very much for participating", <line 1>
"Please call the experimenter.", <line 2> "[Experimenter: Press Esc to
Save the data]";
```

K. DMDX Script used in Lexical Decision task with Repetition Priming in Current Study

LIST 1

```
<ep> <azk> <ContinuousRun> <Delay 158> <FrameDuration 250> <Timeout 2500> <id "Keyboard"> <mr +Space> <MapNegativeResponse "+Left Alt"> <MapPositiveResponse "+Right Alt"> <vm 1280,1024,1024,32,60> <eop>
```

```
0 <line -8>  "LIST 1", <line -6>  "Instructions ",<line -4>  "You will hear short words that",<line -2> "may or may not be real words in Mandarin Chinese",<line 0>  "Press the RIGHT (yes) key if it is a real existing word in Mandarin, ",<line 2> "Press the LEFT (no) key if it is NOT a real word in Mandarin.",<line 4>  " Please answer as accurately and as QUICKLY as possible.",<line 8> "Press SPACEBAR to start a short practice with feedback";
```

```
999 <ms% 2000> "+"  /;

-300000011 <cfb " CORRECT "> <wfb " INCORRECT "> <tlfb "Too slow"> <wav 2> "DMNfiao4" * /;

+300000001 <cfb " CORRECT "> <wfb " INCORRECT"> <tlfb "Too slow"> <wav 2> "MWTchi1" * /;

-300000012 <cfb " CORRECT "> <wfb " INCORRECT "> <tlfb "Too slow"> <wav 2>"DDNgie3" * /;

+300000002 <cfb " CORRECT "> <wfb " INCORRECT "> <tlfb "Too slow"> <wav 2> "MWTzhu1" * /;

+300000003 <cfb " CORRECT "> <wfb " INCORRECT "> <tlfb "Too slow"> <wav 2> "MWCdui4" * /;

+300000004 <cfb " CORRECT "> <wfb " INCORRECT "> <tlfb "Too slow"> <wav 2> "MWVshuo1" * /;

+300000005 <cfb " CORRECT "> <wfb " INCORRECT "> <tlfb "Too slow"> <wav 2> "DDWdiah4" * /;

-300000009 <cfb " CORRECT "> <wfb " INCORRECT "> <tlfb "Too slow"> <wav 2> "MNCfui4" * /;

-300000010 <cfb " CORRECT "> <wfb " INCORRECT "> <tlfb "Too slow"> <wav 2> "MNClua4" * /;

+300000006 <cfb " CORRECT "> <wfb " INCORRECT "> <tlfb "Too slow"> <wav 2> "DDWshi4" * /;
```

-300000007 <cfb " CORRECT "> <wfb " INCORRECT "> <tlfb "Too slow"> <wav 2> "MNTbe1" * /;

-300000008 <cfb " CORRECT "> <wfb " INCORRECT "> <tlfb "Too slow"> <wav 2> "MNTsua1" * /;

0 <line -3> "**Ready for the real experiment?** The feedback is now turned off", <line -1> "Respond **as fast as you can without making mistakes**", <line 1> "If you make a mistake, don't worry and keep going!", <line 4>"**...press SPACEBAR when ready**";

999 <nfb> <ms% 2000> /;

```
+1411411    <ms% 350> "+" / <wav 2> "MWTtin2a" * /;
-8028559    <ms% 350> "+" / <wav 2> "DMNzhun1" * /;
+1312318    <ms% 350> "+" / <wav 2> "MWCmai4a" * /;
-8023759    <ms% 350> "+" / <wav 2> "DMNbuan3" * /;
+1027859    <ms% 350> "+" / <wav 2> "DDWhuo3" * /;
+1312718    <ms% 350> "+" / <wav 2> "MWVdi4a" * /;
-8020159    <ms% 350> "+" / <wav 2> "DMNfai3" * /;
-8029659    <ms% 350> "+" / <wav 2> "DMNlen3" * /;
+1311711    <ms% 350> "+" / <wav 2> "MWTqia2a" * /;
+1026259    <ms% 350> "+" / <wav 2> "DMWdeng3" * /;
+1411421    <ms% 350> "+" / <wav 2> "MWTtin2b" * /;
+1312518    <ms% 350> "+" / <wav 2> "MWVche2a" * /;
-8314318    <ms% 350> "+" / <wav 2> "MNCtiu4a" * /;
-8021259    <ms% 350> "+" / <wav 2> "DDNdon2" * /;
+1412418    <ms% 350> "+" / <wav 2> "MWCzha3a" * /;
+1411811    <ms% 350> "+" / <wav 2> "MWTqi2a" * /;
+1026759    <ms% 350> "+" / <wav 2> "DDWzhon1" * /;
+1312728    <ms% 350> "+" / <wav 2> "MWVda4b" * /;
-8414418    <ms% 350> "+" / <wav 2> "MNCgia3a" * /;
-8020359    <ms% 350> "+" / <wav 2> "DDNkei3" * /;
+1311721    <ms% 350> "+" / <wav 2> "MWTqia1b" * /;
+1312528    <ms% 350> "+" / <wav 2> "MWVcha2b" * /;
-8029859    <ms% 350> "+" / <wav 2> "DMNnui3" * /;
+1312328    <ms% 350> "+" / <wav 2> "MWCzai4b" * /;
-8314118    <ms% 350> "+" / <wav 2> "MNCfai4a" * /;
+1027559    <ms% 350> "+" / <wav 2> "DDWjie2" * /;
-8314328    <ms% 350> "+" / <wav 2> "MNCsiu4b" * /;
+1022159    <ms% 350> "+" / <wav 2> "DMWzou3" * /;
-8023259    <ms% 350> "+" / <wav 2> "DMNzhun1" * /;
+1412428    <ms% 350> "+" / <wav 2> "MWCzha3b" * /;
+1026959    <ms% 350> "+" / <wav 2> "DDWshen1" * /;
-8414428    <ms% 350> "+" / <wav 2> "MNCgia3b" * /;
-8020859    <ms% 350> "+" / <wav 2> "DDNpe4" * /;
-8314128    <ms% 350> "+" / <wav 2> "MNCyai4b" * /;
-8029959    <ms% 350> "+" / <wav 2> "DMNbon3" * /;
+1411821    <ms% 350> "+" / <wav 2> "MWTqi2b" * /;
-8023859    <ms% 350> "+" / <wav 2> "DMNden3" * /;
+1022759    <ms% 350> "+" / <wav 2> "DDWjie2" * /;
```

```
0 <line 0> "Take a break. Press SPACEBAR when ready.";

999 <ms% 2000> "+" /;

-8023459    <ms% 350> "+" / <wav 2> "DMNfie2" * /;
+1028159    <ms% 350> "+" / <wav 2> "DDWwen4" * /;
+1311311    <ms% 350> "+" / <wav 2> "MWThe2a" * /;
+1412618    <ms% 350> "+" / <wav 2> "MWChao3a" * /;
+1021959    <ms% 350> "+" / <wav 2> "DMWshou3" * /;
+1311511    <ms% 350> "+" / <wav 2> "MWTchu1a" * /;
+1022459    <ms% 350> "+" / <wav 2> "DDWshen1" * /;
-8028859    <ms% 350> "+" / <wav 2> "DDNmun4" * /;
-8023659    <ms% 350> "+" / <wav 2> "DMNzhui2" * /;
-8029159    <ms% 350> "+" / <wav 2> "DDNfun4" * /;
-8020959    <ms% 350> "+" / <wav 2> "DDNkui3" * /;
+1022359    <ms% 350> "+" / <wav 2> "DDWzhon1" * /;
-8413411    <ms% 350> "+" / <wav 2> "MNTyuo2a" * /;
-8023159    <ms% 350> "+" / <wav 2> "DMNfe1" * /;
+1023059    <ms% 350> "+" / <wav 2> "DDWwen4" * /;
+1027959    <ms% 350> "+" / <wav 2> "DDWgan3" * /;
-8414218    <ms% 350> "+" / <wav 2> "MNCcou1a" * /;
+1311521    <ms% 350> "+" / <wav 2> "MWTchu2b" * /;
+1027659    <ms% 350> "+" / <wav 2> "DDWxue2" * /;
-8024359    <ms% 350> "+" / <wav 2> "DDNkei4" * /;
+1311321    <ms% 350> "+" / <wav 2> "MWThe1b" * /;
+1026359    <ms% 350> "+" / <wav 2> "DMWzou3" * /;
-8021359    <ms% 350> "+" / <wav 2> "DDNfie1" * /;
+1412628    <ms% 350> "+" / <wav 2> "MWChao3b" * /;
-8413421    <ms% 350> "+" / <wav 2> "MNTyuo2b" * /;
-8028459    <ms% 350> "+" / <wav 2> "DMNmun1" * /;
+1026459    <ms% 350> "+" / <wav 2> "DMWhua4" * /;
-8029559    <ms% 350> "+" / <wav 2> "DMNbuan3" * /;
-8414228    <ms% 350> "+" / <wav 2> "MNCcou1b" * /;
-8024659    <ms% 350> "+" / <wav 2> "DDNfie1" * /;
-8024059    <ms% 350> "+" / <wav 2> "DMNfai3" * /;
+1027059    <ms% 350> "+" / <wav 2> "DDWshe4" * /;
+1022559    <ms% 350> "+" / <wav 2> "DDWshan1" * /;
+1021759    <ms% 350> "+" / <wav 2> "DMWnan2" * /;
-8023959    <ms% 350> "+" / <wav 2> "DMNbon3" * /;
-8024259    <ms% 350> "+" / <wav 2> "DDNlen4" * /;
-8021159    <ms% 350> "+" / <wav 2> "DDNnui2" * /;
-8020759    <ms% 350> "+" / <wav 2> "DDNkei4" * /;

0 <line 0> "Take a break. Press SPACEBAR when ready.";

999 <ms% 2000> "+" /;

+1025359    <ms% 350> "+" / <wav 2> "DMWxin1" * /;
-8313511    <ms% 350> "+" / <wav 2> "MNTmiu1a" * /;
-8020559    <ms% 350> "+" / <wav 2> "DDNlen4" * /;
-8313711    <ms% 350> "+" / <wav 2> "MNTpou2a" * /;
+1411611    <ms% 350> "+" / <wav 2> "MWTcha1a" * /;
-8029059    <ms% 350> "+" / <wav 2> "DMNsian2" * /;
-8023359    <ms% 350> "+" / <wav 2> "DMNtiu1" * /;
+1027259    <ms% 350> "+" / <wav 2> "DDWzuo4" * /;
-8414818    <ms% 350> "+" / <wav 2> "MNVsin2a" * /;
+1022059    <ms% 350> "+" / <wav 2> "DMWshui3" * /;
```

```
+1025559    <ms% 350> "+" / <wav 2> "DMWnan2" * /;
-8313521    <ms% 350> "+" / <wav 2> "MNTmiu2b" * /;
-8020059    <ms% 350> "+" / <wav 2> "DMNfi3" * /;
+1412818    <ms% 350> "+" / <wav 2> "MWVjie1a" * /;
-8313721    <ms% 350> "+" / <wav 2> "MNTpou1b" * /;
+1411621    <ms% 350> "+" / <wav 2> "MWTcha1b" * /;
+1025859    <ms% 350> "+" / <wav 2> "DMWgou3" * /;
+1028259    <ms% 350> "+" / <wav 2> "DDWdinn4" * /;
+1022259    <ms% 350> "+" / <wav 2> "DMWdao4" * /;
-8314518    <ms% 350> "+" / <wav 2> "MNVbua3a" * /;
-8414828    <ms% 350> "+" / <wav 2> "MNVsin2b" * /;
-8020459    <ms% 350> "+" / <wav 2> "DDNfie4" * /;
+1412828    <ms% 350> "+" / <wav 2> "MWVjie1b" * /;
+1025259    <ms% 350> "+" / <wav 2> "DMWgao1" * /;
-8023559    <ms% 350> "+" / <wav 2> "DDNfun4" * /;
-8028759    <ms% 350> "+" / <wav 2> "DMNtiu1" * /;
-8029359    <ms% 350> "+" / <wav 2> "DMNzhui2" * /;
+1022959    <ms% 350> "+" / <wav 2> "DDWgan3" * /;
-8021059    <ms% 350> "+" / <wav 2> "DDNten2" * /;
-8313311    <ms% 350> "+" / <wav 2> "MNTgii2a" * /;
+1027459    <ms% 350> "+" / <wav 2> "DDWshi2" * /;
+1025459    <ms% 350> "+" / <wav 2> "DMWtou2" * /;
+1026659    <ms% 350> "+" / <wav 2> "DDWfang4" * /;
-8413611    <ms% 350> "+" / <wav 2> "MNTfia1a" * /;
+1027759    <ms% 350> "+" / <wav 2> "DDWlao3" * /;
-8028359    <ms% 350> "+" / <wav 2> "DMNfe1" * /;
-8314528    <ms% 350> "+" / <wav 2> "MNVbui3b" * /;
-8313321    <ms% 350> "+" / <wav 2> "MNTgii1b" * /;
-8020659    <ms% 350> "+" / <wav 2> "DDNton4" * /;
+1028059    <ms% 350> "+" / <wav 2> "DDWyan3" * /;
+1026559    <ms% 350> "+" / <wav 2> "DMWdao4" * /;
-8021459    <ms% 350> "+" / <wav 2> "DDNfian1" * /;
-8028959    <ms% 350> "+" / <wav 2> "DMNfie2" * /;

0 <line 0> "Take a break. Press SPACEBAR when ready.";

999 <ms% 2000> "+" /;

-8029459    <ms% 350> "+" / <wav 2> "DMNgiao2" * /;
+1412218    <ms% 350> "+" / <wav 2> "MWVgui4a" * /;
-8313111    <ms% 350> "+" / <wav 2> "MNThin1a" * /;
+1411211    <ms% 350> "+" / <wav 2> "MWTtia1a" * /;
+1026859    <ms% 350> "+" / <wav 2> "DDWshng1" * /;
-8314718    <ms% 350> "+" / <wav 2> "MNVnia1a" * /;
+1026159    <ms% 350> "+" / <wav 2> "DMWshui3" * /;
-8024559    <ms% 350> "+" / <wav 2> "DDNnui2" * /;
+1021659    <ms% 350> "+" / <wav 2> "DMWxin1" * /;
-8413621    <ms% 350> "+" / <wav 2> "MNTfia1b" * /;
+1412228    <ms% 350> "+" / <wav 2> "MWVgui4b" * /;
+1311111    <ms% 350> "+" / <wav 2> "MWTtan1a" * /;
+1022659    <ms% 350> "+" / <wav 2> "DDWyin1" * /;
-8313121    <ms% 350> "+" / <wav 2> "MNThin2b" * /;
-8029759    <ms% 350> "+" / <wav 2> "DMNden3" * /;
-8020259    <ms% 350> "+" / <wav 2> "DDNbe3" * /;
+1411221    <ms% 350> "+" / <wav 2> "MWTtia1b" * /;
-8314728    <ms% 350> "+" / <wav 2> "MNVnua1b" * /;
+1312118    <ms% 350> "+" / <wav 2> "MWCfan4a" * /;
```

```
+1027159    <ms% 350> "+" / <wav 2> "DDWshan1" * /;
+1022859    <ms% 350> "+" / <wav 2> "DDWlao3" * /;
-8028659    <ms% 350> "+" / <wav 2> "DDNson4" * /;
+1021859    <ms% 350> "+" / <wav 2> "DMWxue3" * /;
-8029259    <ms% 350> "+" / <wav 2> "DMNyui2" * /;
+1021559    <ms% 350> "+" / <wav 2> "DMWshan1" * /;
-8413211    <ms% 350> "+" / <wav 2> "MNTnui1a" * /;
+1312128    <ms% 350> "+" / <wav 2> "MWCban4b" * /;
+1311121    <ms% 350> "+" / <wav 2> "MWTtan2b" * /;
+1025159    <ms% 350> "+" / <wav 2> "DMWshan1" * /;
+1025759    <ms% 350> "+" / <wav 2> "DMWxue3" * /;
+1025659    <ms% 350> "+" / <wav 2> "DMWxian3" * /;
-8413811    <ms% 350> "+" / <wav 2> "MNTtua2a" * /;
-8414618    <ms% 350> "+" / <wav 2> "MNVse1a" * /;
-8024459    <ms% 350> "+" / <wav 2> "DDNkui3" * /;
+1027359    <ms% 350> "+" / <wav 2> "DDWyin1" * /;
+1026059    <ms% 350> "+" / <wav 2> "DMWzhi3" * /;
+1025959    <ms% 350> "+" / <wav 2> "DMWshou3" * /;
-8024159    <ms% 350> "+" / <wav 2> "DDNkei3" * /;
-8413221    <ms% 350> "+" / <wav 2> "MNTnui1b" * /;
-8414628    <ms% 350> "+" / <wav 2> "MNVse1b" * /;
-8413821    <ms% 350> "+" / <wav 2> "MNTtua2b" * /;
```

