A Cross-Language Investigation of Phonetic and Phonological Processing of Lexical Tone

by

Xianghua Wu

M.A. (Linguistics), University of Victoria, 2006
B.A., Beijing Language and Culture University, 1995

Dissertation Submitted in Partial Fulfillment
of the Requirements for the Degree of
Doctor of Philosophy

in the
Department of Linguistics
Faculty of Arts and Social Sciences

SIMON FRASER UNIVERSITY

Summer 2012

Approval

Name: **Xianghua Wu**

Degree: **Doctor of Philosophy**

Title of Thesis: ***A cross-language investigation of phonetic and phonological processing of lexical tone***

Examining Committee:

Chair: **Dr. Chung-hye Han**, Associate Professor

Dr. Murray Munro
Senior Supervisor
Professor

Dr. Yue Wang
Supervisor
Associate Professor

Dr. Christian Guilbault
Internal Examiner
Assistant Professor
Department or French

Dr. Valter Ciocca
External Examiner
Professor, School of Audiology and Speech Sciences
The University of British Columbia

Date Defended/Approved: May 24, 2012

Partial Copyright Licence

SFU

SFU SIMON FRASER UNIVERSITY
THINKING OF THE WORLD

STATEMENT OF ETHICS APPROVAL

The author, whose name appears on the title page of this work, has obtained, for the research described in this work, either:

(a) Human research ethics approval from the Simon Fraser University Office of Research Ethics,

or

(b) Advance approval of the animal care protocol from the University Animal Care Committee of Simon Fraser University;

or has conducted the research

(c) as a co-investigator, collaborator or research assistant in a research project approved in advance,

or

(d) as a member of a course approved in advance for minimal risk human research, by the Office of Research Ethics.

A copy of the approval letter has been filed at the Theses Office of the University Library at the time of submission of this thesis or project.

The original application for approval and letter of approval are filed with the relevant offices. Inquiries may be directed to those authorities.

Simon Fraser University Library
Simon Fraser University
Burnaby, BC, Canada

Last update: Spring 2010

Abstract

In an investigation of how lexical tone is perceived and processed at the phonetic and phonological levels, listeners from diverse language backgrounds participated in three perceptual studies. In the first, native Mandarin and Thai listeners assimilated non-native tones to their native tone categories. Results indicated that occurrence of a lower-level phonetic and a higher-level phonological assimilation process was related to listeners' tone experience, as inexperienced listeners recognized only the phonetic distinctions, whereas experienced listeners were sensitive to both the phonetic and phonological distinctions between native and non-native tone categories.

In the second study, native Mandarin, Thai and English listeners participated in a forced-choice tone perception test in which they identified the four Mandarin tone categories. Identification accuracy and confusion patterns revealed that previous tone experience predicted tone perception at the phonetic and phonological levels. Better performance was demonstrated for native than non-native, and experienced than inexperienced listeners. Experienced Thai listeners also showed more native-like performance than experienced English listeners. Tone 2 and Tone 3 were the most confusable tone pair for all but the inexperienced English listeners. Lexical information from the carrier words was also found to help Mandarin and English listeners recognize difficult tones.

In the third study, participants from the second one completed a dichotic listening test assessing tone lateralization in the brain. The results demonstrated a strong influence of acoustic properties, as tones with dynamic F0 contours were lateralized to the left hemisphere while those with flatter F0 contours were lateralized to the right hemisphere. Meanwhile, native and non-native tone experience was associated with a larger degree of left hemisphere activation for Mandarin and experienced Thai listeners relative to those in the remaining groups.

In summary, these three studies indicate tone perception and processing at both the phonetic and phonological levels. In relation to tone experience, inexperienced listeners may attach more importance to phonetic variation while experienced listeners are sensitive to both the phonetic and phonological differences. In terms of theoretical contributions, tone perception results extend the current models of speech perception to the suprasegmental level while tone lateralization results provide evidence supporting the acoustic and functional hypotheses.

Keywords: speech perception; assimilation; dichotic listening; tone; Mandarin; Thai

To Professors Linghua Yue and Xianghe Ni

Acknowledgements

There are many people whom I would like to thank for encouraging and helping me throughout the dissertation process. First and foremost, I would like to thank my senior supervisor, Dr. Murray Munro, for his guidance, support and patience. I greatly appreciate Murray's clear teaching, insightful advice, and willingness to give feedback. He has always been there since I started my PhD program, helping me to improve each paper, presentation, grant proposal, and job application. I am indebted to Dr. Yue Wang for being my co-supervisor, and also for providing me with opportunities to work at her Language and Brain Lab (LABlab). Yue's diligence and professionalism has inspired me during difficult periods. I would also like to extend my thanks to Dr. Christian Guilbault for agreeing to serve on my committee. Special thanks go to Dr. Valter Ciocca for agreeing to be my External Examiner. His research and kindness has inspired me throughout my writing process. I am thankful to Stephanie Gabriel and Dr. Ray Koopman from Simon Fraser University (SFU) and Dr. Ian Walker from the University of Bath for their help with statistics. I am grateful to Dr. Lorna Fadden and Dr. Allison Benner for editing and proofreading my dissertation. I want to thank more than 150 participants who took part in this study in Bangkok, Beijing and Vancouver. I also owe a huge debt of gratitude to my friends and colleagues in the respective cities for their help with recruitment of participants. Without the financial help from the Social Sciences and Humanities Research Council (SSHRC) and a series of fellowships and scholarships awarded by SFU, this dissertation would have not been possible. I extend my warm thanks to the faculty and staff members in the Department of Linguistics at SFU for their support, care and assistance throughout these years. The Language Training Institute at SFU has given me the chance to hone my teaching skills and foster deep friendships with a group of wonderful colleagues and students. I am deeply grateful to my fellow students and friends in Vancouver for their encouragement. Last but not least I would like to thank my family, especially my parents for their deepest love and support.

Table of Contents

List of Tables

List of Figures

1. Introduction

Humans are capable of perceiving a wide array of speech sounds. Among these sounds, some are more perceivable as they contrast in one's native phonological system, whereas others are more difficult to recognize because they are unfamiliar in one's first language. These speech sounds can be incorporated into segmental or suprasegmental inventories. Among the suprasegmentals, tone (or lexical tone) has been widely studied, given its importance at both the phonetic and phonological levels (Abramson, 1978; Bent, Bradlow, & Wright, 2006; Burnham, Francis, Webster, Luksaneeyanawin, Attapaiboon, Lacerda, & Kelle, 1996; Gandour, 1978, 1983; Huang, 2001, 2004; Shen & Lin, 1991; Wang, Jongman, & Sereno, 2003). At the phonetic level, tone is realized as fundamental frequency (F0), with F0 height and contour being the primary acoustic correlates. At the phonological level, tone is imposed on monosyllables to contrast word meanings in tone languages, such as Mandarin (or Mandarin Chinese), Thai and Cantonese. The phonetic properties and phonological functions of tone have attracted considerable interest from researchers. For instance, neurolinguists have examined the extent to which tone processing is driven by the left versus the right hemisphere, in order to establish the phonetic and phonological nature of tone (Gandour, 2006). Child language researchers have also proposed that tones are acquired at the phonetic and phonological levels in the very early period of life (Li & Thompson, 1977). Second language researchers have also evaluated the effects of perceptual training on tone acquisition by learners (Wang, Spence, Jongman, & Sereno, 1999).

1.1 The goals of this dissertation

This dissertation investigates tone perception at the phonetic and phonological levels by native and non-native listeners to address three main goals. The first goal is to determine the similarities and differences between tone perception and segmental perception. Perception of segmental inventories, i.e., consonants and vowels, has been studied extensively at the phonetic and phonological levels. Recent models of cross-language speech perception, such as the Perceptual Assimilation Model (Best, 1995; Best, McRoberts, & Goodell, 2001; Best, McRoberts, & Sithole, 1988, Best & Tyler, 2007) and the Speech Learning Model (Flege, 1995, 2007) have been developed to account for the effects of phonetic similarity and phonological correspondence between native (L1) and non-native (L2) categories on the perception of segmental inventories. Thus, recent tone perception research often seeks evidence to evaluate and expand these models of cross-language speech perception (Bent, 2005; Hallé, Chang, & Best, 2004; So & Best, 2010). The second goal of this study is to investigate the effects of phonetic properties and phonological functions on the cerebral hemispheric processing of tone. Some researchers have considered whether tone is processed in the same manner as other speech prosodic phenomena, such as intonation, because the acoustic cues by which it is conveyed are similar. However, tone differs from other prosodic patterns in its linguistic use. In the linguistic domain, intonation is used to express pragmatic differences or syntactical modes (e.g., statement vs. question). Tone, however, is used to mark semantic contrasts at the lexical level. Thus, tone processing may be affected by its acoustic properties and phonological status (Gandour, Dzemidzic et al., 2003). Further investigation of this issue can also provide data to support the two prevailing theoretical hypotheses on the hemispheric processing of pitch patterns: *the acoustic hypothesis* which states that pitch processing is dependent on acoustic cues and *the functional hypothesis* which proposes that pitch processing is determined by linguistic functional load. The third purpose is to examine the influence of L1 and L2 experience on tone processing at the phonetic and phonological levels. Examination of tone perception by listeners from different language backgrounds can reveal the nature of their sensitivity to the acoustic correlates and phonological function of tone categories (Gandour, 1983; Guion & Pederson, 2007; Huang, 2001; Lee, Vakoch, & Wurm, 1996; Lueng, 2008). Four levels of language experience are of particular interest to the current project: (1) L1 is the target tone language; (2) L2 is the target tone language; (3) L1 is a non-target tone language; (4) L1 is a non-tone language. As tone functions at the phonological level only in tone languages, it is likely that L1 tone experience is transferable to the perception of L2 tones, while in non-tone languages, tone patterns may be associated with phonetically similar L1 intonational patterns (Francis, Ciocca, Ma, & Fenn, 2008). On the other hand, L2 experience may improve listeners' perception of tone and lead to native-like phonological perception (Guion & Pederson, 2007; Guion & Wayland, 2004; Wang, Spence, Jongman, & Sereno, 2003).

In the following sections, current models of cross-language speech perception are introduced. Then, the tone systems in Mandarin and Thai and the phonetic correspondence between Mandarin tones and English intonational patterns, are described. Lastly, an overview of this project is provided.

1.2 Models of cross-language speech perception

In this section, the Perceptual Assimilation Model (PAM) and the Speech Learning Model (SLM) will be reviewed. The PAM was developed in two stages: the first stage focused on segmental perception by non-native listeners with no knowledge of L2 (inexperienced listeners) while later stage focused on L2 learners (experienced listeners). Consequently, the PAM has two versions: the PAM (Best, 1995, Best et al., 1988, 2001) and the PAM-L2 (Best & Tyler, 2007). In the following sections, the PAM is reviewed first, followed by the SLM and the PAM-L2. The PAM provides a theoretical account of how inexperienced listeners perceive non-native phonetic categories based on L1 and L2 phonetic correspondence. The SLM attempts to represent the segmental perception and production of L2 learners with varying L2 experience at different time intervals. Finally, the PAM-L2 predicts segmental discrimination at the phonological level by experienced listeners.

1.2.1 Perceptual Assimilation Model

The Perceptual Assimilation Model (PAM) was developed by Best and colleagues (Best, 1995; Best et al., 1988, 2001) with a focus on the discrimination of non-native phonetic contrasts by inexperienced listeners or naïve listeners with no exposure to a non-native language. As the PAM is built on the Direct Realist Theory (Goldstein & Fowler, 2003; Fowler, 1996), phonetic features in the model represent articulatory gestures. Many researchers have adopted PAM to account for the relationship between perceptual performance and phonetic-acoustic similarity between

L1 and L2 categories (Cebrian, 2006; Hallé et al., 2004; Levy, 2009; Nishi, Strange, Akahane-Yamada, Kubo, & Trent-Brown, 2008; Strange, Bohn, Nishi, & Trent, 2005); thus, the concept of phonetic realization has been extended to refer to a series of acoustic phonetic details.

The PAM posits that inexperienced listeners assimilate non-native sounds to the most articulatorily similar sound in their native system. In other words, inexperienced listeners' performance can be predicted by phonetic relationships between L1 and L2 sounds, such that phonetically similar sounds tend to be perceived as the same sounds and phonetically dissimilar sounds tend to be perceived as different ones. Moreover, in cases where phonetic distinctions also mark L2 phonological differences, the PAM proposes that phonetic-phonological assimilation occurs solely for experienced listeners, as inexperienced listeners are unaware of L2 phonological categories. In other words, inexperienced listeners recognize phonological distinctions only in their native language.

According to this model, perceptual assimilation falls along a continuum in which listeners can assimilate an L2 phone to L1 phonetic categories as a categorized or uncategorized speech sound; and in which some L2 sounds may be "nonassimilable," thus considered as non-speech sounds. One well-known example of a "nonassimilable" pattern is English listeners' perception of Zulu click consonants as non-speech sounds (Best et al., 1988; c.f., Best & Avery, 1999). Based on this continuum, the PAM posits the following six assimilation patterns and the degrees of predicted difficulty in discriminating pair-wise L2 categories for each of these patterns.

1) Two Category (TC). If two L2 phones are assimilated to two different L1 categories, discrimination of the two phones will be very good to excellent;
2) Single Category (SC). If two L2 phones are judged as equally "good" instances of a single native category, they will be difficult to differentiate;
3) Category Goodness (CG). If two L2 phones differ in their goodness as instances of a single L1 category, intermediate levels of perceptual difficulty will be observed;
4) Nonassimilable (NA). When two L2 phones are perceived as non-speech sounds, discrimination should be good to very good;

5) Both Uncategorizable (UU). A non-native phone which is not assigned to any L1 category may be considered an "uncategorizable" speech sound. When paired with another "uncategorizable" phone (UU), discrimination of the UU phones will range from very poor to moderate.
6) Uncategorized versus Categorized (UC). When an "uncategorizable" sound and a "categorizable" phone (i.e., a phone that is sufficiently phonetically similar to be categorized as an instance of a native category) are presented in a pair, discrimination will be relatively easy.

1.2.2 Speech Learning Model (SLM)

The Speech Learning Model (SLM), developed by Flege and his colleagues (Flege, 1987, 1995, 2007; Flege, Schirru, & Mackay, 2003) provides a theoretical account of category assimilation for listeners with varying degrees of L2 experience. The SLM postulates that the formation of L2 phonological categories is similar to that of L1 phonological categories; thus, phonological category formation remains possible over the life span. Perceptual assimilation of L2 categories to L1 categories can be observed before the L2 phonological system is fully established. In other words, the fact that perceptual assimilation occurs implies that an L2 phonological category is associated with an L1 category due to their phonetic relationship. In general, phonetically dissimilar categories are less likely to be assimilated than phonetically similar categories. Moreover, the theory implies that the phonetic distance between L1 and L2 categories is determined by perceptual processes rather than physical realizations. One hypothesis of the SLM is that the mechanism of equivalence classification may lead to the assimilation of L2 categories to L1 categories, even when phonetic differences between L1 and L2 categories can be perceived. According to the SLM, L2 phones may be assimilated to L1 categories as "identical", "similar" or "new" (Flege, 1987). Phonetic similarity varies among the three types. "Identical assimilation" occurs when two phones share the same phonetic properties, such as the vowel /ɪ/ in American English and German (Strange, Bohn, Nishi, & Trent, 2004). A "similar" L2 phone may be assimilated to its phonetic counterpart in the L1, but the differences between the two phones may be detectable, such as the vowel /u/ in American English and French (the English /u/ is characterized by a higher and more variable second formant frequency (F2) than its French counterpart) (Hillenbrand, Clark, & Nearey, 2001). In contrast, a "new" L2 phone has no phonetic counterpart in the L1 system, but is nevertheless consistently assimilated to an L1 phone, in particular by less experienced listeners. For example, naïve American English listeners may categorize the French front rounded vowel /y/ as the American English back rounded vowel /u/, despite the considerable phonetic difference between the two sounds (Levy, 2009).

1.2.3 Perceptual Assimilation Model for learners (PAM-L2)

The PAM-L2, an extended version of the PAM, predicts discrimination patterns for listeners who are experienced in an L2, in particular for listeners with extensive experience in the L2 environment (Best & Tyler, 2007). Unlike the SLM, the PAM-L2 proposes both a higher-level phonological and a lower-level phonetic assimilation process for experienced listeners. In other words, experienced listeners can become sensitive to both phonetic and phonological distinctions between L1 and L2 categories. Unlike phonetic assimilation, phonological assimilation refers to a process in which L2 phonological categories are related to L1 phonological categories due to equivalent lexical functions. The PAM-L2 also posits that lexical function plays a more important role in phonological processing. Thus, listeners may assimilate L2 categories to L1 phonological categories despite perceptible phonetic distinctions.

The PAM-L2 proposes the following four phonological assimilation cases, along with the predicted degree of difficulty in discriminating L2 phonological categories for each case. The model also predicts the phonetic relationship between L1 and L2 categories for some cases of assimilation.

1) One L2 phonological category is assimilated to one L1 phonological category. Discrimination of the assimilated L2 category from other L2 categories will be easy. At the phonetic level, L1 and L2 categories can be perceived as very similar or very dissimilar. For example, the French vowels /i, ɛ, a, u/ are consistently assimilated to their phonetic counterparts /i, ɛ, a, u/ in English by American English listeners who have intensive experience with French (Levy, 2009; Strange, Bohn, Trent, & Nishi, 2004).

2) Two L2 phonological categories are assimilated to one L1 phonological category, but with unequal perceptual distances. Similar to the TC assimilation pattern predicted by the PAM, the difference between the two L2 categories is easily detected. Phonological assimilation is expected to happen between the better L2 exemplar and the L1 category. As suggested, the better L2 exemplar also serves as the phonetic equivalent to the L1 category relative to the deviant one. For example, the American English vowels /i:/ and /ɪ/ are both assimilated to the Japanese vowel /i/ by experienced Japanese listeners of English, but Japanese listeners distinguish between the English /i:/ and the Japanese /i/ more easily than they do between the English /ɪ/ and the Japanese /i/ (Nishi, Strange, Akahane-Yamada, Kubo, &Trent-Brown, 2008).
3) Two L2 phonological categories are equally assimilated to one L1 phonological category as poor or good instances of the category. Discrimination between the L2 categories is difficult because phones have been assimilated to a single category. The phonetic relationship between L1 and L2 categories is not explicated, but it is likely the L1 and L2 categories share phonetic properties to some extent. For example, the Australian English vowel /o:/ and /əʉ/ are both assimilated to the Japanese vowel /ou/ by advanced Japanese learners of English, while the perception of /əʉ/ as /ou/ is comparable to that of /o:/ as /ou/ (Bundgaard-Nielsen, Best, & Tyler, 2011).
4) An L2 phonological category is not assimilated to any L1 category, and is thus "uncategorized" in PAM terms. Due to a lack of phonological correspondence with L1, phonetic similarity tends to play an important role in the formation of L2 categories in this case. For example, if L2 phones are perceived to be more similar to different sets of L1 phonological categories, the functional difference between the L2 phones may eventually be recognized. Also in Bundgaard-Nielsen et al. (2011), five Australian English vowels (/o:, æ, æɔ, əʉ, ɜ:/) are defined as "uncategorized" in relation to the Japanese vowel inventory. Nevertheless, more experienced Japanese listeners reveal high consistency in associating these "uncategorized" vowels with native vowel categories relative to less experienced Japanese listeners.

1.3 Lexical tones in Mandarin and Thai

Lexical tones are realized on monosyllables to signal word meanings in Mandarin and Thai (Chao, 1948; Zsiga, 2007). There are four tones in Mandarin and five in Thai, which are imposed on monosyllables to signal word meanings (see Table 1.1). For example, the syllable /lɑu/ with the four Mandarin tones superimposed stands for *to dredge* (lɑu^1, 捞), *jail* (lɑu^2, 牢), *old* (lɑu^3, 老), and *flood* (lɑu^4, 涝); with the five Thai tones superimposed on it, the same syllable means *roughly* (เลา), *group* (เหล่า), *to tell* (เหล้า), *stable* (เล้า) and *to sharpen* (เหลา), respectively. The most important acoustic correlates of lexical tones are F0 height and contour (Abramson, 1975; Lin, 1988), though duration, intensity, vowel quality and voice quality can all affect tone perception to some extent (Belotel-Grenie & Grenie, 1997; Howie, 1976; Lin & Repp, 1989; Vance, 1976). Tones accompany every word in Mandarin and Thai. Thus like consonants and vowels, tones perform a phonemic function in these languages.

Mandarin and Thai tones are shown in Table 1.1, along with their acoustic descriptions. For the sake of convenience, in the current project, the four Mandarin tones are labelled as MT1 (Mandarin Tone 1), MT2, MT3 and MT4, referring to the high level, rising, falling rising and falling tones, respectively (Chao, 1948). The five Thai tones are similarly labelled as TT1, TT2, TT3, TT4 and TT5, to represent mid falling, low falling, high falling, high rising and low falling rising tones, respectively (Ladefoged, 2001). On a 5-point scale with 1 corresponding to the lowest pitch and 5 the highest, the Mandarin tones have been numerically represented as 55 (MT1), 35 (MT2), 214 (MT3), and 51 (MT4) to show changes in F0 height and contour (Chao, 1948). Likewise, the Thai tones are 32 (TT1), 21 (TT2), 51 (TT3), 45 (TT4) and 214 (TT5). The F0 properties of each tone, however, may vary among individual speakers (Gandour, Potisuk, Ponglorpisit, & Dechongkit, 1991; Zsiga, 2007). A typical example showing varied phonetic realizations of tone is that of the rising MT2, which is often produced as a falling rising tone that is similar to MT3, indicating a more complex contour shape (Blitcher, Diehl, & Cohen, 1990; Shen & Lin, 1991).

Table 1.1. Comparison of Mandarin and Thai tones

Language	Number	Name	Label	F0	5-point scale description
Mandarin	4	Tone 1	MT1	high level	55
		Tone 2	MT2	rising	35
		Tone 3	MT3	falling rising	214
		Tone 4	MT4	falling	51
Thai	5	Samand tone	TT1	mid falling	32
		Aek tone	TT2	low falling	21
		To tone	TT3	high falling	51
		Tree tone	TT4	high rising	45
		Jattawa tone	TT5	low falling rising	214

As pointed out earlier, Mandarin employs tone sandhi, a phonological phenomenon where one tone category is realized as another tone category at the word level (Duanmu, 2007). Tone sandhi exists in some tone languages, such as Mandarin and Cantonese (Chen, 2000; Yu, 2007), but not in others, such as Thai (Gandour, 1983), where tones are stable and unaffected by tonal phonotactics. In Mandarin, MT3 is associated with two tone sandhi rules; thus, the phenomenon is also sometimes referred to as T3 sandhi (Duanmu, 2007). First, MT3 (214) is realized as MT2 (35) in a two-tone sequence where MT3 precedes another MT3. Second, MT3 (214) is realized as a partial MT3 (21), i.e., a low falling tone, in cases where it precedes the other three Mandarin tones, MT1, MT2 and MT4 (Lin, 1996). As shown in Table 1.1 above, the most phonetically similar tones in Thai and Mandarin are MT4 and TT3, and MT3 and TT5; and the low falling tone produced in some instances of Mandarin tone sandhi corresponds to TT2.

1.4 English intonational patterns and correspondence to Mandarin tones

Unlike Mandarin and Thai tones, which are associated with a single syllable, lexical prosodic features are realized contrastively on two-syllable sequences in English, with one syllable being stressed and the other unstressed. The stressed syllable is primarily characterized as having a higher pitch than the unstressed syllable. However, due to the difference in prosodic domain, it is unlikely that English word stress would be assimilated to Mandarin or Thai tones in the same way that Thai tones would be assimilated to Mandarin tones (or vice versa). As a result, tone researchers focus on the correspondence between English intonational patterns and tones (e.g., Bent, 2005; Francis et al., 2008).

English has a large number of intonational patterns. Pierrehumbert & Hirschberg (1990) enumerated 22 English intonational patterns that are produced from combinations of six pitch accents, two phrase accents, and two boundary tones. These intonational patterns can be found not only in North American English, but also in British English (Francis et al., 2008; Ladd, 2008). Some patterns are used in short utterances, including monosyllabic words, to indicate grammatical contrasts or pragmatic meanings. Among these patterns, the rising and falling contours are categorically distinctive, as the former indicates standard yes-no questions and the latter indicates affirmative/negative statements. These two classes can be further divided according to the pitch value in the initial or final position (Grabe, Rosner, García-Albea, & Zhou, 2003; Wells, 2006). For example, Wells (2006) further classified falling contours into high fall and low fall, which differ in terms of the pitch value at the beginning of a given contour. High fall and low fall are pragmatically categorized based on degree of emotional involvement. For example, a high fall is associated with greater interest, while a low fall connotes lack of interest (p. 218). A falling rising contour also exists in English, carrying various pragmatic implications, such as partial correction or negation (Halliday & Greaves, 2008; Pierrehumbert & Steele, 1989; Wells, 2006). The English level tone is not as high as the Mandarin high level tone, and carries a mid level contour that is used in some types of interjection (Wells, 2006).

Therefore, all four Mandarin tone categories have phonetic counterparts in English (see Table 1.2 below), despite the fact that the English intonational patterns are used at the grammatical or pragmatic levels, while the Mandarin tones function as phonemes.

Table 1.2. ***Grammatical/pragmatic functions of English intonational patterns and comparison to Mandarin tones***

English intonation		Mandarin tone
Function	**Phonetic description**	
interjection	mid level	MT1 (high level)
yes-no question	rising	MT2 (high rising)
partial correction or negation	falling rising	MT3 (falling rising)
declarative statement	falling	MT4 (high falling)

1.5 Overview

This dissertation comprises three experimental studies on tone perception/processing which are linked by the research goals. The first (Chapter 2) is a tone assimilation study, aiming to discover how tone categories in one language are perceptually associated with tone categories in another language, and whether listeners rely on phonetic or phonological cues in tone assimilation. Native listeners of two tone languages, Mandarin and Thai, participated in this study. These listeners also differed in their L2 tone experience. The second study (Chapter 3) consists of a tone identification experiment in which listeners identified the four Mandarin tones. This study was conducted to determine the extent to which listeners employ their previous knowledge of the linguistic usage of pitch patterns when identifying native or non-native tone categories. Native listeners of Mandarin, Thai and English participated in this study; the non-Mandarin listeners also differed in their L2 experience with Mandarin. The third study (Chapter 4) is a dichotic tone perception study investigating phonetic and phonological effects of tone on the hemispheric processing of the four Mandarin tones. Listeners identified the tones presented simultaneously to the left and right ears. The difference in the performance of the two ears reflects which hemisphere is more activated. Participants in the third study were the same as those who participated in the tone identification study.

2. Tone assimilation by native listeners of Mandarin and Thai

2.1. Introduction

Cross-language perceptual assimilation has been widely examined in previous studies at the segmental level, stimulating discussion on what factors assimilate L2 phonetic/phonological categories to their L1 phonetic/phonological counterparts (Levy, 2009; Nishi, Strange, Akahane-Yamada, Kubo, &, Trent-Brown, 2008; Schmidt, 1996; Strange, 1999; Strange, Bohn, Trent, & Nishi, 2004). However, few studies have focused on perceptual assimilation at the suprasegmental level. This study, therefore, investigates the extent to which L2 tone categories are assimilated to L1 tone categories. Tone, like segments, functions as a phoneme; thus it is expected that the findings from segmental assimilation studies can be extended to tone assimilation. However, tone assimilation may be a very different process from segmental assimilation in that the primary acoustic correlate of tone is F0 (Lin & Repp, 1989; Tong, Francis, & Gandour, 2007; Tseng, 1981) whereas the perception of segments is based primarily on spectral and temporal cues in speech (Lieberman & Blumstein, 1988). Therefore, the aim of the current study is to explore the similarities and differences between segmental assimilation and tone assimilation within the framework of recent models of cross-language speech perception.

2.2. Issues related to perceptual assimilation

Three central issues in recent models of cross-language perceptual assimilation are of particular interest to the current study. One is the existence of two levels of assimilation: phonetic assimilation and phonological assimilation. These two levels are related to a major difference between the Perceptual Assimilation Model (PAM, Best, 1995; Best et al., 1988, 2001; Best & Tyler, 2007) and the Speech Learning Model

(SLM, Flege, 1995, 2007). The PAM proposes the existence of both levels of assimilation while the SLM characterizes category assimilation only at the phonetic level. The second issue is the effect of language experience on perceptual assimilation. The PAM predicts phonetic assimilation only for listeners without L2 experience (inexperienced listeners) and both phonetic and phonological assimilation for those with L2 experience (experienced listeners). In relation to the first issue, the SLM postulates phonetic assimilation for listeners with varying L2 experience. The third issue concerns the relationship between phonetic similarity and perceptual assimilation. According to the PAM, phonetic similarity between L1 and L2 phones is predictive of phonetic assimilation but not phonological assimilation. The SLM posits that despite the importance of phonetic similarity, assimilation of phonetic categories can occur not only between phonetically similar phones in L1 and L2, but also between dissimilar ones.

2.2.1. *Phonetic assimilation*

Phonetic assimilation refers to a process in which listeners attend to the phonetic similarities between incoming L2 speech sounds and L1 phonetic categories (Best et al., 1988). As suggested by recent models of cross-language perception, both inexperienced and experienced listeners can be expected to demonstrate perceptual assimilation at the phonetic level. Some studies have revealed that L2 phonetic categories are assimilated to their phonetic equivalents in L1 relating to vowels (e.g., Nishi et al., 2008; Levy, 2009), consonants (e.g., Best, McRoberts & Goodell, 2001), and pitch patterns (e.g., Bent et al., 2006; Leung, 2008; So & Best, 2010). For example, in the perception of German/French vowels both inexperienced and experienced American English (AmE) listeners assimilated German/ French vowel categories /i, ɛ, a, u/ to the AmE vowels /i, ɛ, a, u/ (Levy, 2009; Strange et al., 2004. At the suprasegmental level, assimilation of Mandarin tones to Cantonese tones was investigated for Cantonese speakers with Mandarin experience in Leung (2008) and those without Mandarin experience in So & Best (2010). Results of the two studies indicate that Mandarin tones were identified as their phonetic equivalents in Cantonese, such as the high level tones and the rising tones, indicating effects of phonetic similarity on tone perception.

Additionally, segmental and suprasegmental studies have also found that the assimilated phonetic categories differ in the degree of phonetic similarity. For example, Strange et al. (2004) noted that the German vowel /ɪ/ was assimilated to the American English /ɪ/ by native American English speakers, despite the fact that the phonetic dissimilarities are greater between these vowels than those between /i, ɛ, a, u/ in L1 and L2. Results of the above two tone studies also showed that Cantonese listeners identified Mandarin falling rising tones as Cantonese rising tones, suggesting that phonetic assimilation may be induced by partial common features, such as a rising contour.

Lastly, previous studies have revealed that phonetic assimilation may occur between dissimilar phonetic categories (Iverson, Kuhl, Yamada, Diesch, Tohkura, Ketterman, & Siebert, 2003; Levy, 2009; Strange et al., 2004). It has been suggested that listeners' misuse of phonetic cues may cause L2 phonetic categories to be misperceived as L1 categories. A well-known example is the confusion between the English consonants /ɹ/ and /l/ for Japanese listeners (e.g., Iverson et al., 2003). A well-accepted explanation is that the English liquids /ɹ/ and /l/ differ primarily in F3, a phonetic feature not attended to by Japanese listeners. The confusion occurs when Japanese listeners attempt to use a familiar phonetic cue, F2, to distinguish these two English consonants. The misuse of perceptual cues has also been found to cause poor discrimination between phonetically distinctive tone categories, especially for listeners without tone experience. In a multidimensional scaling (MDS) study, Gandour (1983) used F0 height and contour as the acoustic cues to examine the perceptual distance between 19 synthesized pitch contrasts created from a naturally produced tone imposed on the monosyllable /wa/. The listeners were monolinguals of four tone languages (Cantonese, Mandarin, Taiwanese and Thai) and English. Language-specific discrimination was detected as the listeners attached greater importance to either F0 height or F0 contour. The tone language listeners were more sensitive to F0 contour than the non-tone language listeners and the Thai listeners were more sensitive to F0 contour than the Mandarin listeners. These findings have also been supported by some more recent studies (e.g. Francis et al., 2008; Guion & Pederson, 2007; Huang & Johnson, 2011).

2.2.2. *Phonological assimilation*

In contrast, phonological assimilation refers to a perceptual process in which listeners assimilate L2 phonological categories to L1 phonological categories that share the same lexical functions. In previous studies, phonological assimilation has been observed in three main ways. In the first pattern, L1 phonological categories are replaced by L2 categories without changing L1 word meanings (Best & Tyler, 2007). The second pattern revealed in previous research is poor discrimination between allophonic variations that serve as the same phoneme (e.g., Boomershine, Hall, Hume, & Johnson, 2008; Huang, 2001). In the third pattern, more experienced listeners relate "unassimilable" L2 phonological categories to L1 phonological categories (e.g., Bundgaard-Nielsen, Best, & Tyler, 2011).

The phenomenon of limited discriminability of allophones demonstrates speech assimilation at the phonological level, as two functionally equivalent phonetic categories tend to be more difficult to discriminate. At the segmental level, for example, Boomershine et al. (2008) found that allophonic contrasts (e.g., /d/ and /r/) were less distinguishable than phonemic contrasts (e.g., /d/ and /ð/) for native English speakers; and that response time was longer in the discrimination of the allophonic contrasts than the phonemic contrasts. At the suprasegmental level, Huang (2001) noted a significantly smaller perceptual distance between two Mandarin allophonic tones, the rising Tone 2 and falling rising Tone 3, than between other tone pairings. In her study, the Mandarin listeners made judgements on pairwise native tone contrasts in a discrimination task. The response time was significantly longer when the listeners discriminated between Tone 2 and Tone 3. This result was accounted for by Tone 3 sandhi rules in Mandarin, i.e., Tone 3 is realized as Tone 2 when it is followed by another Tone 3 at the word level (Duanmu, 2007).

Cross-language tone assimilation was examined by Leung (2008) and So & Best (2010). Like Huang (2001), these two studies also observed tone assimilation due to tone sandhi rules. In Cantonese, the high level and high falling tones are allophonic variations (Yu, 2007). In So & Best (2010), Cantonese listeners without Mandarin experience assimilated Mandarin high level and high falling tones to the Cantonese high level tone. The same result was also observed by Leung (2008) for Cantonese listeners with Mandarin experience. These findings thus revealed phonological tone assimilation for both inexperienced and experienced listeners, contrary to the predictions of the PAM. Cantonese listeners' high sensitivity to phonological tone changes may be accounted for by their rich L1 tone system (Lee, 1999). A recent study indicated that Cantonese listeners exhibit greater sensitivity to the acoustic distinctions and engage in phonological processing more strongly than their Mandarin counterparts, especially in the "speech" condition (Zheng, Minett, Peng, & Wang, 2010). It was suggested a "denser" L1 tone system in Cantonese may reinforce L2 tone perception at both the phonetic and phonological levels (c.f., Bohn & Best, 2012).

2.2.3. Summary

In summary, previous studies have revealed a cue-based process of phonetic assimilation and a function-based process of phonological assimilation. Phonetic assimilation can be predicted by phonetic similarities between L1 and L2 phones, although some are more phonetically similar than others. Moreover, the use of perceptual cues may affect phonetic assimilation. At the functional level, phonetic similarity is no longer the primary factor that directs assimilation process to a higher phonological level. Instead, lexical function associates an L2 phonological category with its counterparts in L1. Language experience is related to assimilation processes at the phonetic and phonological levels. In terms of phonological assimilation, some previous findings are not consistent with predictions of the PAM, suggesting that inexperienced listeners are also sensitive to the phonological correspondence between L1 and L2 phonological categories.

2.3. The current study

The current study investigates tone assimilation at the phonetic and phonological levels by native speakers of Mandarin and Thai. These listeners differed not only in L1 but also in L2 tone experience. Compared to inexperienced listeners who had no Mandarin experience, experienced listeners had 0.5-2 years of experience with Mandarin (for the experienced Thai listeners) or Thai (for the experienced Mandarin listeners).

The experiments employed a cross-language perceptual assimilation task, also known as category mapping task, which has been used widely for segmental assimilation (Levy, 2009; Nishi et al., 2008; Schmidt, 1996; Strange, 2007; Strange et al., 2004). The assimilation task consists of a mapping portion and a rating portion, which are performed sequentially for each token. In the mapping portion, listeners choose a native phonological category which sounds most similar to the incoming non-native sound; then, in the rating portion, they rate the similarity between the native and non-native sounds for the degree of goodness-of-fit on a Likert scale. Three research questions are addressed and predictions are made.

1) **Does tone assimilation occur at the phonetic and phonological levels?** Based on previous findings (e.g., Huang, 2001; Lueng, 2008; So & Best, 2010), it is expected that tone assimilation occurs at both phonetic and phonological levels. In particular, phonological tone assimilation will be demonstrated through Mandarin tone sandhi rules which lead a falling rising tone to be assimilated to a rising tone or a low falling tone and vice versa. The phonetic and phonological assimilation patterns are assumed to fit into "categorizable" or "assimilable" types given the similarities between Mandarin and Thai tone systems.
2) **Does language experience affect phonetic and phonological tone assimilation?**
   It is expected that both experienced and inexperienced listeners will assimilate L2 tones to L1 tone categories at the phonetic level and Thai listeners may attach more importance to F0 contour than Mandarin listeners (Gandour, 1983). Since neither Mandarin nor Thai has an especially dense tone system (c.f., Cantonese), it is expected that only experienced listeners will assimilate L2 tones to L1 tone categories at the phonological level. Inexperienced listeners are expected to assimilate L2 tones to L1 tone categories at the phonetic level. For inexperienced Mandarin listeners, it is likely that Mandarin tones may be assimilated to Thai tone categories based on Mandarin tone sandhi.
3) **Is phonetic similarity predictive of phonetic/phonological assimilation?** According to previous studies (e.g., So & Best, 2010), it is expected that phonetic similarity is predictive of phonetic assimilation but not phonological assimilation. Both Mandarin and Thai listeners are expected to assimilate L2 tones to L1 tone

categories with similar F0 height and contour at the phonetic level (Gandour, 1983). At the phonological level, only tones related to Tone 3 sandhi are expected to be assimilated to dissimilar L1 tone categories. For the rest of tones, phonetic similarity between L1 and L2 tones still predicts assimilation patterns.

2.4. Acoustic analysis

The main purpose of the acoustic analysis was to discover the actual acoustic properties of the Mandarin and Thai tones, in particular those related to F0 height and contour. A set of acoustic parameters were measured, including F0 at five points (0%, 25%, 50%, 75% and 100%) on tone contours, maximum F0, minimum F0, average F0, location of the turning point (i.e. where the pitch movement changed from falling to rising), and pitch range of falling and rising contours.

2.4.1. Method

2.4.1.1. Talkers

The stimuli were produced by two native speakers each of Mandarin and Thai. The Mandarin speakers, one male and one female, were born and raised in Beijing, and were aged 20 and 22, respectively. At the time of recording, the male speaker had 1.5 years of residence in English-speaking countries and the female speaker, 2.5 years. The Thai speakers, also one male and one female, were born and raised in Bangkok, were aged 23 and 20, respectively. The male Thai speaker had 4 years of residence and the female speaker 2 years of residence in Canada or the United States. The four speakers reported no experience with other tone languages and all reported normal speaking and hearing ability. They were all undergraduate students of Simon Fraser University in Canada.

2.4.1.2. Stimuli

A variety of syllables (see Table 2.2) and a hummed /m/ were produced with the four Mandarin and the five Thai tones to create a total of 99 citation tokens (44 Mandarin and 55 Thai). These tokens covered all the stimulus words selected for the three perceptual tasks of the current project, i.e., the tone assimilation task, the tone identification task, and the dichotic tone listening task. Of the 12 syllables, eight (/paa, pju, tuo, fej, kjɛn, kaa, lau, siau/) were produced with both Mandarin and Thai tones. Of the remaining four, two were produced with Mandarin tones only (/pi, fan/), and the other two were produced with Thai tones only (/fəŋ, maa/). The stimulus words carried varying syllable structures, including CV(V), CVC, CCV(V), and CCVC. The CV(V) syllables were produced consistently with a long vowel (VV) with the Thai tones, but not with the Mandarin tones, as vowel duration is phonemic in Thai, but not in Mandarin.

The familiarity of the real words (see Table 2.1) was rated on a five-point Likert scale by five native speakers each of Mandarin and Thai, all of whom were undergraduate students: 1 stood for "least familiar" and 5 for "most familiar". All items and instructions were printed in Mandarin characters or Thai letters, and the raters circled one of the five numbers according to their familiarity with these words. The mean ratings showed a consistent high-level familiarity with the native words: M = 4.4 for the Thai words (SD = 0.9); M = 4.8 for the Mandarin words (SD = 0.3). Moreover, none of the Mandarin or Thai speakers rated /pju/ and /kjɛn/ as legal syllables in their respective native languages.

Table 2.1. Syllables produced with Mandarin or Thai tones for acoustic analysis

Syllable structure	Real word (Mandarin & Thai)	Real word (Mandarin & English)	Real word (Mandarin)	Real word (Thai)	Nonsense word
CV(V)	/lau, pa(a), ma(a)/	/pi/	/ fej, tuo/	/khaa/	
CVC		/fan/	/fəŋ/		
CCV(V)				/sjau/	/pju/
CCVC					/kjɛn/

2.4.1.3. Procedure

The recording was conducted in a sound-absorbent room at the Language and Brain Lab of Simon Fraser University, using the Presonus Digital Audio 24 B27/96K Firewire Recording Interface and a Shure KSM 109 microphone, with a sampling rate of 44.1 kHz, 16 bit. To ensure that each tone was fully produced in a natural utterance, tones in Mandarin and Thai were produced in the final position of a carrier sentence (Duanmu, 2007). The Mandarin words were read within the carrier sentence "我读一个

_____" (I read a _____), with the target word appearing in the final underscored position. Similarly, the Thai words were produced within the frame "ผมอยากบอกว่า _______" by the male speaker and in the frame "ฉันอยากบอกว่า _______" by the female speaker. Both sentences mean "I want to read __________." The only difference between the sentences is the first person pronoun ("I") in the subject position corresponding to the different genders (ผม for male and ฉัน for female). Each sentence was recorded six times (3 repetitions x 2 speakers). Before the formal recording, the speakers practiced all the stimulus words in the carrier sentence several times to ensure a natural, fluent production.

The target words were then excised from the digitized speech files and mean-energy intensity was normalized to 70 dB in Praat (Version 5.1.05, Boersma & Weenink, 2009) for each stimulus word. These words were then played via AKG headsets at a comfortable level to native raters, two each of Mandarin and Thai, who evaluated the goodness of the tone and syllable production. From among the most highly-rated words, one exemplar of each target tone was selected for the acoustic analysis and the perception experiments.

2.4.1.4. Measurements

In total, 88 Mandarin items (11 syllables x 4 Mandarin tone x 2 speakers) and 110 Thai items (11 syllables x 5 Thai tones x 2 speakers) were further modified for acoustic analysis. Three acoustic modifications were performed, including duration normalization, pitch normalization, and creaky voice normalization.

2.4.1.4.1. Duration normalization

The duration of each tone was normalized using Praat (Version 5.1.05, Boersma & Weenink, 2009) to adjust for differences in length caused by speaking rates and syllable structures. In order to maintain all F0 information from the signal, the longest pitch contour of each tone category was first selected (Wang, Jongman, & Sereno, 2003). Of the four Mandarin tones, the selected MT1 was 420 ms; MT2 460 ms; MT3 580 ms; and MT4 360 ms. Of the five Thai tones, the selected tones were: TT1600 ms; TT2 540 ms; TT3 530 ms; TT4 580 ms; and TT5 600 ms. Compared to the Mandarin tones, the Thai tones had a longer average duration (Mandarin 455 ms vs. Thai 570 ms).

Lengthening of tonal duration was achieved using the enhance-pitch-synchronous overlap-and-add (PSOLA) function in Praat, which manipulates the duration of a sound signal without altering the pitch of a speech component (see also Bent, 2005). Once the duration was lengthened, new pitch tracks were extracted automatically by means of the autocorrelation method in Praat. Each tracked pitch point was measured at a 10 ms interval to generate, for example, 51 pitch points for a 500 ms-tone or 55 pitch points for a 540 ms-tone, and so on.

2.4.1.4.2. Pitch measurement & normalization

The F0 values of each tone were measured and normalized in Praat to accommodate the difference in pitch range among the talkers, especially between the males and the females. Pitch settings were adjusted for the male and female speakers because naturally, the average pitch of the female speakers was higher than that of the male speakers (Boersma & Weenink, 2009). As a result, in the normalization of pitch values, the pitch range was set to 70-400 Hz for the female speakers and to 50-300 Hz for the male speakers.

F0 values in Hertz (Hz) were collected for each pitch point at 10 ms intervals from the automatically generated pitch track in Praat. Each F0 value was then converted from Hz to T value by means of the following formula (Ladd, Silverman, Tolkmitt, Bergmann, & Scherer, 1985; Peabody & Seneff, 2009; Rose, 1987; Wang et al., 2003):

$$(1) \mathrm{T} = \left[\frac{\log X - \log L}{(\log H - \log L)}\right] * 5,$$

where X is the pitch (in Hz) at any given point, L is the lowest pitch and H the highest pitch (in Hz) produced by a given talker. T value ranged from 0 to 5, corresponding to the 5-point scale proposed by Chao (1948) to account for the Mandarin tones. According to this formula, 0 refers to the lowest pitch (when log X = log L) and 5 the highest pitch (when log X = log H) in the production of a given talker.

2.4.1.4.3. Creaky voice normalization

Some pitch tracks on a tone contour were absent due to the occurrence of glottalized phonation (or creaky voice) (Gerratt & Kreiman, 2001). Two types of creaky voice appeared in the tone productions, both in Mandarin and Thai. The first type appeared with cycles alternating in amplitude and/or frequency. Pitch values were recovered for this type of creaks by measuring the cycles manually (see Figure 2.1). Another type of creaky voice, vocal fry creaks, co-occurred with extremely low frequency, showing an aperiodic waveform (or irregular phonation) from which it is not possible to obtain a pitch value for the period of each cycle (Ishi, Sakakibara, Ishiguro, & Hagita, 2008). The F0 value was replaced with a zero for the occurrence of vocal fry creaks (c.f., Bent, 2005). Creaky voice has been found on some Mandarin tones in naturally produced speech (Keating & Kuo, 2010), but its appearance is not mandatory (c.f. Wang et al., 2003). Despite the fact that creaky voice is associated with low frequency, such as MT3 (Davison, 1991), it occurs with some higher frequency tones, such as MT2 (Belotel-Grenié & Grenié, 1997; also Yu, 2010). Furthermore, the portion of creaky voice varies from tone to tone. For instance, MT3, consist of creaky voice 68% of the time, while MT2 consists of creaky voice only 8%, on average (Yu, 2010). In the current research, creaky voice was found in the Mandarin tones MT3 and MT4, as well as in the three Thai tones, TT2, TT3 and TT5.

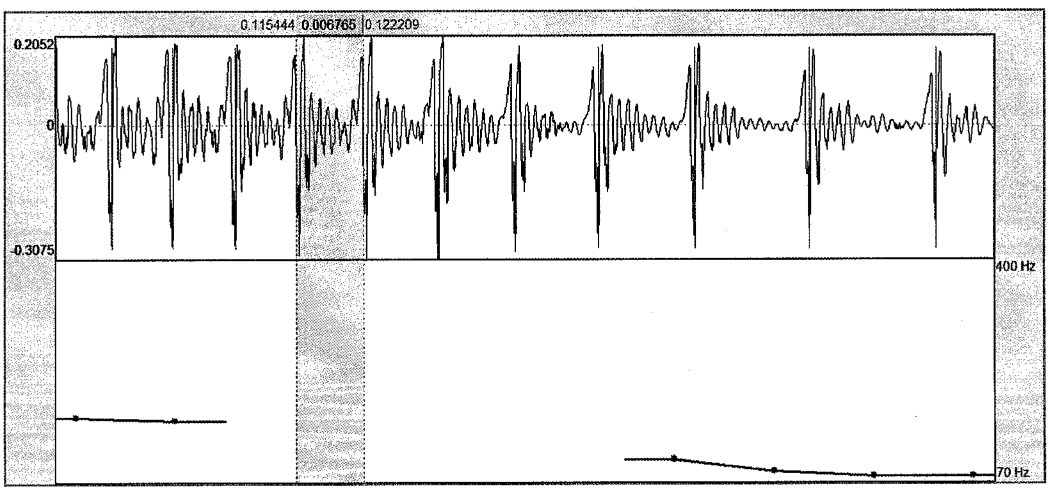


Figure 2.1. ***Top: Waveform of MT3 with the syllable /lau/ produced by a female native speaker. The pink area shows that the pitch value can be recovered by measuring the length of a single period (T) and calculating through the formula f0 = 1/T (about 148 Hz in this case). Bottom: The blue line indicates the pitch track, which is missing with the occurrence of creaky voice.***

To quantify the similarity of F0 height and contour between tone categories in Mandarin and Thai, the following seven parameters (c.f., Bent, 2005; Gandour, Potisuk, Dechongkit, & Suvit, 1992; Khouw & Ciocca, 2007; Wang et al., 2003) were extracted for each tone exemplar: (1) F0 values at five evenly separated points on a tone contour, i.e., 0% (start F0), 25%, 50%, 75% and 100% (end F0); (2) maximum F0; (3) minimum F0; (4) average F0; (5) amount of falling range (in T value); (6) amount of rising range (in T value); and (7) location (in %) of the turning point for the falling rising tones. These measures were selected because it has been suggested that they determine the primary acoustic characters of tone (e.g., Gandour, 1983; Keating & Kuo, 2010) and thus affect tone perception from different perspectives, such as assimilation or hemispheric processing.

2.4.2. Results

Figure 2.2 shows the F0 trajectory of each tone in Mandarin and Thai, based on T values distributed at a 10% interval along each tone contour. Based on the parameters described above, the four Mandarin and five Thai tones were classified into the following four types: (1) level tone: MT1 and TT1; (2) falling tone: MT4, TT2 and TT3; (3) rising tone: MT2 and TT4; and (4) falling rising tone: MT3 and TT5. Table 2.2 provides the mean acoustic measurements of the seven parameters for the nine tones in each tone type.

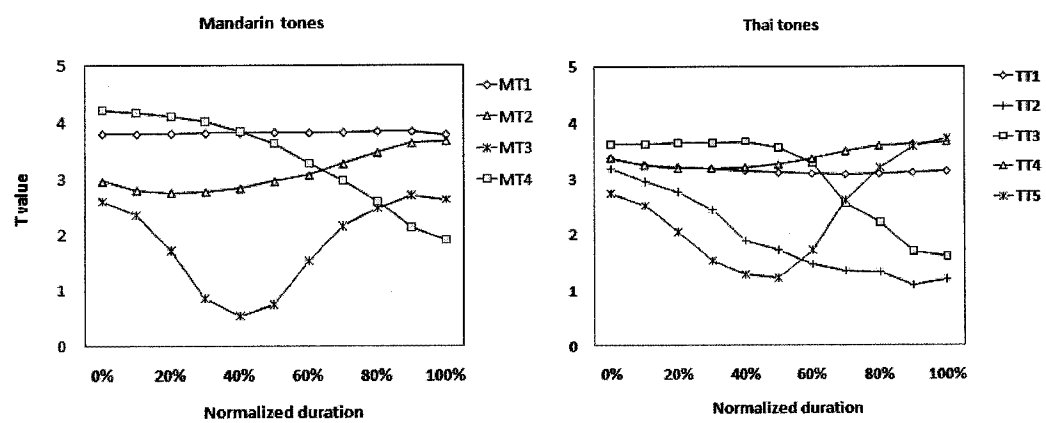


Figure 2.2. ***Mandarin and Thai tones produced by native speakers. Pitch height is shown in T value.***

MT1 and TT1 have a level pattern due to a negligible pitch range (MT1: 0.2 and TT1: 0.4). Nevertheless, MT1 and TT1 differ in F0 values. As shown in Table 2.2, the mean F0 values at the five designated points on tone contour, the maximum F0, the minimum F0, and the average F0 are all higher for MT1 relative to TT1. For this reason, MT1 is realized as a high level tone, while TT1 is a mid level tone (given that the average F0s of MT1 and TT1 are 3.8 and 3.2 out of 5, respectively).