```
0 <line -2> "The End...thank you for participating", ?<line 0> "Please call
the experimenter!", <line 2> "[Experimenter: Press Esc to Save the data]";
```

LIST 2

```
<ep> <azk> <ContinuousRun> <Delay 158> <FrameDuration 250> <Timeout 2500> <id
"Keyboard"> <mr +Space> <MapNegativeResponse "+Left Alt">
<MapPositiveResponse "+Right Alt"> <vm 1280,1024,1024,32,60> <eop>
```

```
0 <line -8>  "LIST 2", <line -6>  "Instructions ",<line -4>  "You will
hear short words that",<line -2> "may or may not be real words in Mandarin Chinese",<line
0>  "Press the RIGHT (yes) key if it is a real existing word in Mandarin, ",<line 2>
"Press the LEFT (no) key if it is NOT a real word in Mandarin.",<line 4>  " Please answer
as accurately and as QUICKLY as possible.",<line 8> "Press SPACEBAR to start a short
practice with feedback";
```

```
999 <ms% 2000> "+"  /;
```

```
-300000011 <cfb " CORRECT "> <wfb " INCORRECT "> <tlfb "Too
slow"> <wav 2> "DMNfiao4" * /;

+300000001 <cfb " CORRECT "> <wfb " INCORRECT"> <tlfb "Too
slow"> <wav 2> "MWTchi1" * /;

-300000012 <cfb " CORRECT "> <wfb " INCORRECT "> <tlfb "Too
slow"> <wav 2>"DDNgie3" * /;
```

```
+300000002 <cfb " CORRECT "> <wfb " INCORRECT "> <tlfb "Too slow"> <wav 2> "MWTzhu1" * /;

+300000003 <cfb " CORRECT "> <wfb " INCORRECT "> <tlfb "Too slow"> <wav 2> "MWCdui4" * /;

+300000004 <cfb " CORRECT "> <wfb " INCORRECT "> <tlfb "Too slow"> <wav 2> "MWVshuo1" * /;

+300000005 <cfb " CORRECT "> <wfb " INCORRECT "> <tlfb "Too slow"> <wav 2> "DDWdiah4" * /;

-300000009 <cfb " CORRECT "> <wfb " INCORRECT "> <tlfb "Too slow"> <wav 2> "MNCfui4" * /;

-300000010 <cfb " CORRECT "> <wfb " INCORRECT "> <tlfb "Too slow"> <wav 2> "MNClua4" * /;

+300000006 <cfb " CORRECT "> <wfb " INCORRECT "> <tlfb "Too slow"> <wav 2> "DDWshi4" * /;

-300000007 <cfb " CORRECT "> <wfb " INCORRECT "> <tlfb "Too slow"> <wav 2> "MNTbe1" * /;

-300000008 <cfb " CORRECT "> <wfb " INCORRECT "> <tlfb "Too slow"> <wav 2> "MNTsua1" * /;


0 <line -3> "Ready for the real experiment? The feedback is now turned off", <line -1> "Respond as fast as you can without making mistakes", <line 1> "If you make a mistake, don't worry and keep going!", <line 4>"...press SPACEBAR when ready";

999 <nfb> <ms% 2000>  /;


-8024159    <ms% 350> "+" / <wav 2> "DDNkei3" * /;
+1022059    <ms% 350> "+" / <wav 2> "DMWshui3" * /;
-8029159    <ms% 350> "+" / <wav 2> "DDNfun4" * /;
-8020159    <ms% 350> "+" / <wav 2> "DMNfai3" * /;
-8021259    <ms% 350> "+" / <wav 2> "DDNdon2" * /;
+1028259    <ms% 350> "+" / <wav 2> "DDWdinn4" * /;
-8023459    <ms% 350> "+" / <wav 2> "DMNfie2" * /;
-8028759    <ms% 350> "+" / <wav 2> "DMNtiu1" * /;
-8021459    <ms% 350> "+" / <wav 2> "DDNfian1" * /;
-8023259    <ms% 350> "+" / <wav 2> "DMNzhun1" * /;
+1022459    <ms% 350> "+" / <wav 2> "DDWshen1" * /;
-8020759    <ms% 350> "+" / <wav 2> "DDNkei4" * /;
+1312418    <ms% 350> "+" / <wav 2> "MWCsha3a" * /;
-8414718    <ms% 350> "+" / <wav 2> "MNVnia1a" * /;
-8028359    <ms% 350> "+" / <wav 2> "DMNfe1" * /;
```

```
+1027259    <ms% 350> "+" / <wav 2> "DDWzuo4" * /;
-8313211    <ms% 350> "+" / <wav 2> "MNTnui2a" * /;
+1026559    <ms% 350> "+" / <wav 2> "DMWdao4" * /;
+1021859    <ms% 350> "+" / <wav 2> "DMWxue3" * /;
+1025459    <ms% 350> "+" / <wav 2> "DMWtou2" * /;
-8414118    <ms% 350> "+" / <wav 2> "MNCfai4a" * /;
-8024359    <ms% 350> "+" / <wav 2> "DDNkei4" * /;
-8414728    <ms% 350> "+" / <wav 2> "MNVnia1b" * /;
+1027759    <ms% 350> "+" / <wav 2> "DDWlao3" * /;
+1411511    <ms% 350> "+" / <wav 2> "MWTchu1a" * /;
+1311211    <ms% 350> "+" / <wav 2> "MWTtia2a" * /;
+1027959    <ms% 350> "+" / <wav 2> "DDWgan3" * /;
+1026459    <ms% 350> "+" / <wav 2> "DMWhua4" * /;
-8313221    <ms% 350> "+" / <wav 2> "MNTnui1b" * /;
+1025259    <ms% 350> "+" / <wav 2> "DMWgao1" * /;
+1412318    <ms% 350> "+" / <wav 2> "MWCmai4a" * /;
+1312428    <ms% 350> "+" / <wav 2> "MWCzha3b" * /;
-8414128    <ms% 350> "+" / <wav 2> "MNCfai4b" * /;
+1412518    <ms% 350> "+" / <wav 2> "MWVche2a" * /;
+1311221    <ms% 350> "+" / <wav 2> "MWTtia1b" * /;
-8024659    <ms% 350> "+" / <wav 2> "DDNfie1" * /;
+1412328    <ms% 350> "+" / <wav 2> "MWCmai4b" * /;
-8028559    <ms% 350> "+" / <wav 2> "DMNzhun1" * /;
+1411521    <ms% 350> "+" / <wav 2> "MWTchu1b" * /;

0 <line 0> "Take a break. Press SPACEBAR when ready.";

999 <ms% 2000> "+" /;

-8029459    <ms% 350> "+" / <wav 2> "DMNgiao2" * /;
+1312218    <ms% 350> "+" / <wav 2> "MWVguo4a" * /;
-8314618    <ms% 350> "+" / <wav 2> "MNVsou1a" * /;
+1022559    <ms% 350> "+" / <wav 2> "DDWshan1" * /;
+1027859    <ms% 350> "+" / <wav 2> "DDWhuo3" * /;
+1022759    <ms% 350> "+" / <wav 2> "DDWjie2" * /;
+1311811    <ms% 350> "+" / <wav 2> "MWTqi1a" * /;
+1025859    <ms% 350> "+" / <wav 2> "DMWgou3" * /;
-8021159    <ms% 350> "+" / <wav 2> "DDNnui2" * /;
+1412528    <ms% 350> "+" / <wav 2> "MWVche2b" * /;
+1023059    <ms% 350> "+" / <wav 2> "DDWwen4" * /;
-8020259    <ms% 350> "+" / <wav 2> "DDNbe3" * /;
-8020559    <ms% 350> "+" / <wav 2> "DDNlen4" * /;
-8029259    <ms% 350> "+" / <wav 2> "DMNyui2" * /;
+1022359    <ms% 350> "+" / <wav 2> "DDWzhon1" * /;
-8314628    <ms% 350> "+" / <wav 2> "MNVse1b" * /;
-8029559    <ms% 350> "+" / <wav 2> "DMNbuan3" * /;
+1027359    <ms% 350> "+" / <wav 2> "DDWyin1" * /;
+1021559    <ms% 350> "+" / <wav 2> "DMWshan1" * /;
+1311821    <ms% 350> "+" / <wav 2> "MWTqi2b" * /;
+1027059    <ms% 350> "+" / <wav 2> "DDWshe4" * /;
+1312228    <ms% 350> "+" / <wav 2> "MWVgui4b" * /;
+1028059    <ms% 350> "+" / <wav 2> "DDWyan3" * /;
-8023559    <ms% 350> "+" / <wav 2> "DDNfun4" * /;
-8028959    <ms% 350> "+" / <wav 2> "DMNfie2" * /;
+1311611    <ms% 350> "+" / <wav 2> "MWTcha2a" * /;
+1411311    <ms% 350> "+" / <wav 2> "MWThe2a" * /;
-8023659    <ms% 350> "+" / <wav 2> "DMNzhui2" * /;
```

```
+1027159    <ms% 350> "+" / <wav 2> "DDWshan1" * /;
-8414318    <ms% 350> "+" / <wav 2> "MNCtiu4a" * /;
+1412118    <ms% 350> "+" / <wav 2> "MWCfan4a" * /;
-8028459    <ms% 350> "+" / <wav 2> "DMNmun1" * /;
-8029659    <ms% 350> "+" / <wav 2> "DMNlen3" * /;
-8021359    <ms% 350> "+" / <wav 2> "DDNfie1" * /;
-8023959    <ms% 350> "+" / <wav 2> "DMNbon3" * /;
+1026959    <ms% 350> "+" / <wav 2> "DDWshen1" * /;
-8029859    <ms% 350> "+" / <wav 2> "DMNnui3" * /;
+1411321    <ms% 350> "+" / <wav 2> "MWThe2b" * /;
-8023359    <ms% 350> "+" / <wav 2> "DMNtiu1" * /;
+1311621    <ms% 350> "+" / <wav 2> "MWTcha1b" * /;
+1026659    <ms% 350> "+" / <wav 2> "DDWfang4" * /;
-8414328    <ms% 350> "+" / <wav 2> "MNCtiu4b" * /;
+1412128    <ms% 350> "+" / <wav 2> "MWCfan4b" * /;

0 <line 0> "Take a break. Press SPACEBAR when ready.";

999 <ms% 2000> "+" /;

+1021759    <ms% 350> "+" / <wav 2> "DMWnan2" * /;
+1022959    <ms% 350> "+" / <wav 2> "DDWgan3" * /;
+1022259    <ms% 350> "+" / <wav 2> "DMWdao4" * /;
+1026059    <ms% 350> "+" / <wav 2> "DMWzhi3" * /;
+1025659    <ms% 350> "+" / <wav 2> "DMWxian3" * /;
-8020059    <ms% 350> "+" / <wav 2> "DMNfi3" * /;
+1411111    <ms% 350> "+" / <wav 2> "MWTtan1a" * /;
-8313811    <ms% 350> "+" / <wav 2> "MNTtua1a" * /;
+1027659    <ms% 350> "+" / <wav 2> "DDWxue2" * /;
+1027459    <ms% 350> "+" / <wav 2> "DDWshi2" * /;
+1022659    <ms% 350> "+" / <wav 2> "DDWyin1" * /;
-8020659    <ms% 350> "+" / <wav 2> "DDNton4" * /;
+1026159    <ms% 350> "+" / <wav 2> "DMWshui3" * /;
-8024459    <ms% 350> "+" / <wav 2> "DDNkui3" * /;
-8024059    <ms% 350> "+" / <wav 2> "DMNfai3" * /;
-8024259    <ms% 350> "+" / <wav 2> "DDNlen4" * /;
-8020359    <ms% 350> "+" / <wav 2> "DDNkei3" * /;
+1411121    <ms% 350> "+" / <wav 2> "MWTtan1b" * /;
-8020459    <ms% 350> "+" / <wav 2> "DDNfie4" * /;
-8028859    <ms% 350> "+" / <wav 2> "DDNmun4" * /;
+1312618    <ms% 350> "+" / <wav 2> "MWCbao3a" * /;
-8020859    <ms% 350> "+" / <wav 2> "DDNpe4" * /;
-8313821    <ms% 350> "+" / <wav 2> "MNTtua2b" * /;
-8314818    <ms% 350> "+" / <wav 2> "MNVsen2a" * /;
+1311411    <ms% 350> "+" / <wav 2> "MWTtin1a" * /;
+1412718    <ms% 350> "+" / <wav 2> "MWVdi4a" * /;
+1022859    <ms% 350> "+" / <wav 2> "DDWlao3" * /;
-8413711    <ms% 350> "+" / <wav 2> "MNTpou2a" * /;
+1026859    <ms% 350> "+" / <wav 2> "DDWshng1" * /;
+1021959    <ms% 350> "+" / <wav 2> "DMWshou3" * /;
-8413111    <ms% 350> "+" / <wav 2> "MNThin1a" * /;
+1312818    <ms% 350> "+" / <wav 2> "MWVjia1a" * /;
+1022159    <ms% 350> "+" / <wav 2> "DMWzou3" * /;
+1312628    <ms% 350> "+" / <wav 2> "MWChao3b" * /;
-8313411    <ms% 350> "+" / <wav 2> "MNTyuo1a" * /;
-8029959    <ms% 350> "+" / <wav 2> "DMNbon3" * /;
+1412728    <ms% 350> "+" / <wav 2> "MWVdi4b" * /;
```

```
+1311421    <ms% 350> "+" / <wav 2> "MWTtin2b" * /;
-8314828    <ms% 350> "+" / <wav 2> "MNVsin2b" * /;
-8413721    <ms% 350> "+" / <wav 2> "MNTpou2b" * /;
+1026359    <ms% 350> "+" / <wav 2> "DMWzou3" * /;
+1312828    <ms% 350> "+" / <wav 2> "MWVjie1b" * /;
-8413121    <ms% 350> "+" / <wav 2> "MNThin1b" * /;

0 <line 0> "Take a break. Press SPACEBAR when ready.";

999 <ms% 2000> "+" /;

-8028659    <ms% 350> "+" / <wav 2> "DDNson4" * /;
+1025159    <ms% 350> "+" / <wav 2> "DMWshan1" * /;
+1025359    <ms% 350> "+" / <wav 2> "DMWxin1" * /;
-8020959    <ms% 350> "+" / <wav 2> "DDNkui3" * /;
-8313421    <ms% 350> "+" / <wav 2> "MNTyuo2b" * /;
+1411711    <ms% 350> "+" / <wav 2> "MWTqia2a" * /;
-8413511    <ms% 350> "+" / <wav 2> "MNTmiu1a" * /;
-8314218    <ms% 350> "+" / <wav 2> "MNCfou1a" * /;
-8414518    <ms% 350> "+" / <wav 2> "MNVbua3a" * /;
+1026759    <ms% 350> "+" / <wav 2> "DDWzhon1" * /;
+1026259    <ms% 350> "+" / <wav 2> "DMWdeng3" * /;
+1021659    <ms% 350> "+" / <wav 2> "DMWxin1" * /;
-8021059    <ms% 350> "+" / <wav 2> "DDNten2" * /;
-8413311    <ms% 350> "+" / <wav 2> "MNTgii2a" * /;
+1027559    <ms% 350> "+" / <wav 2> "DDWjie2" * /;
+1025759    <ms% 350> "+" / <wav 2> "DMWxue3" * /;
+1411721    <ms% 350> "+" / <wav 2> "MWTqia2b" * /;
-8029059    <ms% 350> "+" / <wav 2> "DMNsian2" * /;
-8023859    <ms% 350> "+" / <wav 2> "DMNden3" * /;
-8029759    <ms% 350> "+" / <wav 2> "DMNden3" * /;
-8413521    <ms% 350> "+" / <wav 2> "MNTmiu1b" * /;
-8414528    <ms% 350> "+" / <wav 2> "MNVbua3b" * /;
+1025559    <ms% 350> "+" / <wav 2> "DMWnan2" * /;
-8313611    <ms% 350> "+" / <wav 2> "MNTfia2a" * /;
-8314418    <ms% 350> "+" / <wav 2> "MNCkia3a" * /;
-8314228    <ms% 350> "+" / <wav 2> "MNCcou1b" * /;
-8023159    <ms% 350> "+" / <wav 2> "DMNfe1" * /;
+1025959    <ms% 350> "+" / <wav 2> "DMWshou3" * /;
-8413321    <ms% 350> "+" / <wav 2> "MNTgii2b" * /;
+1028159    <ms% 350> "+" / <wav 2> "DDWwen4" * /;
-8023759    <ms% 350> "+" / <wav 2> "DMNbuan3" * /;
-8024559    <ms% 350> "+" / <wav 2> "DDNnui2" * /;
-8029359    <ms% 350> "+" / <wav 2> "DMNzhui2" * /;
-8313621    <ms% 350> "+" / <wav 2> "MNTfia1b" * /;
-8314428    <ms% 350> "+" / <wav 2> "MNCgia3b" * /;