There are three falling tones, MT4, TT2 and TT3. The pitch values of MT4 and TT3 are closer and higher relative to TT2, including the F0 values at the five points, maximum F0, minimum F0 and average F0. The falling range is more similar between MT4 and TT3, demonstrating a steeper falling slope (MT4: 2.7; TT2: 2.6; TT3: 2.2). Due to the difference in F0 height, in particular the starting F0, TT2 can be subcategorized as a low falling tone.

MT3 and TT5 carry a complex falling rising (dipping) contour. A falling rising tone is characterized by a falling slope followed by a rising one. The falling and rising ranges are distinctive, with relatively high T values for both tones (MT3: falling, 2; rising, 2.4, and TT5: falling, 2; rising, 2.8). MT3 and TT5 are also very close in terms of the location of the turning point, despite TT5 showing a higher average F0.

The rising tones MT2 and TT4 both start with a short and slightly falling slope before moving upwards. Compared to the falling rising tones, the amounts of the falling range are nearly negligible, while those of the rising ranges are also fairly indistinct (MT2: falling 0.3; rising: 1 and TT4: falling 0.2; rising, 0.6).

Table 2.2. Mean acoustic measurements for Mandarin and Thai tones. The pitch value is shown in terms of T.

Tone type		Pitch values					Max. F0	Min. F0	Avg. F0	Turning point	Falling range	Rising range
	Tone	0%	25%	50%	75%	100%						
Level	MT1	3.8	3.8	3.8	3.8	3.8	3.9	3.7	3.8			
	TT1	3.4	3.2	3.1	3.1	3.1	3.4	3	3.2			
Falling	MT4	4.2	4.1	3.6	2.6	1.9	4.3	1.6	3.3		2.7	
	TT3	3.6	3.7	3.6	2.3	1.6	3.8	1.6	3		2.2	
	TT2	3.2	2.6	1.7	1.7	1.2	3.5	0.9	2.1		2.6	
Falling rising	MT3	2.6	1.3	0.7	2.3	2.6	3.1	0.5	1.9	46%	2	2.4
	TT5	2.8	1.6	1.2	2.9	3.7	3.8	1	2.4	46%	2	2.8
Rising	MT2	3	2.8	3	3.3	3.5	3.7	2.7	3.1	22%	0.3	1
	TT4	3.4	3.2	3.3	3.5	3.7	3.8	3	3.4	23%	0.2	0.6

In summary, the results of the acoustic analysis indicate the acoustic correspondence between the Mandarin and Thai tones. According to the measures relating to F0 height and contour, these tones were grouped into four tone types, i.e., level tones, MT1 and TT1; falling tones, MT4, TT3, and TT2; rising tones, MT2 and TT4; and falling rising tones, MT3 and TT5. The results are consistent with the numerical descriptions by Ladefoged (2001), except for TT1. The acoustic analysis results were used as an index to interpret the results from the assimilation experiments in the current study.

2.5. Perceptual assimilation

2.5.1. Method

2.5.1.1. Listeners

In total, 72 listeners participated in the assimilation experiment. Of the 36 Mandarin listeners, 20 had no prior experience with other tone languages (inexperienced listeners) and the remaining 16 had instruction and immersion experience with Thai (experienced listeners). The inexperienced Mandarin listeners were born and raised in the Beijing area. They were undergraduate students of several Beijing universities at the time of testing, with a mean age of 20 years (range: 19-21), and comprised nine males and 11 females. This group of listeners is referred to as IM (Inexperienced Mandarin) in the current study. The experienced Mandarin listeners studied Thai as a major subject for two years in China at the Guangxi University for Nationalities and for three months at the University of the Thai Chamber of Commerce in Bangkok, and had a mean age of 21 years (range: 19-24), comprising four males and 12 females. They listeners are labelled as EM (Experienced Mandarin) in the current study.

The other 36 participants were native listeners of Thai. Twenty had no prior experience with Mandarin or other tone languages. They are referred as IT (Inexperienced Thai) listeners in this dissertation. All were undergraduate students of the University of the Thai Chamber of Commerce in Bangkok, with a mean age of 22 years (range: 19-24), and comprised eight males and 12 females. Another 16 native listeners of Thai had 0.5-2 years of Mandarin learning and immersion experience as registered

CSL (Chinese as a second language) students at Beijing College of Chinese Language and Culture. This group consisted of 10 males and 6 females, with a mean age of 23 years (range: 21-27). They are labelled as ET (Experienced Thai) listeners in this study.

All 72 listeners reported normal hearing and speech ability, and all were right-handed and had very limited or no musical training. The experienced listeners reported more use of L2 than L1 during the stay in the target language-speaking countries. On average, they studied the target language for 5-6 hrs per week, with the target language being the teaching language.

2.5.1.2. Stimuli

The stimuli used for the assimilation experiment were described in the acoustic analysis section 2.4.1.2. They were produced by two native speakers of each language. Of the 10 syllables shown in Table 2.2, /tuo/, /fej/, /kha(a)/, /sjau/ as well as the hummed /m/ produced with the Mandarin and Thai tones were used for the assimilation task. The duration was equalized for each tone category in the two languages using the enhance-pitch-synchronous overlap-and-add (PSOLA) function in Praat. The intensity was normalized at 70 dB across all stimuli.

2.5.1.3. Procedure

The assimilation experiment was conducted in quiet rooms in Bangkok for the IT and EM listeners and in Beijing for the IM and ET listeners. They listened to the stimulus words presented via AKG headsets, receiving the signal from a personal laptop. The experiment was run on Paradigm software, Beta version (Paradigm, 2009). Instructions in the native language of the listeners were displayed on the screen of the laptop. The perceptual assimilation task had a mapping portion followed by a rating portion for each item. In the mapping portion, listeners heard a non-native tone and then chose the native tone category which sounded most similar to the presented tone by pressing one of the labelled keys. Immediately afterward, they rated the degree of similarity between the native and non-native tones on a 5-point Likert scale, with 1 representing "least similar" and 5 "most similar". To familiarize listeners with the process, in a practice session, 10 non-native words were randomly presented for mapping and rating. In the following experimental session, 80 Mandarin words (5 syllables x 2 speakers x 4 tones x

2 blocks) and 100 Thai words (5 syllables x 2 speakers x 5 tones x 2 blocks) were presented randomly to the Thai and Mandarin listeners, respectively. This experiment lasted for approximately 20 min. Both the mapping and rating tasks were self-paced, but the listeners were instructed to respond as quickly as possible. Listeners took a self-paced break between the blocks.

2.5.1.4. Data analysis

Data collected from the assimilation task were analyzed to obtain the following three measures: percentage of response, degree of diversity, and correlation between mapping and rating.

Percentage of response indicates how frequently an L2 tone category is being mapped onto an L1 tone category and "modal response" refers to the highest percentage of response. Thus, modal response reveals the most common associations between L1 and L2 tone categories. The sum of percentages of modal and non-modal responses was 100% for an input tone as a consequence of a forced-choice task.

Degree of response diversity (K') was calculated using the following formula (Koopman, personal communication; Simpson, 1949):

$$(2)\ K' = \frac{1}{\sum_{i=1}^{R} P_i^2}$$

In tone assimilation, R is the number of L1 tone categories and P_i is the percentage of responses showing an L2 category (i) being assimilated to L1 tone categories. Minimum diversity (K' = 1) indicates an L2 tone category being mapped only to one L1 tone category, while maximum diversity represents an L2 category being equally mapped to all L1 tone categories; thus K' is equal to the number of L1 tone categories (e.g., five for the Thai group and four for the Mandarin group). Also, according to Krumhansl (1978), if one object within a stimulus set is similar to many objects, it should be less similar to a particular object. Therefore, in tone assimilation, if K' reaches the maximum diversity, an L2 tone category is less similar to the modal response L1 tone category; if K' is close to the minimum diversity, an L2 category is very similar to the modal response L1 tone category. For instance, in the case of four tone categories in L1, if the percentage of responses of an L2 tone being assimilated to the four L1 tone categories is 25%, K' reaches its maximum level, thus is not more similar to any of the L1 tones.

The average goodness rating was also calculated for each mapping pattern, followed by a measurement of the correlation between the mean percentage of responses and the mean ratings. It was expected that the ratings would be positively correlated with the mapping responses.

2.5.2. Results

Percentage of responses, degree of response diversity and relationship between mapping and rating are provided first for the assimilation of Mandarin tones to Thai tones by the inexperienced and experienced Thai listeners and then for the assimilation of Thai tones to Mandarin tones by the inexperienced and experienced Mandarin listeners.

2.5.2.1. Assimilation of Mandarin tones to Thai tones

2.5.2.1.1. Percentage of response

Table 2.3 shows the mean percentage of mapping responses and ratings for Mandarin tones being mapped to Thai tone categories by the inexperienced (IT) and experienced (ET) Thai listeners. According to modal responses (highlighted in bold), MT1 was mapped onto TT1 by both IT and ET listeners with relatively high ratings (IT: 4 and ET: 4.2), despite a difference in percentage of responses (IT: 59% vs. ET: 96%). Moderate consistency in responses was also observed for MT4, which was mapped to TT3 by both groups (IT: 55% vs. ET: 72%) with a high mean rating (4 in both cases). MT2 and MT3, however, were mapped to different Thai tones by the IT and ET listeners. The IT listeners mapped MT2 mainly to TT4 (mapping: 61%, rating: 3.6) and MT3 mainly to TT5 (mapping: 51%, rating: 4), while the ET listeners mapped MT2 primarily to TT5 (mapping: 53%, rating: 3.8) and MT3 to TT2 (mapping: 45%, rating: 3.5). Nevertheless, the ET listeners also showed a notable percentage of responses when mapping MT2 to TT4 (26.6%) and MT3 to TT5 (29.4%).

Table 2.3. Percentage of response and similarity ratings for Mandarin tones being mapped onto Thai tones by IT and ET listeners

		TT1		TT2		TT3		TT4		TT5	
		Mapping	Rating	Mapping	Rating	Mapping	Rating	Mapping	Rating	Mapping	Rating
IT	MT1	**59.0**	4.0	5.0	3.0	5.5	3.5	27.8	3.4	2.8	3.1
	MT2	3.5	3.0	5.0	2.9	8.0	3.6	**61.3**	3.6	22.3	3.8
	MT3	3.0	2.7	26.8	3.4	9.3	3.3	10.5	3.3	**50.5**	4.0
	MT4	3.5	3.0	18.8	3.6	**54.5**	4.0	20.8	3.5	2.5	3.0
ET	MT1	**96.3**	4.2	0	n.a.	1.8	3.0	1.9	3.0	0	n.a.
	MT2	4.4	4.4	6.3	3.5	9.4	3.6	26.6	3.5	**53.4**	3.8
	MT3	5.3	4.2	**45.3**	3.4	14.7	3.8	5.3	2.6	29.4	3.6
	MT4	2.8	4.1	6.6	3.9	**72.2**	4.0	15.9	3.5	2.5	3.1

2.5.2.1.2. Degree of diversity

Table 2.3 above provides an overview of tone assimilation by the IT and ET listeners. As indicated, some Mandarin tones were assimilated to one or two Thai tone categories, but others were assimilated to a greater number of Thai tone categories. For example, although both IT and ET listeners perceived MT1 as most similar to TT1, the IT listeners assimilated MT1 into all the native tones, while the ET listeners mapped MT1 only to TT1, TT3 and TT4. In addition, the results also show that IT listeners mapped Mandarin tones to some Thai tones with fairly high percentages of response, such as the mappings of MT1 to TT4, MT2 to TT5, MT3 to TT2, and MT4 to TT4 and TT2. To quantify the variance in distribution, degree of mapping diversity (K') was calculated using Formula (2) above and presented in Figure 2.3 below for the IT and ET listeners.

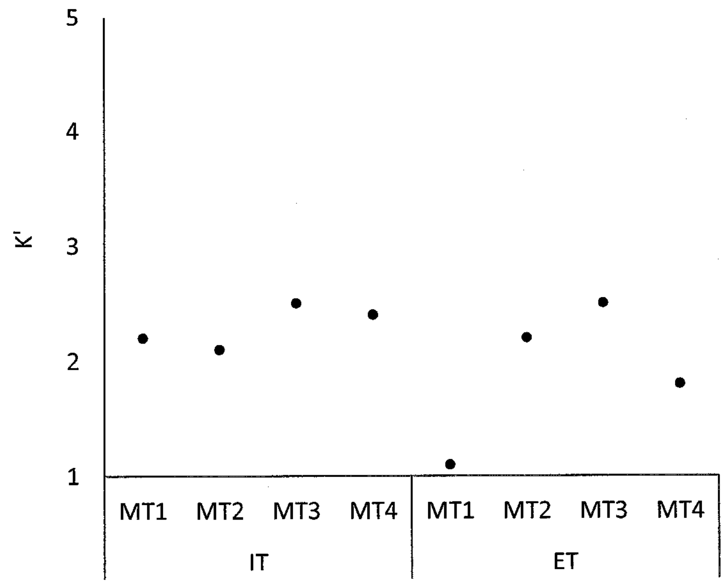


Figure 2.3. ***Mean degree of diversity for Mandarin tones being mapped onto Thai tones by IT and ET listeners***

A two-factor mixed-design ANOVA was conducted using K' as the dependent variable, experience group (IT and ET) as a between-groups factor, and input tone (MT1, MT2, MT3 and MT4) as a repeated measure. Main effects of group, $F\,(1, 34) = 7.2$, $p < .011$ and tone, $F\,(3, 102) = 11.6$, $p < .0001$, as well as interaction of group and tone, were significant, $F\,(3, 102) = 8.2$, $p < .0001$. To simplify the analysis of tone effects on K' for individual groups, follow-up one-factor ANOVAs were performed using K' as the dependent variable and tone as the independent variable for the IT and ET groups. A significant effect of tone was detected for the ET group, $F\,(3, 45) = 25.1$, $p < .0001$, but not for the IT group $F\,(3, 57) = 1.1$, $p = .35$. Results of Bonferroni-adjusted post hoc analyses for the ET group indicated significant differences in K' between MT1 and MT2 ($p < .0001$), MT1 and MT3 ($p < .0001$), MT1 and MT4 ($p < .0001$), as well as between MT3 and MT4 ($p = .006$), showing a lower diversity of mapping for MT1 compared to the other three tones and MT4 compared to MT3 for the ET listeners. To discover the effects of group on K' for individual tones, one-factor ANOVAs were performed for MT1, MT2, MT3 and MT4. Significant between-group differences were observed for MT1, $F\,(1, 35) = 27.9$, $p < .0001$ and MT4, $F\,(1, 35) = 8.5$, $p = .006$, though not for MT2, $F\,(1, 35) = .27$, $p = .61$ or MT3 $F\,(1, 35) = .03$, $p = .87$, indicating lower diversity of mapping of MT1 and MT4 onto Thai tone categories by the ET listeners.

2.5.2.1.3. Relationship between mapping and rating

Table 2.3 shows variable relationships between the percent of mapping responses and the corresponding ratings, as some mappings with a high percentage of response also had a high rating score. For example, the ET listeners assimilated MT1 to TT1 as the modal response, while the similarity between the two tones was also rated highly (4.2). However, in other cases, the mapping and rating results were less consistent. For example, MT4 was assimilated to TT1 with a very low percentage of response (2.8%), but the two tones were rated as highly similar (4.1). A bivariate correlation analysis was then conducted between the average mapping responses for each assimilation pattern and the average rating scores. Correlation analysis indicated a moderate, but significant relationship between the two measures for the Thai listeners, r (38) = .53, p = .001 (Figure 2.4).

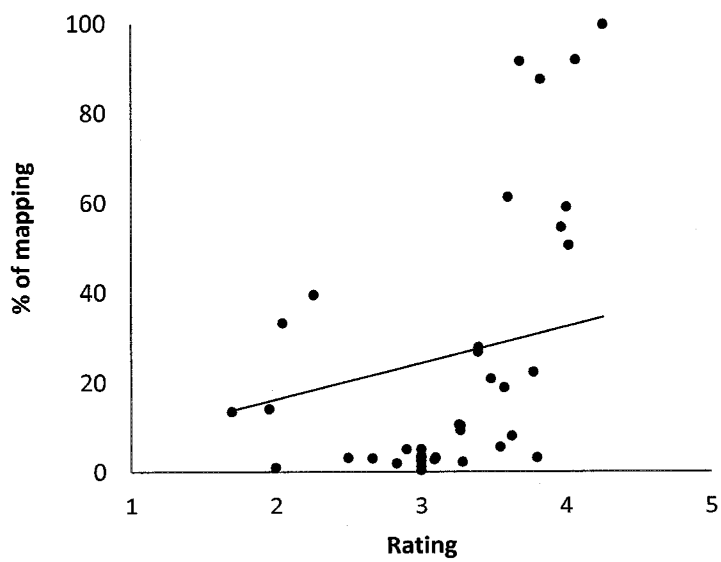


Figure 2.4. ***Relationship between mapping and similarity rating for Thai listeners***

2.5.2.2. Assimilation of Thai tones to Mandarin tones

2.5.2.2.1. Percentage of responses

Table 2.4 presents the modal and non-modal responses of the Thai tones being assimilated to the Mandarin tone categories. With regard to the modal response (marked in bold), TT1 was mapped to MT1 and TT3 to MT4 by the Mandarin listeners of both groups with relatively high ratings (TT1→MT1, IM: 99%, 4; EM: 99%, 4.2 and TT3→MT4, IM: 95%, 4 and EM: 95%, 4.2). In addition, the listeners also mapped TT4 to MT2 (IM: 84%, EM: 55%), but with relatively lower ratings (IM: 3.3, EM: 2.7). Inconsistency between the two groups was observed for TT2 and TT5, as TT2 was mapped to MT4 (mapping: 84%, rating: 3.7) and TT5 to MT3 (mapping: 78%, rating: 3.7) by the IM listeners; whereas TT2 was mapped to MT3 (mapping: 88%, rating: 4.1) and TT5 to MT2 (mapping: 69%, rating: 3.7) by the EM listeners. The table also shows that TT5 was mapped to MT3 with a notable percentage of response (30.3%) by the EM listeners.

Table 2.4. Percentage of response and similarity ratings for Thai tones being mapped onto Mandarin tones by IM and EM listeners

		MT1		MT2		MT3		MT4	
		Mapping	Rating	Mapping	Rating	Mapping	Rating	Mapping	Rating
IM	TT1	**98.8**	4.0	1.0	1.8	0.3	4.0	0	n.a.
	TT2	2.8	2.5	1.5	2.3	12.0	3.0	**83.8**	3.7
	TT3	4.3	3.2	0	n.a.	1.0	1.8	**94.8**	4.0
	TT4	15.3	2.7	**83.8**	3.3	1.0	3.0	0	n.a.
	TT5	0	n.a.	22.5	3.4	**77.5**	3.7	0	n.a.
EM	TT1	**99.1**	4.1	0.3	4.0	0.6	2.0	0	n.a
	TT2	1.3	3.0	3.4	4.2	**87.8**	4.1	7.5	4.0
	TT3	0	n.a.	0.3	3.0	5.0	3.6	**94.7**	4.2
	TT4	20.9	1.8	**54.7**	2.7	10.6	2.9	13.8	2.0
	TT5	0.3	2.0	**69.4**	3.7	30.3	3.5	0	n.a.

2.5.2.2.2. *Degree of diversity*

The K' scores were obtained via Formula (2) for each Thai tone being mapped to Mandarin tone categories; the results are displayed in Figure 2.5. A mixed-design ANOVA was computed, using K' as the dependent variable, experience group (IM and EM) as the between-groups factor, and input tone (TT1, TT2, TT3, TT4 and TT5) as the repeated measure. The main effect of tone proved significant, $F(4, 136) = 16.5$, $p < .0001$, whereas the effect of group $F(1, 34) = .49$, $p = .49$ and the interaction between group and tone were not significant, $F(4, 136) = 1.6$, $p = .17$. Further Bonferroni-adjusted post hoc analyses indicated significantly less diversity for TT1 than TT2 ($p = .001$), TT4 ($p < .0001$) and TT5 ($p < .0001$) and also significantly less diversity for TT3 than TT4 ($p = .001$) and TT5 ($p < .0001$).

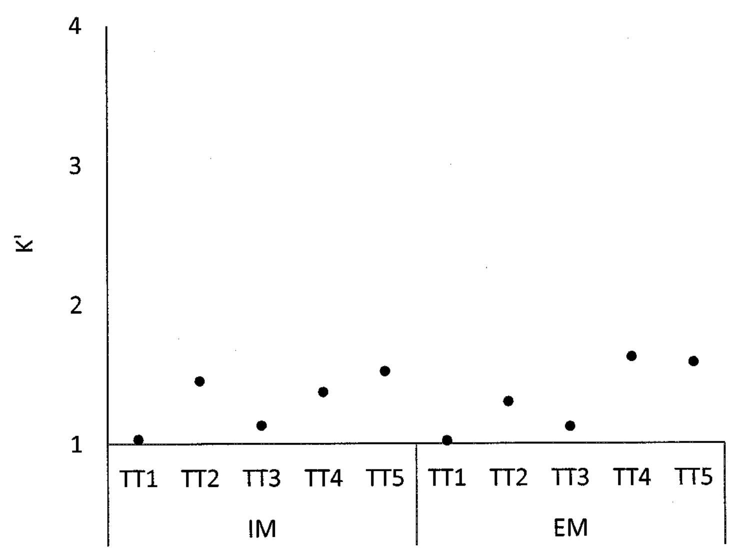


Figure 2.5. Mean degree of diversity for Thai tones being mapped onto Mandarin tones by IM and EM listeners

2.5.2.2.3. *Relationship between mapping and rating*

For the same reason noted above, a bivariate correlation analysis was performed to examine the relationship between the percentage of mapping response and the corresponding ratings for the Mandarin listeners' performance. Similarly, a moderate but significant correlation was observed between the two measures, $r(32) = .54$, $p = .001$ (Figure 2.6).

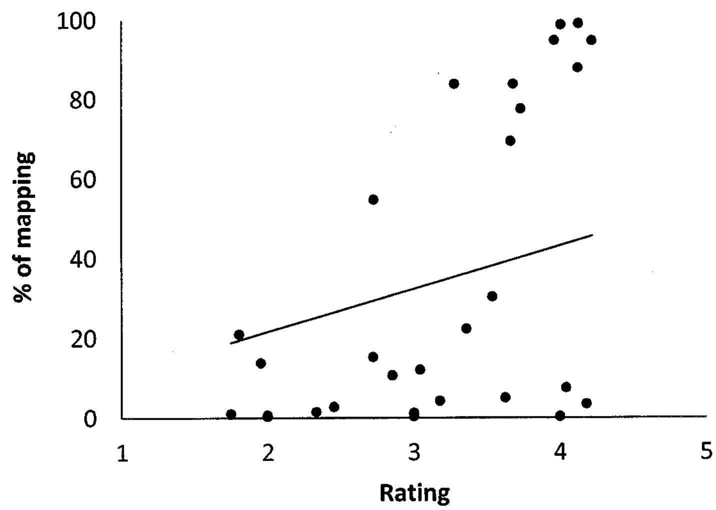


Figure 2.6. ***Relationship between mapping and similarity rating for Mandarin listeners***

2.5.2.3. Summary

To summarize, modal responses indicate that the perceived relationships between Mandarin and Thai tone categories are based on both phonetic similarity and phonological function. The results show that both inexperienced and experienced listeners assimilated some of the L2 tones to the phonetically most similar L1 tone categories. As shown in Table 2.5, the ET and IT listeners assimilated MT1 to TT1 and MT4 to TT3, and the EM and IM listeners also assimilated TT1 to MT1, TT3 to MT4 and TT4 to MT2. A difference was also observed between the inexperienced and experienced listeners. For the Thai listeners, MT2 was assimilated to TT4 by the IT listeners, but to TT5 by the ET listeners, while MT3 was assimilated to TT5 by the IT listeners, but to TT2 by the ET listeners. Similarly, for the Mandarin listeners, TT2 was assimilated to MT4 by the IM listeners, but to MT3 by the EM listeners; TT5 was assimilated to MT3 by the IM listeners and to MT2 by the EM listeners. Additionally, degree of diversity also revealed differences between the inexperienced and experienced listeners in the Thai group but not in the Mandarin group. The ET listeners showed lower degree of diversity when mapping MT1 to TT1 and MT4 to TT3 than the IT listeners. In terms of individual tone patterns, level tones (MT1 and TT1) and high falling tones (MT4 and TT3) showed lower diversity relative to other types of tones, in particular for the ET, IM and EM listeners.

Table 2.5. Modal responses for Mandarin and Thai tones by IT, ET, IM and EM listeners

Input tone	Input tone type	Modal response			
		IT	ET	IM	EM
MT1	Level	TT1	TT1		
MT2	Rising	TT4	TT5		
MT3	Falling rising	TT5	TT2		
MT4	Falling	TT3	TT3		
TT1	Level			MT1	MT1
TT2	Falling			MT4	MT3
TT3	Falling			MT4	MT4
TT4	Rising			MT2	MT2
TT5	Falling rising			MT3	MT2

The results also indicate moderate but significantly positive correlations between rating scores and percentage of mapping responses for the Mandarin and Thai listeners, suggesting the reliability of the experimental design, as a high percentage of mapping responses was generally associated with high rating scores.

2.6. Discussion

In response to the three research questions (section 2.5), the findings reported here demonstrate tone assimilation at both the phonetic and phonological levels (Question 1). The occurrence of the two levels of assimilation was also found to be related to the listeners' L2 experience (Question 2). In other words, L2 experience led the experienced listeners to assimilate non-native tones to native tone categories at both the phonetic and phonological levels, while lack of L2 experience caused inexperienced listeners to assimilate non-native tones to native tone categories solely at the phonetic level. Moreover, phonetic similarity was found to be predictive of tone assimilation, in particular for the inexperienced listeners (Question 3). The following sections discuss the effects of L2 experience and phonetic similarity on tone assimilation at the phonetic and phonological levels, and the implications of these findings for current models of cross-language perception.

2.6.1. *Phonetic assimilation*

In terms of phonetic assimilation, as predicted, L2 tones were assimilated to phonetically similar L1 tones, in particular by the inexperienced listeners. In accordance with the acoustic analysis results (section 2.6.2), L1 and L2 tones with the most similar acoustic features tended to be perceived to be most similar. Nevertheless, if an L1 tone has only one phonetic counterpart in L2, despite some acoustic differences, the L2 tone was still assimilated to the L1 tone. For example, although Mandarin and Thai level tones differ in F0 height, all four groups perceived these two tones to be most similar. In contrast, if an L1 tone has two phonetic counterparts in L2, assimilation was in favor of the closer match, i.e., Category Goodness in PAM terms. For example, the Thai listeners appeared to assimilate the Mandarin high falling tone to their L1 high falling tone rather than to L1 low falling tone. The Mandarin listeners, however, assimilated both Thai high falling tone and Thai low falling tone to Mandarin high falling tone, showing a comparable percentage of responses and degree of diversity. These results may suggest distinctive degrees of sensitivity to F0 height and contour for the Mandarin and Thai listeners (c.f., Gandour, 1983). That is, the Mandarin listeners were more tolerant of (or less sensitive to) F0 height differences and paid more attention to the F0 contour, while the Thai listeners, in particular the inexperienced listeners, were sensitive to both F0 height and contour. As a result, two tones differing in F0 height tended to be perceived as less similar (e.g., level tones) or even dissimilar (e.g., the high falling tone and the low falling tone) by the IT listeners.

The phonetic tone assimilation patterns also demonstrated the effect of L2 experience, in particular for the Thai listeners. Across the tones, the ET listeners' performance was less variable than the IT listeners', as indicated by percentage of responses and degree of diversity. For individual tones, the ET listeners' choice of phonetic equivalents between level tones and falling tones was more focused than the IT listeners' choice. This discrepancy may be explained by the effect of L2 experience, as learning experience has been suggested to induce native-like use of perceptual cues in the discrimination of L2 tone categories (Guion & Pederson, 2007). Like the Mandarin listeners in both groups, the ET listeners showed more tolerance for differences in F0 height than the IT listeners, for example, in the assimilation of L2 level tone to L1 level tone. The PAM-L2 posits that perceptual assimilation by the experienced listeners occurs at both the phonetic and phonological levels, a prediction that was supported by the current study. Thus, the ET listeners' performance may have reflected their recognition of both phonetic and phonological correspondences between L1 and L2 tones. In other words, Mandarin high level tone and Mandarin high falling tone were selected by the ET listeners as Thai mid level tone and high falling tone, due to their similar phonetic features and their lexical function. Moreover, Thai rising tone did not correspond to any Mandarin tone categories for the ET listeners, suggesting that they did not associate this tone with Mandarin tone categories at both the phonetic and phonological levels.

A between group difference was not observed for the Mandarin listeners. As mentioned earlier, this finding was probably due to the Mandarin listeners' focus on F0 contour, such that differences in F0 height were neglected, especially by the IM listeners. In addition to the level and falling tones, the Thai rising tone was also assimilated to the Mandarin rising tone at the phonetic level by the Mandarin listeners. However, the EM listeners perceived the Thai rising tone as less similar to the Mandarin rising tone relative to the IM listeners. This finding echoes the ET listeners' performance, showing that the experienced listeners were less inclined to relate the Thai rising tone to Mandarin tone categories.

2.6.2. Phonological assimilation

Another main finding of the current study is that the experienced listeners exhibited a phonological tone assimilation process, in which L2 tone categories were assimilated to L1 tone categories with the same lexical function. As shown in previous studies, tone sandhi rules often link L1 tone categories with L2 tone categories at the phonological level for native and non-native listeners (Huang, 2001; Leung, 2008; So & Best, 2010). In the current study, phonological assimilation was shown in two respects, one of which is related, and one of which is unrelated, to tone sandhi rules. When tone assimilation was related to tone sandhi, phonetic similarity became less important. For example, an L2 rising tone or a low falling tone was assimilated to an L1 falling rising tone. In contrast, when tone assimilation was unrelated to tone sandhi, L1 and L2 tone categories were phonetically similar or identical. For example, an L1 mid level tone was assimilated to an L2 high level tone, or an L1 high falling tone was assimilated to an L2 high falling tone. These two patterns also differentiate phonetic tone assimilation from phonological tone assimilation, as phonetic assimilation exhibits only the patterns that are not related to tone sandhi.

In the current study, the IM listeners did not appear to assimilate L2 tones to L1 tone categories at the phonological level, even though they were familiar with tone sandhi rules in L1 and were able to detect the phonetic similarities between L1 and L2 tone categories. Instead, the IM listeners showed similar assimilation patterns to the IT listeners (Table 2.6) who had no knowledge of tone sandhi rules in Mandarin. These findings suggest that phonological tone assimilation is independent of phonetic tone assimilation in a cross-language context for the IM listeners (c.f., Huang, 2001). Therefore, phonetic similarities between L1 and L2 tone categories do not transfer to the phonological level. Moreover, the fact that the IM listeners did not show tone assimilation at the phonological level does not support previous findings (So & Best, 2010; Zheng et al., 2010). This result suggests that enhanced tone perception may occur when inexperienced listeners speak an L1 tone language with a large number of tone categories and rich tone changes (e.g., Cantonese).

By contrast, the performance of the experienced listeners demonstrated the effects of Mandarin tone sandhi on tone assimilation. When the input tone was a falling rising tone, the ET listeners associated it with a low falling tone and the EM listeners associated it with a rising tone; when the input tone was a rising tone, the ET listeners mapped it to a falling rising tone; when the input tone was a low falling tone, the EM listeners assimilated it to a falling rising tone (Table 2.5). The performance of the ET listeners also indicates that the phonological use of tone categories, including tone sandhi rules, can be acquired by L2 listeners. In addition, Thai rising tone was not perceived as the phonological equivalent of falling rising tone by the ET and EM listeners, although it was perceived to be phonetically most similar to the Mandarin rising tone. Once again, the experienced listeners indicated a combination of phonetic and phonological assimilation processes relating to Thai rising tone. The reason why Thai rising tone appeared to be an exception to the experienced listeners needs to be explored in future research.

2.6.3. *Implications for models of cross-language speech perception*

The main purpose of the current study is to examine to what extent recent models of cross-language perception can be applied to tone assimilation with regard to three issues: (1) the existence of two levels of assimilation; (2) the effect of language experience on tone assimilation; and (3) the relationship between phonetic similarity and tone assimilation.

First, the current study supports the PAM's assumption of the existence of a lower-level, cue-based assimilation process (phonetic assimilation) and a higher-level, function-based assimilation process (phonological assimilation). Even before the formation of L2 phonological categories is completed, learners are able to associate L1 and L2 categories based on their lexical functions through study and living experience with L2.

Second, L2 experience is essential to a higher-level assimilation process, in particular for listeners whose L1 tone system is of a moderate size (e.g., Mandarin and Thai). In contrast, a lower-level assimilation process is due to listeners' sensitivity to acoustic distinctions and is unrelated to the functional use of tone categories. In agreement with the PAM, the current study presents a lower-level assimilation process for inexperienced listeners and both a lower- and higher-level assimilation process for experienced listeners.

Third, tone assimilation associates L2 tones with phonetically "identical", "similar" and "new" L1 tone categories in support of the SLM (Flege, 1995). However, unlike the SLM's predictions, the current findings indicate that perceptual assimilation is not limited to the phonetic level, as "new" (dissimilar) tone categories in L1 and L2 may be associated only at the phonological level on the basis of tone sandhi rules, whereas "identical" and "similar" tone categories in L1 and L2 can be related to both phonetic and phonological levels, e.g., L2 high falling tone being assimilated to L1 high falling tone and L2 mid level tone being mapped to L1 high level tone.

2.7. Concluding remarks

To conclude, the current study reveals tone assimilation at the phonetic level for inexperienced listeners, but tone assimilation at the phonetic and phonological levels for experienced listeners. Thus, experienced listeners' assimilation involves more complex perceptual processes, which focus not only on the phonetic properties, but also the phonological status, of L1 and L2 tone categories.

Tone sandhi is found to be tied to tone assimilation at the phonological level. However, as tone sandhi does not exist in Thai, the current study could examine tone assimilation only in relation to Mandarin tone sandhi rules. In future work, researchers may consider investigating tone assimilation by employing two tone systems that both carry allophonic tone contrasts, such as Mandarin and Cantonese (Lee, 1999; Yu, 2007). Additionally, phonological tone assimilation may also be detected by replacing L1 tone categories with L2 tone categories on a number of real words in L1 (Best & Tyler, 2007). Given only a difference in tone, if the replacement does not affect word meanings, such a finding may suggest that the two tone categories share the same phonological function.

Future research may also investigate phonetic and phonological category assimilation using other lexical prosodic features, such as Japanese pitch accent and English word stress. As these features are realized with distinctive phonetic properties and phonological functions relative to tone (Wu, Tu, & Wang, 2011), it is worth comparing the effects of phonetic and phonological factors on the assimilation patterns for different types of lexical prosodic features. Moreover, since perceptual assimilation is linked with listeners' L2 experience, further work on prosodic assimilation may incorporate L2 experience at different stages, allowing for a quasi-longitudinal observation about the establishment of an L2 phonological system or the formation of L2 phonological categories.

3. Identification of Mandarin tones

3.1. Background

To process meaningful words successfully in tone languages, listeners must focus their attention not only on the segmental inventories, but also on a suprasegmental feature: tone (Brown-Schmidt & Canseco-Gonzalez, 2004). Tone is characterized uniquely in terms of its acoustic properties and phonological status. First, it is linked with monosyllables to distinguish word meanings. Second, F0 height and contour are the primary perceptual cues which contribute to over 90% of tone identification (Abramson, 1997; Lin, 1988).Third, like segmental inventories, multiple tones can be complementarily distributed as allophonic variations in some languages, such as Mandarin and Cantonese (Duanmu, 2007; Yu, 2007). These characteristics may cause tone to be perceived in different manners according to listeners' experience with it. Therefore, the main purpose of the current study is to investigate the effects of acoustic and phonological factors on tone perception on listeners from different prosodic backgrounds. In addition, as lexical status of word stimuli has been found to affect the phonemic perception of segments (Lieberman, 1963) and of suprasegmental features (Kleber & Niebuhr, 2010), its effects on tone perception are also examined in this study. In the following sections, previous research on tone perception is reviewed, focusing on the effects of language experience, acoustic properties, phonological status, and lexical status of carrier words.

3.1.1. The effects of language experience

Tone perception is experience-sensitive (Blicher, Diehl, & Cohen, 1990; Guion & Pederson, 2007; Lee et al., 1996; Wang, 1976; Wang et al., 1999, 2001; Wayland & Guion, 2004). Previous findings indicate that L1 exposure to a tone system leads tone to be perceived categorically, while lack of exposure tends to result in difficulty recognizing tone categories, especially for non-tone language speakers (Hallé et al., 2004; Wu & Lin, 2008). Studies have also found that exposure to an L2 tone system induces a more native-like perception pattern in terms of accuracy and use of perceptual cues (e.g., Guion & Pederson, 2007; Wayland & Guion, 2004). For example, in Guion & Pederson (2007), native Mandarin listeners, advanced English learners of Mandarin, and English listeners without tone experience discriminated tone patterns differing in F0 height and contour. In their study, Mandarin listeners and advanced English learners chose F0 contour as the primary perceptual cue, while English listeners without tone experience attached more importance to F0 height at the extremities (i.e., beginning or end point) of a tone.

In addition, researchers have investigated to what extent L1 tone background affects L2 tone perception. Some results indicate that L1 tone experience enhances perception of L2 tones (Lee et al., 1996; Wayland & Guion, 2004). For example, Wayland & Guion (2004) discovered that Mandarin monolingual listeners were better than their English counterparts in the perception of two tones (mid and low) in Thai before and after participating in several short laboratory training sessions. The findings of their study suggest that L1 tone experience is transferable to the perception of L2 tones. Lee et al.'s (1996) study provides two different findings. One is that Cantonese speakers who had no Mandarin experience outperformed their English counterparts in the perception of Mandarin tones. The other is that Mandarin speakers who had no Cantonese experience showed comparable performance relative to their English counterparts. In another study examining the perception of Cantonese tones (Francis et al., 2008), overall performance also indicated little difference between Mandarin and English listeners, suggesting no advantage of L1 experience for the Mandarin listeners in the perception of Cantonese tones.

Lastly, previous studies have also compared tone language speakers and non-tone language speakers for their sensitivity to non-speech pitch contrasts (Bent et al., 2006; Stagray & Downs, 1993). For example, Bent et al. (2006) presented Mandarin tones and synthesized tone patterns imposed on monosyllables to monolingual Mandarin and English listeners. Their results indicated that Mandarin listeners were better at discriminating and identifying native tones than their English counterparts. Their perception of the non-speech pitch patterns, however, showed more errors relative to the English listeners.

To summarize, previous studies demonstrate the effects of exposure to target tone system on tone perception and also note inconsistency in terms of whether and how L1 tone experience may be transferred to L2 tone perception. However, previous research has not considered whether L1 prosodic experience may be transferred to L2 tone perception by advanced learners from tone and non-tone language backgrounds.

3.1.2. *Phonetic and phonological effects*

The acoustic properties and phonological status of tone have been found to affect tone perception (Francis et al., 2003, 2008; Hallé et al., 2004; Huang, 2001, 2004; Shen & Lin, 1991). Research on the effects of acoustic properties indicates that sensitivity to F0 correlates may lead to different perception patterns. For tone language speakers, acoustic similarities between L1 and L2 tone categories tend to enhance L2 tone perception. As shown in Chapter 2, both Mandarin and Thai listeners, irrespective of their L2 experience, appear to assimilate L2 tones to L1 tone categories with the most similar F0 height and contour. Thus, acoustic similarities can be essential for tone language speakers in recognizing L2 tone categories. For non-tone language speakers, the acoustic correspondence between L1 prosodic patterns and L2 tones may also play an important role in the identification of L2 tone categories. In a recent study, Francis et al. (2008) used a tone identification task to investigate the perception of six Cantonese tones by native listeners of Mandarin and English. Their findings demonstrate the effect of acoustic correspondence on the performance of Mandarin and English listeners. For instance, Mandarin listeners were good at identifying Cantonese high level, mid level and rising tones, as they correspond to Mandarin high level and rising tones in terms of F0 height and contour. For English listeners, similar F0 realizations also helped them identify the Cantonese high rising tone and low rising tone, probably because both tones correspond to the yes-no question pattern in English. Francis et al.'s (2008) findings also echo Gandour (1983), demonstrating English speakers' sensitivity to F0 height: their identification of three level tones in Cantonese was excellent.

In contrast, acoustic similarities between tones in the same language may cause discrimination difficulty for native and non-native listeners. For example, Mandarin Tone 2 and Tone 3 are realized with a similar dipping contour (Blicher et al., 1990; Moore & Jongman 1997; Shen & Lin, 1991), even though Tone 2 is conventionally referred to as a rising tone (Chao, 1968). Previous studies have revealed these two tones as the most confusable tone pair in Mandarin for both native (e.g., Chuang & Hiki, 1972; Zue, 1976) and non-native listeners (e.g., Kiriloff, 1969; Wang et al., 1999). Some tone researchers suggest that the ratio of falling and rising slopes to the whole tone contour and the timing of the turning point are crucial in the discrimination of these two tones (e.g., Shen & Lin, 1991). Level F0 contour has also been found to cause perceptual difficulty for Thai speakers in categorizing three native tones: mid, low and high. Abramson (1973) examined Thai listeners' identification and discrimination of these three level tones using a categorization tone perception test. The result showed a within-category effect among these three tones, indicating that perceptual confusion arose from the same level tone contour.

In addition to acoustic properties, the phonological status of tone categories may affect tone perception for native and experienced listeners. The results reported in Chapter 2 of this dissertation reveal that phonological similarity leads one tone to be assimilated to its phonological equivalent(s). A few studies have investigated the role of phonological status in tone perception (Huang, 2001; Leung, 2008; So & Best, 2010). In agreement with previous findings on the phonological perception of segmental inventories (Boomershine et al., 2008; Kazanina, Philips, & Idsardi, 2006; Pegg & Werker, 1997; Peperkamp et al., 2003; Whalen, Best, & Irwin, 1997), tone categories are difficult to discriminate if they are allophonic variations. Tone sandhi has been found to produce phonological effects on tone perception. For example, Mandarin Tone 2 is an allophone of Tone 3, in that Tone 3 is realized as Tone 2 when it is followed by another Tone 3 at the word level. Huang (2001) found that discrimination of Tone 2 and Tone 3 was more difficult than discrimination of other Mandarin tone pairings for native listeners (see also the review in Chapter 2). Similarly, Cantonese high falling tone is also realized as an allophonic variant of high level tone (Chen, 2000; Yu, 2007). Cross-language studies, such as Lueng (2008) and So & Best (2010), have observed that Mandarin high level tone and high falling tone were assimilated to Cantonese high level tone by Cantonese listeners (see also the review in Chapter 2). The performance of Mandarin and Cantonese listeners supports the view that phonological equivalence causes confusion between allophonic tone pairs.

In summary, acoustic correspondence between L1 prosodic patterns and L2 tone inventories has been found to enhance listeners' perception of L2 tones. Nevertheless, acoustic similarities may impede listeners' discrimination between tone categories from the same language. Moreover, listeners' choice of perceptual cues (F0 height or contour) can influence their perception of L2 tones. The phonological status of tone categories may make allophonic tone variations difficult to differentiate by tone language speakers. However, there has been no data indicating that phonological similarity affects tone perception by non-tone language speakers who have L2 experience.

3.1.3. The effects of lexical status

Lexical status refers to the semantic situation with which tone is associated. For example, Mandarin Tone 1 imposed on the syllable *ma*[1] yields a meaningful word (*mother*), but the same tone imposed on *bu*[1] generates a nonsense word. The influence of lexical information on tone perception has been studied to determine whether tone perception is improved by the semantic information conveyed in real words and to determine whether tone perception is confounded with word perception. Previous studies have yielded contradictory findings on the effects of lexical status. For example, Lee et al. (1996) tested Cantonese, Mandarin and English listeners for their discrimination of Cantonese and Mandarin tones imposed on real and nonsense words in these respective languages. Cantonese listeners performed better in the perception of native tones for real words than nonsense words, but the same result was not observed for Mandarin listeners in the perception of Mandarin tones. Instead, Mandarin listeners showed a ceiling effect, which may have inhibited any difference appearing between real words and nonsense words. Furthermore, no difference was observed between real and nonsense words for English listeners in the perception of Cantonese or Mandarin tones. For Cantonese listeners' perception of native tones, Lee et al. (1996) suggested that semantic information played a role in helping listeners to distinguish native tones.

In a series of three experiments, Cutler and Chen (1997) used a spoken-word recognition task to examine to what extent tone processing was related to lexical status. The participants included native speakers of Cantonese and Dutch, a tone language and non-tone language, respectively. In the first experiment, Cantonese listeners were presented with a disyllabic real word in Cantonese and a number of nonsense words that differed from the real word by tone, onset, or rime of the second syllable. They were asked to judge whether or not these words were real words in Cantonese. The researchers found that listeners often accepted a nonsense word as a real word when the only difference between the two words was tone. In the second and third experiments, the Cantonese and Dutch speakers made same-different judgements on syllables presented in the first experiment. For the Cantonese listeners, longer reaction time and lower discrimination accuracy were observed for syllables differing in tone, but no difference in tone perception was found for real and nonsense words. The Dutch listeners showed difficulty in discriminating inconsistent syllable pairs due to tone or segmental differences. These findings suggest that tone processing is probably independent of lexical status for both tone and non-tone language speakers.

Taken together, these research findings suggest that tone language speakers may benefit from lexical status in the perception of native tone. However, given the contradictory findings, it is probable that lexical status functions only when tone discrimination is not extremely easy. Despite the aid of semantic information, tone perception is not subject to word perception. In other words, tone perception is pre-lexical, occurring at a stage separate from lexical perception.

3.2. The current study

The main purpose of the current study is to explore to what extent acoustic correlates and phonological status affect tone perception by listeners from diverse tone backgrounds, including native listeners of Mandarin, Thai and English. In an identification task, listeners identified the four Mandarin tones imposed on monosyllabic real and nonsense words in Mandarin. The non-Mandarin listeners also differed in their L2 Mandarin experience: those with and without Mandarin experience are referred as experienced and inexperienced listeners, respectively. The current study aims to answer four main research questions on the effects of language experience, acoustic and phonological properties, and lexical status on tone perception.

(1) **What effects does language experience have on the identification of Mandarin tones?**

Previous studies have indicated that tone perception is experience-sensitive. Better performance is therefore expected for native listeners than non-native listeners; for experienced listeners than inexperienced listeners; and for Thai listeners than their English counterparts, whether experienced or inexperienced.

(2) **To what extent do acoustic properties affect tone identification by listeners in each group?**

It is expected that the most confusing pair, Tone 2 and Tone 3, will be identified poorly by both native and non-native listeners due to their similarity in F0 contour. Thai listeners may have more difficulty differentiating Tone 2 and Tone 3 because of their dependence on F0 contour in tone perception. It is also predicted that English listeners will have difficulty identifying Tone 1, Tone 2 and Tone 4, which contain a high F0 on the tone contour, given English listeners' tendency to use F0 height to discriminate tone patterns (e.g., Guion & Pederson, 2007). Additionally, English speakers have also reported that Tone 1 and Tone 4 are the most confusing tone pair, not Tone 2 and Tone 3 (Wang et al., 1999; Shen, 1989). Thus, it is also expected that Tone 1 and Tone 4 will be more poorly identified by the English listeners.

(3) **To what extent does phonological status affect tone identification?**

For Mandarin and experienced listeners, it is assumed that allophonic variations, i.e., that involving Tone 2 and Tone 3, will be identified poorly, as these tones serve the same phonological function (owing to tone sandhi) in Mandarin. The influence of tone allophones is not expected to be shown in identification patterns of inexperienced listeners, due to their lack of exposure to Mandarin and to any tone language in which tone sandhi is employed. As Tone 2 and Tone 3 have been found to be confusable at both the acoustic and phonological levels, it is expected that the experienced listeners will exhibit more errors in the identification of Tone 2 and Tone 3 relative to their inexperienced counterparts.

(4) **To what extent does lexical status affect tone identification?**

Of the two types of stimulus words, real words provide a semantically meaningful context, while nonsense words do not. Therefore, if tone perception is prelexical, no difference is expected for real words versus nonsense words. However, if tone perception is subject to word perception, listeners will perform better for real words than nonsense words. In addition, the experienced Thai and English listeners might benefit from the lexical information conveyed in real words, due to their acquired knowledge of Mandarin.

3.3. Method

3.3.1. Listeners

Listeners were recruited from three cities: Beijing, Bangkok and Vancouver. Based on their previous experience with Mandarin, the participants were classified into five groups: NM (Native Mandarin), IT (Inexperienced Thai), ET (Experienced Thai), IE (Inexperienced English) and EE (Experienced English); there were 20 people in each group. Participants were all right-handed young adults, and reported normal hearing and speech ability. None had any formal linguistic or musical training (i.e., less than 6 years' musical training; see Wong, Skoe, Russo, Dees, & Kraus, 2007).

The NM listeners were recruited from Beijing. This group comprised 9 males and 11 females, with a mean age of 20 years (range: 19-21). They were born and raised in Beijing and had no experience with other tone languages or any regional languages. They were all undergraduate university students in Beijing learning English as a foreign language. This group of listeners also participated in the assimilation experiment presented in Chapter 2 (in which they were labeled as IM listeners).

The IT listeners were recruited from Bangkok, and comprised 5 males and 15 females, with a mean age of 21 years (range: 19-29). They were born and raised in Bangkok, with no previous knowledge of other tone languages or any regional languages. At the time of testing, they were all attending Chulalongkorn University in Bangkok as undergraduate or graduate students. They were comparable in terms of language experience to the IT listeners who participated in the assimilation experiment.

Of the 20 ET listeners, 16 also participated in the assimilation experiment. All 20 were recruited from the same college in Beijing. All were native speakers of Thai, with 0.5-2 years of Mandarin study and immersion experience in Beijing. The 20 ET listeners included 11 males and 9 females, with a mean age of 25 years (range: 20-32).

Listeners in the IE group were native speakers of Canadian English or American English. This group included 7 males and 13 females, whose mean age was 24 years (range: 19-37). They were recruited from two universities in Vancouver (Simon Fraser University and the University of British Columbia), and had no prior knowledge of Mandarin or other tone languages. At the time of testing, they were attending university as undergraduate or graduate students or working as staff members.

The EE listeners' level of experience with Mandarin was similar to that of their Thai counterparts in the ET group. At the time of testing, they were attending university in Beijing and had 0.5-2 years of study and immersion in Mandarin. They were native speakers of North American English, comprised of 11 males and 9 females, with a mean age of 23 years (range: 19–27).

The ET and EE listeners also reported more L2 use than L1 use in the target-language-speaking country based on the hours of Mandarin lessons they took per week and on the amount of daily L1 and L2 use. It must be noted that the actual number of participants recruited was more than 100. As this experiment was conducted as a pre-dichotic test (see Chapter 4 for the dichotic test), data collected from those participants who failed to pass the test at a 60% correct level were not included in the rest of the study.

3.3.2. Stimuli

The stimulus words were a combination of real and nonsense words in Mandarin. Eight syllables with the four Mandarin tones superimposed resulted in 32 real words, while the other four syllables formed 16 nonsense words (see Table 3.1). These words were produced by the two native speakers of Mandarin described in Chapter 2. The real words are meaningful and commonly used in Mandarin. Their familiarity was rated by Mandarin native speakers (see Chapter 2). In addition, the syllables /lau, pa(a), ma(a), sjau / with the Thai tones imposed were also rated as commonly used real words in Thai by Thai native speakers (see Chapter 2), while two English native speakers also concurred that the syllables /pi, fan, hu, pai/ correspond to the English words *bee, fan, who and pie*, respectively. In contrast, /pju, kjɛn, dy, fiau/ were rated as non-native syllables by the raters regardless of their L1. As shown in Table 3.1, /lau, pa(a), pi, fan, pju, kjɛn/ with the Mandarin tones imposed were assigned to the identification test. The remaining syllables, i.e., /ma(a), sjau, hu, pai/ and /dy, fiau/ were used in a tone familiarization session for the inexperienced Thai and English listeners to familiarize them with the pitch changes (F0 height and contour) of the four Mandarin tones with their numerical labels, i.e., Tone 1, Tone 2, Tone 3 and Tone 4.