0 <line -2> "The End...thank you for participating", ?<line 0> "Please call
the experimenter!", <line 2> "[Experimenter: Press Esc to Save the data]";
```

LIST 3

```
<ep> <azk> <ContinuousRun> <Delay 158> <FrameDuration 250> <Timeout 2500> <id "Keyboard"> <mr +Space> <MapNegativeResponse "+Left Alt"> <MapPositiveResponse "+Right Alt"> <vm 1280,1024,1024,32,60> <eop>
```

```
0 <line -8>  "LIST 3", <line -6>  "Instructions ",<line -4>  ”You will hear short words that”,<line -2> “may or may not be real words in Mandarin Chinese”,<line 0>  “Press the RIGHT (yes) key if it is a real existing word in Mandarin, ”,<line 2> “Press the LEFT (no) key if it is NOT a real word in Mandarin.”,<line 4>  “ Please answer as accurately and as QUICKLY as possible.”,<line 8> “Press SPACEBAR to start a short practice with feedback”;
```

```
999 <ms% 2000> "+"  /;

-300000011 <cfb " CORRECT "> <wfb " INCORRECT "> <tlfb "Too slow"> <wav 2> "DMNfiao4" * /;

+300000001 <cfb " CORRECT "> <wfb " INCORRECT"> <tlfb "Too slow"> <wav 2> "MWTchi1" * /;

-300000012 <cfb " CORRECT "> <wfb " INCORRECT "> <tlfb "Too slow"> <wav 2>"DDNgie3" * /;

+300000002 <cfb " CORRECT "> <wfb " INCORRECT "> <tlfb "Too slow"> <wav 2> "MWTzhu1" * /;

+300000003 <cfb " CORRECT "> <wfb " INCORRECT "> <tlfb "Too slow"> <wav 2> "MWCdui4" * /;

+300000004 <cfb " CORRECT "> <wfb " INCORRECT "> <tlfb "Too slow"> <wav 2> "MWVshuo1" * /;

+300000005 <cfb " CORRECT "> <wfb " INCORRECT "> <tlfb "Too slow"> <wav 2> "DDWdiah4" * /;

-300000009 <cfb " CORRECT "> <wfb " INCORRECT "> <tlfb "Too slow"> <wav 2> "MNCfui4" * /;

-300000010 <cfb " CORRECT "> <wfb " INCORRECT "> <tlfb "Too slow"> <wav 2> "MNClua4" * /;

+300000006 <cfb " CORRECT "> <wfb " INCORRECT "> <tlfb "Too slow"> <wav 2> "DDWshi4" * /;

-300000007 <cfb " CORRECT "> <wfb " INCORRECT "> <tlfb "Too slow"> <wav 2> "MNTbe1" * /;

-300000008 <cfb " CORRECT "> <wfb " INCORRECT "> <tlfb "Too
```

```
slow"> <wav 2> "MNTsua1" * /;


0 <line -3> "Ready for the real experiment? The feedback is now turned off",
<line -1> "Respond as fast as you can without making mistakes", <line 1> "If you
make a mistake, don't worry and keep going!", <line 4>"...press SPACEBAR when
ready";

999 <nfb> <ms% 2000>  /;



+1025559    <ms% 350> "+" / <wav 2> "DMWnan2" * /;
-8029059    <ms% 350> "+" / <wav 2> "DMNsian2" * /;
-8024659    <ms% 350> "+" / <wav 2> "DDNfie1" * /;
-8029859    <ms% 350> "+" / <wav 2> "DMNnui3" * /;
+1311311    <ms% 350> "+" / <wav 2> "MWThe1a" * /;
+1022559    <ms% 350> "+" / <wav 2> "DDWshan1" * /;
+1027359    <ms% 350> "+" / <wav 2> "DDWyin1" * /;
+1026159    <ms% 350> "+" / <wav 2> "DMWshui3" * /;
-8020659    <ms% 350> "+" / <wav 2> "DDNton4" * /;
-8313311    <ms% 350> "+" / <wav 2> "MNTgii1a" * /;
+1411411    <ms% 350> "+" / <wav 2> "MWTtin1a" * /;
+1411211    <ms% 350> "+" / <wav 2> "MWTtia2a" * /;
+1027259    <ms% 350> "+" / <wav 2> "DDWzuo4" * /;
-8020259    <ms% 350> "+" / <wav 2> "DDNbe3" * /;
+1412218    <ms% 350> "+" / <wav 2> "MWVguo4a" * /;
+1311321    <ms% 350> "+" / <wav 2> "MWThe2b" * /;
-8021259    <ms% 350> "+" / <wav 2> "DDNdon2" * /;
-8029659    <ms% 350> "+" / <wav 2> "DMNlen3" * /;
-8021059    <ms% 350> "+" / <wav 2> "DDNten2" * /;
-8313321    <ms% 350> "+" / <wav 2> "MNTgii2b" * /;
-8023959    <ms% 350> "+" / <wav 2> "DMNbon3" * /;
-8029459    <ms% 350> "+" / <wav 2> "DMNgiao2" * /;
+1411221    <ms% 350> "+" / <wav 2> "MWTtia2b" * /;
+1026059    <ms% 350> "+" / <wav 2> "DMWzhi3" * /;
+1021659    <ms% 350> "+" / <wav 2> "DMWxin1" * /;
+1028259    <ms% 350> "+" / <wav 2> "DDWdinn4" * /;
-8020059    <ms% 350> "+" / <wav 2> "DMNfi3" * /;
+1412228    <ms% 350> "+" / <wav 2> "MWVguo4b" * /;
+1411421    <ms% 350> "+" / <wav 2> "MWTtin1b" * /;
-8029159    <ms% 350> "+" / <wav 2> "DDNfun4" * /;
+1412418    <ms% 350> "+" / <wav 2> "MWCsha3a" * /;
+1021759    <ms% 350> "+" / <wav 2> "DMWnan2" * /;
-8023859    <ms% 350> "+" / <wav 2> "DMNden3" * /;
+1026859    <ms% 350> "+" / <wav 2> "DDWshng1" * /;
+1028159    <ms% 350> "+" / <wav 2> "DDWwen4" * /;
-8028959    <ms% 350> "+" / <wav 2> "DMNfie2" * /;
+1412428    <ms% 350> "+" / <wav 2> "MWCsha3b" * /;

0 <line 0> "Take a break. Press SPACEBAR when ready.";

999 <ms% 2000> "+" /;

-8023359    <ms% 350> "+" / <wav 2> "DMNtiu1" * /;
+1027959    <ms% 350> "+" / <wav 2> "DDWgan3" * /;
```

```
-8413611    <ms% 350> "+" / <wav 2> "MNTfia2a" * /;
-8313711    <ms% 350> "+" / <wav 2> "MNTpou1a" * /;
+1028059    <ms% 350> "+" / <wav 2> "DDWyan3" * /;
+1026459    <ms% 350> "+" / <wav 2> "DMWhua4" * /;
+1025659    <ms% 350> "+" / <wav 2> "DMWxian3" * /;
-8020459    <ms% 350> "+" / <wav 2> "DDNfie4" * /;
+1412618    <ms% 350> "+" / <wav 2> "MWCbao3a" * /;
+1027159    <ms% 350> "+" / <wav 2> "DDWshan1" * /;
+1023059    <ms% 350> "+" / <wav 2> "DDWwen4" * /;
+1022159    <ms% 350> "+" / <wav 2> "DMWzou3" * /;
+1311111    <ms% 350> "+" / <wav 2> "MWTtan2a" * /;
+1022359    <ms% 350> "+" / <wav 2> "DDWzhon1" * /;
-8028659    <ms% 350> "+" / <wav 2> "DDNson4" * /;
-8413621    <ms% 350> "+" / <wav 2> "MNTfia2b" * /;
-8413411    <ms% 350> "+" / <wav 2> "MNTyuo1a" * /;
-8028459    <ms% 350> "+" / <wav 2> "DMNmun1" * /;
+1025459    <ms% 350> "+" / <wav 2> "DMWtou2" * /;
-8313721    <ms% 350> "+" / <wav 2> "MNTpou2b" * /;
+1025359    <ms% 350> "+" / <wav 2> "DMWxin1" * /;
+1311711    <ms% 350> "+" / <wav 2> "MWTqia1a" * /;
+1412628    <ms% 350> "+" / <wav 2> "MWCbao3b" * /;
+1021559    <ms% 350> "+" / <wav 2> "DMWshan1" * /;
-8021159    <ms% 350> "+" / <wav 2> "DDNnui2" * /;
-8413421    <ms% 350> "+" / <wav 2> "MNTyuo1b" * /;
+1022959    <ms% 350> "+" / <wav 2> "DDWgan3" * /;
+1312518    <ms% 350> "+" / <wav 2> "MWVcha2a" * /;
-8414818    <ms% 350> "+" / <wav 2> "MNVsen2a" * /;
+1311121    <ms% 350> "+" / <wav 2> "MWTtan1b" * /;
+1311721    <ms% 350> "+" / <wav 2> "MWTqia2b" * /;
-8020359    <ms% 350> "+" / <wav 2> "DDNkei3" * /;
+1026359    <ms% 350> "+" / <wav 2> "DMWzou3" * /;
-8028559    <ms% 350> "+" / <wav 2> "DMNzhun1" * /;
-8314318    <ms% 350> "+" / <wav 2> "MNCsiu4a" * /;
-8024259    <ms% 350> "+" / <wav 2> "DDNlen4" * /;
+1027459    <ms% 350> "+" / <wav 2> "DDWshi2" * /;
+1022459    <ms% 350> "+" / <wav 2> "DDWshen1" * /;
+1312528    <ms% 350> "+" / <wav 2> "MWVche2b" * /;
-8024559    <ms% 350> "+" / <wav 2> "DDNnui2" * /;
+1026759    <ms% 350> "+" / <wav 2> "DDWzhon1" * /;
-8029359    <ms% 350> "+" / <wav 2> "DMNzhui2" * /;
+1022859    <ms% 350> "+" / <wav 2> "DDWlao3" * /;
-8414828    <ms% 350> "+" / <wav 2> "MNVsen2b" * /;
-8314328    <ms% 350> "+" / <wav 2> "MNCtiu4b" * /;

0 <line 0> "Take a break. Press SPACEBAR when ready.";

999 <ms% 2000> "+" /;

+1025159    <ms% 350> "+" / <wav 2> "DMWshan1" * /;
-8024059    <ms% 350> "+" / <wav 2> "DMNfai3" * /;
-8023459    <ms% 350> "+" / <wav 2> "DMNfie2" * /;
+1026559    <ms% 350> "+" / <wav 2> "DMWdao4" * /;
-8414218    <ms% 350> "+" / <wav 2> "MNCfou1a" * /;
-8314118    <ms% 350> "+" / <wav 2> "MNCyai4a" * /;
+1026959    <ms% 350> "+" / <wav 2> "DDWshen1" * /;
-8020159    <ms% 350> "+" / <wav 2> "DMNfai3" * /;
+1026659    <ms% 350> "+" / <wav 2> "DDWfang4" * /;
```

```
+1022259    <ms% 350> "+" / <wav 2> "DMWdao4" * /;
-8413211    <ms% 350> "+" / <wav 2> "MNTnui2a" * /;
-8024159    <ms% 350> "+" / <wav 2> "DDNkei3" * /;
-8023759    <ms% 350> "+" / <wav 2> "DMNbuan3" * /;
+1312118    <ms% 350> "+" / <wav 2> "MWCban4a" * /;
-8414418    <ms% 350> "+" / <wav 2> "MNCkia3a" * /;
-8024359    <ms% 350> "+" / <wav 2> "DDNkei4" * /;
-8314718    <ms% 350> "+" / <wav 2> "MNVnua1a" * /;
+1026259    <ms% 350> "+" / <wav 2> "DMWdeng3" * /;
+1412818    <ms% 350> "+" / <wav 2> "MWVjia1a" * /;
-8414228    <ms% 350> "+" / <wav 2> "MNCfou1b" * /;
-8314128    <ms% 350> "+" / <wav 2> "MNCfai4b" * /;
+1311511    <ms% 350> "+" / <wav 2> "MWTchu2a" * /;
-8413221    <ms% 350> "+" / <wav 2> "MNTnui2b" * /;
-8023559    <ms% 350> "+" / <wav 2> "DDNfun4" * /;
-8028359    <ms% 350> "+" / <wav 2> "DMNfe1" * /;
-8313511    <ms% 350> "+" / <wav 2> "MNTmiu2a" * /;
+1312128    <ms% 350> "+" / <wav 2> "MWCfan4b" * /;
+1027559    <ms% 350> "+" / <wav 2> "DDWjie2" * /;
+1027859    <ms% 350> "+" / <wav 2> "DDWhuo3" * /;
-8029959    <ms% 350> "+" / <wav 2> "DMNbon3" * /;
-8314728    <ms% 350> "+" / <wav 2> "MNVnia1b" * /;
+1025259    <ms% 350> "+" / <wav 2> "DMWgao1" * /;
+1412828    <ms% 350> "+" / <wav 2> "MWVjia1b" * /;
-8313111    <ms% 350> "+" / <wav 2> "MNThin2a" * /;
-8414428    <ms% 350> "+" / <wav 2> "MNCkia3b" * /;
-8029759    <ms% 350> "+" / <wav 2> "DMNden3" * /;
+1027659    <ms% 350> "+" / <wav 2> "DDWxue2" * /;
-8313521    <ms% 350> "+" / <wav 2> "MNTmiu1b" * /;
+1311521    <ms% 350> "+" / <wav 2> "MWTchu1b" * /;

0 <line 0> "Take a break. Press SPACEBAR when ready.";

999 <ms% 2000> "+" /;

+1021959    <ms% 350> "+" / <wav 2> "DMWshou3" * /;
+1022059    <ms% 350> "+" / <wav 2> "DMWshui3" * /;
+1312718    <ms% 350> "+" / <wav 2> "MWVda4a" * /;
+1027759    <ms% 350> "+" / <wav 2> "DDWlao3" * /;
+1411611    <ms% 350> "+" / <wav 2> "MWTcha2a" * /;
-8023259    <ms% 350> "+" / <wav 2> "DMNzhun1" * /;
-8020859    <ms% 350> "+" / <wav 2> "DDNpe4" * /;
-8029259    <ms% 350> "+" / <wav 2> "DMNyui2" * /;
+1312318    <ms% 350> "+" / <wav 2> "MWCzai4a" * /;
-8023159    <ms% 350> "+" / <wav 2> "DMNfe1" * /;
+1022759    <ms% 350> "+" / <wav 2> "DDWjie2" * /;
+1021859    <ms% 350> "+" / <wav 2> "DMWxue3" * /;
-8313121    <ms% 350> "+" / <wav 2> "MNThin1b" * /;
-8020559    <ms% 350> "+" / <wav 2> "DDNlen4" * /;
+1027059    <ms% 350> "+" / <wav 2> "DDWshe4" * /;
+1025759    <ms% 350> "+" / <wav 2> "DMWxue3" * /;
-8021459    <ms% 350> "+" / <wav 2> "DDNfian1" * /;
-8413811    <ms% 350> "+" / <wav 2> "MNTtua1a" * /;
-8029559    <ms% 350> "+" / <wav 2> "DMNbuan3" * /;
+1025859    <ms% 350> "+" / <wav 2> "DMWgou3" * /;
-8028759    <ms% 350> "+" / <wav 2> "DMNtiu1" * /;
+1312728    <ms% 350> "+" / <wav 2> "MWVdi4b" * /;
```