Table 3.1. Syllables for the familiarization and test sessions of the pre-dichotic identification test

	Real word		Nonsense word	
	Familiarization	Identification	Familiarization	Identification
NM	-	/lau, pa(a)/	-	/pju, kjɛn/
ET	-	/lau, pa(a)/	-	/pju, kjɛn/
IT	/ma(a), sjau/	/lau, pa(a)/	/dy, fiau/	/pju, kjɛn/
EE	-	/pi, fan/	-	/pju, kjɛn/
IE	/hu, pai/	/pi, fan/	/dy, fiau/	/pju, kjɛn/

3.3.3. Procedure

This experiment was conducted in quiet rooms in Beijing for the NM, ET and EE listeners; in Bangkok for the IT listeners; and in Vancouver for the IE listeners. All the stimuli were presented binaurally via AKG or Logitech acoustical headsets, receiving the signal from a personal laptop. This experiment was run on E-Prime Version 1.0 (Schneider, Eschman, & Zuccolotto, 2002) individually for the listeners; instructions were displayed on the screen in listeners' native language.

The inexperienced listeners participated in a 15-minute tone familiarization session before the real identification test. The design of the familiarization session was similar to that of the real test, except that the listeners were provided with feedback in the training session. First, each auditory Mandarin token was presented and the listeners identified the tone by matching it to one of the patterns shown on the screen (Figure 3.1). They were asked to press one of the corresponding key on the keyboard of the laptop.

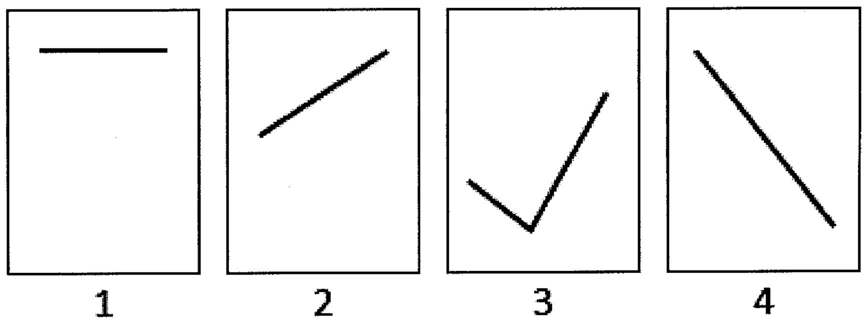


Figure 3.1. ***Pitch movements of Mandarin tones along with their numerical labels***

Each target tone was presented in two different patterns: singly or followed by a falling tone. The familiarization session consisted of four blocks, comprising a total of 128 stimuli (4 syllables x 4 blocks x 2 speakers x 4 tones). In the first two blocks, listeners heard each target tone singly, such as *ma*^{1}, but in the third and fourth block, the target tone was presented followed by a falling tone in a two-tone sequence, such as *ma*^{1}*ma*^{4}. The falling tone was employed as a distracter to increase the difficulty of the task. Previous research has shown that tone perception in citation form is likely to result in a ceiling effect, especially for native speakers (Lee et al., 1996; Wayland & Guion, 2004). Therefore, some acoustical modifications, such as mixing tones with Gaussian noise, are sometimes made in tone perception studies to increase the difficulty of the task (Wang et al., 2004).

In the current study, presentation of target tones in a context was intended to increase the difficulty of the tone identification task, as the human memory span is time-bound (Humstone, 1919; McCabe, 2008). Hence, a delayed recall of the target tone requires a longer reaction time and higher discriminability (Lee et al., 1996). To assess whether or not a tone sequence would increase task difficulty, a pilot study was conducted. Four native Mandarin listeners, 3 males and 1 female, with a mean age of 20 years (range: 19-21) identified the Mandarin tones in the two contexts. Their performance (mean accuracy rate) was better and showed a ceiling effect when the tones were presented singly (98%), but much poorer when the tones were presented in a two-tone sequence followed by a falling tone (75%). In addition, the falling tone was selected as the distracter because this tone exists in all the three languages as a commonly used pitch pattern, and is not involved in any tone changes at the phonological level, i.e., tone sandhi.

The tone identification test was carried out in the same way as the familiarization session, except that each target tone was presented only in a sequence and no feedback was provided. Eighty stimulus words (5 syllables x 4 tones x 2 speakers x 2 blocks) were presented in a random order. This session lasted approximately seven minutes. The time for each response was two seconds and there was a self-paced break between blocks.

Participants who scored higher than 60% on average (well above the chance level of 25%) on the identification test continued on and took the dichotic listening test (see Chapter 4). This standard (60% correct response) was applied to both native and non-native listeners, so that the results of the dichotic listening test could be discussed based on a comparable criterion.

3.4. Results

3.4.1. Tone identification

A three-factor mixed design ANOVA was conducted with group (NM, ET, IT, EE, and IE) as the between-subjects factor and tone (Tone 1, Tone 2, Tone 3, and Tone 4) and word type (real word and nonsense word) as within-subjects factors. Results are summarized in Table 3.2, indicating the significant main effects of tone, word type and group, and the significant interactions of tone x group, tone x word type, word type x group, and tone x word type x group. The size of the group effect (η_p^2 =.602) was especially large, suggesting a strong difference among groups (Wilkinson, 1999).

Table 3.2. ***The source, F, p values and size of effect for ANOVAs on the identification test.***

Source	F	Sig.	Partial Eta Squared (η_p^2)
Tone	(3, 285)=18.2	<.0001	.161
Word type	(1, 95)=26.8	<.0001	.220
Group	(4,95)=35.9	<.0001	.602
Tone x Group	(12, 285)=8.4	<.0001	.261
Word type x Group	(4,95)=6.9	<.0001	.225
Word type x Tone	(3,285)=15.9	<.0001	.144
Word type x Tone x Group	(12,285)=4.4	<.0001	.156

3.4.1.1. Main effects

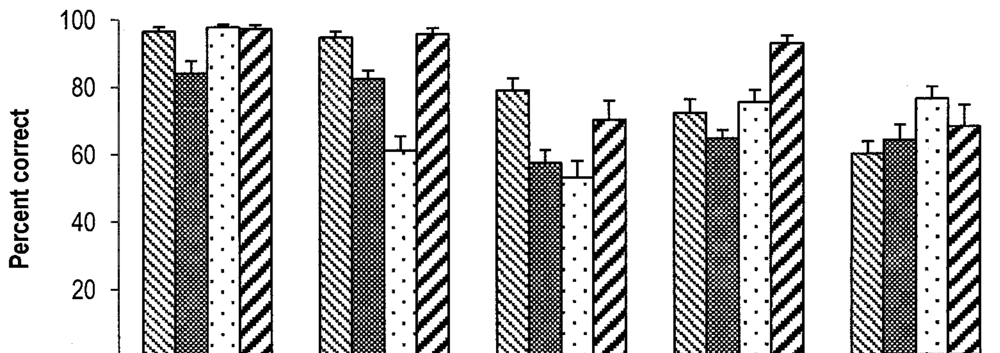
First, the main effect of group was significant. Results of Bonferroni post hoc analyses indicated better performance of the Mandarin listeners (NM) than the non-Mandarin listeners and of the experienced listeners (ET and EE) than the non-experienced ones (IT and IE). Moreover, no significant difference was detected between the ET and EE groups or the IT and IE groups (NM, 94%; ET, 83%; IT, 65%; EE, 76%; IE, 67%). A set of p values exhibited significant differences between the NM and ET groups ($p = .004$), EE ($p < .0001$), IT ($p < .0001$), or IE ($p < .0001$), as well as between the ET and IT groups ($p < .0001$) or IE ($p < .0001$) and between the EE and IT groups ($p = .001$) or IE ($p = .018$).

For the main effect of tone, Bonferroni post hoc analysis revealed higher identification accuracy for Tone 1 and Tone 4 than for Tone 2 and Tone 3 ($p < .0001$, except $p = .004$ for Tone 1 and Tone 3), but no difference between Tone 1 and Tone 4 or between Tone 2 and Tone 3 (Tone 1, 80%; Tone 2, 70%; Tone 3, 72%; and Tone 4, 85%).

The main effect of word type was also significant. The tones on the real words were identified more accurately than those on the nonsense words ($p < .0001$) (real words: 79% and nonsense words: 75%).

3.4.1.2. Interaction of group x tone

Figure 3.2 illustrates the interaction of group x tone. Further one-factor ANOVA analyses using tone as the within-subjects factor evaluated the main effect of tone for each group, NM, $F(3, 57) = 10.5$, $p < .0001$, $\eta_p^2 = .355$; ET, $F(3, 57) = 40.9$, $p < .0001$, $\eta_p^2 = .683$; IT, $F(3, 57) = 7.4$, $p < .0001$, $\eta_p^2 = .281$; and ET, $F(3, 57) = 11.4$, $p < .0001$, $\eta_p^2 = .413$, except the IE group, $F(3, 57) = 2.3$, $p = .087$. Post hoc analysis (Bonferroni adjustment) further revealed that for NM listeners, Tone 2 was identified less accurately than all other tones (Tone 1, $p = .023$; Tone 3, $p = .001$; Tone 4, $p = .008$). Among the ET listeners, Tone 2 and Tone 3 were more poorly identified than Tone 1 (Tone 2, $p = .003$; Tone 3, $p < .0001$) and Tone 4 (Tone 2, $p = .001$; Tone 3, $p < .0001$); and Tone 3 was more poorly identified than Tone 2 ($p = .001$). For the IT listeners, Tone 2 and Tone 3 were more poorly identified than Tone 1 (Tone 2, $p < .0001$; Tone 3, $p = .001$). For the EE listeners, Tone 1, Tone 2 and Tone 3 were identified less accurately than Tone 4 (Tone 1, $p = .001$; Tone 2, $p < .0001$; Tone 3, $p = .005$).



	NM	ET	IT	EE	IE
Tone 1	96	95	79	72	60
Tone 2	84	82	57	64	64
Tone 3	98	61	53	75	76
Tone 4	97	96	70	93	68

Figure 3.2. ***Interaction of group x tone. Identification accuracy is compared between groups for each tone category.***

Responses for each tone were also broken down by group for analysis (Figure 3.2). One-factor ANOVA analyses showed a group effect for all the four tones: Tone 1, $F(4, 95) = 25.01$, $p < .0001$; Tone 2, $F(4, 95) = 11.6$, $p < .0001$; Tone 3, $F(4, 95) = 20.5$, $p < .0001$; and Tone 4, $F(4, 95) = 12.72$, $p < .0001$. Post hoc analysis (Bonferroni) further indicated the pairwise differences. The difference between the NM and ET groups in the identification of Tone 1, Tone 2 and Tone 4 was not significant; nor was the difference between the NM and EE groups in the identification of Tone 4 (p = 1). The ET listeners were better at identifying Tone 1 and Tone 2 than the EE listeners (Tone 1, $p < .0001$; Tone 2, $p = .006$) and the ET and EE groups' perception of Tone 3 and Tone 4 showed no difference. The IT listeners performed better than the IE listeners in the perception of Tone 1 ($p < .0001$), while the IE listeners performed better than the IT listeners in the perception of Tone 3 ($p < .0001$). There was no difference between the two groups in the perception of Tone 2 and Tone 4. Between the experienced and inexperienced listeners in the same L1 group, the ET listeners showed better performance for Tone 1 ($p = .005$), Tone 2 ($p < .0001$) and Tone 4 ($p < .0001$), but not for Tone 3 (p = 1), while the EE listeners showed no difference from the IE listeners in the perception of Tone 1 ($p = .065$), Tone 2 (p = 1) and Tone 3 (p = 1), but better performance in the perception of Tone 4 ($p < .0001$). Table 3.3 summarizes the results, showing the relative performance of one group versus another in the identification of individual tones.

Table 3.3. ***Interaction of group x tone. Paired groups are shown in the left, with the bold letter indicating the better-performing group. Tones which led the difference in identification between the groups are shown at right.***

Paired group		Tone			
NM	ET			Tone 3	
NM	EE	Tone 1	Tone 2	Tone 3	
NM	IT	Tone 1	Tone 2	Tone 3	Tone 4
NM	IE	Tone 1	Tone 2	Tone 3	Tone 4
ET	EE	Tone 1	Tone 2		
ET	IT	Tone 1	Tone 2		Tone 4
ET	IE	Tone 1	Tone 2		Tone 4
EE	IT			Tone 3	Tone 4
EE	IE				Tone 4
IE	IT			Tone 3	
IT	IE	Tone 1			

3.4.1.3. Interaction of group x word type

Results of the multiple-factor repeated measures ANOVA analysis indicated a significant interaction of word type x group (Figure 3.3). To examine the effect of word type, further one-factor ANOVA analyses were performed for each group. The Mandarin and English listeners (both groups) showed better performance when the tones were imposed on real words, while the Thai listeners' performance was not affected by word type, NM: $F(1, 19) = 7.6$, $p = .012$, $\eta_p^2 = .287$ (real words 95%, nonsense words 92%) ; ET: $F(1, 19) = .67$, $p = .424$, $\eta_p^2 = .034$ (real words 84%, nonsense words 82%); IT: $F(1, 19) = .36$, $p = .558$, $\eta_p^2 = .018$ (real words 64%, nonsense words 65%); EE: $F(1, 19) = 25.4$, $p < .0001$, $\eta_p^2 = .572$ (real words 81%, nonsense words 70%); IE: $F(1, 19) = 14.2$, $p = .001$, $\eta_p^2 = .427$ (real words 71%, nonsense words 63%).

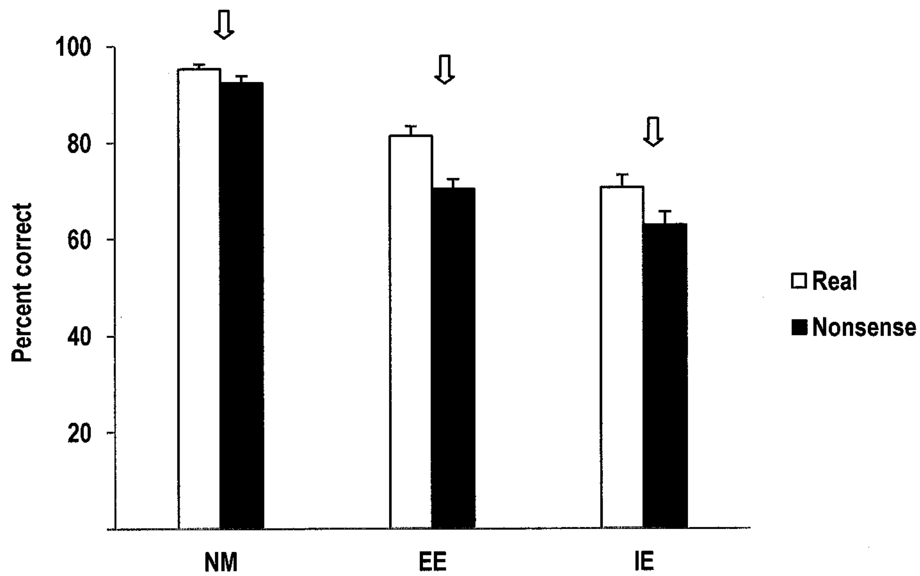


Figure 3.3. ***Interaction of group x word type. The arrows indicate significantly better performance for real words at p < .05.***

One-way ANOVA analyses with group as the between-subjects factor were conducted for tones imposed on real and nonsense words separately. A significant group effect was observed for tones imposed on both types of words: real words, $F(4, 95) = 30.1$, $p < .0001$; nonsense words, $F(4, 95) = 31.3$, $p < .0001$. The Bonferroni post hoc analyses indicated similar results to the main effect of group, based on overall performance, i.e., the native speakers were better than the non-native speakers, and the experienced speakers were better than the inexperienced speakers, while no difference was detected between the Thai and English speakers with similar Mandarin experience. For the nonsense words, however, the EE group's performance was inferior to that of the ET group (ET: 83% vs. EE: 70%, $p = .002$), but showed no difference when compared to that of the IE group ($p = .191$) or IT group ($p = 1$).

3.4.1.4. Interaction of tone x word type

The effect of tone was analyzed for tones on real and nonsense words. A one-way ANOVA with tone being the within-subjects factor revealed a tone effect for both word types: real words, $F(3, 297) = 17.7$, $p < .0001$, $\eta_p^2 = .152$; nonsense words, $F(3, 197) = 11.6$, $p < .0001$, $\eta_p^2 = .105$. The adjusted Bonferroni post hoc analyses indicated that Tone 2 and Tone 3 were more poorly identified than Tone 1 and Tone 4 ($p < .0001$, between Tone 2 and Tone 4, $p = .003$), while no difference was shown between Tone 1 and Tone 4 or between Tone 2 and Tone 3 when these tones were imposed on real words. For the nonsense words, responses to Tone 4 were more accurate than responses to the other three tones (Tone 1, $p = .006$; Tone 2, $p < .0001$; Tone 3, $p = .016$).

A one-factor ANOVA using word type as the within-subjects factor revealed a significant effect of word type on the identification of Tone 1 and Tone 2, but not on Tone 3 and Tone 4: Tone 1, $F(1, 99) = 21.9$, $p < .0001$, $\eta_p^2 = .227$; Tone 2, $F(1, 99) = 14.9$, $p < .0001$, $\eta_p^2 = .131$. In addition, Tone 1 and Tone 2 were better identified when they were imposed on the real words than on the nonsense words, i.e., Tone 1, 87% vs. 74% and Tone 2, 75% vs. 66%.

3.4.1.5. Interaction of tone x group x word type

To determine the effect of word type on the interaction of tone x group, separate one-factor ANOVA analyses were conducted for each language group by tone. A significant main effect of word type was revealed for Tone 1 for NM, [$F(1, 19) = 5.4$, $p = .031$, $\eta_p^2 = .223$], EE [$F(1, 19) = 22.6$, $p < .0001$, $\eta_p^2 = .543$], and IE [$F(1, 19) = 26.6$, $p < .0001$, $\eta_p^2 = .583$] listeners; and also for Tone 2 for NM listeners [Tone 2, $F(1, 19) = 8$, $p = .011$, $\eta_p^2 = .296$] and for Tone 3 for the IT listeners [$F(1, 19) = 12.9$, $p = .002$, $\eta_p^2 = .4$]. No effect of word type for Tone 4 was shown for any group. Furthermore, as the p values referred to Bonferroni adjustment for multiple comparisons at two levels, a p value of .031 (> .025) indicated no significant effect of word type on Tone 1 for the NM listeners. For the EE and IE listeners, post hoc analyses revealed that Tone 1 was identified more accurately on the real words than on the nonsense words (EE: real, 87%; nonsense, 57%, $p < .0001$ and IE: real, 76%; nonsense, 44%, $p < .0001$). For the NM listeners, Tone 2 was better identified on the real words than on the nonsense words (real, 88%; nonsense, 80%, $p < .011$). For the IT listeners, Tone 3 was identified more accurately on the nonsense words than on the real words (real, 44%; nonsense, 62%, $p = .002$).

In summary, the overall performance showed significant effects of group, tone and word type. The NM, ET and EE listeners were better at identifying Mandarin tones than the IT and IE listeners, and Tone 2 and Tone 3 appeared more difficult to recognize than Tone 1 and Tone 4. Additionally, tones on real words were more accurately identified than those on nonsense words. However, further analyses indicated the ET listeners were comparable to the NM ones in the identification of most Mandarin tone categories, though this level of performance was not seen for the EE listeners despite their similar Mandarin experience. Tone 2 or Tone 3 was most difficult for the NM, ET, IT and EE listeners. Tone 1 was also difficult to recognize for the EE listeners. Moreover, the four tones presented similar difficulty for the IE listeners. Word type (i.e., lexical status) affected tone identification only for the NM, EE and IE listeners and only in the identification of difficult tones.

3.4.2. *Confusions*

In Tables 3.4-3.8, the percentages of identification data are presented for each listener group. Mean confusion was calculated by averaging the confusion rates for the inconsistent responses, such as Tone 1 being identified as Tone 2, Tone 3 or Tone 4 in Table 3.4 (Tables 3.5-3.8 are based on this same procedure). Identification responses for real and nonsense words were presented separately for close examination of any influence of lexical status.

Table 3.4. ***Tone confusion matrix of Mandarin tone identification by the Mandarin speakers (n = 240 items). Correct responses are shown in bold.***

Token	Identified as									
	Tone 1		Tone 2		Tone 3		Tone 4		Mean error	
	real	nonsense	real	nonsense	real	nonsense	real	nonsense	real	nonsense
Tone 1	**94.4**	**95**	1.6	2.5	0	0.4	0	0.4	0.5	1.1
Tone 2	1.6	0.4	**84.7**	**80**	8.4	15	1.6	0.8	3.9	5.4
Tone 3	0	0	0.4	0.4	**94**	**97.9**	1.6	0.8	0.5	0.4
Tone 4	0	2.1	2.4	0.4	0	0	**94**	**96.7**	0.8	0.8

The mean confusion shows excellent performance for the NM listeners, except for Tone 2. This result is consistent with the findings for identification accuracy reported above (see Figure 3.2). As can be seen, the poorer performance for Tone 2 is the result of moderate difficulty in distinguishing Tone 2 from Tone 3, in particular when it is imposed on the nonsense words (real words 8.4%; nonsense words 15%). In contrast, the percentage that Tone 3 is misidentified as Tone 2 is negligible.

Table 3.5. ***Tone confusion matrix of Mandarin tone identification by the experienced Thai speakers (n = 240 items). Correct responses are shown in bold.***

Token	Identified as									
	Tone 1		Tone 2		Tone 3		Tone 4		Mean error	
	real	nonsense	real	nonsense	real	nonsense	real	nonsense	real	nonsense
Tone 1	**97.8**	**90.6**	0.9	3.6	0	0.4	0	4	0.3	2.7
Tone 2	0.4	1.8	**85.3**	**78.6**	9.4	13.8	1.3	1.8	3.7	5.8
Tone 3	1.3	0.4	38.8	32.1	**55.8**	**64.7**	0.4	0.9	10.2	11.1
Tone 4	0.4	1.8	0.9	0.4	2.2	0.4	**95.5**	**96.4**	1.2	0.9

The ET listeners appear to have had moderate to great difficulty in discriminating Tone 2 and Tone 3, as shown by the mean confusion. In contrast to the NM listeners, the ET listeners identified Tone 3 as Tone 2 more frequently (real words 38.8%; nonsense words 32.1%), but identified Tone 2 as Tone 3 at a comparable level to the NM listeners (real words 9.4%; nonsense words 13.8%). Likewise, the ET listeners had the least difficulty in identifying Tone 1 and Tone 4.

Table 3.6. ***Tone confusion matrix of Mandarin tone identification by the inexperienced Thai speakers (n = 240 items). Correct responses are shown in bold.***

Token	Identified as									
	Tone 1		Tone 2		Tone 3		Tone 4		Mean error	
	real	nonsense	real	nonsense	real	nonsense	real	nonsense	real	nonsense
Tone 1	**75.6**	**80.6**	8.1	6.3	5.6	5.6	5.6	5.6	6.5	5.8
Tone 2	5.6	2.5	**62.5**	**50.0**	21.9	31.9	8.1	7.5	11.9	14.0
Tone 3	5.0	1.3	38.1	30.6	**43.1**	**61.9**	8.8	3.1	13.3	11.7
Tone 4	5.6	5.6	10.6	12.5	6.3	11.3	**73.1**	**66.9**	7.5	9.8

The IT listeners showed variable identification of the four Mandarin tones, and significant difficulty in distinguishing Tone 2 and Tone 3. In contrast to the ET listeners, the IT listeners identified Tone 2 as Tone 3 more frequently (real words 21.9%; nonsense words 31.9%), while they made errors in identifying Tone 3 as Tone 2 at a comparable level (real words 38.1%; nonsense words 30.6%). The IT listeners identified Tone 1 and Tone 4 more consistently, whereas misidentification was still distributed among the other three tones with higher mean confusion rates for Tone 4 than Tone 1.

Table 3.7. ***Tone confusion matrix of Mandarin tone identification by the experienced English speakers (n = 240 items). Correct responses are shown in bold.***

Token	Identified as									
	Tone 1		Tone 2		Tone 3		Tone 4		Mean error	
	real	nonsense	real	nonsense	real	nonsense	real	nonsense	real	nonsense
Tone 1	**86.9**	**56.9**	8.1	18.1	0.0	1.3	3.1	22.5	3.8	14.0
Tone 2	4.4	6.3	**69.4**	**59.4**	23.8	17.5	0.6	12.5	9.6	12.1
Tone 3	0.6	1.3	16.9	15.6	**79.4**	**70.6**	0.6	8.8	4.8	8.5
Tone 4	6.9	0.6	1.3	1.3	0.6	0.0	**90.0**	**95.0**	2.9	0.6

The EE listeners were very accurate in identifying Tone 4, but not in identifying the other three tones. Each of the three tones was misidentified as a particular tone more frequently than the others: Tone 1 as Tone 4 on nonsense words (22.5%), Tone 2 as Tone 3 on real (23.8%) and nonsense (17.5%) words, and Tone 3 as Tone 2 on real (16.9%) and nonsense (15.6%) words. Misidentification of Tone 2 as Tone 4 was also fairly frequent on nonsense words (12.5%).

Table 3.8. Tone confusion matrix of Mandarin tone identification by the inexperienced English speakers (n = 240 items). Correct responses are shown in bold.

Token	Identified as									
	Tone 1		Tone 2		Tone 3		Tone 4		Mean error	
	real	nonsense	real	nonsense	real	nonsense	real	nonsense	real	nonsense
Tone 1	**75.0**	**43.8**	18.8	38.1	0.0	2.5	5.6	15.0	8.1	18.5
Tone 2	15.6	10.0	**67.5**	**59.4**	8.1	15.0	7.5	14.4	10.4	13.1
Tone 3	3.1	5.0	14.4	6.9	**71.9**	**79.4**	7.5	7.5	7.5	6.5
Tone 4	26.9	13.8	4.4	6.3	0.6	5.6	**66.3**	**69.4**	10.6	8.5

The IE listeners exhibited more variable identification of the four Mandarin tones than the IT listeners. As shown in Table 3.10, the most confusing tone pairs for the IE listeners were Tone 1 and Tone 2 (real words 18.8%; nonsense words 38.1%) and Tone 4 and Tone 1 (real words, 26.9%; nonsense words, 13.8%). In addition, Tone 2 and Tone 3, as well as Tone 2 and Tone 4, also showed moderate confusion rates (Tone 2 as Tone 3: real words 8.1%; nonsense words 15%; Tone 2 as Tone 4: real words 7.5%; nonsense words 14.4%; Tone 3 as Tone 2: real word 14.4%; nonsense words 6.9%).

To summarize, Tone 2 was the most difficult tone for the NM listeners to identify and it was often confused with Tone 3. Tone 2 and Tone 3 were very difficult to distinguish by the Thai listeners in the ET and IT groups and therefore very difficult to identify. Tone 1 was the most difficult tone for the English listeners, in particular when it was imposed on nonsense words. Among the English listeners, Tone 1 was mostly confused with Tone 2 and Tone 4.