```
+1411811    <ms% 350> "+" / <wav 2> "MWTqi1a" * /;
+1411621    <ms% 350> "+" / <wav 2> "MWTcha2b" * /;
+1312328    <ms% 350> "+" / <wav 2> "MWCmai4b" * /;
-8414618    <ms% 350> "+" / <wav 2> "MNVsou1a" * /;
+1025959    <ms% 350> "+" / <wav 2> "DMWshou3" * /;
-8024459    <ms% 350> "+" / <wav 2> "DDNkui3" * /;
-8314518    <ms% 350> "+" / <wav 2> "MNVbui3a" * /;
-8020759    <ms% 350> "+" / <wav 2> "DDNkei4" * /;
-8021359    <ms% 350> "+" / <wav 2> "DDNfie1" * /;
-8023659    <ms% 350> "+" / <wav 2> "DMNzhui2" * /;
-8413821    <ms% 350> "+" / <wav 2> "MNTtua1b" * /;
-8414628    <ms% 350> "+" / <wav 2> "MNVsou1b" * /;
+1022659    <ms% 350> "+" / <wav 2> "DDWyin1" * /;
-8028859    <ms% 350> "+" / <wav 2> "DDNmun4" * /;
+1411821    <ms% 350> "+" / <wav 2> "MWTqi1b" * /;
-8314528    <ms% 350> "+" / <wav 2> "MNVbua3b" * /;
-8020959    <ms% 350> "+" / <wav 2> "DDNkui3" * /;
```

```
0 <line -2> "The End...thank you for participating", ?<line 0> "Please call the experimenter!", <line 2> "[Experimenter: Press Esc to Save the data]";
```

LIST 4

```
<ep> <azk> <ContinuousRun> <Delay 158> <FrameDuration 250> <Timeout 2500> <id "Keyboard"> <mr +Space> <MapNegativeResponse "+Left Alt"> <MapPositiveResponse "+Right Alt"> <vm 1280,1024,1024,32,60> <eop>
```

```
0 <line -8>  "LIST 4", <line -6>  "Instructions”,<line -4>  ”You will hear short words that”,<line -2> “may or may not be real words in Mandarin Chinese”,<line 0>  “Press the RIGHT (yes) key if it is a real existing word in Mandarin, ”,<line 2> “Press the LEFT (no) key if it is NOT a real word in Mandarin.”,<line 4>  “ Please answer as accurately and as QUICKLY as possible.”,<line 8> “Press SPACEBAR to start a short practice with feedback”;
```

```
999 <ms% 2000> "+"  /;
```

```
-300000011 <cfb " CORRECT "> <wfb " INCORRECT "> <tlfb "Too slow"> <wav 2> "DMNfiao4" * /;

+300000001 <cfb " CORRECT "> <wfb " INCORRECT"> <tlfb "Too slow"> <wav 2> "MWTchi1" * /;

-300000012 <cfb " CORRECT "> <wfb " INCORRECT "> <tlfb "Too slow"> <wav 2>"DDNgie3" * /;

+300000002 <cfb " CORRECT "> <wfb " INCORRECT "> <tlfb "Too slow"> <wav 2> "MWTzhu1" * /;
```

```
+300000003 <cfb " CORRECT "> <wfb " INCORRECT "> <tlfb "Too slow"> <wav 2> "MWCdui4" * /;

+300000004 <cfb " CORRECT "> <wfb " INCORRECT "> <tlfb "Too slow"> <wav 2> "MWVshuo1" * /;

+300000005 <cfb " CORRECT "> <wfb " INCORRECT "> <tlfb "Too slow"> <wav 2> "DDWdiah4" * /;

-300000009 <cfb " CORRECT "> <wfb " INCORRECT "> <tlfb "Too slow"> <wav 2> "MNCfui4" * /;

-300000010 <cfb " CORRECT "> <wfb " INCORRECT "> <tlfb "Too slow"> <wav 2> "MNClua4" * /;

+300000006 <cfb " CORRECT "> <wfb " INCORRECT "> <tlfb "Too slow"> <wav 2> "DDWshi4" * /;

-300000007 <cfb " CORRECT "> <wfb " INCORRECT "> <tlfb "Too slow"> <wav 2> "MNTbe1" * /;

-300000008 <cfb " CORRECT "> <wfb " INCORRECT "> <tlfb "Too slow"> <wav 2> "MNTsua1" * /;


0 <line -3> "Ready for the real experiment? The feedback is now turned off", <line -1> "Respond as fast as you can without making mistakes", <line 1> "If you make a mistake, don't worry and keep going!", <line 4>"...press SPACEBAR when ready";

999 <nfb> <ms% 2000>  /;


+1025259    <ms% 350> "+" / <wav 2> "DMWgao1" * /;
+1411511    <ms% 350> "+" / <wav 2> "MWTchu2a" * /;
-8028659    <ms% 350> "+" / <wav 2> "DDNson4" * /;
-8020359    <ms% 350> "+" / <wav 2> "DDNkei3" * /;
-8020759    <ms% 350> "+" / <wav 2> "DDNkei4" * /;
+1027059    <ms% 350> "+" / <wav 2> "DDWshe4" * /;
-8414118    <ms% 350> "+" / <wav 2> "MNCyai4a" * /;
-8029459    <ms% 350> "+" / <wav 2> "DMNgiao2" * /;
-8028559    <ms% 350> "+" / <wav 2> "DMNzhun1" * /;
-8029859    <ms% 350> "+" / <wav 2> "DMNnui3" * /;
+1411521    <ms% 350> "+" / <wav 2> "MWTchu2b" * /;
-8021359    <ms% 350> "+" / <wav 2> "DDNfie1" * /;
-8021059    <ms% 350> "+" / <wav 2> "DDNten2" * /;
-8313211    <ms% 350> "+" / <wav 2> "MNTnui1a" * /;
-8314818    <ms% 350> "+" / <wav 2> "MNVsin2a" * /;
+1311211    <ms% 350> "+" / <wav 2> "MWTtia1a" * /;
-8414128    <ms% 350> "+" / <wav 2> "MNCyai4b" * /;
-8021459    <ms% 350> "+" / <wav 2> "DDNfian1" * /;
```

```
+1022359    <ms% 350> "+" / <wav 2> "DDWzhon1" * /;
+1025759    <ms% 350> "+" / <wav 2> "DMWxue3" * /;
+1027259    <ms% 350> "+" / <wav 2> "DDWzuo4" * /;
+1412518    <ms% 350> "+" / <wav 2> "MWVcha2a" * /;
+1027359    <ms% 350> "+" / <wav 2> "DDWyin1" * /;
+1027559    <ms% 350> "+" / <wav 2> "DDWjie2" * /;
+1026859    <ms% 350> "+" / <wav 2> "DDWshng1" * /;
-8029959    <ms% 350> "+" / <wav 2> "DMNbon3" * /;
+1411111    <ms% 350> "+" / <wav 2> "MWTtan2a" * /;
-8314828    <ms% 350> "+" / <wav 2> "MNVsen2b" * /;
-8313221    <ms% 350> "+" / <wav 2> "MNTnui2b" * /;
-8029359    <ms% 350> "+" / <wav 2> "DMNzhui2" * /;
+1028159    <ms% 350> "+" / <wav 2> "DDWwen4" * /;
+1311221    <ms% 350> "+" / <wav 2> "MWTtia2b" * /;
+1411311    <ms% 350> "+" / <wav 2> "MWThe1a" * /;
+1025859    <ms% 350> "+" / <wav 2> "DMWgou3" * /;
+1026659    <ms% 350> "+" / <wav 2> "DDWfang4" * /;
+1311611    <ms% 350> "+" / <wav 2> "MWTcha1a" * /;
+1027159    <ms% 350> "+" / <wav 2> "DDWshan1" * /;
+1412528    <ms% 350> "+" / <wav 2> "MWVcha2b" * /;
-8024359    <ms% 350> "+" / <wav 2> "DDNkei4" * /;
-8029059    <ms% 350> "+" / <wav 2> "DMNsian2" * /;
+1411121    <ms% 350> "+" / <wav 2> "MWTtan2b" * /;
+1411321    <ms% 350> "+" / <wav 2> "MWThe1b" * /;
-8028859    <ms% 350> "+" / <wav 2> "DDNmun4" * /;
-8023459    <ms% 350> "+" / <wav 2> "DMNfie2" * /;
+1311621    <ms% 350> "+" / <wav 2> "MWTcha2b" * /;

0 <line 0> "Take a break. Press SPACEBAR when ready.";

999 <ms% 2000> "+" /;

+1022059    <ms% 350> "+" / <wav 2> "DMWshui3" * /;
+1022459    <ms% 350> "+" / <wav 2> "DDWshen1" * /;
-8020159    <ms% 350> "+" / <wav 2> "DMNfai3" * /;
-8029559    <ms% 350> "+" / <wav 2> "DMNbuan3" * /;
+1312818    <ms% 350> "+" / <wav 2> "MWVjie1a" * /;
-8313411    <ms% 350> "+" / <wav 2> "MNTyuo2a" * /;
+1311811    <ms% 350> "+" / <wav 2> "MWTqi2a" * /;
-8023259    <ms% 350> "+" / <wav 2> "DMNzhun1" * /;
+1311411    <ms% 350> "+" / <wav 2> "MWTtin2a" * /;
-8023859    <ms% 350> "+" / <wav 2> "DMNden3" * /;
+1026359    <ms% 350> "+" / <wav 2> "DMWzou3" * /;
-8413311    <ms% 350> "+" / <wav 2> "MNTgii1a" * /;
+1022759    <ms% 350> "+" / <wav 2> "DDWjie2" * /;
-8314418    <ms% 350> "+" / <wav 2> "MNCgia3a" * /;
+1021559    <ms% 350> "+" / <wav 2> "DMWshan1" * /;
+1022659    <ms% 350> "+" / <wav 2> "DDWyin1" * /;
+1311821    <ms% 350> "+" / <wav 2> "MWTqi1b" * /;
-8024159    <ms% 350> "+" / <wav 2> "DDNkei3" * /;
+1312828    <ms% 350> "+" / <wav 2> "MWVjia1b" * /;
-8414518    <ms% 350> "+" / <wav 2> "MNVbui3a" * /;
-8024259    <ms% 350> "+" / <wav 2> "DDNlen4" * /;
-8313421    <ms% 350> "+" / <wav 2> "MNTyuo1b" * /;
+1311421    <ms% 350> "+" / <wav 2> "MWTtin1b" * /;
+1025659    <ms% 350> "+" / <wav 2> "DMWxian3" * /;
-8023959    <ms% 350> "+" / <wav 2> "DMNbon3" * /;
```

```
+1312618    <ms% 350> "+" / <wav 2> "MWChao3a" * /;
+1021859    <ms% 350> "+" / <wav 2> "DMWxue3" * /;
+1022959    <ms% 350> "+" / <wav 2> "DDWgan3" * /;
-8028759    <ms% 350> "+" / <wav 2> "DMNtiu1" * /;
-8413321    <ms% 350> "+" / <wav 2> "MNTgii1b" * /;
-8414718    <ms% 350> "+" / <wav 2> "MNVnua1a" * /;
-8024459    <ms% 350> "+" / <wav 2> "DDNkui3" * /;
-8314428    <ms% 350> "+" / <wav 2> "MNCkia3b" * /;
+1025959    <ms% 350> "+" / <wav 2> "DMWshou3" * /;
+1023059    <ms% 350> "+" / <wav 2> "DDWwen4" * /;
-8024659    <ms% 350> "+" / <wav 2> "DDNfie1" * /;
-8020459    <ms% 350> "+" / <wav 2> "DDNfie4" * /;
-8414528    <ms% 350> "+" / <wav 2> "MNVbui3b" * /;
+1022259    <ms% 350> "+" / <wav 2> "DMWdao4" * /;
+1312628    <ms% 350> "+" / <wav 2> "MWCbao3b" * /;
-8414728    <ms% 350> "+" / <wav 2> "MNVnua1b" * /;

0 <line 0> "Take a break. Press SPACEBAR when ready.";

999 <ms% 2000> "+" /;

+1026959    <ms% 350> "+" / <wav 2> "DDWshen1" * /;
-8023659    <ms% 350> "+" / <wav 2> "DMNzhui2" * /;
-8020659    <ms% 350> "+" / <wav 2> "DDNton4" * /;
+1026559    <ms% 350> "+" / <wav 2> "DMWdao4" * /;
+1027659    <ms% 350> "+" / <wav 2> "DDWxue2" * /;
-8029659    <ms% 350> "+" / <wav 2> "DMNlen3" * /;
-8020959    <ms% 350> "+" / <wav 2> "DDNkui3" * /;
-8413511    <ms% 350> "+" / <wav 2> "MNTmiu2a" * /;
+1025459    <ms% 350> "+" / <wav 2> "DMWtou2" * /;
+1027459    <ms% 350> "+" / <wav 2> "DDWshi2" * /;
-8029259    <ms% 350> "+" / <wav 2> "DMNyui2" * /;
+1026459    <ms% 350> "+" / <wav 2> "DMWhua4" * /;
-8413111    <ms% 350> "+" / <wav 2> "MNThin2a" * /;
+1412318    <ms% 350> "+" / <wav 2> "MWCzai4a" * /;
-8413521    <ms% 350> "+" / <wav 2> "MNTmiu2b" * /;
-8313611    <ms% 350> "+" / <wav 2> "MNTfia1a" * /;
+1312218    <ms% 350> "+" / <wav 2> "MWVgui4a" * /;
+1026259    <ms% 350> "+" / <wav 2> "DMWdeng3" * /;
-8314618    <ms% 350> "+" / <wav 2> "MNVse1a" * /;
+1021959    <ms% 350> "+" / <wav 2> "DMWshou3" * /;
+1312418    <ms% 350> "+" / <wav 2> "MWCzha3a" * /;
-8023159    <ms% 350> "+" / <wav 2> "DMNfe1" * /;
+1022159    <ms% 350> "+" / <wav 2> "DMWzou3" * /;
+1412328    <ms% 350> "+" / <wav 2> "MWCzai4b" * /;
-8413711    <ms% 350> "+" / <wav 2> "MNTpou1a" * /;
+1312228    <ms% 350> "+" / <wav 2> "MWVguo4b" * /;
+1412118    <ms% 350> "+" / <wav 2> "MWCban4a" * /;
-8313621    <ms% 350> "+" / <wav 2> "MNTfia2b" * /;
-8314218    <ms% 350> "+" / <wav 2> "MNCcou1a" * /;
+1312428    <ms% 350> "+" / <wav 2> "MWCsha3b" * /;
-8413121    <ms% 350> "+" / <wav 2> "MNThin2b" * /;
+1021659    <ms% 350> "+" / <wav 2> "DMWxin1" * /;
-8028959    <ms% 350> "+" / <wav 2> "DMNfie2" * /;
-8413721    <ms% 350> "+" / <wav 2> "MNTpou1b" * /;
-8314628    <ms% 350> "+" / <wav 2> "MNVsou1b" * /;
-8023759    <ms% 350> "+" / <wav 2> "DMNbuan3" * /;
```

```
-8314228    <ms% 350> "+" / <wav 2> "MNCfou1b" * /;
+1412128    <ms% 350> "+" / <wav 2> "MWCban4b" * /;

0 <line 0> "Take a break. Press SPACEBAR when ready.";

999 <ms% 2000> "+" /;

-8020259    <ms% 350> "+" / <wav 2> "DDNbe3" * /;
-8029759    <ms% 350> "+" / <wav 2> "DMNden3" * /;
-8414318    <ms% 350> "+" / <wav 2> "MNCsiu4a" * /;
-8021159    <ms% 350> "+" / <wav 2> "DDNnui2" * /;
-8021259    <ms% 350> "+" / <wav 2> "DDNdon2" * /;
+1412718    <ms% 350> "+" / <wav 2> "MWVda4a" * /;
+1022559    <ms% 350> "+" / <wav 2> "DDWshan1" * /;
-8313811    <ms% 350> "+" / <wav 2> "MNTtua2a" * /;
+1025559    <ms% 350> "+" / <wav 2> "DMWnan2" * /;
+1027859    <ms% 350> "+" / <wav 2> "DDWhuo3" * /;
+1025159    <ms% 350> "+" / <wav 2> "DMWshan1" * /;
-8414328    <ms% 350> "+" / <wav 2> "MNCsiu4b" * /;
+1025359    <ms% 350> "+" / <wav 2> "DMWxin1" * /;
-8020859    <ms% 350> "+" / <wav 2> "DDNpe4" * /;
+1026059    <ms% 350> "+" / <wav 2> "DMWzhi3" * /;
-8024059    <ms% 350> "+" / <wav 2> "DMNfai3" * /;
-8313821    <ms% 350> "+" / <wav 2> "MNTtua1b" * /;
-8028459    <ms% 350> "+" / <wav 2> "DMNmun1" * /;
+1028259    <ms% 350> "+" / <wav 2> "DDWdinn4" * /;
-8023559    <ms% 350> "+" / <wav 2> "DDNfun4" * /;
-8029159    <ms% 350> "+" / <wav 2> "DDNfun4" * /;
+1411711    <ms% 350> "+" / <wav 2> "MWTqia1a" * /;
+1412728    <ms% 350> "+" / <wav 2> "MWVda4b" * /;
+1028059    <ms% 350> "+" / <wav 2> "DDWyan3" * /;
-8020559    <ms% 350> "+" / <wav 2> "DDNlen4" * /;
-8024559    <ms% 350> "+" / <wav 2> "DDNnui2" * /;
-8023359    <ms% 350> "+" / <wav 2> "DMNtiu1" * /;
+1021759    <ms% 350> "+" / <wav 2> "DMWnan2" * /;
+1026159    <ms% 350> "+" / <wav 2> "DMWshui3" * /;
+1026759    <ms% 350> "+" / <wav 2> "DDWzhon1" * /;
-8020059    <ms% 350> "+" / <wav 2> "DMNfi3" * /;
+1022859    <ms% 350> "+" / <wav 2> "DDWlao3" * /;
+1027959    <ms% 350> "+" / <wav 2> "DDWgan3" * /;
+1027759    <ms% 350> "+" / <wav 2> "DDWlao3" * /;
-8028359    <ms% 350> "+" / <wav 2> "DMNfe1" * /;
+1411721    <ms% 350> "+" / <wav 2> "MWTqia1b" * /;


0 <line -2> "The End...thank you for participating", ?<line 0> "Please call
the experimenter!", <line 2> "[Experimenter: Press Esc to Save the data]";
```

L. Praat Script used in Phonological Proficiency Task used in Current Study

"ooTextFile"

"ExperimentMFC 6"

blank while playing? <no>

stimuli are sounds? <yes>

"" ".wav"

carrier phrase "" ""

initial silence duration 0.5 seconds

medial silence duration 0 seconds

final silence duration 0 seconds

69 different stimuli

```
"L005N" ""
"L005Q" ""
"L005R" ""
"L025N" ""
"L025Q" ""
"L025R" ""
"L032N" ""
"L032Q" ""
"L032R" ""
"L035N" ""
"L035Q" ""
"L035R" ""
"L039N" ""
"L039Q" ""
"L039R" ""
"L054N" ""
"L054Q" ""
"L054R" ""
"L064N" ""
"L064Q" ""
"L064R" ""
"L080N" ""
"L080Q" ""
```

"L080R" ""
"L083N" ""
"L083Q" ""
"L083R" ""
"L087N" ""
"L087Q" ""
"L087R" ""
"L098N" ""
"L098Q" ""
"L098R" ""
"L100N" ""
"L100Q" ""
"L100R" ""
"L106N" ""
"L106Q" ""
"L106R" ""
"L109N" ""
"L109Q" ""
"L109R" ""
"L115N" ""
"L115Q" ""
"L115R" ""
"L117N" ""
"L117Q" ""
"L117R" ""
"L118N" ""
"L118Q" ""
"L118R" ""
"L119N" ""
"L119Q" ""
"L119R" ""
"L120N" ""
"L120Q" ""
"L120R" ""
"L121N" ""
"L121Q" ""
"L121R" ""
"L122N" ""
"L122Q" ""
"L122R" ""
"L123N" ""
"L123Q" ""
"L123R" ""
"L135N" ""
"L135Q" ""
"L135R" ""

1 replications per stimulus

break after every 69 stimuli

<PermuteBalancedNoDoublets>

"Click to start rating samples"

"How native-like is the pronunciation?

Give your rating on a scale from 1 (poor) to 7 (native-like)."

"You can have a short break if you like. Click to proceed."

"That finishes all the samples to rate. Thank you very much!"

1000 replays

replay button 0.3 0.7 0.55 0.7 "Click here to play the sound again" ""

ok button 0 0 0 0 "" ""

oops button 0.3 0.7 0.05 0.15 "Oops! Click here to change answer" ""

responses are sounds? <no> "" "" "" "" 0 0 0

7 response categories

0.111 0.222 0.75 0.85 "1 (poor)" 24 "" "1"

0.222 0.333 0.75 0.85 "2" 24 "" "2"

0.333 0.444 0.75 0.85 "3" 24 "" "3"

0.444 0.555 0.75 0.85 "4" 24 "" "4"

0.555 0.666 0.75 0.85 "5" 24 "" "5"

0.666 0.777 0.75 0.85 "6" 24 "" "6"

0.777 0.888 0.75 0.85 "7" 24 "" "7"

0 goodness categories

M. Questions and Reading for Phonological Proficiency Task

Simplified script 简体字

Please introduce yourself.
请介绍你自己。
Qǐng jièshào nǐ zìjǐ

What did you do last weekend?
你上个周末做了什么？
Nǐ shàng gè zhōumò zuòle shénme?

What do you do every day?
你每天都做什么？
Nǐ měitiān dōu zuò shénme?

Why do you study Mandarin?
你为什么学习中文？
Nǐ wèishéme xuéxí zhōngwén?

Please count from 1 to 20.
请从一数到20。
Qǐng cóng yī shǔ dào 20.

Please read these sentences outloud:
请大声读出这些句子
Qǐng dàshēng dú chū zhèxiē jùzi

Pinyin

Nǐhǎo. Wǒ de míngzi shì Dàwèi. Wǒ shì cóng měiguó lái de. Xiànzài wǒ shì dàxuéshēng. Wǒ māma shì yīshēng. Wǒ bàba shì lǎoshī. Wǒ yě yǒu yīgè mèimei. Tā hái zài gāozhōng dú shū. Míngnián, wǒ yào qù Táiwān xué Zhōngwén. Bìyè yǐhòu, wǒ yào zài Zhōngguó zuò shēngyi

你好。我的名字是大卫。我是从美国来的。现在我是大学生。我母是医生。我父是老师。我也有一个妹妹。她还在高中读书。明年，我要去台湾学中文。毕业以后，我要在中国做生意。

Traditional script 繁體字

Please introduce yourself.
請介紹你自己。
Qǐng jièshào nǐ zìjǐ

What did you do last weekend?

你上個週末做了什麼？。
Nǐ shàng gè zhōumò zuòle shénme?

What do you do every day?
你每天都做什麼？
Nǐ měitiān dōu zuò shénme?

Why do you study Mandarin?
你為什麼學習中文？
Nǐ wèishéme xuéxí zhōngwén?

Please count from 1 to 20.
請從一數到20。
Qǐng cóng yī shǔ dào 20.

Please read these sentences outloud:
請大聲讀出這些句子
Qǐng dàshēng dú chū zhèxiē jùzi

Pinyin
Nǐhǎo. Wǒ de míngzi shì Dàwèi. Wǒ shì cóng měiguó lái de. Xiànzài wǒ shì dàxuéshēng. Wǒ māma shì yīshēng. Wǒ bàba shì lǎoshī. Wǒ yě yǒu yīgè mèimei. Tā hái zài gāozhōng dú shū. Míngnián, wǒ yào qù Táiwān xué Zhōngwén. Bìyè yǐhòu, wǒ yào zài Zhōngguó zuò shēngyi

你好。我的名字是大衛。我是從美國來的。現在我是大學生。我媽媽是醫生。我爸爸是老師。我也有一個妹妹。她還在高中讀書。明年，我要去台灣學中文。畢業以後，我要在中國做生意。

Translation
Hello. My name is David. I am from the US. I am now a university student. My mother is a doctor. My father is a teacher. I also have a younger sister. She still studies in high school. Next year I want to go to Taiwan to study Mandarin. After I graduate, I want to do business in China.

N. Recruitment flyer samples

Earn $10 and my gratitude. Listen to some cool foreign language sounds!

Who?

Anyone who is a native speaker of Japanese or Korean or Thai AND does not also speak natively or with high proficiency another Asian language (e.g., Japanese, Korean, Mandarin, Cantonese, Thai or Vietnamese). You must also be an undergraduate or graduate student at Indiana University. I can test two people at a time, so bring a friend!

What?

You will listen to some foreign sounds and fill out a questionnaire about your language background.

Where?

Second Language Psycholinguistics Laboratory at **Weatherly Hall Room 220** (at 400 N. Sunrise Drive, north of the School of Education Building and a little southeast of Teter Quadrangle)

When?

You can come at almost any time, including the weekend. Please make an appointment via email ahead of time. I am testing all summer. It takes in total 30-35 minutes.

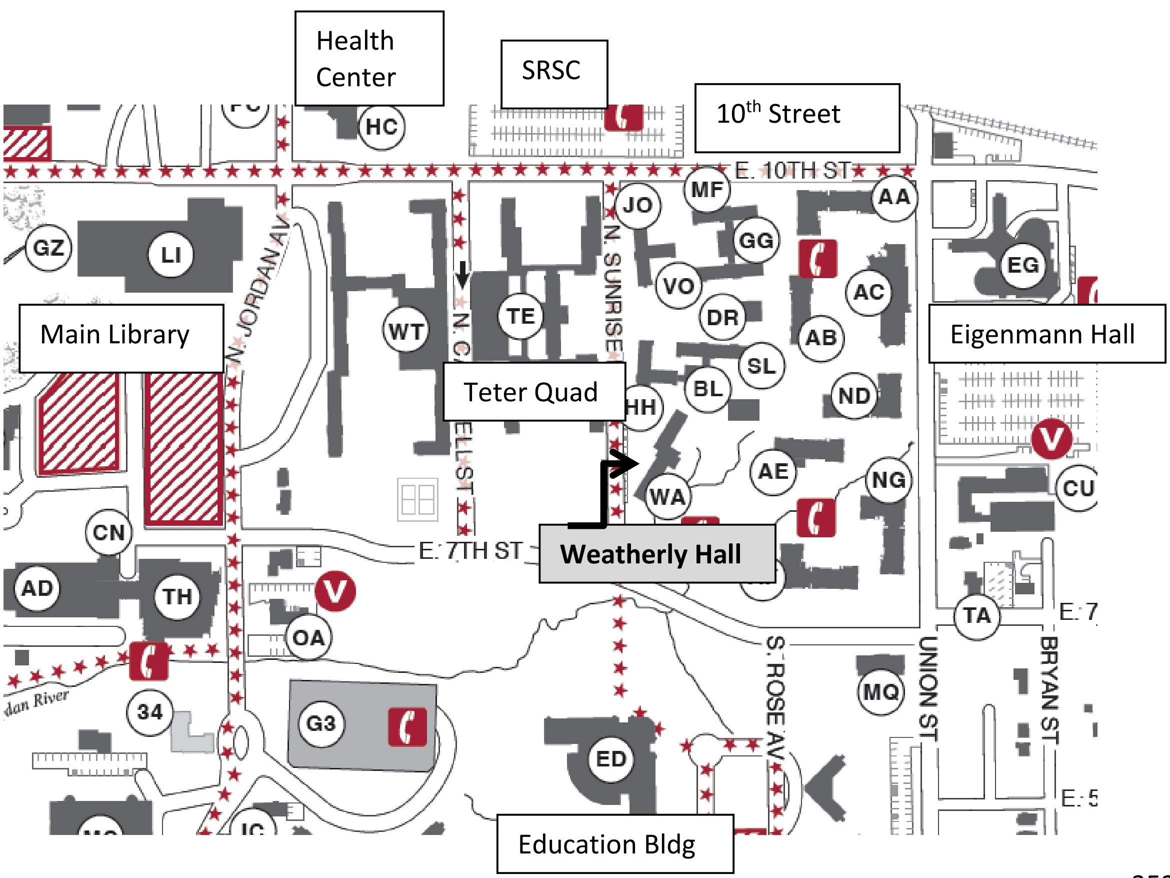
Why?

To study how people hear foreign language sounds.

Contact:

If you wish to participate, please contact Vance Schaefer at vkschaef@indiana.edu

Thank you very much.



Earn $20 and my gratitude just by listening to Mandarin Chinese!

Who?

Anyone who is a native speaker of Mandarin. You must also be an undergraduate or graduate student at Indiana University. I can test two people at a time, so bring a friend!

What?

You will listen to some foreign language sounds, and Mandarin and fill out a questionnaire about your language background.

Where?

Second Language Psycholinguistics Laboratory at **Weatherly Hall Room 220** (at 400 N. Sunrise Drive, north of the School of Education Building and a little southeast of Teter Quadrangle)

When?

You can come at almost any time, including the weekend. Please make an appointment via email ahead of time. I am testing all summer. It takes in total 60-65 minutes.

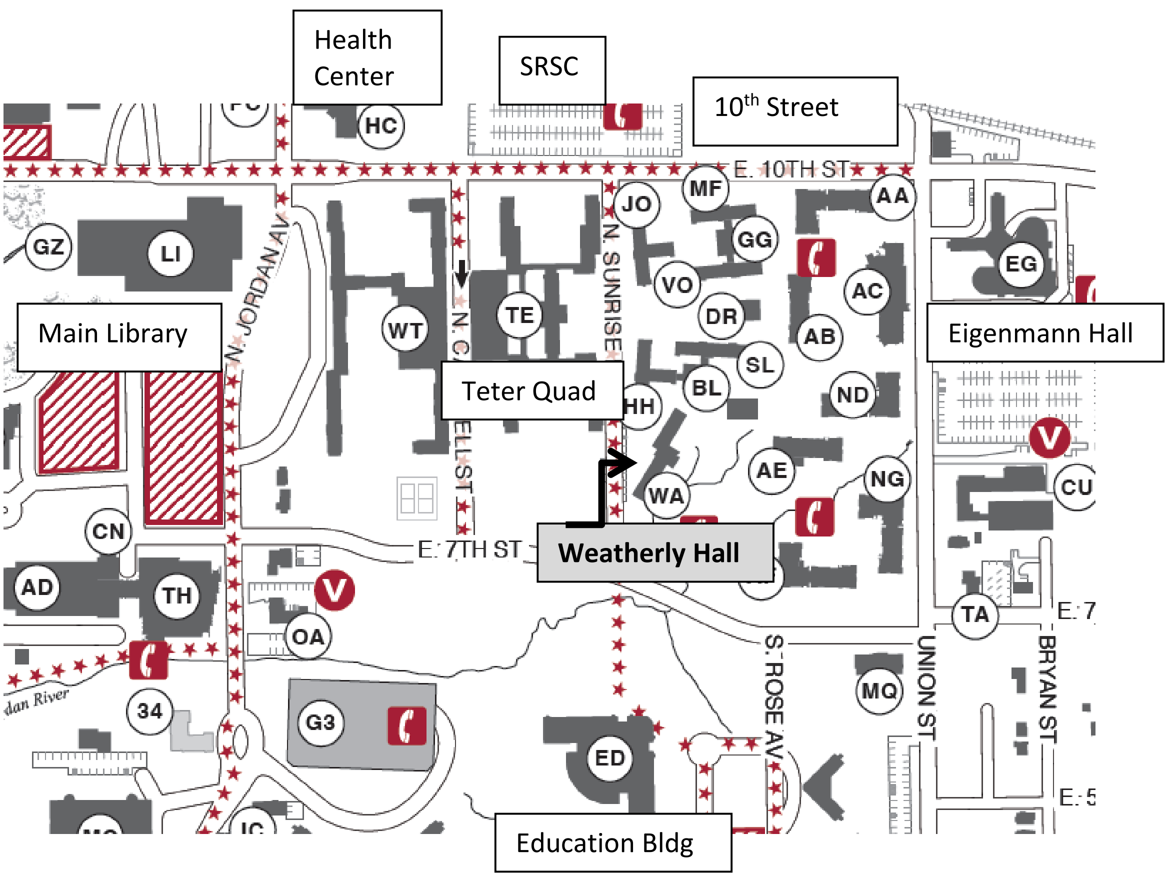
Why?

To study how people hear foreign language sounds.

Contact:

If you wish to participate, please contact Vance Schaefer at vkschaef@indiana.edu

Thank you very much.



O. Results for L2 Mandarin learners for three tasks of ABX monosyllabic, ABX disyllabic and Lexical decision tasks

Accuracy and Reaction Time Results for each L2 Mandarin Learner and L1 Mandarin Speaker for Mandarin Lexical Decision task with Repetition Priming Compared to Both the ABX Monosyllabic and Disyllabic Thai Tasks

	Lexical decision task							ABX monosyllabic task				ABX disyllabic task			
Subject (L=L2 Mandarin Learner; M = L1 Mandarin speaker)	Tone vs Segment	Minimal Pair vs Same for real words	Mean of 1st reaction time (ms)	Mean of 2nd reaction time (ms)	Difference between 1st and 2nd reaction times	Real word vs non-word accuracy score (%)	List*	Tone vs Segment	Accuracy (%)	Reaction time (ms)		Tone vs Segment	Accuracy (%)	Reaction time (ms)	
L005	Tone	Minimal Pair	1112.86	952.22	160.64	89.4	1RB	tone	68.8	correct	1319.46	tone	80	correct	1323.37
		Same	1270.30	1247.86	22.44					incorrect	2165.07			incorrect	2246.41
	Segment	Minimal Pair	1181.80	1176.69	5.11			segment	87.5	correct	1057.84	segment	95	correct	1371.48
		Same	1211.37	1066.63	144.74					incorrect	2173.33			incorrect	2500.00
L025	Tone	Minimal Pair	1161.80	1281.09	-119.29	91.3	2RB	tone	78.8	correct	999.56	tone	80	correct	1451.09
		Same	914.65	923.39	-8.74					incorrect	1440.95			incorrect	1708.76
	Segment	Minimal Pair	1030.81	1104.92	-8.74			segment	95	correct	787.74	segment	92.5	correct	993.21
		Same	998.55	935.23	63.32					incorrect	1520.53			incorrect	784.02
L032	Tone	Minimal Pair	1355.15	1328.85	26.3	95.6	3RB	tone	35	correct	1494.26	tone	47.5	correct	1549.31
		Same	1417.38	1200.56	216.82					incorrect	1418.54			incorrect	1640.07
	Segment	Minimal Pair	1415.09	1340.76	74.33			segment	27.5	correct	1278.23	segment	87.5	correct	1533.38
		Same	1455.66	1286.71	168.95					incorrect	1308.93			incorrect	2174.98
L035	Tone	Minimal Pair	960.38	1053.04	-92.66	92.5	4RB	tone	95	correct	1126.71	tone	100	correct	1006.16
		Same	962.89	782.20	180.69					incorrect	1884.90			incorrect	NA
	Segment	Minimal Pair	985.80	915.59	70.21			segment	100	correct	732.92	segment	100	correct	824.30
		Same	971.96	846.95	125.01					incorrect	NA			incorrect	NA
L039	Tone	Minimal Pair	1106.90	1106.28	0.62	93.1	1RB	tone	80	correct	1353.19	tone	87.5	correct	1348.89
		Same	1192.86	1150.48	42.38					incorrect	1812.20			incorrect	1584.08
	Segment	Minimal Pair	1137.44	1116.43	21.01			segment	98.8	correct	1007.29	segment	100	correct	1147.30
		Same	1049.16	1080.45	-31.29					incorrect	2500.00			incorrect	NA
L054	Tone	Minimal Pair	1078.99	1158.73	-79.74	96.3	2RB	tone	90	correct	1003.26	tone	87.5	correct	1212.66
		Same	917.20	753.94	163.26					incorrect	1638.86			incorrect	1763.94
	Segment	Minimal Pair	956.87	918.90	37.97			segment	97.5	correct	763.68	segment	97.5	correct	928.19
		Same	937.06	828.35	108.71					incorrect	1269.48			incorrect	1084.22
L064	Tone	Minimal Pair	1021.99	995.69	26.3	88.8	3RB	tone	86.3	correct	1031.99	tone	92.5	correct	1260.46
		Same	1394.35	926.83	467.52					incorrect	1505.65			incorrect	1265.47

	Segment	Minimal Pair	1142.15	1115.79	26.36			segment	93.8	correct	882.49	segment	90	correct	1211.12
		Same	1108.98	1031.99	76.99					incorrect	1433.21			incorrect	2075.85
L080	Tone	Minimal Pair	1333.75	1589.55	-255.8	93.1	4RB	tone	98.8	correct	1014.90	tone	90	correct	1496.54
		Same	1299.74	1178.99	120.75					incorrect	878.75			incorrect	2073.99
	Segment	Minimal Pair	1423.51	1521.21	-97.7			segment	98.8	correct	923.44	segment	95	correct	1393.46
		Same	1387.13	1424.88	-37.75					incorrect	946.92			incorrect	1937.76
L087	Tone	Minimal Pair	1639.25	1511.05	128.2	93.8	2RB	tone	90	correct	1324.63	tone	97.5	correct	1357.18
		Same	1276.74	1317.44	-40.7					incorrect	2204.34			incorrect	2022.74
	Segment	Minimal Pair	1545.05	1395.04	150.01			segment	100	correct	1154.79	segment	100	correct	1137.99
		Same	1286.79	1180.62	106.17					incorrect	NA			incorrect	NA
L100	Tone	Minimal Pair	1365.35	1421.87	-56.52	92.5	4LB	tone	82.5	correct	1324.63	tone	77.5	correct	1193.59
		Same	950.19	873.39	76.8					incorrect	2204.34			incorrect	1393.62
	Segment	Minimal Pair	953.10	922.80	30.3			segment	95	correct	866.01	segment	95	correct	835.41
		Same	860.10	864.53	-4.43					incorrect	1491.50			incorrect	1761.78
L106	Tone	Minimal Pair	1122.25	1149.63	-27.38	89.9	1RB	tone	87.5	correct	1240.65	tone	85	correct	1337.64
		Same	1230.68	1129.14	101.54					incorrect	1814.94			incorrect	2119.15
	Segment	Minimal Pair	1158.03	1172.96	-14.93			segment	93.8	correct	1043.91	segment	100	correct	1248.66
		Same	1164.28	1067.54	96.74					incorrect	2210.12			incorrect	NA
L109	Tone	Minimal Pair	1809.74	1725.44	84.3	91.9	2RB	tone	95	correct	1052.28	tone	100	correct	1325.81
		Same	1377.93	1334.64	43.29					incorrect	1914.88			incorrect	NA
	Segment	Minimal Pair	1559.17	1557.16	2.01			segment	98.8	correct	922.45	segment	97.5	correct	966.83
		Same	1558.21	1267.05	291.16					incorrect	642.08			incorrect	2500.00
L118	Tone	Minimal Pair	1099.40	1014.72	84.68	94.4	1LB	tone	83.8	correct	1198.39	tone	80	correct	1389.19
		Same	1275.14	1136.26	138.88					incorrect	1430.40			incorrect	1461.74
	Segment	Minimal Pair	1150.80	1153.71	-2.91			segment	95	correct	1063.80	segment	97.5	correct	1137.39
		Same	1258.90	1062.17	196.73					incorrect	1328.09			incorrect	1263.50
L119	Tone	Minimal Pair	1269.01	1025.50	243.51	91.3	2RB	tone	62.5	correct	1256.87	tone	80	correct	1586.39
		Same	961.37	1006.07	-44.7					incorrect	1341.00			incorrect	1921.58
	Segment	Minimal Pair	1236.76	1066.69	170.07			segment	95	correct	1203.29	segment	95	correct	1461.24