3.5. Discussion

In this study, identification accuracy and confusion matrices demonstrate the influence of L1 and L2 experience on acoustic and phonological tone perception. In the following discussion, the current results are interpreted in relation to the proposed hypotheses concerning the interaction of language experience with acoustic properties, phonological status and lexical status in the perception of Mandarin tones.

3.5.1. The effects of language experience

As predicted, the current findings indicate significant effects of L1 and L2 tone experience. Native tone experience enhances listeners' perception of tone categories, supporting previous studies (e.g., Hallé et al., 2004). Furthermore, L2 experience with a tone system also helps listeners in the identification of L2 tone categories (Wayland & Guion, 2003). Accordingly, a lack of experience with a tone system inhibits listeners' perception of the tone categories in that system (Guion & Pederson, 2007). One widely discussed question is whether or not L1 tone experience is transferable to the perception of L2 tones (Francis et al., 2008). In the current study, the IT and IE listeners showed no difference in the identification of Mandarin tones. Thus, the IT listeners did not benefit from their L1 tone experience relative to the IE listeners. However, in comparison with the EE listeners, the ET listeners exhibited more native-like performance, suggesting a more categorical perception for this group that may be due to the establishment of some Mandarin tone categories in their phonological system. Nevertheless, this finding may be explained by the similarities between the Mandarin and Thai tone systems with regard to the number of tones and the phonetic correspondence between the tone categories. Further research is needed to examine the effects of learners' L1 tone experience on L2 tone perception in cases where the L1 and L2 tone systems differ in the number of tone categories and acoustic correspondences (e.g., Yoruba, with three level tones, and Cantonese, with six level or contour tones).

3.5.2. The effects of acoustic properties

The current findings indicate that Tone 2 and Tone 3 are indeed, as predicted, the most confusing tone pair in Mandarin. It was expected that these two tones would be poorly discriminated by the native and non-native listeners due to their similar F0 contour. The native listeners showed poorer performance for Tone 2 but not for Tone 3, and the confusion matrix indicated that the native listeners significantly misidentified Tone 2 as Tone 3, but not Tone 3 as Tone 2. Thus, it is likely that the native listeners' poorer identification of Tone 2 was not caused by the F0 similarity between Tone 2 and Tone 3. The non-native listeners also showed difficulty in discriminating between Tone 2 and Tone 3, particularly the Thai listeners, whether experienced or inexperienced. As predicted, the Thai listeners' confusion between Tone 2 and Tone 3 may be accounted for by their choice of F0 contour as the primary perceptual cue (Gandour, 1983). Furthermore, the English listeners in both the experienced and inexperienced groups exhibited difficulty discriminating Tone 1, Tone 2 and Tone 4, supporting previous findings (Francis et al., 2008; Guion & Pederson, 2007). The confusion matrices also indicate that these tones were more confusable when they were imposed on nonsense words, suggesting an interaction between acoustic properties and lexical status.

3.5.3. *The effects of phonological status*

It was expected that the native listeners and the experienced listeners would show poorer discrimination between the allophonic variations, Tone 2 and Tone 3. This prediction is partially supported by the current findings. The native listeners identified Tone 2 more poorly than the other tones and the confusion matrix further indicated that Tone 2 was assimilated to Tone 3. The fact that Tone 2 is more likely to be perceived as Tone 3 is likely due to the fact that Tone 3 is the basic form; the sandhi rule changes Tone 3 into Tone 2. Because Tone 3 remains the basic form in the listeners' lexicon, they tend to misperceive Tone 2 as Tone 3. This result is also consistent with the assimilation findings from Chapter 2, which indicated that falling rising tone (Tone 3) was assimilated to low falling tone, although rising tone (Tone 2) appeared to be assimilated to falling rising tone. Thus, the poor perception of Tone 2 by the native listeners may be due to its phonological equivalence to Tone 3, rather than to acoustic similarities in F0.

The experienced listeners were expected to demonstrate higher confusion between Tone 2 and Tone 3 relative to their inexperienced counterparts. This prediction was not supported by the Thai listeners' performance, since the IT listeners showed more difficulty than the ET listeners in identifying Tone 2, and misidentified Tone 2 as Tone 3 more frequently than the ET listeners. This outcome may be attributable to the IT listeners' higher dependence on F0 contour in tone perception, while the ET listeners may have attached less importance to F0 contour, based on their L2 experience with Mandarin (Gandour 1983; Guion & Pederson, 2007). In contrast, the EE listeners demonstrated higher confusion between Tone 2 and Tone 3 than the IE listeners. This pattern may be due to their Mandarin learning experience, which possibly induced a native-like perceptual cue to tone perception. In other words, the EE listeners might use F0 contour in the identification of Mandarin tones, while the IE listeners might rely more on F0 height (Guion & Pederson, 2007). Although the experienced listeners' perception of Tone 2 and Tone 3 reflects the influence of L2 experience on the use of perceptual cues, it is still likely that the phonological relationship between Tone 2 and Tone 3 affected their performance, for they showed lower accuracy and higher confusion in the perception of Tone 2 and Tone 3 relative to other tones or tone pairings.

3.5.4. The effects of lexical status

In agreement with Cutler & Chen (1997), the current findings do not support the view that lexical status affects tone perception since the difference in tone identification due to lexical status was seen for only a few tones and for only some groups. Previous research has found that native listeners of some tone languages, such as Cantonese, may use semantic information to help recognize native tone categories; however, this pattern was not seen among speakers of other tone language speakers, such as Mandarin, due to a ceiling effect (Lee et al., 1996). Results of this study indicate that the native and non-native listeners may indeed benefit from semantic information in the perception of difficult tones. For example, identification accuracy was higher for real words in the perception of Tone 2 by the Mandarin listeners in the perception of Tone 1 by the experienced and inexperienced English listeners. Most likely, the Thai listeners showed no preference for (Mandarin) real words because they did not recognize the stimulus words as real words in Thai, even though the syllables exist in Thai and each Mandarin tone is assimilated to Thai tones. This may also suggest that recognition of a word in a tone language is determined not only by the F0 realization of tone categories. Other phonetic factors, such as voice quality or vowel duration, need to be considered. In contrast, the fact that the English listeners may have recognized the syllables /pi/ and /fan/ with Mandarin tones imposed as the English words *bee* (or *pee* if they could abandon their need for aspiration in that context) and *fan* may have helped the English listeners identify the difficult Tone 1. In agreement with Lee et al. (1996), Mandarin listeners showed a ceiling effect in the identification of Tone 1, Tone 3 and Tone 4 (mean accuracy = 97%) and thus no effect of lexical status emerged. The difference between the current study and Lee et al. (1996) is that this study found interactions of tone pattern and lexical status, while Lee et al.'s study accounted for the contextual effect in a more global manner. Both studies suggest that task difficulty may induce the use of semantic information during tone perception.

3.5.5. *Implications for models of cross-language speech perception*

The current findings suggest that recent models of cross-language speech perception, including the PAM (including the PAM-L2) and the SLM, may be employed to account for the perception of lexical tone by non-native listeners with and without L2 experience.

First, the PAM predicts three patterns for the assimilation of L2 phonetic categories to L1 phonetic categories by inexperienced listeners. Some L2 categories can be assimilated into native categories, some can be assimilated, but not categorized into native categories; and some categories cannot be assimilated to any L1 speech sounds. In the current study, Mandarin tones were likely to be perceived as "assimiliable" speech sounds by the IT and IE listeners, as each Mandarin tone is associated with a Thai tone or an English intonational pattern. The IT and IE listeners' performance, in the PAM terms, may be classified into "Two Category" and "Both Uncategorizable", respectively. For example, it might be predicted that Mandarin Tone 1 and Tone 4 should be assimilated to the Thai mid level and falling tones by the IT listeners, thus resulting in good discrimination. For the IE listeners, the "Both Uncategorizable" assimilation pattern may explain the poor discrimination between these two tones for this group. However, what the PAM does not predict, but that is borne out in the current study, is the effect of listeners' preference for F0 correlates on the perception of L2 tones. For example, even though Tone 2 and Tone 3 correspond phonetically to the Thai rising and falling rising tones, respectively, the confusion caused by the similar F0 contour may have caused the IT listeners to have difficulty in discriminating these two tones. Therefore, the predicted excellent discrimination for the assimilation pattern was not shown by the IT listeners.

For the experienced listeners, the current findings support the SLM, showing that it is possible for adults to learn new phonetic categories. Moreover, the experienced listeners' performance was also more native-like compared to that of the inexperienced listeners. This study also indicates phonological effects on tone perception among the experienced listeners. Accordingly, L2 tone categories may be classified into "assimilable" for the ET listeners and "nonassimilable" for the EE listeners. For the ET listeners, each Mandarin tone category can be associated with Thai tone categories, due to phonological correspondence (see Chapter 2). By contrast, since F0 is not employed to contrast word meanings in English, the EE listeners might have perceived Mandarin tone categories in a "nonassimilable" manner. In accordance with the PAM-L2, the "assimilable" tones, such as Tone 1 and Tone 4, were identified with high accuracy by the ET listeners, indicating excellent discrimination between the two tones. However, like the IT listeners, the ET listeners showed poor discrimination between Tone 2 and Tone 3, even though they correspond phonetically to two different L1 tone categories. This result may be accounted for by listeners' preference for F0 contour as an acoustic cue. However, the ET listeners' confusion between Tone 2 and Tone 3 is probably partly due to the phonological similarity between these two tones. The complexity in the perception of Tone 2 and Tone 3 by the ET listeners appears not to be predicted by the PAM-L2 or the SLM. Nevertheless, in agreement with the PAM-L2, "nonassimilable" L2 tone categories were identified more consistently by the EE listeners than by the IE listeners, in particular Tone 1 and Tone 4.

To summarize, results from the current tone identification study support both the PAM and the SLM, indicating an extension of these models to the suprasegmental level. Some central predictions made by the PAM and the SLM are confirmed by the tone perception patterns inexperienced and experienced listeners, as discussed above. However, this study also yielded some findings that are not predicted or accounted for by these models, such as the perception of Tone 2 and Tone 3 by the Thai listeners. Therefore, models of cross-language tone perception must incorporate the influence of allophonic tone changes in L2, and listeners' preference for F0 correlates, into their predictions for, listeners' performance in L2 tone perception.

3.6. Conclusion

The current research supports previous studies (Francis et al., 2008; Hallé et al., 2004) showing that non-native speakers, including those without previous tone experience, are able to recognize L2 tone categories. Depending on listeners' language experience, L2 tone categories may be perceived at the phonetic or phonological level. Experienced listeners' tone perception suggests that the establishment of an L2 tone system is a long-term process. However, L1 tone experience may help in the formation of the L2 tone system, as the connections between L1 and L2 tone categories are relatively easier to make at both the phonetic and phonological levels.

Contextual influence may also affect tone perception. Lexical status, under some circumstances, has been found to help listeners to recognize native tone categories (Lee et al., 1996). The current study echoes this finding, based on the performance of the Mandarin and English listeners in the identification of difficult tones. The fact that English listeners benefited from semantic information in the identification of tone patterns may support the view that non-tone language speakers process foreign tones with reference to their intonational patterns (Ladd, 2008).

4. Hemispheric processing of Mandarin tones

4.1. Background

This study investigates hemispheric tone processing within a dichotic listening paradigm. The main purpose of this study is to explore how tone is processed in the human brain due to acoustic properties and phonological functions. Speech lateralization research has revealed that in the processing of linguistic stimuli, one or the other of the left or right hemisphere predominates, although both hemispheres are usually activated to some degree simultaneously. As reviewed earlier, tone is a type of prosodic signal with high linguistic function. Thus, it is reasonable to assume that the processing of tone is no different from the processing of other linguistically relevant stimuli, such as segments. However, it has been argued that acoustic properties affect hemispheric asymmetry in the processing of speech signals (Tartter, 1988). The prosodic nature of tone, therefore, may lead to right rather than left hemispheric specialization, as in the processing of musical melodies (Zatorre, Evans, & Meyer, 1994). To meet the purpose of the current project, it is of great interest to probe tone processing due to acoustic correlates and phonological status, as the former associates tone with prosodic patterns, thus indicating rightward asymmetry, while the latter represents tone as a linguistic unit, hence suggesting leftward asymmetry. Recently, these two conflicting notions have been discussed within the framework of the two following hypotheses: the acoustic hypothesis and the functional hypothesis (reviews see Wong, 2002; Zatorre & Gandour, 2008). The acoustic hypothesis postulates that tone processing is no different from prosodic processing, and is dependent uniquely upon changes in acoustic cues (e.g., Ivry & Lebby, 1993). By contrast, the functional hypothesis argues for the primacy of the linguistic function of tone, and thus proposes that tone processing is determined by its phonemic status (e.g., Hsieh, Gandour, Wong, & Hutchins, 2001). Some researchers suggest combining the two accounts, such that hemispheric specialization of tone processing is linked not only with higher-level phonological processing, but also with lower-level acoustic processing (e.g., Gandour et al., 2003; Zatorre & Gandour, 2008). These hypotheses provide ideal theoretical accounts for the current research, as processing of tone at the prosodic level can support the acoustic hypothesis, while processing at the phonological level can support the functional hypothesis. If the results of the current study demonstrate that tone is processed at both levels, a combined account should be considered. In the following sections, these accounts are first reviewed with related studies.

4.1.1. Functional hypothesis

The functional hypothesis predicts that tone processing is lateralized to the left hemisphere due to its high linguistic functional load. According to Van Lancker (1980), tone carries higher linguistic functional load, in contrast to other features of speech prosody, such as word stress or intonation. Functional load "refers to the extent of contrastivity between linguistic units (e.g., phonemic distinction) as well as a measure of the number of minimal pairs for a given contrast, gauging the frequency with which two features contrast" (Wu et al., 2011: p. 2; see also King, 1967; Surendran & Niyogi, 2006). According to this definition, tone has a high functional load, as every monosyllabic word in a tone language carries a tone to signal word meanings. By contrast, word stress is used to make grammatical distinctions for a small number of words, such as in English (Cutler, Dahan, & Van Donselaar, 1997), while intonation is usually employed to express emotions or to indicate sentence types (Cruttenden, 1997). Given its high functional load, tone processing may be specialized to the left hemisphere, which is the dominant hemisphere for speech perception and comprehension (Fedorenko, Behr, & Kanwisher, 2011; Geschwind, 1979).

Some researchers have noted that function-based tone processing needs to be considered in relation to language experience and syllabic (semantic) context (Wong, 2002; Zatorre & Gandour, 2008). First, language experience determines whether or not tone is linguistically significant to listeners. In terms of L1 experience, tone is phonologically contrastive to native listeners, but not to non-native listeners. As a result, tone processing has been found left-lateralized for native listeners but right-lateralized or non-lateralized for non-native listeners (for studies of Mandarin tone processing, see Hsieh et al., 2001; Wang, Behne, Jongman, & Sereno, 2004; Wang, Jongman, &

Sereno, 2001; for Thai: Gandour, Wong, & Hutchins, 1998; Van Lancker, 1973, 1978; for Cantonese: Ho, 2010). Wang et al. (2001) employed a dichotic listening method to test Mandarin and inexperienced English listeners for their tone processing patterns. Four Mandarin tones with four Mandarin syllables comprising 16 real words were presented to the native and non-native listeners in a two-response dichotic paradigm. Mandarin tones were lateralized to the left hemisphere for the Mandarin listeners and processed bilaterally for the English listeners. In another study, Hsieh et al. (2001) tested Mandarin and inexperienced English listeners for their perception of tone and segmental inventories using positron emission tomography (PET). Mandarin tones were imposed on both speech and non-speech tokens. The results revealed left hemisphere specialization for the Mandarin listeners in the processing of native tones, regardless of the speech mode of the carrier tokens. In contrast, the English listeners showed no hemisphere specialization for the processing of speech tones and right hemisphere specialization for the processing of non-speech tones. In other studies, non-native listeners with L1 tone experience have also shown a lesser degree of involvement of the left hemisphere relative to native listeners (Gandour, Wong, Hsieh, Weinzapfel, Van Lancker & Hutchins, 2000; Gandour, Wong, Lowe, Dzemidzic, Satthamnuwong, Tong & Li, 2002; Wang et al., 2004). In Gandour et al.'s (2000, 2002) studies, Thai tones were presented to native listeners of Thai and Mandarin and to English listeners. The results revealed leftward asymmetry only for the Thai listeners, while the Mandarin and English listeners exhibited more activation in the right hemisphere. The distinction in performance between the Thai and Mandarin listeners implies that tone lateralization may be affected by the phonological status of tone categories in the listeners' L1. Wang et al.'s (2004) dichotic listening study on Mandarin tone processing revealed left asymmetry for the native listeners and bilateral processing for English and Norwegian listeners (Norwegian employs tone changes to discriminate word meanings). Nevertheless, leftward asymmetry is not the only outcome seen in tone processing studies of native listeners. Previous studies have also demonstrated bilateral processing patterns for native listeners, indicating even distribution of activated regions in the left and right hemispheres (Baudoin-Chial, 1986; Galaburda, 1995).

Very few studies have probed the effect of L2 learning experience on the cortical processing of tone (Senero & Wang, 2007; Zatorre & Gandour, 2008). Following up on research on brain activity after non-native phonetic learning (e.g., Golestani & Zatorre, 2004), Zatorre and Gandour (2008) postulated that tone learning experience should redistribute activated areas in the left and right hemispheres. For example, they suggested that after L2 tone learning, cortical activation should be detected in the left posterior superior temporal gyrus and adjacent regions, as well as additional cortical areas within the right inferior frontal gyrus (p. 1099). Their hypothesis was supported by Wang, Sereno, Jongman and Hirsch (2003). In Wang et al.'s (2003) study, English beginning learners of Mandarin Chinese received brief tone training, in which they learned to identify the four Mandarin tones with feedback. The findings supported the hypothesis of Zatorre and Gandour (2008): fMRI scans exhibited additional activation in the left hemisphere superior temporal gyrus and in the right hemisphere inferior frontal gyrus only after the training was provided. Nevertheless, despite the observed cortical modification, no hemisphere dominance was evident among the English learners. In a dichotic listening study, Ke (1992) compared L1 and L2 speech processing of monolingual English listeners and those with three years of Mandarin learning experience. The results indicated left hemisphere dominance for the monolingual listeners in the processing of L1 stimuli, but no hemispheric difference for the bilingual listeners in the processing of L1 and L2 stimuli, suggesting that speech lateralization can be modified by L2 learning experience or tone learning experience.

Some critics of the functional tone processing hypothesis have argued that tone processing is confounded with lexical processing, as tone information is encapsulated in a syllabic structure consisting of suprasegmental and segmental tiers (Cutler & Clifton, 1999; Levelt, 1999; Poeppel et al., 2004; Wong, 2002). In response to this argument, some previous research has attempted to separate tone from lexical status by superimposing tone on either meaningless words (e.g., Xu, Gandour, & Francis, 2006) or non-speech tokens (e.g., Hsieh et al., 2001). These studies indicate that lexical status does not affect tone lateralization patterns. For example, Hsieh et al. (2001) observed the same left hemisphere dominance (LHD) pattern in the native listeners' processing of Mandarin tones, whether they were imposed on speech (real words) or non-speech (low-pass filtered) tokens. In Xu et al. (2006), Thai and Mandarin tones were imposed on Mandarin syllables, resulting in nonsense words in Thai, but real words in Mandarin. These stimuli were then presented to Mandarin and Thai listeners. Analysis of fMRI scans indicated a region of interest (ROI) in the left hemisphere that was activated only when the Mandarin and Thai listeners were presented with native tones. However, no hemispheric specialization was observed for listeners in the processing of either native or non-native tones, indicating a bilateral processing pattern. Nevertheless, some other tone processing studies have demonstrated different hemisphere lateralization patterns in the processing of speech and non-speech tones. For example, Gandour et al. (2003) presented Mandarin listeners with native tones imposed on speech words and non-speech hums (created in Praat). The averaged fMRI scans demonstrate significantly larger activated areas in the left hemisphere for the speech tones than the non-speech hummed tones. Van Lancker (1978) also obtained a similar result, except that the hummed tones were produced by a native speaker.

In summary, many previous studies support the functional hypothesis, indicating that tone processing exhibits left hemisphere dominance. However, given the complexity introduced by factors such as L1 and L2 tone experience, or the lexical status to which tone is attached, more research is necessary to account for the full range of lateralization patterns found in previous tone processing research.

4.1.2. Acoustic hypothesis

Unlike the functional hypothesis, which emphasizes the difference in linguistic functional load carried by tone versus other prosodic features, the acoustic hypothesis focuses on distinctions in particular pitch-related acoustic cues, such as prosodic domain (Poeppel, 2003) or frequency range (Ivry & Roberson, 1998). As reviewed in Zatorre and Gandour (2008), outside the speech domain, the right auditory cortex is responsible for prosodic processing under specific circumstances (e.g., Binder, Frost, Hammeke, Cox, Rao, & Prieto, 1997; Patel & Balaban, 2001; Penhune, Zatorre, & Evans, 1998). Previous research has revealed that inside the speech domain, the processing of linguistic pitch contrasts is also driven primarily by the right hemisphere. For example, Bradvik, Dravins, Holtas, Rosen, Ryding, and Ingvar (1991) found that Swedish subjects with right hemisphere damage (RHD) had greater difficulty identifying contrastive stress or prosodic patterns at the phrasal level compared to those with no damage. Gandour, Tong, Wong, Talavage, Dzemidzic, Xu, & Lowe (2004) and Shipley-Brown, Dingwall, Berlin, Yeni-Komshian, & Gordon-Salant (1988) found rightward asymmetry in native speakers' perception of linguistic intonation through fMRI scans and a dichotic listening paradigm, respectively. Zhang, Shu, Zhou, Wang, & Li (2010) created synthesized intonation and speech rhyme materials from real speech utterances. Despite variations in acoustic realization (i.e., intonation is expressed mainly by fundamental frequency, while speech rhyme combines duration, pitch, pause, and intensity), these two prosodic features were found to be right-lateralized to the auditory association cortex. Zhang et al.'s (2010) finding is also consistent with that of Poeppel (2003) in suggesting that the right hemisphere is sensitive to longer prosodic windows (above 150-250 ms). Poeppel predicted that at the syllable level, tone falls into a longer window, and should hence be lateralized to the right hemisphere (see also Gandour et al., 2003).

Recently, researchers have paid much attention to the processing of individual categories inside a prosodic unit. Some studies have found not only rightward asymmetry, but also leftward asymmetry or bilateral processing in processing distinctive acoustic realizations. For example, in a dichotic listening study on Japanese pitch accent, Wu et al. (2011) tested native speakers of Japanese, Mandarin Chinese, and English for their identification of three disyllabic Japanese pitch patterns: H*L (high-accent-low), LH*, and LH. While overall performance showed no hemispheric preference, listeners in all three language groups showed right hemisphere preference for processing the H*L pattern, left hemisphere preference for LH*, and no hemisphere dominance for LH, indicating that different acoustic distinctions be processed in different areas of the brain. Ho (2010) examined hemispheric lateralization of Cantonese tones in an event-related potential (ERP) study. Native speakers' ERP amplitudes were compared between level and contour tones. The results demonstrated a left lateralization for all the tones, but a greater left lateralization effect for the contour tones. In Wang et al.'s (2004) dichotic study of Mandarin tone lateralization, a larger number of listeners showed left hemisphere dominance for Mandarin Tone 3, which carries a more complex tone contour (falling rising), than for other Mandarin tone categories. A higher degree of left hemisphere involvement in the processing of Mandarin Tone 3 is also reflected in tone production studies. For example, in a study of the production of Mandarin tones by speakers with left hemisphere damage (LHD), Tone 3 was responsible for over half of the tonal confusions (see Gandour, 2006, for a review). Broadly speaking, researchers have found that LHD patients have greater difficulty producing tones with dynamic F0 contours. Thai, high rising and falling rising tones and Mandarin high rising, falling rising and high falling tones are among the tones that LHD patient find hard to produce (Gandour, 2006; Liang & Heuven, 2004). Taken together, the performance of normal listeners and LHD patients in the perception and production of Mandarin and Thai tones echo the findings of Ho (2010), suggesting a greater left hemisphere bias for contour tones than level tones. This pattern has also been accounted for within the categorical model of tone perception which concludes that level tones tend to be perceived within-category while contour tones tend to be perceived across-category (e.g., Abramson, 1973; Francis et al., 2003). In terms of hemispheric processing, more neural activity in the left hemisphere is observed in the perception of across-category than within-category phonetic contrasts (Dehaene-Lambertz, Dupoux, & Gout, 2000; Jacquemot, Pallier, LeBihan, Dehaene, & Dupoux, 2003; Poeppel, Guillemin, Thompson, Fritz, Bavelier, & Braun, 2004). For example, Jacquemot et al., (2003) presented Japanese and French speakers with acoustic and phonological contrasts differing in vowel duration. The short-long vowel contrast was phonological for the Japanese speakers but not for the French speakers. fMRI scans indicated more activated areas in the left hemisphere for the Japanese speakers but not the French ones. Therefore, previous research suggests that the left hemisphere lateralization of tones with dynamic F0 is explained by the greater tendency for such tones to be perceived categorically. However, some findings may not support this view. For example, in a study reviewed by Gandour (2006), some Cantonese LHD patients showed no difference in the production of the level and contour tones.

To summarize, not all previous studies indicate rightward lateralization for tone/pitch processing. Rather, different acoustic realizations may cause tone/pitch to be processed in the left or right side of the brain, or to exhibit no hemispheric dominance. Therefore, in extending the acoustic hypothesis to tone or pitch processing, further research should take into account the responses of the left and right hemispheres to different acoustic features of individual tone patterns.

4.1.3. A combined account

Although the functional and acoustic hypotheses are supported by a great deal of empirical evidence, many researchers suggest that the strengths and weaknesses of the two hypotheses call for an alternative hypothesis on language processing (e.g., Wong, 2002; Zatorre & Gandour, 2008). For example, Zatorre and Gandour (2008) suggest a combined model, in which the acoustic model can account for early stages in processing of speech and non-speech signals, and the functional hypothesis can explain later stages of processing that involve more abstract knowledge, such as linguistic function. This combined (or more complete) model can explain hemispheric processing at different time intervals on the basis of stimulus characteristics. Similar alternative hypotheses have been proposed to account for the interplay of linguistic and acoustic aspects of pitch processing suggested by some other studies (e.g., Gandour et al., 2003; Wu et al., 2011). For example, Gandour et al. (2003) examined hemispheric processing of Mandarin tones and intonation by native listeners. In a working hypothesis, the researchers proposed that the left-lateralized processing of tone and the bilateral processing of intonation were the outcome of an interaction between functional load and prosodic domain (an acoustic feature). In other words, the acoustic processing of a longer prosodic domain led intonation to be processed predominantly in the right hemisphere, whereas the functional processing of phonological features caused tone to be lateralized to the left hemisphere.