		Same	1049.62	1174.34	-124.72					incorrect	1434.14			incorrect	2233.92
L120	Tone	Minimal Pair	1090.14	1023.82	66.32	92.5	3RB	tone	91.3	correct	962.09	tone	82.5	correct	1233.43
		Same	1201.15	1009.60	191.55					incorrect	1534.55			incorrect	1794.97
	Segment	Minimal Pair	1180.78	942.47	238.31			segment	97.5	correct	738.57	segment	100	correct	849.01
		Same	970.65	949.62	21.03					incorrect	741.82			incorrect	NA
L122	Tone	Minimal Pair	874.99	1014.25	-139.26	94.4	1RB	tone	96.3	correct	880.83	tone	85	correct	1063.28
		Same	1103.90	873.41	230.49					incorrect	1093.61			incorrect	1169.97
	Segment	Minimal Pair	1021.44	1028.37	-6.93			segment	98.8	correct	680.55	segment	100	correct	806.84
		Same	968.09	1018.01	-49.92					incorrect	673.10			incorrect	NA
L123	Tone	Minimal Pair	1406.03	1438.26	-32.23	91.3	2RB	tone	96.3	correct	1031.37	tone	97.5	correct	1366.24
		Same	984.28	1072.96	-88.68					incorrect	2500.00			incorrect	2500.00
	Segment	Minimal Pair	1156.47	1198.38	-41.91			segment	91.3	correct	805.19	segment	100	correct	1049.07
		Same	900.34	1010.84	-110.5					incorrect	1846.38			incorrect	NA
L135	Tone	Minimal Pair	847.17	690.78	156.39	91.9	3RV	tone	83.8	correct	886.45	tone	85	correct	928.26

		Same	976.70	868.09	108.61					incorrect	938.89			incorrect	894.31
	Segment	Minimal Pair	880.04	831.32	48.72			segment	97.5	correct	747.98	segment	97.5	correct	805.70
		Same	1048.07	808.74	239.33					incorrect	1518.89			incorrect	475.03
M003	Tone	Minimal Pair	964.60	985.70	-21.1	95	1RV	tone	80	correct	1097.34	tone	85	correct	1347.81
		Same	967.29	909.61	57.68					incorrect	2249.00			incorrect	1968.49
	Segment	Minimal Pair	866.91	1111.94	-245.03			segment	92.5	correct	947.99	segment	95	correct	1043.85
		Same	1014.76	1356.00	-341.24					incorrect	2062.55			incorrect	2500.00
M004	Tone	Minimal Pair	1186.73	1162.14	24.59	88.1	2RB	tone	85	correct	944.01	tone	100	correct	1189.96
		Same	1101.79	929.05	172.74					incorrect	1518.89			incorrect	NA
	Segment	Minimal Pair	1118.12	893.10	225.02			segment	92.5	correct	961.40	segment	97.5	correct	1022.28
		Same	1140.12	1339.42	-199.3					incorrect	1774.34			incorrect	547.27
M007	Tone	Minimal Pair	1080.82	1052.60	28.22	95	2RB	tone	75	correct	1217.54	tone	90	correct	1271.57
		Same	1069.69	922.44	147.25					incorrect	1478.54			incorrect	1611.60
	Segment	Minimal Pair	1006.24	884.25	121.99			segment	98.8	correct	857.81	segment	100	correct	1100.10
		Same	884.53	1102.19	-217.66					incorrect	1089.74			incorrect	NA
M010	Tone	Minimal Pair	1536.02	1648.41	-112.39	96.9	3LV	tone	85	correct	1294.95	tone	82.5	correct	1482.48
		Same	1495.66	1049.74	445.92					incorrect	1996.89			incorrect	2292.91
	Segment	Minimal Pair	1107.90	1135.26	-27.36			segment	87.5	correct	1188.74	segment	100	correct	1447.90
		Same	1093.78	1566.47	-472.69					incorrect	1702.51			incorrect	NA
M011	Tone	Minimal Pair	1642.96	1483.64	159.32	93.1	4RB	tone	88.8	correct	1294.95	tone	97.5	correct	1320.87
		Same	1244.89	1124.18	120.71					incorrect	1996.89			incorrect	1509.03
	Segment	Minimal Pair	1324.24	1470.02	-145.78			segment	97.5	correct	962.31	segment	97.5	correct	1138.62
		Same	1504.38	1480.74	23.64					incorrect	1292.85			incorrect	2500.00
M012	Tone	Minimal Pair	912.77	1383.22	-470.45	94.4	1RB	tone	68.8	correct	1211.29	tone	80	correct	1356.47
		Same	1170.34	1175.91	-5.57					incorrect	1767.75			incorrect	1438.76
	Segment	Minimal Pair	1146.99	1105.94	41.05			segment	91.3	correct	1138.65	segment	92.5	correct	1237.95
		Same	1178.25	1145.55	32.7					incorrect	1942.85			incorrect	1588.73
M015	Tone	Minimal Pair	1316.90	1929.66	-612.76	90.6	2RB	tone	75	correct	1297.27	tone	60	correct	1583.57
		Same	1383.50	1188.26	195.24					incorrect	2200.14			incorrect	2337.84
	Segment	Minimal Pair	1369.18	1503.21	-134.03			segment	78.8	correct	1380.73	segment	95	correct	1367.30
		Same	1174.00	1256.24	-82.24					incorrect	2096.25			incorrect	2500.00
M024	Tone	Minimal Pair	919.23	1255.71	-336.48	97.5	4RB	tone	85	correct	1150.47	tone	70	correct	1334.29
		Same	947.34	1010.60	-63.26					incorrect	1408.50			incorrect	1488.98
	Segment	Minimal Pair	961.65	1099.43	-137.78			segment	95	correct	959.64	segment	90	correct	1242.06
		Same	1120.03	1008.35	111.68					incorrect	1635,65			incorrect	1896.34
M030	Tone	Minimal Pair	1006.95	1011.65	-4.7	98.8	4RB	tone	95	correct	1069.95	tone	100	correct	1287.62
		Same	1026.06	997.36	28.7					incorrect	2036.42			incorrect	NA
	Segment	Minimal Pair	926.11	907.90	18.21			segment	97.5	correct	954.71	segment	100	correct	1196.28
		Same	935.53	924.99	10.54					incorrect	1227.32			incorrect	NA
M033	Tone	Minimal Pair	966.66	949.90	16.76	96.9	2RB	tone	91.3	correct	1031.22	tone	97.5	correct	1303.02
		Same	1130.50	762.36	368.14					incorrect	1691.81			incorrect	2500.00
	Segment	Minimal Pair	931.21	829.22	101.99			segment	95	correct	905.81	segment	100	correct	1214.08
		Same	845.75	835.50	10.25					incorrect	1310.52			incorrect	NA

M036	Tone	Minimal Pair	1858.51	1388.01	470.5	77.4	3RB	tone	57.5	correct	1387.31	tone	37.5	correct	1634.76
		Same	1500.88	1319.67	181.21					incorrect	1809.79			incorrect	1746.12
	Segment	Minimal Pair	1341.12	1132.56	208.56			segment	67.5	correct	1217.31	segment	60	correct	1393.84
		Same	1702.97	1113.06	589.91					incorrect	2008.51			incorrect	1755.60
M041	Tone	Minimal Pair	1001.61	1127.28	-125.67	95	3RB	tone	97.5	correct	835.45	tone	90	correct	1112.44
		Same	986.73	867.25	119.48					incorrect	1212.11			incorrect	1261.23
	Segment	Minimal Pair	848.10	893.76	-45.66			segment	97.5	correct	807.94	segment	100	correct	1004.32
		Same	981.27	816.28	164.99					incorrect	672.88			incorrect	NA
M046	Tone	Minimal Pair	1098.10	1187.05	-88.95	91.3	1RB	tone	78.8	correct	1249.89	tone	87.5	correct	1353.51
		Same	987.33	1012.85	-25.52					incorrect	1510.06			incorrect	1504.88
	Segment	Minimal Pair	1015.19	1254.07	-238.88			segment	87.5	correct	1197.58	segment	97.5	correct	1321.75
		Same	1063.09	1014.33	48.76					incorrect	1265.25			incorrect	1461.01
M047	Tone	Minimal Pair	1432.06	1306.03	126.03	92.5	2RV	tone	78.8	correct	1126.67	tone	77.5	correct	1246.87
		Same	1331.29	1366.19	-34.9					incorrect	2225.01			incorrect	2197.03
	Segment	Minimal Pair	1053.91	1223.04	-169.13			segment	97.5	correct	959.69	segment	97.5	correct	1187.76
		Same	975.09	1231.82	-256.73					incorrect	2216.21			incorrect	2500.00
M049	Tone	Minimal Pair	1028.50	954.21	74.29	96.9	3RB	tone	97.5	correct	1252.00	tone	92.5	correct	1472.81
		Same	1012.68	917.19	95.49					incorrect	1634.55			incorrect	1561.81
	Segment	Minimal Pair	1056.67	1034.98	21.69			segment	93.8	correct	1151.58	segment	100	correct	1356.17
		Same	1032.79	942.00	90.79					incorrect	1432.17			incorrect	NA
M051	Tone	Minimal Pair	970.06	1174.33	-204.27	90.6	4LV	tone	96.3	correct	942.67	tone	97.5	correct	1205.79
		Same	1153.86	887.28	266.58					incorrect	2456.56			incorrect	2500.00
	Segment	Minimal Pair	974.25	788.67	185.58			segment	98.8	correct	867.92	segment	100	correct	1195.92
		Same	956.98	1051.17	-94.19					incorrect	652.33			incorrect	NA
M052	Tone	Minimal Pair	916.77	1168.13	-251.36	95.6	1RB	tone	88.8	correct	1031.78	tone	82.5	correct	1364.70
		Same	1321.70	1039.22	282.48					incorrect	1411.38			incorrect	2117.97
	Segment	Minimal Pair	967.36	998.24	-30.88			segment	97.5	correct	1032.64	segment	87.5	correct	1284.88
		Same	1301.89	941.15	360.74					incorrect	1086.16			incorrect	2109.15
M055	Tone	Minimal Pair	1617.01	1116.78	500.23	95	2RB	tone	76.3	correct	1242.52	tone	70	correct	1552.31
		Same	1027.54	992.26	35.28					incorrect	1893.91			incorrect	1453.01
	Segment	Minimal Pair	932.05	906.03	26.02			segment	91.3	correct	987.73	segment	95	correct	1174.86
		Same	1007.83	774.37	233.46					incorrect	1572.75			incorrect	2253.61
M058	Tone	Minimal Pair	897.63	921.25	-23.62	95.6	3RV	tone	92.5	correct	1007.90	tone	92.5	correct	1134.75
		Same	1220.68	1010.74	209.94					incorrect	1281.65			incorrect	1313.47
	Segment	Minimal Pair	935.98	1063.35	-127.37			segment	97.5	correct	825.81	segment	100	correct	1039.07
		Same	1038.38	959.33	79.05					incorrect	1488.60			incorrect	NA
M059	Tone	Minimal Pair	1349.68	1166.32	183.36	89.4	4RB	tone	95	correct	1029.21	tone	65	correct	1216.35
		Same	1319.49	1190.69	128.8					incorrect	1489.01			incorrect	1785.82
	Segment	Minimal Pair	1312.84	1166.61	146.23			segment	97.5	correct	878.18	segment	90	correct	1079.25
		Same	1192.79	1101.22	91.57					incorrect	1875.50			incorrect	2110.30
M063	Tone	Minimal Pair	1038.58	1091.64	-53.06	96.3	1RV	tone	77.5	correct	1168.84	tone	87.5	correct	1122.92
		Same	1079.55	1035.07	44.48					incorrect	1744.13			incorrect	1716.97
	Segment	Minimal Pair	1011.51	1025.01	-13.5			segment	85	correct	1035.73	segment	100	correct	1073.76

		Same	1014.08	1223.28	-209.2					incorrect	1836.43			incorrect	NA
M065	Tone	Minimal Pair	1520.59	1263.68	256.91	96.9	2RV	tone	86.3	correct	1149.35	tone	87.5	correct	1363.15
		Same	1280.36	1567.26	-286.9					incorrect	1857.07			incorrect	2087.53
	Segment	Minimal Pair	1466.53	1367.14	99.39			segment	97.5	correct	980.11	segment	100	correct	1200.57
		Same	1232.28	1360.92	-128.64					incorrect	960.21			incorrect	NA
M066	Tone	Minimal Pair	1177.94	1183.61	-5.67	98.8	3RB	tone	87.5	correct	1416.81	tone	97.5	correct	1368.88
		Same	1142.44	1230.92	-88.48					incorrect	2247.29			incorrect	1683.32
	Segment	Minimal Pair	1297.25	1214.90	82.35			segment	95	correct	1145.50	segment	100	correct	1227.54
		Same	1348.61	1227.66	120.95					incorrect	2249.07			incorrect	NA
M067	Tone	Minimal Pair	1254.39	1113.14	141.25	92.5	4RV	tone	91.3	correct	1072.13	tone	92.5	correct	1114.55
		Same	877.35	1238.79	-361.44					incorrect	2457.13			incorrect	2131.45
	Segment	Minimal Pair	982.95	878.97	103.98			segment	98.8	correct	1001.73	segment	100	correct	1159.60
		Same	824.83	831.10	-6.27					incorrect	2500.00			incorrect	NA
M068	Tone	Minimal Pair	878.99	966.33	-87.34	95	1RB	tone	87.5	correct	979.31	tone	92.5	correct	1266.27
		Same	1072.67	939.15	133.52					incorrect	1087.78			incorrect	1615.67
	Segment	Minimal Pair	790.88	857.81	-66.93			segment	98.8	correct	829.98	segment	97.5	correct	1148.07
		Same	899.17	934.44	-35.27					incorrect	798.16			incorrect	1207.09
M069	Tone	Minimal Pair	1249.45	1311.37	-61.92	95	2RV	tone	77.5	correct	1341.35	tone	77.5	correct	1399.54
		Same	995.50	986.37	9.13					incorrect	2017.87			incorrect	2255.20
	Segment	Minimal Pair	1472.88	920.62	552.26			segment	90	correct	1034.56	segment	95	correct	1234.87
		Same	1095.42	1329.17	-233.75					incorrect	1737.24			incorrect	2500.00
M070	Tone	Minimal Pair	1319.83	1161.50	158.33	96.3	3RB	tone	98.8	correct	887.85	tone	95	correct	1192.07
		Same	1110.78	997.83	112.95					incorrect	930.79			incorrect	2105.31
	Segment	Minimal Pair	1303.72	1141.34	162.38			segment	98.8	correct	755.03	segment	100	correct	992.90
		Same	1463.74	1194.75	268.99					incorrect	454.43			incorrect	NA
M071	Tone	Minimal Pair	1194.60	1501.33	-306.73	95.6	4RV	tone	77.5	correct	1256.69	tone	87.5	correct	1584.19
		Same	1228.72	1106.19	122.53					incorrect	1950.54			incorrect	2206.23
	Segment	Minimal Pair	1468.99	899.58	569.41			segment	96.3	correct	1058.59	segment	92.5	correct	1388.77
		Same	1106.03	1028.29	77.74					incorrect	2185.53			incorrect	2500.00
M072	Tone	Minimal Pair	1173.41	1092.77	80.64	96.9	1RB	tone	90	correct	1314.87	tone	90	correct	1508.07
		Same	1062.26	1023.81	38.45					incorrect	2086.14			incorrect	2100.63
	Segment	Minimal Pair	1156.86	1303.76	-146.9			segment	97.5	correct	1206.11	segment	92.5	correct	1363.53
		Same	1204.05	1307.33	-103.28					incorrect	1172.59			incorrect	2449.57
M075	Tone	Minimal Pair	1231.46	1111.26	120.2	88.8	2RB	tone	66.3	correct	1239.41	tone	57.5	correct	1576.09
		Same	1254.31	915.92	338.39					incorrect	1409.88			incorrect	1480.00
	Segment	Minimal Pair	1020.64	1178.63	-157.99			segment	80	correct	1065.56	segment	77.5	correct	1342.91
		Same	794.42	900.82	-106.4					incorrect	1218.89			incorrect	1548.19
M090	Tone	Minimal Pair	913.94	899.66	14.28	91.9	3RB	tone	78.8	correct	1091.80	tone	65	correct	1319.96
		Same	1151.08	1031.61	119.47					incorrect	2163.78			incorrect	2191.87
	Segment	Minimal Pair	960.71	1142.47	-181.76			segment	86.3	correct	907.88	segment	92.5	correct	1115.74
		Same	1113.07	968.10	144.97					incorrect	2249.29			incorrect	2005.65

*List: R=right-handed; L=Left-handed; B=computer 1; V=computer 2

P. Score report sample sent to L2 Mandarin Learners

Score report for Learners of Mandarin as Second Language (L2)

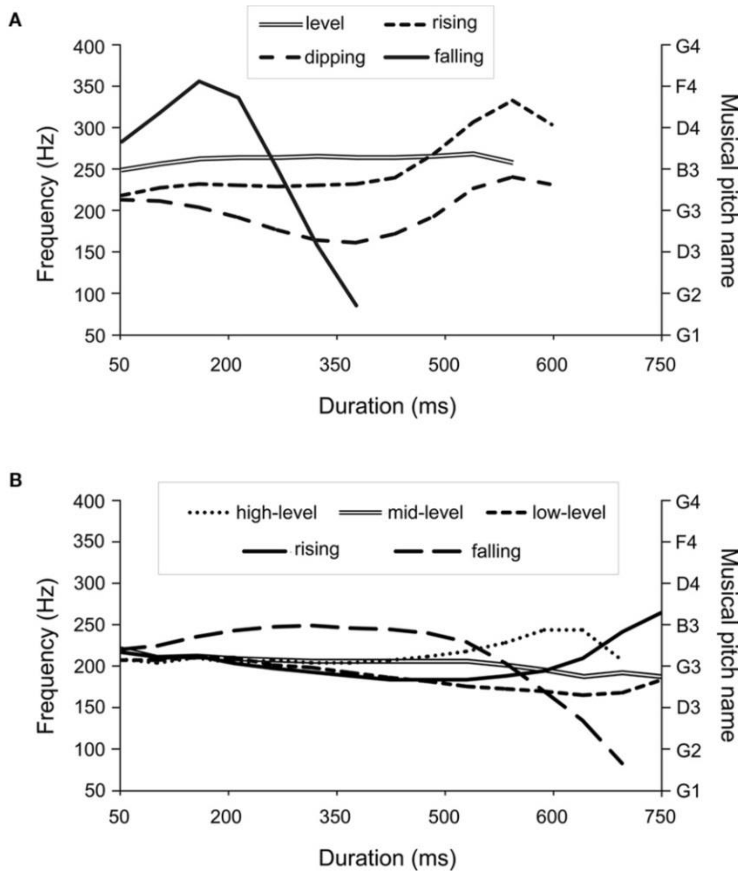
Name: xxx email: xxx

Thank you again for participating in my study. My apologies for this late report of your scores, but it has only been recently that I have finished analyzing the data.

You kindly participated in four tasks:

A. Matching similar monosyllabic stimuli featuring made-up words based on the Thai language comparing either tones or vowels/consonants.
B. Matching similar disyllabic stimuli featuring made-up words based on the Thai language comparing either tones or vowels/consonants.
C. Deciding whether stimuli were real words or made-up words in Mandarin Chinese.
D. Answering questions, counting to 20 and reading a passage in Mandarin Chinese, which were recorded in a soundbooth.

A comparison of Mandarin tones and Thai tone as spoken by a female in each language pronouncing [ma].



Source: Tillman, B., Burnham, D., Nguyen, S., Grimault, N., Gosselin, N., & Peretz, I. (2011). Congenital Amusia (or tone deafness) interferes with pitch processing in tone languages. *Frontiers in Psychology Vol. 20*: 1-15.

1) <u>ABX monosyllabic task:</u>

You listened to made-up words composed of one syllable using Thai tones. You heard three words in a row and determined whether the second word you heard of three was similar to the first word or the third word. The words either differed by tone (e.g., [naa] rising tone vs [naa] mid tone vs [naa] mid tone) or by vowel/consonant (e.g., [naa] mid tone vs [naa] mid tone vs [maa] mid tone).

We measured your accuracy rates and your reaction times in matching the same words. For Mandarin learners, we expect higher accuracy rates for the words differing by one vowel or consonant than for the words differing by tone as vowels/consonants should be easier to perceive. We also expect faster reaction times for words differing by one vowel or consonant than for words differing by tone.

Also, on the tone comparisons, we expect L2 Mandarin learners to outperform L1 English speakers who have not learned Mandarin or another tonal language. Looking at your performance on the individual tone comparisons may show you which tones in Thai and possibly in Mandarin are more difficult/easier for you to perceive or distinguish from another.

Results for the Monosyllabic task using Thai tones on non-words

Target	Your scores		L2 Mandarin Learner group average		L1 Mandarin Native-Speaker group average		L1 English Monolingual group average	
	Accuracy (%)	RT (ms)	Accuracy (%) (SD)	RT (ms) (CI)	Accuracy (%)(SD)	RT (ms) (CI)	Accuracy (%)(SD)	RT (ms) (CI)
Overall tone	00	000	86.3 (2.3)	1053 (989; 1120)	85 (1.6)	1088 (1040; 1137)	79.5 (2.6)	1170 (1108; 1236)
Overall segments (vowel or consonant differences)	00	000	96.1 (0.8)	855 (790; 926)	93.6 (1.0)	951 (906; 997)	93.5 (1.4)	991 (923; 1063)
Individual tone comparisons								
Falling vs Rising	00	000	92.7 (2.1)	948 (891; 1009)	94.2 (1.4)	960 (890; 1037)	90.6 (2.6)	1064 (988; 1146)
Low vs Falling	00	000	93.4 (2.4)	988 (929; 1052)	90.4 (2.2)	949 (878; 1025)	86.5 (3.7)	1071 (994; 1154)

Low vs Rising	00	000	78.7 (5.5)	1170 (1097; 1247)	75.0 (3.2)	1206 (1114; 1304)	64.6 (4.3)	1303 (1205; 1410)
Mid vs Falling	00	000	85.3 (2.8)	1091 (1023; 1164)	86.3 (2.5)	1018 (929; 1116)	83.9 (3.4)	1146 (1064; 1236)
Mid vs Rising	00	000	89.7 (2.4)	1122 (1052; 1194)	92.5 (1.7)	1070 (994; 1151)	84.9 (3.0)	1211 (1124; 1305)
High vs Falling	00	000	91.9 (2.5)	1050 (986; 1119)	92.1 (2.1)	1020 (943; 1103)	80.7 (4.1)	1102 (1022; 1189)
High vs Rising	00	000	89.0 (3.1)	1047 (982; 1114)	85.0 (2.9)	1025 (959; 1096)	77.6 (4.2)	1139 (1055; 1228)
Low vs Mid	00	000	77.2 (4.8)	1202 (1125; 1285)	71.7 (3.6)	1156 (1076; 1242)	72.4 (4.4)	1255 (1162; 1355)
Low vs High	00	000	89.0 (3.1)	1099 (1033; 1172)	84.6 (2.8)	1020 (953; 1092)	79.7 (4.4)	1146 (1063; 1236)
Mid vs High	00	000	75.7 (4.2)	1180 (1107; 1259)	77.9 (3.2)	1202 (1118; 1292)	74.5 (3.2)	1373 (1272; 1482)
Individual segmental comparisons								
Falling	00	000	97.8 (1.0)	951 (897; 1005)	94.2 (1.4)	853 (783; 929)	93.5 (1.7)	985 (920; 1054)
Rising	00	000	94.5 (1.6)	940 (887; 995)	92.3 (1.4)	856 (795; 921)	93.5 (1.5)	999 (933; 1069)

Low	00	000	97.1 (0.9)	938 (885; 993)	93.3 (1.3)	870 (788; 960)	93.2 (1.6)	963 (900; 1031)
Mid	00	000	94.5 (1.3)	953 (900; 1007)	94.4 (1.0)	850 (785; 920)	92.7 (2.0)	1029 (961; 1101)
High	00	000	96.7 (1.2)	951 (897; 1005)	94.0 (1.4)	850 (784; 921)	94.8 (1.4)	979 (915; 1047)

Note: SE = standard error.CI = Confidence interval (lower; upper).
RTs appear to be "rounded up" as the RTs were logged during calculation.

2) <u>ABX disyllabic task:</u>

You listened to made-up words composed of two syllables using Thai tones. You heard three words in a row and determined whether the second word you heard of three was similar to the first word or the third word. The first or second syllable of the word either differed by tone (e.g., [naa] rising tone vs [naa] mid tone vs [naa] mid tone) or by vowel/consonant (e.g., [naa] mid tone vs [naa] mid tone vs [maa] mid tone).

Again, we measured your accuracy rates and your reaction times in matching the same words. For Mandarin learners, we expect higher accuracy rates for the words differing by one vowel or consonant than for the words differing by tone as vowels/consonants should be easier to perceive. We also expect faster reaction times for words differing by one vowel or consonant than for words differing by tone.

Results for the Disyllabic task using Thai tones on non-words

Target	Your scores		L2 Mandarin Learner group average		L1 Mandarin Native-Speaker group average		L1 English Monolingual group average	
	Accuracy (%)	RT (ms)	Accuracy (%) (SD)	RT (ms) (CI)	Accuracy (%)(SD)	RT (ms) (CI)	Accuracy (%)(SD)	RT (ms) (CI)
Overall tone	00	000	87.5 (1.8)	1265 (1179; 1357)	85.9 (2.0)	1294 (1227; 1363)	77.5 (3.1)	1368 (1287; 1454)
Overall segments (vowel or consonant differences)	00	000	97.2 (0.7)	1026 (957; 1100)	96.7 (0.7)	1168 (1109; 1231)	96.6 (0.7)	1140 (1073; 1210)
Individual tone comparisons								
Low+High vs High+Low	00	000	95.6 (1.8)	1190 (1099; 1288)	92.2 (2.4)	1197 (1127; 1271)	80.7 (3.5)	1301 (1214; 1396)
Low+Mid vs High+Low	00	000	91.2 (2.5)	1192 (1101; 1291)	95.3 (1.5)	1197 (1127; 1274)	81.8 (3.5)	1297 (1210; 1391)
Low+High vs Mid+High	00	000	72.1 (4.2)	1313 (1209; 1427)	66.4 (4.3)	1384 (1297; 1476)	63.0 (5.1)	1459 (1356; 1569)
Low+Mid vs	00	000	87.5 (2.5)	1423 (1314; 1541)	85.3 (3.3)	1454 (1365; 1542)	80.2 (3.5)	1484 (1384; 1592)

Low+High								
Low+Mid vs Mid+High	00	000	91.2 (2.3)	1240 (1146; 1343)	90.1 (2.1)	1309 (1231; 1390)	81.8 (3.6)	1332 (1243; 1429)
Individual segmental comparisons								
High+Falling	00	000	95.6 (1.4)	1018 (941; 1102)	97.4 (1.1)	1165 (1097; 1236)	94.8 (1.8)	1121 (1047; 1200)
High+Low	00	000	97.1 (1.3)	1059 (979; 1145)	97.0 (1.2)	1191 (1122; 1266)	96.9 (1.1)	1152 (1076; 1233)
Low+High	00	000	97.8 (1.2)	1030 (952; 1113)	97.4 (0.9)	1151 (1085; 1221)	97.9 (1.0)	1127 (1053; 1206)
Low+Mid	00	000	97.1 (1.7)	999 (923; 1080)	96.6 (1.2)	1168 (1099; 1239)	95.8 (1.4)	1143 (1068; 1223)
Mid+High	00	000	98.5 (1.0)	1027 (950; 1111)	95.3 (1.3)	1162 (1094; 1234)	97.4 (1.0)	1155 (1079; 1236)

Note: SE = standard error.CI = Confidence interval (lower; upper).
RTs appear to be "rounded up" as the RTs were logged during calculation.

3) <u>Lexical Decision task:</u>

You listened to both real words and made-up words composed of one or two syllables using Mandarin tones. You determined whether the word you heard was a real or made-up word in Mandarin. We measured your accuracy rates in answering whether the word was real (YES) or made-up (NO). We also measured your reaction times in answering.

A few real words were repeated twice and we measured the difference between the two reaction times. Native speakers react faster when they hear the same real word the second time. So, we expect that proficient non-native speakers would do the same. Accordingly, native speakers would not react faster to hearing [ma] tone 3 ‘horse’ after hearing [ma] tone 1 ‘mother’ since these are two different words. A less proficient learner of Mandarin, however, may “subconsciously” react faster as such a learner might not have a strong representation of tone linked with this word, i.e., they hear [ma] tone 3 and [ma] tone 1 and consider the two stimuli to be the same word even though the two words are different since the tones of the two are different.

In short, you should have faster reaction times on hearing the same word the second time (i.e., larger positive difference between the RTs for the same word), but similar reaction times on both the first and second hearing of a minimal pair as the two words are different (i.e., close to “zero” difference between the RTs for the minimal pair words). However, we are testing whether this is indeed the case for Mandarin tones, so we might not necessarily see these results.

Your accuracy rate in guessing that the stimuli you heard were real words/made-up words in Mandarin Chinese is: 89.4%.

Real word reaction times (ms) for Tone and Segmental Same and Minimal Pairs for L1 Mandarin native speakers and L2 Mandarin Learners for Lexical Decision task with Repetition Priming Showing First and Second Repetition times

	Tone			Segmentals		
Pairings	Your average RTs	L2 Learners	L1 Mandarin	Your average RTs	L2 Learners	L1 Mandarin
Same/ first	000	1116.863 (84.1; 91.0)	1122.018 (65.2; 69.2)	000	1093.956 (82.4; 91.8)	1071.519 (62.3; 66.1)
Same/ second	000	1018.591 (76.7; 82.9)	1032.761 (60.0; 63.7)	000	1023.293 (77.1; 83.3)	1064.143 (61.8; 65.7)
Same word difference	000	98.272	89.257	000	70.663	7.376

Minimal Pair first	000	1164.126 (87.7; 94.8)	1142.878 (68.9; 70.5)	000	1140.250 (85.9; 92.9)	1066.596 (62.0; 65.8)
Minimal Pair second	000	1142.878 (86.1; 93.1)	1145.513 (66.6; 70.7)	000	1109.175 (83.5; 90.3)	1051.962 (61.1; 64.9)
Minimal Pair difference	000	21.248	-2.635	000	31.075	14.634

Note: CI = confidence interval (lower; upper).

**A minimal pair would be words that differ by one vowel or one consonant or tone, e.g., [ma] tone 3 and [ma] tone 1 are minimal pairs differing by tone 3 vs 1 while [ma]tone 4 and [ba] tone 4 are minimal pairs differing by the consonants [m] vs [b].*

4) <u>Phonological Proficiency task:</u>

You were recorded answering a few simple questions, counting from 1-20, and reading a passage. Your recording was spliced into three samples: counting from 1-10, 7 seconds of your answer to why you study Mandarin and reading the line "*Míngnián, wǒ yào qù Táiwān xué Zhōngwén. Bìyè yǐhòu, wǒ yào zài Zhōngguó zuò shēngyi*" (Next year I want to go to Taiwan to study Mandarin. After I graduate, I want to do business in China.). 明年，我要去台湾学中文。毕业以后，我要在中国做生意。明年，我要去台灣學中文。畢業以後，我要在中國做生意。).

Your three samples were then rated by three native speakers of Mandarin (one from Mainland China and two from Taiwan). The three raters rated each sample from 1 to 7 with 1 being non-native-like and 7 being native-like. The samples from all participants were randomized. Raters were also told to rate pronunciation in terms of vowels, consonants, and tone, and to ignore fluency, pausing, grammar, odd meanings, possible poor splicing, etc. As such, any long pauses beyond 400ms were cut down to 400ms to prevent raters from conflating fluency with pronunciation.