Taken together, the research findings described above indicate that the acoustic hypothesis can be employed to explain the effects of acoustic distinctions on tone processing, while the functional hypothesis may be applied to account for the processing of tone due to its phonemic function. In terms of lateralization pattern, the functional account may severally predict that tone is lateralized to the left hemisphere; the acoustic account, however, associates tone with varying lateralization patterns according to F0 changes. As pointed out, a combined account may contribute to the explanation of the overall picture of tone processing given that tone is realized at both the phonetic and phonological levels.

4.2. The current study

The main goal of the current study is to investigate the extent to which acoustic correlates and phonological status affect hemispheric tone processing. At the acoustic-phonetic level, the current study focuses on the effects of F0 height and contour on tone processing, while at the phonological level, the effects of L1 and L2 experience as well as lexical status are examined. Native speakers of Mandarin, Thai, and English participated in a dichotic listening test in which they identified the four Mandarin tone categories. The non-Mandarin listeners differed in their L2 experience with Mandarin: those with and without Mandarin experience are referred to as experienced listeners and inexperienced listeners, respectively.

Based on previous findings on the effects of acoustic properties and linguistic function on hemispheric specialization for lexical tones (e.g., Hsieh et al., 2001; Gandour et al., 2003; Xu et al, 2006; Wang et al., 2004), the current study addresses three research questions. These research questions and associated predictions are listed below.

(1) **Do acoustic properties affect hemispheric lateralization for tone processing?**

Based on previous findings on the acoustic effects of F0 height and contour of individual tone categories (e.g., Abramson, 1973; Ho, 2010; Wang et al., 2004), it is expected that the contour tones in Mandarin, in particular the falling rising Tone 3, will be more left-lateralized than the high level Tone 1. This prediction is also made because level tones tend to be perceived non-categorically, thus causing less cortical activation, especially in the left hemisphere. In contrast, contour tones are associated with greater activation in the left hemisphere, for they are more likely to be perceived categorically due to their dynamic acoustic distinctions (Dehaene-Lambertz et al., 2000; Jacquemot et al., 2003; Poeppel et al., 2004).

(2) **Does L1 and L2 experience affect tone processing?**

Both L1 and L2 experience involves the linguistic use of tone categories; thus, left hemisphere dominance or a higher degree of left hemisphere involvement is expected for the native and experienced listeners relative to the non-native and inexperienced listeners, respectively. For the experienced listeners, since additional activated areas have been observed in both left and right hemispheres, better performance is expected for both ears relative to the inexperienced listeners (Golestani & Zatorre, 2004; Wang et al., 2003). Additionally, processing patterns are expected to be the same for the inexperienced Thai and inexperienced English listeners, given that the Mandarin tone system is non-native to the listeners in both groups (Gandour et al., 2003; Wang et al., 2004). It is assumed that these listeners will show comparable involvement of the left and right hemispheres or right hemisphere dominance.

(3) **Does lexical status affect hemispheric tone processing?**
This question is addressed because, given the phonological function of tones, tone lateralization patterns may differ as a function of stimulus type (Poeppel et al., 2004; Wong, 2002). If tone processing is affected by the semantic information of the stimulus words, it is expected that tone imposed on real words is more left-lateralized than tone imposed on nonsense words. If, conversely, tone processing is independent of lexical processing (i.e., prelexical), no difference in lateralization pattern is expected between real words and nonsense words (Hsieh et al., 2001; Xu et al., 2006).

4.3. Method

4.3.1. Listeners

All 100 listeners from the identification study (see Chapter 3) participated in the dichotic listening study. These participants were grouped into native Mandarin (NM), experienced Thai (ET), inexperienced Thai (IT), experienced English (EE) and inexperienced English (IE) groups, with 20 in each group.

4.3.2. Stimuli

The testing stimuli presented to the listeners included the four Mandarin tones imposed on two syllables to form eight monosyllabic real words for each group, and imposed on two other syllables to form eight nonsense words for all the groups[1]. These

[1] As in the identification test, the hummed tones were presented in the same manner as the tones imposed on the real and nonsense words in the dichotic listening test. Data collected from the hums will be reported and discussed in a future work.

syllables were identical to those used in the identification test (see Table 3.1 of Chapter 3): the syllables /lau, pa(a)/ were used to form real words for the Mandarin and Thai groups; the syllables /pi, fan/ were used to create real words for the English groups; and the syllables /pju, kjɛn/ were used to form nonsense words for all groups.

Each target tone was presented in the same manner as it was in the identification task, i.e., a target tone was followed by a falling tone, to the left and right ears. Target tones were presented in a two-tone sequence to increase task difficulty and to avoid any ceiling effect. Ninety-six dichotic pairs (4 syllables x 12 pairing patterns x 2 talkers) were created, such that in each pair, the two tokens had the same segmental components but different tone patterns for the target tones. For example, ba^1 *"eight"* and ba^4 *"father"*, were paired as of ba^1-ba^4 and ba^4-ba^4. Dichotic pairs were constructed and edited using Audacity 1.2.6; one word in each pair was imported into the left channel and the other into the right channel. Each pair was normalized to 70 db for intensity using Praat (Boersma & Weenink, 2009). Aside from 40 ms of silence between the target tone and the falling tone, i.e., 20 ms at the end of the target tone and 20 ms at the beginning of the following tone, no additional interval was inserted between the two tones. The duration of each token was also equalized across the four tones for each syllable, using the same method introduced in Chapter 2. Thus, the durations of the four tones were identical for each syllable. In other words, the tones presented to the listeners differed in pitch, but not in intensity or length.

4.3.3. Procedure

The experimental design follows the dichotic listening paradigm, a common and reliable method for examining hemispheric performance in the perception of verbal materials (Hugdahl, 1995; Kimura, 1961, 1967; Voyer &Techentin, 2009). Dichotic listening is a process in which listeners are presented with two different sounds simultaneously to the left and right ears. As speech signals are delivered through enhanced contralateral auditory pathways, better performance of the right ear (right ear advantage, or REA) relative to the left ear indicates more active involvement of the left hemisphere (Bryden, 1988; Hugdahl, 2003; Penner, Schläfli, Opwis, & Hugdahl, 2009). In addition to REA, listeners may exhibit better performance of the left ear (left ear advantage, or LEA) or comparable performance of the left and right ears (no ear advantage, or NEA). In a dichotic listening test, listeners can be instructed to respond to one (Brancucci, Anselmo, Martello, & Tommasi, 2008; Techentin & Voyer, 2009) or two stimuli (Millay, Roeser, & Godfrey, 1977; Wang et al., 2001, 2004; Wu et al., 2011) with or without directed auditory attention. The current study followed Wang et al. (2001, 2004), using a two-response paradigm in which listeners' attention is directed to the stimuli in both ears (see also Millay, Roeser, & Godfrey, 1977; Van Lancker, 1978). The listeners responded to the target tones presented to the left and right ears by pressing one of the numbers 1-4 on the keyboard. To eliminate response order bias, half of the listeners responded first to the stimulus in the left ear and then to the one in the right ear, while the other half responded in the opposite order. The response time for each stimulus was 2 s.

The 96 stimuli were assigned into three blocks semi-randomly following each response order such that in one block, 32 randomized stimuli were played to the listeners, followed by a short self-paced break. A familiarization session was also provided before the real test with 24 dichotic pairs (1 syllable x 12 pairing patterns x 2 speakers), which were presented in the same response order as the real test. The syllable selected for the practice was *ma*, which was not used in the real test. The dichotic test lasted approximately 50 min for each participant.

4.4. Results

4.4.1. Perceptual accuracy

The percentage of correct responses was examined through a mixed design analysis of variance (ANOVA) using tone (Tone 1, Tone 2, Tone 3, Tone 4), word type (real word, nonsense word), and ear (left ear, right ear) as the within-subjects factors and group (NM, ET, IT, EE, IE) as the between-subject factor. Since the current study focuses on the patterns of hemispheric specialization for tone processing, only the main effect of ear factor or the interaction of ear and other factors are reported and discussed in this section.

Table 4.1 lists the results of the main effect of ear and the interactions of ear and the other three factors, including tone, word type, and group. As can be seen, no main ear effect was observed, while the interactions of ear and group, and ear and tone were significant with a large effect size, i.e., $\eta_p^2 > .1$ (Cohen, 1988). The interactions of ear x tone x word type, and ear x word type x group x tone were also significant, but with smaller effect sizes of .065 and .08, respectively.

Table 4.1. The ear-related source, F, p values and size of effect for ANOVAs on percent of correct responses (significant results are highlighted.)

Source	F	Sig.	Partial Eta Squared (η_p^2)
Ear	(1, 95)= 0.21	.654	.002
Ear x Group	(4, 95)=3.1	**.019**	.116
Ear x Word type	(1, 95)=0.86	.356	.009
Ear x Word type x Group	(4, 95)=1.17	.33	.047
Ear x Tone	(3, 285)=30.36	**<.0001**	.242
Ear x Tone x Group	(12, 285)=1.3	.21	.052
Ear x Tone x Word type	(3, 285)=6.6	**<.0001**	.065
Ear x Word type x Group x Tone	(12, 285)=2.1	**.019**	.08

4.4.1.1. Interaction of ear x group

The interaction of ear and group was significant, $F(4, 95) = 3.1$, $p = .019$. To examine ear effect for each group, a one-factor ANOVA with ear as the within-subjects factor was performed for each group. The results, however, did not show a significant difference between the left and right ears for any of the five groups: NM, F (1, 19) = 1.6, $p = .226$, $\eta_p^2 = .076$; ET, $F(1, 19) = 3$, $p = .097$, $\eta_p^2 = .138$; IT, $F(1, 19) = 2.1$, $p = .162$, $\eta_p^2 = .1$; EE, $F(1, 19) = 1.6$, $p = .222$, $\eta_p^2 = .077$; IE, $F(1, 19) = 4$, $p = .06$, $\eta_p^2 = .174$. Furthermore, to explore the effect of group on the performance of the left and right ears, a set of one-factor ANOVAs with group as the between-subjects factor was conducted for each ear. As shown in Figure 4.1, the effect of group on right ear performance was significant, $F(4, 95) = 13.7$, $p < .0001$. Bonferroni adjusted post hoc analyses showed a significant difference in the right ear performance between the NM listeners and the IT, EE, and IE listeners ($p < .0001$), but not between the NM and the ET listeners ($p = .078$). Additionally, the ET listeners' performance was better than that of the IT and IE listeners (ET vs. IT, $p = .006$; ET vs. IE, $p = .003$), but comparable to that of the EE listeners ($p = .522$). Moreover, the effect of group was also significant for left ear performance, F (4,

95) = 8.8, $p < .0001$. Post hoc (Bonferroni) analyses indicated better performance for the Mandarin speakers than for listeners in the four other groups (NM vs. ET: $p = .028$; NM vs. IT: $p < .0001$; NM vs. EE: $p = .03$; NM vs. IE: $p < .0001$), while no difference in left ear performance was observed among the four non-native groups (Figure 4.1).

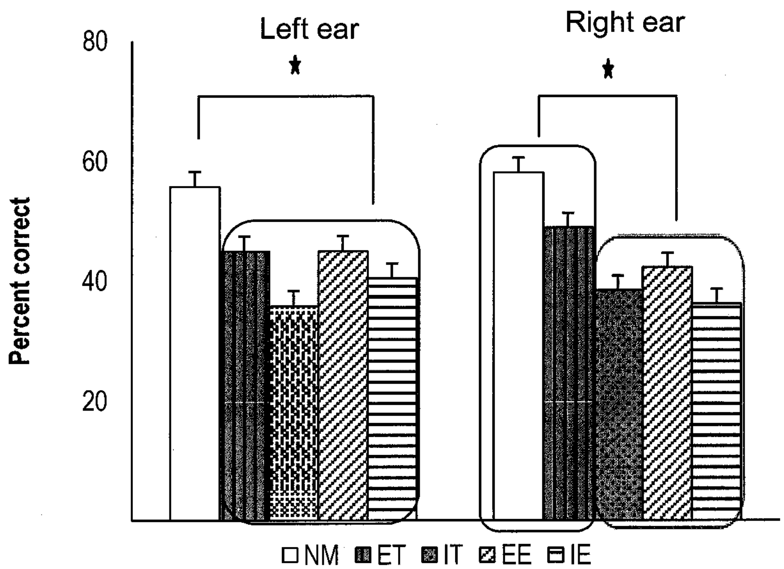


Figure 4.1. ***Interaction of ear x group. The stars indicate a significant difference between the native and non-native listeners at p < .05. The difference in right-ear performance between NM and ET groups is not significant.***

4.4.1.2. Interaction of ear x tone

Ear and tone also showed significant interactions. A one-factor ANOVA with ear as the within-subjects factor showed LEA for Tone 1, $F(1, 99) = 21.7$, $p < .0001$, $\eta_p^2 = .18$, and Tone 2, $F(1, 99) = 17.8$, $p < .0001$, $\eta_p^2 = .152$, but showed REA for Tone 3, $F(1, 99) = 31.5$, $p < .0001$, $\eta_p^2 = .241$, and Tone 4, $F(1, 99) = 17.6$, $p < .0001$, $\eta_p^2 = .151$.The percentage of correct responses is presented in Figure 4.2.

In order to determine the difference in tone identification in the left and right ears, a one-factor ANOVA using tone as the within-subjects factor was conducted across groups. The result showed a significant difference between the tones for the left ear, $F(3, 297) = 17.2$, $p < .0001$, $\eta_p^2 = .147$, and the right ear, $F(3, 297) = 14.3$, $p < .0001$, $\eta_p^2 = .126$. Post hoc tests further indicated that in the left ear, Tone 1 was identified significantly more accurately than Tone 2, Tone 3, and Tone 4 (Tone 1 vs. Tone 2 and Tone 3, p = .001; Tone 4, p < .0001), while Tone 2 was also identified more accurately than Tone 4 (p = .002). In the right ear, performance for Tone 3 was better than for Tone 1, Tone 2, and Tone 4 (Tone 3 vs. Tone 1, p = .002, Tone 2, p < .0001, Tone 4, p = .017) and performance for Tone 4 was also better than for Tone 2 (p = .01) (see Figure 4.2).

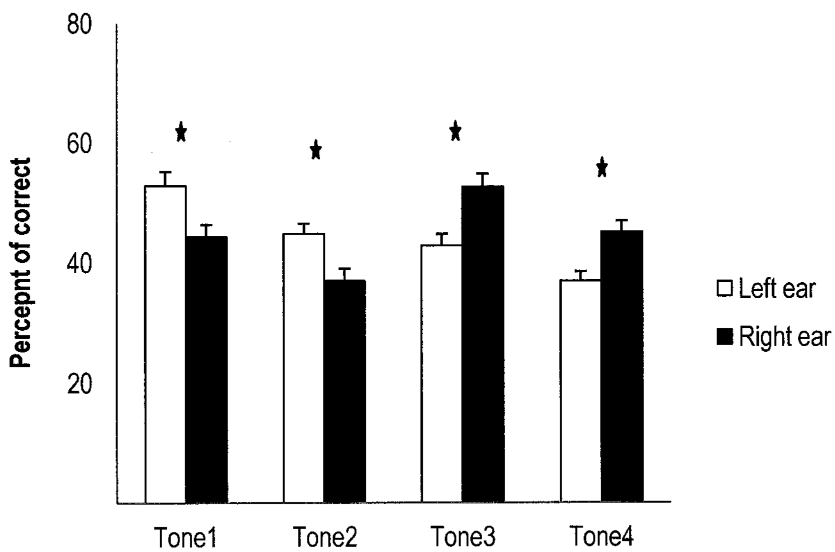


Figure 4.2. ***Interaction of ear x tone. The stars indicate a significant difference at p < .05.***

4.4.1.3. Interaction of ear x tone x word type

The interaction of ear x tone x word type was significant, but with a small effect size, η_p^2 = .065, indicating a weak interaction among these three factors (Wilkinson, 1999). To examine the role of word type on the lateralization of Mandarin tones, a set of one-factor repeated measure ANOVAs with ear as the within-subjects factor was performed for each tone imposed on the real words and the nonsense words. The results indicate similar ear advantage (EA) patterns for the four tones regardless of word type. That is, Tone 1 and Tone 2 were better identified in the left ear while Tone 3 and Tone 4 were better identified in the right ear when they were imposed on both real and nonsense words (see Table 4.2).

Table 4.2. Mean percentage of correct response, F, p value and size of effect for ANOVAs as a function of word type x tone x ear (standard deviations are in parentheses and significant results are highlighted.)

			% correct	F	p	Partial Eta Squared (η_p^2)
Tone 1	Real	Left	54 (24)	(1,99) = 6.5	.013	.061
		Right	48 (21)			
	Nonsense	Left	52 (27)	(1,99) = 24.6	**<.0001**	.199
		Right	41 (23)			
Tone 2	Real	Left	45 (20)	(1,99) = 9.6	**.003**	.088
		Right	38 (23)			
	Nonsense	Left	45 (20)	(1,99) = 12.1	**<.0001**	.117
		Right	36 (22)			
Tone 3	Real	Left	48 (22)	(1,99) = 5.6	.020	.053
		Right	53 (24)			
	Nonsense	Left	38 (23)	(1,99) = 43.8	**<.0001**	.307
		Right	52 (23)			
Tone 4	Real	Left	37 (18)	(1,99) = 21.5	**<.0001**	.178
		Right	48 (22)			
	Nonsense	Left	37 (19)	(1,99) = 4.03	.047	.039
		Right	42 (21)			

4.4.2. Distribution of ear preference

In addition to mean perceptual accuracy, data were also examined in terms of the distribution of ear preference, that is, in terms of the number of listeners showing LEA, REA, or NEA. This analysis was performed using Pearson's chi-square ($\chi 2$) analysis. A three-way contingency table was created in SPSS (SPSS Inc., Chicago, IL) with ear preference as the column variable, tone category as the row variable, and language group as the layer variable. The results indicate a significant association between tone pattern and distribution of ear preference for the listeners in the NM group, $\chi 2$ (6) = 21.2, p = .002; ET, $\chi 2$ (6) = 15.7, p = .016; IT, $\chi 2$ (6) = 13, p = .042, and EE group, $\chi 2$ (6) = 13.1, p = .041, but not for listeners in the IE group, $\chi 2$ (6) = 9.4, p = .132. Thus, further analysis was performed to examine the distribution of ear preference in the processing of individual tone patterns for the NM, ET, IT, and EE groups. As illustrated in Table 4.3, the processing patterns were consistent with those observed for each tone through perceptual accuracy, i.e., LEA for Tone 1 and Tone 2 and REA for Tone 3 and Tone 4. For Tone 1, more listeners in the NM, ET, and EE groups showed LEA than those showing REA or NEA, while for Tone 2, more listeners in the NM, IT, and EE groups exhibited LEA than REA or NEA. For Tone 3 and Tone 4, REA listeners outnumbered LEA or NEA listeners in all four groups.

Table 4.3. ***The number of listeners showing left ear advantage (LEA), right ear advantage (REA), or no ear advantage (NEA) for NM, ET, IT, and EE groups in the perception of the four Mandarin tones. For each group, n is 20. The distribution results are shown by x^2 and p values (outnumbered EA pattern and significant results are highlighted.).***

		NEA	LEA	REA	X^2	p
Tone 1	NM	1	**14**	5	13.3	**0.00**
	ET	2	**13**	5	9.7	**0.01**
	IT	0	**12**	8	0.8	0.37
	EE	2	**13**	5	9.7	**0.01**
Tone 2	NM	1	**11**	8	7.9	**0.02**
	ET	3	**11**	6	4.9	0.09
	IT	1	**12**	7	9.1	**0.01**
	EE	1	**13**	6	10.9	**0.00**
Tone 3	NM	1	1	**18**	28.9	**0.00**
	ET	3	3	**14**	12.1	**0.00**
	IT	1	4	**15**	16.3	**0.00**
	EE	2	6	**12**	7.6	**0.02**
Tone 4	NM	1	6	**13**	10.9	**0.00**
	ET	1	6	**13**	10.9	**0.00**
	IT	2	5	**13**	9.7	**0.01**
	EE	1	5	**14**	13.3	**0.00**

4.4.3. Summary

To summarize, the overall performance exhibited an ear difference in the processing of individual tone categories across groups. Both native and non-native listeners showed LEA for Tone 1 and Tone 2 and REA for Tone 3 and Tone 4. However, no ear difference was observed among groups or between word types, indicating bilateral processing of Mandarin tones by the native and non-native listeners and for the real and nonsense words. Nevertheless, right ear performance showed a group difference, in particular between the native and non-native listeners and between the ET listeners and other non-native listeners. The results indicate that the NM listeners were better at identifying native tone categories presented to the right ear than the non-Mandarin listeners, except the ET listeners. In contrast, left ear performance was better for the native listeners than the non-native listeners in all four groups. Furthermore, the distribution of ear preference also revealed a group difference, as ear preference was associated with tone patterns in the NM, ET, IT, and EE groups, but not in the IE listeners.

4.5. Discussion

In this section, results of the current study are discussed in relation to the theoretical hypotheses with regard to the effects of acoustic properties and linguistic functions on hemispheric specialization for tone processing.

4.5.1. Acoustic processing

The current study exhibits right lateralization for the high level Tone 1 and left hemisphere specialization of two contour tones: the falling rising Tone 3 and falling Tone 4. This result supports the hypothesis predicting right hemisphere dominance for level tones and left hemisphere dominance for contour tones. However, another contour tone, the rising Tone 2, was found to be right-lateralized, like the high level tone. The acoustic hypothesis states that pitch processing is more commonly lateralized to the right hemisphere (Zatorre & Gandour, 2008). However, previous findings also revealed differences in hemispheric specialization due to acoustic distinctions between individual pitch patterns (Gandour, 2006; Ho, 2010; Wang et al., 2004; Wu et al., 2011). Although the current findings do not show left hemisphere dominance for native listeners in the processing of tone patterns (e.g., Ho, 2010; Wang et al., 2004), they do suggest that the left hemisphere is more sensitive to dynamic tone changes than the right hemisphere. As for the rising Tone 2, it is probably its flatter contour that induces right hemisphere lateralization, exhibiting a similar processing pattern to that seen for the high level Tone 1. As shown in Table 2.3 of Chapter 2, the average falling and rising ranges for Tone 2 were 0.3 and 1 (in T value, the same hereafter), respectively. In contrast, the falling and rising ranges for Tone 3 were 2 and 2.4, respectively, showing distinctive falling and rising contours, while the falling range of Tone 4 was 2.7, reflecting a steeper falling slope. Moreover, the relatively higher F0 may also explain why Tone 1 and Tone 2 were processed more actively in the right hemisphere because cerebral responses to higher frequency were found to be weighted toward the right (Boemio, Fromm, Braun, & Poeppel, 2005; Zatorre & Belin, 2001). In Mandarin, Tone 1 is realized as a high level tone and Tone 2 as a rising tone. Compared to the low falling rising Tone 3 and high falling Tone 4, Tone 1 and Tone 2 carry a higher average F0. Both explanations support the acoustic hypothesis indicating the influence of acoustic changes on tone processing.

4.5.2. *Functional processing*

The current results also indicate a group difference in lateralization patterns for tone processing, though not the difference that was expected. Left hemisphere dominance was predicted for the NM listeners and the experienced listeners and right hemisphere dominance or bilateral processing was expected for the inexperienced listeners. However, in the current study, no ear difference was shown for any of the groups; instead, a group difference was observed only in monaural performance. The Mandarin listeners' left ear performance was better than that of the non-Mandarin listeners, indicating a higher degree of right hemisphere involvement. In the right ear, however, the Mandarin listeners showed comparable accuracy relative to the ET listeners, while outperforming the non-Mandarin listeners in the other groups. Hence, with respect to right ear performance, the ET listeners were similar to the Mandarin listeners, indicating a high degree of left hemisphere activation. A group difference was also observed in the distribution of ear preference for different tone patterns for the NM, ET, IT, and EE groups, but not for the IE group. In three respects, these findings support the functional hypothesis with regard to the effects of language experience on tone processing. First, the NM listeners showed a higher degree of left hemisphere involvement, suggesting that linguistic processing of tone is more active in this group relative to the non-native listeners. Second, the ET listeners also showed a higher degree of left hemisphere involvement than the non-Mandarin listeners in the other groups, indicating that the interplay of L1 and L2 tone experience may enhance activation in the left hemisphere. Third, the IE listeners showed no association between the distribution of ear preference and tone patterns, which may suggest a reduced sensitivity to tone categories due to a lack of tone experience.

In addition, the current study found no difference in the processing of tones imposed on real words and nonsense words, confirming the prediction that tone processing is independent of lexical status (Hsieh et al., 2001; Xu et al., 2006). This finding also provides evidence that tone lateralization to the left hemisphere, observed in previous studies such as Wang et al. (2001, 2004), stems from the phonological use of tone rather than from the semantic information conveyed by the carrier words (c.f., Wong, 2002); listeners' attention is focused on the tonal information. As a result, semantic processing is not confounded with tone processing.

4.5.3. Acoustic-functional processing

As suggested by the above discussion, the current study supports both the acoustic and functional hypotheses: on the one hand, the acoustic correlates of tone are associated with rightward or leftward asymmetry across groups; on the other hand, native listeners exhibit greater left hemisphere involvement in tone processing than non-native listeners. The interplay of acoustic and linguistic aspects has been noted in previous studies (e.g., Gandour et al., 2003; Wu et al., 2011). Thus, results of the current study, together with the findings of previous research, can be interpreted as support for the combined acoustic-functional account, which has been proposed to explain pitch processing at a lower acoustic and a higher functional level at different time intervals (Gandour et al., 2003; Zatorre & Gandour, 2008). In particular, Zatorre and Gandour (2008) have suggested that functional processing may occur later than acoustic processing. In the current study, the acoustic correlates of tone had a greater influence than linguistic function on tone processing among inexperienced listeners. However, among native listeners and experienced Thai listeners, both acoustic correlates and linguistic function were found to be significant in tone processing. In comparison with previous findings, this may suggest that acoustic factors play a more important role than linguistic function for pitch patterns with a lower functional load (Gandour et al., 2003; Wu et al., 2011; Zhao, Shu, Zhang, Wang, Gong & Li, 2008). For example, Japanese pitch accent was found to be processed at the acoustic level to a greater extent by native and non-native listeners (Wu et al., 2011). In another study, both Mandarin and English listeners showed dependence on the prosodic domain in the processing of Mandarin intonation patterns, which are also associated with a lower functional load (Gandour et al., 2003).