	Numbers			Question			Reading			Total Average Score (higher = more proficient)
Evaluator	1	2	3	1	2	3	1	2	3	

Your scores	0	0	0	0	0	0	0	0	0	0
Overall average	4.3	4.1	4.3	2.9	4.1	4.4	3	4.4	3.9	3.94

Overall, we conclude that learners of Mandarin can and do learn tones, which is good news. In their perception of Thai tone, learners of Mandarin performed more accurately and faster than native speakers of English who have not learned Mandarin or any other tone language. Also, on the Mandarin tones, learners of Mandarin generally reacted more quickly upon hearing a word a second time and did not react quickly to stimuli that were different words that varied by tone. This is all good news as it means that you have been able to learn tone. Congratulations!

If you should have any questions, please feel free to email me.

Again, thank you very much for participating in my study!

vance schaefer

Candidate's Curriculum Vitae

Vance Ken Schaefer

CURRICULUM VITAE

801 E. Hunter Ave. #15, Bloomington, IN 47401
email: vkschaef@indiana.edu
(206) 992-4161
http://vanceschaefer.weebly.com

EDUCATION

PhD candidate (degree expected 2015) Second Language Studies, Department of Second Language Studies, Indiana University, Bloomington, Indiana
Coursework: L2 phonology, phonology, phonetics, second language pedagogy
Dissertation: *Influence of the first and second language on the perception of Thai tones*
Minor: Japanese

M. A. 2009. TESOL & Applied Linguistics, Department of Second Language Studies, Indiana University, Bloomington, Indiana
Coursework: second language acquisition, ESL/EFL pedagogy, applied linguistics.

M. A. 2009. Japanese, Department of East Asian Languages and Cultures, Indiana University, Bloomington, Indiana
Master's essay: *An optimality theoretical approach to resolving vowel hiatus in the Classical Japanese verbal paradigm*.
Coursework: Japanese linguistics, JFL pedagogy, Classical Japanese, translation, research methodology

M. A. 1996. International Studies, Graduate School of International Studies, Kobe University, Kobe, Japan
Master's thesis: *Japan's Refugee Policy: The Japanese Government's Policy and Position toward Indochinese Refugees in the 70s and 80s.*
Coursework: Japanese overseas aid & foreign policy, human rights law.

B. A. 1986. Linguistics, Northwestern University, Evanston, Illinois
Subfield certification in practical (TESOL) and historical linguistics.
Coursework: second language acquisition, applied linguistics, phonology, American English.

Junior-year abroad, 1984-85. Waseda University, Tokyo, Japan (Earlham College Program),
Coursework: Japanese language, linguistics, culture.

PROFESSIONAL EXPERIENCE

Teaching Experience at Indiana University

2009-2014 2006-2007	Associate Instructor, English Language Improvement Program (ELIP), Department of Second Language Studies, Indiana University, Bloomington, Indiana Courses for matriculated students: - *Pronunciation* - *Speaking Fluency Development* - *Academic Listening* - *Academic Literacy Development*
2006-2014	Associate Instructor, Intensive English Program (IEP), Department of Second Language Studies, Indiana University, Bloomington, Indiana Courses for college preparation students: - *Variations in the English Language* (2015, 2014) - *Pronunciation* (2015, 2014) - *The Art of Storytelling (2015)* - *Development of the English Language* (2013) - *IELTS* (2014, 2013, 2011) - *Translation and Interpretation* (2014, 2013) - *Academic Presentations* (2012) - *Exploring Translation through Harry Potter* (2011) - *World Englishes* (2010) - *Shakespeare and English* (2014) - *Communication, Grammar, Reading, Writing*
2008 (summer) 2009 (summer)	Instructor, International Associate Instructor Workshop, Department of Chemistry, Indiana University, Bloomington, Indiana Intensive 2-week workshop (4-plus hours per day) focusing on pronunciation and teaching methods
2005-2006, 2009-2011	Associate Instructor, Third-year Japanese, Department of East Asian Languages and Cultures, Indiana University, Bloomington, Indiana Textbooks: *An Integrated Approach to Intermediate Japanese, Tobira Gateway to Advanced Japanese Learning through Content and Multimedia*
2003-2005, 2007-2009	Associate Instructor, Second-year Japanese, Department of East Asian Languages and Cultures, Indiana University, Bloomington, Indiana Textbooks: *Nakama 2, Genki II*
2004 (summer)	Instructor, Upward Bound high school Japanese, Indiana University,

Bloomington, Indiana
Textbook: *Adventures in Japanese 1*

2005 (summer), 2006 (summer) — Co-instructor, Japanese Language & Culture Workshop for Indiana University staff and faculty, East Asian Studies Center, Indiana, University, Bloomington, Indiana

Overseas Teaching Experience

2001-2003 — Instructor, Kaplan, Bangkok, Thailand
TOEFL listening/writing, GRE writing
1999-2000 — Instructor, Excel Language Learning Centre (ELLC), Taipei, Taiwan
IELTS preparation courses, writing, business English
1999 — English teacher, Cornell American English School, Taipei, Taiwan.
English for children (5-9 years old) - singing, playing games, reading stories.
1991-1994 — Part-time instructor YMCA – Osaka, Japan
General English for high school/junior high/elementary students
1989-1991 — Instructor, AEON, Kobe, Japan
Conversation, all levels
1988-1989 — Instructor AEON, Yonago, Japan
Conversation, all levels
Fall, 1987 — Instructor, American Language Center, Casablanca, Morocco
Intensive beginning, advance classes, including TOEFL courses
1986-1987 — *Mombusho* English Fellow (precursor to the JET program), Japanese Ministry of Education– Utsunomiya, Japan
English conversation at prefectural high schools, assistant to the Prefectural High School English Language Supervisor

Other employment

2008-2014 — Interviewer, TEPAIC (Test of English Proficiency for International Associate Instructor Candidates), Department of Second Language Studies, Indiana University, Bloomington, Indiana
2009-2014 — Interviewer, IEPE (Indiana English Proficiency Exam), Department of Second Language Studies, Indiana University, Bloomington, Indiana
2001-2003 — Japanese-English translator, PALNET, Bangkok, Thailand
1991-1998 — In-house English language editor/translator, Nippon Arts, Osaka, Japan

AWARDS

Indiana University College of Arts and Sciences Dissertation Year Research (Louise McNutt Graduate) Fellowship (2014-2015)
Language Learning Dissertation Grant 2014

Indiana University Bloomington Lieber Memorial Teaching Associate Award - Japanese (2005-6)

Research Interests

L2 phonology (perception of lexically-contrastive pitch), pronunciation teaching, foreign accent, ESL/EFL/JFL pedagogy, World Englishes, contact linguistics, translation, Asian languages

Languages

Native: English
Fluent: Japanese (Japanese Language Proficiency Test, level 1 certification) (including Kansai dialect)
Conversational: Thai, Mandarin, German
Reading: Mandarin, German, French
Grammatical understanding: Classical Japanese, Korean

Computer Software Skills

Praat, Audacity, DMDX, SPSS

Publications

Schaefer, V. & Darcy, I. (submitted). A communicative approach and dialect exposure enhance pitch accent awareness by learners of Japanese.

Schaefer, V. & Darcy, I. (2014). Lexical function of pitch in the first language shapes cross-linguistic
perception of Thai tones. *Laboratory Phonology 5*(4): 489-522.

Schaefer, V. & Darcy, I. (2014). Pitch Prominence matters: Perception of Thai tones by Seoul Korean and Kyungsang Korean speakers. *Proceedings of the International Symposium on the Acquisition of Second Language Speech Concordia Working Papers in Applied Linguistics, 5*, 597-611.

Schaefer, V. & Darcy, I. (2014). Linguistic Prominence of Pitch within the Native Language Determines Accuracy of Tone Processing. In R. T. Miller, K. I. Martin, C. M. Eddington, A. Henery, N. Marcos Miguel, A. M. Tseng, A. Tuninetti, & D. Walter (Eds.), *Selected Proceedings of the 2012 Second Language Research Forum: Building Bridges Between Disciplines* (pp. 1-14). Somerville, MA: Cascadilla Proceedings Project.

Dekydtspotter, L., Miller, A. K., Schaefer, V., Chang, Y., & Kim, O-H. (2010). Clause-Edge Reactivations of Fillers in Processing English as a Second Language. In M. T. Prior, Y. Watanabe, & S-K. Lee (Eds.), *Selected Proceedings of the 2008 Second Language Research*

*Forum: Exploring SLA Perspectives, Positions, and Practic*es (pp. 108-122). Somerville, MA: Cascadilla Proceedings Project.

CONFERENCE PRESENTATIONS

Schaefer, V. (2014, November). Connecting ESL Students with Communities Through Teaching Language Variation. Paper presented at INTESOL 2014, Indianapolis, IN.

Abe, L. & Schaefer, V. (2014, November). From Listening to Presentation: Integrating Oracy Skills for Public Speaking. Paper presented at INTESOL 2014, Indianapolis, IN.

Schaefer, V. & Darcy, I. (2014, September). A communicative approach and dialect exposure enhance pitch accent awareness by learners of Japanese. Paper presented at Pronunciation in Second Language Learning and Teaching (PSLLT) 2014, Santa Barbara, CA.

Schaefer, V. (2013, November). Translating the Way to Improved Fluency, Pronunciation, Accuracy and Awareness. Poster presented at INTESOL 2013, Indianapolis, IN.

Schaefer, V. & Darcy, I. (2013, May). Cross-linguistic perception of Thai tones is shaped by the functional prominence of lexically-contrastive pitch in L1. Paper presented at New Sounds 2013, Montreal, Canada.

Schaefer V. & Darcy, I. (2012, October). Cross-linguistic perception of Thai tones. Paper presented at 31st Second Language Research Forum (SLRF), Pittsburgh, PA.

Schaefer, V. (2012, October). Teaching to the Pronunciation Needs of Mandarin-speaking Students. Paper presented at INTESOL 2012, Indianapolis, IN.

Schaefer, V. & Darcy, I. (2012, July). Cross-linguistic perception of Thai tones. Poster presented at 13th Conference of Laboratory Phonology (LabPhon), Stuttgart, Germany.

Darcy, I., Dekydtspotter, L., Sprouse, R. A., Glover, J., Kaden, C., McGuire, M., Schaefer, V., & Scott, J. H. G. (2009, May). Category formation and lexical encoding of a new contrast. Poster presented to The Acoustical Society of America's 2nd Special Workshop on Speech: Cross-language speech perception and variation in linguistic experience. Portland, OR.

ADDITIONAL LANGUAGE STUDIES

Goethe Institut, Berlin, Germany
Taipei Language Institute, Taipei, Taiwan (ROC)
American University Alumni (AUA) Language Center, Bangkok, Thailand

PROFESSIONAL ASSOCIATION MEMBERSHIPS

American Council on the Teaching of Foreign Languages (ACTFL)
Teachers of English to Speakers of Other Languages (TESOL)
Indiana Teachers of English of Speakers of Other Languages (INTESOL)
American Association of Teachers of Japanese (AATJ)

ADDITIONAL INFORMATION

Dual national of the US and UK

REFERENCES

Linda Abe	abels@indiana.edu	phone: (812) 855-4974
Isabelle Darcy	idarcy@indiana.edu	phone: (812) 855-0033
Rex A. Sprouse	rsprouse@indiana.edu	phone: (812) 855-3248

Linda Abe, PhD
Department of Second Language Studies
1021 E. Third St.
Memorial Hall 312
Indiana University
Bloomington, IN 47405-7005

Isabelle Darcy, PhD
Department of Second Language Studies
1021 E. Third St.
Memorial Hall 301
Indiana University
Bloomington, IN 47405-7005

Rex A. Sprouse, PhD
Department of Germanic Studies
1020 E. Kirkwood Ave.
Ballantine Hall 661
Indiana University
Bloomington, IN 47405-7103