4.6. General discussion

The findings of this study benefit our understanding of tone processing in three respects. First, they show that both L1 and L2 tone experience affect the hemispheric lateralization of tone processing. Second, this study suggests that patterns of hemispheric specialization for tone processing may be predicted by the F0 trajectories of different tones. Lastly, this study provides new evidence for a combined account of tone processing that recognizes the influence of acoustic properties and linguistic function on processing patterns.

In this study, two measures, perceptual accuracy and distribution of ear preference, demonstrated a hierarchical processing pattern for listeners who differ in their L1 and L2 tone experience, with native speakers occupying the highest position in the hierarchy. Echoing previous studies (Gandour et al., 2003; Hsieh et al., 2001; Xu et al., 2006; Wang et al., 2004), native speakers' superior right ear performance suggests greater left hemisphere involvement, and by extension, the presence of higher-level linguistic processing. Non-native listeners who are familiar with the use of tone from their L1 and who have experience with the L2 tone system occupy the second position in the hierarchy. This group of listeners, which has not been studied in previous research on the neural processing of tone, exhibited native-like performance in terms of both right ear performance and sensitivity to F0 contour distinctions. Non-native listeners who were familiar with the tone from their L1, but were inexperienced with the L2 tone system, as well as listeners without L1 tone experience but with L2 tone experience, occupy the third position in the hierarchy. These listeners showed strong sensitivity to acoustic distinctions in tone categories, but little involvement of phonological processing. At the bottom of the hierarchy are listeners without L1 or L2 tone experience. Their processing of tone is limited to the acoustic level, with a much lesser degree of sensitivity to categorical differences between tones.

This study found that the acoustic processing of tones can be predicted by their F0 trajectories, with dynamic F0 being lateralized to the left hemisphere and static F0 to the right hemisphere, supporting previous observations from lesion studies, ERP research, and dichotic listening tasks (Gandour, 2006; Ho, 2010; Wang et al., 2004). Moreover, tones with dynamic F0 tend to be processed more categorically than tones with static F0 (Abramson, 1973). Thus the current findings also support previous research showing that tones which are perceived more categorically are associated with greater neural activation in the left than the right hemisphere (Jacquemot et al., 2003; Poeppel et al., 2004). Nevertheless, arguments may arise due to the rising contour of Mandarin Tone 2. Among the three contour tones in Mandarin, Tone 2 was the only one showing right hemisphere dominance. The finding may be accounted for by the relatively flatter contour of Tone 2. One optional explanation is that the rightward asymmetry in the processing of Tone 1 and Tone 2 was induced by other acoustic factors, such as high frequency (Boemio et al., 2005; Zatorre & Belin, 2001). In Mandarin, Tone 1 is realized as a high level tone and Tone 2 a mid rising tone, thus it is likely the high average F0 that led them to be lateralized to the right hemisphere.

One of the contributions of the current study lies in its support for a combined account of tone processing, which emphasizes not only the influence of acoustic properties, but also linguistic function on hemispheric pitch processing. As noted above, an integrated account (e.g., Gandour et al., 2003; Zatorre & Gandour, 2008) states that pitch processing occurs both at earlier (acoustic) and later (linguistic) stages of hemispheric processing, demonstrating the interplay of acoustic and linguistic aspects. In the current study, the aforementioned hierarchical processing patterns might correspond to different time intervals in hemispheric processing. For example, the native listeners' performance reflects tone processing from earlier to later stages, indicating sensitivity to both acoustic cues and the linguistic function of tone. By contrast, the non-native language listeners with no previous tone experience depend more heavily on acoustic cues in the processing of tone categories, reflecting earlier-stage pitch processing. The interaction of acoustic and linguistic aspects of processing was also observed to some extent through distribution of ear preference, which suggests that tone experience may enhance listeners' ability to categorize tone patterns, an ability that results in consistently different lateralization patterns.

4.7. Concluding remarks

The current study highlights the effects of acoustic properties on tone processing, thus supporting the acoustic hypothesis; this study also supports the functional hypothesis, based on the different performance of the listeners with differing L1 and L2 tone experience. Taken together, the findings of this dichotic study provide evidence for a combined acoustic-functional hypothesis of tone processing.

Further research may take into account the effects of experimental tasks on hemispheric tone processing. Previous studies have pointed out that a set of cognitive factors, such as memory load, focus of attention, and problem-solving strategy, affect hemispheric engagement in pitch processing (Kinsbourne, 1970; Voyer &Techentin, 2009; Zatorre & Gandour, 2008). For example, memory load and general auditory attention have been found to induce rightward asymmetry in pitch processing (Gandour et al., 2004; Plante, Creusere, & Sabin, 2002; Zatorre & Gandour, 2008), while selective attention has been found to cause leftward asymmetry in tone processing (Li, Gandour, Talavage, Wong, Dzemidzic, Lowe, & Tong, 2003).

In addition, task difficulty may influence hemispheric specialization for pitch processing. For example, previous research has shown greater involvement of the right hemisphere with an increase in task difficulty (Aasland & Baum, 2003). In the current study, task difficulty was increased by asking listeners to identify the target tone in a two-tone sequence. Although the modification might not affect binaural performance for native listeners (see Chapter 3), its effect was uncertain in the current dichotic listening task, as judged by native speakers' bilateral processing pattern for tone (c.f., Wang et al., 2001, 2004). Hence, whether or not the greater degree of right-hemisphere involvement in tone processing shown in the current study is the outcome of additional memory load or increased task difficulty needs further investigation in future research.

5. General discussion and future research

This dissertation is comprised of three perceptual studies on tone assimilation, tone identification, and hemispheric lateralization. In this chapter, the main findings of the three studies are summarized and then discussed first in terms of the phonetic level, with a focus on the effects of acoustic properties on tone perception, and second in terms of the phonological level, with a focus on the effects of language experience and phonological status on tone perception. Next, the effect of lexical status on tone perception/processing is discussed, followed by the implications of the findings for recent models of cross-language speech perception. In the final section of the chapter, directions for future research are proposed.

5.1. Summary

5.1.1. Tone assimilation

The tone assimilation study was conducted such that native listeners of Mandarin and Thai assimilated non-native tones to native tone categories. Participants in the four language groups (IM, EM, IT and ET) differed in their L1 and L2 experience. The study revealed that the inexperienced listeners relied solely on phonetic similarities, while the experienced listeners relied not only on phonetic similarities, but also on phonological correspondences when mapping non-native tones to native tone categories.

A set of acoustic measures based on F0 height and contour were computed for each Mandarin and Thai tone. On the basis of these measures, the nine tones (four in Mandarin and five in Thai) were classified into four types: level, rising, falling and falling rising tone. Each Mandarin tone has a Thai equivalent, except for the high falling tone, which is similar to both the Thai high falling and low falling tones (see Table 2.3).

According to modal responses, Mandarin and Thai tones of the same tone type were perceived as most similar by the IT and IM listeners, indicating the influence of phonetic similarities. In addition, the assimilation patterns indicate that non-native tones were assimilated to native tone categories with similar F0 contours, suggesting that F0 contours are more important than F0 height in tone assimilation as a perceptual cue. The effect of phonetic equivalence was also observed for the experienced listeners, in particular between the level tones and the falling tones.

Another significant finding regarding tone assimilation is that the experienced listeners assimilated non-native tones to native tone categories on the basis of their phonological correspondence rather than their phonetic similarities. The assimilation between rising and falling rising tones, as well as between low falling and falling rising tones, was probably due to phonological tone changes relating to Tone 3 sandhi, rather than to phonetic relationships between the tones. According to Tone 3 sandhi rules, rising tone and low falling tone can replace the falling rising tone (Tone 3) at the word level in Mandarin.

Degree of diversity in tone assimilation is also an important measure in addition to modal response. According to degree of diversity, the Thai groups showed higher diversity (or lower similarity) relative to the Mandarin groups (Thai: 2.1 vs. Mandarin: 1.3), in particular the IT listeners (2.3). Level and falling tones were associated with lower diversity relative to other tones, exhibiting higher perceptual similarities to their respective phonetic or phonological counterparts.

5.1.2. Tone identification

The findings of the identification study indicate important influences of L1 and L2 tone experience on the identification of Mandarin tones at the phonetic and phonological levels. This study tested listeners with five different levels of experience with tone. Listeners who were not familiar having the target Mandarin tones participated in a short tone training session before the real test. To increase error rates, target tones were presented in a two-tone sequence in which the target tone was followed by a Mandarin falling tone.

Overall performance revealed the significant effects of language group, tone, and word type, suggesting better performance for native listeners than for non-native listeners, and better performance for experienced listeners relative to inexperienced listeners. Results also showed higher accuracy in the identification of Tone 1 and Tone 4 than of Tone 2 and Tone 3; and higher accuracy for real words than nonsense words.

Listeners' dependence on acoustic cues (F0 height and contour) was consistent with previous observations (Gandour, 1983; Guion & Pederson, 2007; Huang & Johnson, 2010). For example, Thai listeners showed more confusion between tones with similar F0 contours than similar F0 heights, while English listeners demonstrated the opposite performance. However, the fact that Tone 2 and Tone 3 were the most confusable pair, especially for the NM, ET, IT and EE listeners, may indicate the influence of both the phonetic and phonological factors on tone identification: Tone 2 and Tone 3 are not only phonetically similar, but also phonologically interchangeable.

Of the two experienced groups, the ET listeners showed more native-like performance on the identification of Mandarin tones than the EE listeners. The same difference, however, was not observed for the inexperienced listeners in the IT and IE groups. Therefore, it is likely that L1 tone experience may help listeners in the long-term acquisition of new tone categories, such that they show better identification of these L2 tones once they have gained L2 tone experience (c.f., Wayland & Guion, 2004).

The semantic information in stimulus words was found to be beneficial only in a few cases. Those who recognized the real words, i.e., the Mandarin and English listeners, showed higher accuracy for the real words than the nonsense words in the identification of difficult tones: Tone 2 for the Mandarin listeners and Tone 1 for the English listeners. These results suggest that lexical information can help listeners to perceive difficult tones and that non-tone language speakers may associate lexical tones with reference to the intonational patterns in L1 (Ladd, 2008).

5.1.3. Hemispheric tone processing

The third study used a dichotic listening paradigm to test the same listeners from the second study for their recognition of two different tones presented simultaneously, with one tone to the left ear and the other to the right ear. The overall results showed no hemispheric dominance in the processing of Mandarin tones by Mandarin and non-Mandarin listeners. Nevertheless, the findings of this study provide evidence supporting the two pitch processing hypotheses: the acoustic hypothesis and the functional hypothesis.

In support of acoustic hypothesis, the individual tone patterns were found to be processed differently in the left and right ears across listeners or word types: Tone 1 and Tone 2 were better identified in the left ear, while Tone 3 and Tone 4 were better perceived in the right ear. In addition, monaural performance in the left and right ears revealed group difference such that the native listeners outperformed the non-native listeners. In the right ear, however, the ET (but not the EE) listeners' performance was found to be no different from that of the Mandarin listeners, suggesting an interplay of L1 and L2 tone experience in the processing of L2 tone patterns.

5.2. Phonetic and phonological tone perception

Taken together, the three studies reveal phonetic and phonological influences on tone perception in different ways. The phonetic effect is seen among all listeners irrespective of their language experience, and indicates listeners' dependence on acoustic correlates (F0 height or contour) in processing tone. In contrast, the phonological effect is closely related to listeners' language experience. Only the native and experienced listeners are sensitive to the phonological use of tone categories and only the experienced listeners are sensitive to the phonological correspondence between L1 and L2 tone categories.

5.2.1. Phonetic perception of tone

In the tone assimilation study of the current project, the inexperienced listeners were found to assimilate the most phonetically similar tones to native tone categories. Likewise, the experienced listeners also showed sensitivity to phonetic similarities between L1 and L2 tones. The difference between the two groups is probably that the inexperienced listeners' performance should be interpreted in relation to an acoustic mode in which the linguistic use of tone is less influential (Bent, 2005), while the experienced listeners' performance should be discussed with respect to a linguistic mode related to the phonemic use of tone (Best & Tyler, 2007). Accordingly, for the inexperienced listeners, the assimilated non-native tone categories were perceived as phonetic equivalents, while for the experienced listeners, these tones were associated as phonological equivalents.

The tone identification study indicated average identification accuracy at above 60%, even for the IE listeners, showing a relatively high level of sensitivity to the phonetic distinctions between tone categories. Furthermore, tone confusion patterns indicated listeners' dependence on individual acoustic cues stemming from their native language experience. For example, the high confusion rate between Mandarin Tone 2 and Tone 3 for the IT listeners was possibly due to their similar F0 contour, while the high confusion rate between Tone 1 and Tone 4 for the IE listeners likely arose from the high F0 in these two tones (Blitcher et al., 1990; Guion & Pederson, 2007).

The dichotic listening study supports the acoustic hypothesis, indicating the effects of acoustic properties, in particular the realization of F0 height and contour, on the hemispheric processing of individual tone patterns. Both native and non-native listeners showed consistent lateralization patterns for the processing of the four Mandarin tones: Tone 1 and Tone 2 were lateralized to the right hemisphere, while Tone 3 and Tone 4 were lateralized to the left hemisphere. This finding is consistent with those reported in previous studies, in that left hemisphere dominance is associated with tones with more dynamic tone contours, such as the falling rising Tone 3 and falling Tone 4, while right hemisphere dominance is related to level tones, such as the high level Tone 1 (Gandour, 2006; Ho, 2010; Wang et al., 2004). Lateralization patterns in tone processing may also be affected by the average level of frequency. For example, a higher frequency, such as that in the high level Tone 1 and rising Tone 2, tends to induce greater activation in the right hemisphere (Zatorre & Belin, 2001).

5.2.2. Phonological perception of tone

Unlike many other prosodic patterns, tone is characterized by its phonological function in distinguishing word meanings. As phonological function is experience-dependent, it has been found that only native or experienced listeners are sensitive to the phonological use of tone (Huang, 2001; Leung, 2008). In support of this, tone assimilation patterns in this study were found to associate tones with their phonological counterparts only for the experienced listeners of Mandarin and Thai. In Mandarin, allophonic tone variantions, such as those between rising tone and falling rising tone, or between low falling tone and falling rising tone, are more difficult to distinguish for they represent the same phoneme (Huang, 2001; Hume & Johnson, 2003). As a result, falling rising tone was assimilated to low falling tone, or rising tone was assimilated to falling rising tone only by the experienced listeners of Thai and Mandarin. While the inexperienced Mandarin listeners were sensitive to the phonetic similarities between Mandarin and Thai tones, and they were also familiar with Tone 3 sandhi in Mandarin, they did not associate any of the Thai tones as Mandarin tone categories at the phonological level.

The tone identification study indicates the influence of the phonological use of tone in two respects. First, the native Mandarin listeners showed greater confusion stemming from the allophonic contrast between Tone 2 and Tone 3 by identifying Tone 2 as Tone 3. This result supports the proposal that Tone 2 and Tone 3 are difficult to distinguish due to their phonological similarity (Hume & Johnson, 2003). This finding also provides empirical confirmation that Tone 2 tends to be misperceived as Tone 3 due to the allophonic connection. Second, the ET listeners showed high accuracy comparable to that of the native Mandarin listeners in the identification of most of the Mandarin tones (i.e., Tone 1, Tone 2 and Tone 4). In contrast to the EE and the IT listeners, the ET listeners' performance indicated more categorical perception of Mandarin tones, suggesting that some Mandarin tones might have been acquired as phonological categories by the ET listeners but not by the EE listeners (Wang, 1976).

Hemispheric lateralization patterns are also associated with the phonological processing of tone. Previous findings have indicated a higher degree of left hemisphere involvement for pitch patterns with a higher linguistic functional load relative to those with a lower functional load (Gandour et al., 2002, 2003; Wu et al., 2011) and also for native listeners relative to non-native listeners, or experienced listeners relative to inexperienced listeners (Wang et al., 2001, 2004). In the current dichotic listening study, the overall performance indicates bilateral processing of Mandarin tones for all the listeners, regardless of differences in their language experience. Nevertheless, the native and the experienced listeners exhibited greater left hemisphere involvement than the non-native and the inexperienced listeners, respectively. Of the two experienced groups, the ET listeners who had both L1 and L2 tone experience showed comparable performance to the native Mandarin listeners in the right ear. The results of the dichotic listening study, taken together with the results of the identification study described above, suggests that the ET listeners acquired some Mandarin tones within a phonological system.

5.3. The effects of lexical status

The current research also investigated the extent to which tone perception is affected by lexical status to determine whether tone perception/processing is confounded with word recognition/lexical processing (Cutler & Chen, 1997; Cutler & Otake, 1999; Wong, 2002). Accordingly, the stimulus words were divided into real words and nonsense words. The identification study indicates an influence of word type, showing better performance for real words. However, further analysis reveals that the influence was evident for the Mandarin listeners only in the identification of Tone 2 and for the English listeners only in the identification of Tone 1. Both tones were the most difficult to identify for the listeners. Thus, it is likely that listeners' attention is strictly focused on lexical information during tone perception. However, listeners can still detect the difference between word types and use the lexical information to help recognize difficult tones (c.f., Cutler & Chen, 1997; Lee et al., 1996).

The results of the dichotic listening study echo the findings of Hsieh et al. (2001); the hemispheric processing of the Mandarin and English listeners was no different for real words and nonsense words. This finding supports the notion that tone processing is pre-lexical, independent of the lexical status of the stimulus words (Hsieh et al., 2001; Xu et al., 2006). However, the results of the current dichotic study conflict with some previous research on the effect of semantic information on hemispheric specialization for tone processing. Table 5.1 compares five studies (including the current one) in terms of relevant methodological factors (research paradigm and stimulus words) and research results. As shown, the prior findings are consistent with the current study in showing no effect of lexical information on tone processing. However, the results are inconsistent in terms of the connection between tone processing and stimulus type (real words, nonsense words and non-speech stimuli). For example, tone in a non-speech condition is associated with two types of lateralization patterns: no hemisphere dominance (Baudoin-Chial, 1986; Gandour et al., 2003) and left hemisphere dominance (Hsieh et al., 2001). With nonsense words, tones may be processed predominantly in the left hemisphere (Gandour et al., 2003) or equally in both hemispheres (Xu et al., 2006; the current study). Future research using the same research paradigm and all three types of stimulus word is necessary to investigate hemispheric lateralization for tone processing. Such research will help not only to determine the influence of word type on tone processing, but also to determine whether the inconsistent findings of recent studies are the result of different research paradigms.

Table 5.1. Comparison of tone processing research for the relationship of stimulus word type and lateralization pattern (NHD = no hemisphere dominance; LHD = left hemisphere dominance; RHD = right hemisphere dominance)

	Research paradigm	Stimulus word	Lateralization pattern
Baudoin-Chial (1986)	dichotic	Real words	NHD
		Hums	NHD
Gandour et al. (2003)	fMRI	Nonsense words	LHD
		Hums	NHD
Hsieh et al. (2001)	PET	Real words	LHD
		Non-speech	LHD
Xu et al. (2006)	fMRI	Real words (Mandarin)	NHD
		Nonsense words (Thai)	NHD
The current study	dichotic	Real words	NHD
		Nonsense words	NHD

5.4. Implications for recent models of cross-language perception

Previous studies have employed two recent models of cross-language perception, i.e., the PAM (including the PAM-L2) and the SLM, to account for tone perception results (e.g., Bent, 2005; Francis et al., 2008; Hallé et al., 2004; So & Best, 2010). The current project builds on the foundation of these previous studies on the assimilation and identification of tone categories by employing listeners from more diverse language backgrounds.

In agreement with the PAM, the current findings indicate that tone assimilation by inexperienced listeners occurs only at the phonetic level. In terms of the PAM-L2, the current findings indicate that for experienced listeners, tone assimilation is affected by both phonetic properties and phonological functions. The current findings also support and expand the SLM by indicating that learning experience improves tone perception and that listeners who have both L1 and L2 tone experience establish L2 tone categories earlier than those without this background.

The current findings suggest that the PAM may be applied to the suprasegmental level to account for tone perception by inexperienced listeners. In the current assimilation study, the inexperienced listeners showed high sensitivity to the phonetic similarities between Mandarin and Thai tones, and assimilated L2 tones to their phonetic counterparts in L1, reflecting the assimilation patterns for "categorized" contrasts. Additionally, the IE listeners showed more variable identification of the four "uncategorizable" Mandarin tones relative to the listeners in the other groups, echoing the prediction that discrimination of "uncategorizable" phones ranges from poor to moderate (Hallé et al., 2004). However, the PAM may not be predictive in some conditions. For example, the model does not predict IT listeners' difficulty in discriminating Tone 2 and Tone 3. Even though these two tones correspond to the "Two Category" assimilation pattern, which is associated with very easy discrimination, the IT listeners showed high confusion between Tone 2 and Tone 3 instead. Therefore, when using the PAM to explain cross-language tone perception, some factors, such as listeners' choice of perceptual cues, need to be considered.

The current findings also support the PAM-L2, in that experienced listeners perceive tones at both the phonetic and phonological levels. The assimilation study in the current project is probably the first study to provide empirical support for processing tones at these two levels for experienced listeners. Tone 3 sandhi in Mandarin also provides an effective platform to demonstrate the effects of phonological tone changes on tone perception. Furthermore, the inexperienced listeners' performance supports the PAM-L2 in showing that phonological tone perception occurs only among experienced listeners. Unlike the EM listeners, the IM listeners assimilated Thai tones to Mandarin tone categories without being affected by Tone 3 sandhi, even though it exists in their native tone system. Additionally, in agreement with the PAM-L2, the experienced listeners also showed high sensitivity to phonetic similarities, indicating tone assimilation at both the phonetic and phonological levels. Nevertheless, the current findings also echo the PAM-L2 in showing that phonetic similarity is less important than phonological function in the process of phonological assimilation (Best & Tyler, 2007). For example, due to the influence of Tone 3 sandhi, the ET listeners assimilated Mandarin rising tone to Thai native falling rising tone, and Mandarin falling rising tone to Thai low falling tone.

The SLM concerns speech perception by listeners with varying L2 experience. The current findings support the SLM in showing that L2 learning experience improves significantly the perception of L2 phonetic categories. Moreover, in agreement with the SLM, the current results also indicate that category assimilation occurs at the phonetic level for experienced listeners. In addition, in support of the SLM, the current findings indicate tone assimilation patterns for L1 and L2 tone categories that are phonetically identical, similar, and dissimilar (new). The current findings may contribute to the further development of the SLM in showing that category assimilation occurs not only at the phonetic level, but also at the phonological level. The combined findings of the identification and dichotic listening studies suggest that some L2 tone categories may be established earlier than other tone categories. Therefore, it is likely that category assimilation occurs between L1 and L2 phonological categories that have the same phonological status in L1 and L2.

5.5. Future research

The current project investigates native and non-native listeners' sensitivity to the acoustic properties and linguistic function of lexical tones through three tone perception studies on tone assimilation, identification, and hemispheric processing, respectively. In future research, the current findings on tone perception need to be compared with tone production patterns, given that the two may be correlated (Xu et al., 2011; Wang et al., 2003). Tone production has been shown to be experience-specific. For instance, non-tone language speakers tend to produce smaller F0 ranges in tone production relative to tone language speakers (e.g., Ding, Jokisch, Hoffmann, 2010). However, previous studies have not yet explored tone production by speakers from diverse language backgrounds. As the current project indicates differences in tone perception relating to

L1 and L2 experience, it is expected that speakers with the same L1 and L2 experience as the listeners in the current study will exhibit experience-specific production performance.

Future research on tone perception should also systematically vary methodological factors, such as participants, stimuli, and research paradigm. For example, experienced listeners may be classified into early learners (those who began learning the L2 as children) and late learners (those who began learning the L2 as adults) or listeners at different stages of L2 learning. Previous research has found that a number of contextual variables in the presentation of stimuli strongly influence prosodic perception and production (Bent, 2005; Huang & Holt, 2009). Relevant contextual distinctions that may be varied systematically include the length of carrier word/phrase, speech mode, or semantic situation. As an extension of the current findings, future investigations may explore the influence of lexical status and speech mode on the perception and production of other lexical prosodic patterns, such as pitch accent and word stress.

The finding that lexical information can help listeners in the perception of difficult tones may be further explored in phonetic learning research. In the current identification study, lexical information might have helped the Mandarin and English listeners in the perception of the most poorly identified tones. The notion that lexical information helps listeners to distinguish, and speakers to produce, phonetic/phonological features is consistent with some recent findings on pronunciation teaching. In Saito and Lyster (2011), for example, native Japanese speakers received meaning-based instruction to learn the distinctions between the English consonants /ɹ/ and /l/, which are known to be difficult to discriminate and pronounce among Japanese learners. Their findings revealed improved participant performance in the perception and production of the non-native phonetic contrasts. Future research may consider the effects of meaning-based instruction in teaching tone in classroom settings or in training participants in laboratory settings. The findings of such studies would shed light on the potential of meaning-based instruction in teaching pronunciation and listening skills, and would provide further discussion on the significance of learning strategies in applied linguistics and experimental phonetics.

References

Abramson, A. S. (1975). The tones of central Thai: Some perceptual experiments. *Studies in Tai Linguistics in Honor of William J.Gedney. Bangkok: Central Institute of English Language,* 1-16.

Abramson, A. S. (1978). Static and dynamic acoustic cues in distinctive tones. *Language and Speech, 21*(4), 319-325.

Abramson, A. S. (1979). The noncategorical perception of tone categories in Thai. *Frontiers of Speech Communication Research,* 127-134.

Abramson, A. S. (1997). The Thai tonal space. *Southeast Asian Linguistic Studies: In Honour of Vichin Panupong,* 1-10.

Baudoin-Chial, S. (1986). Hemispheric lateralization of modern standard Chinese tone processing. *Journal of Neurolinguistics, 2*(1-2), 189-199.

Belotel-Grenie, A., & Grenie, M. (1997). Types de phonation et tons en Chinois standard. *Cahiers De Linguistique-Asie Orientale, 26*(2), 249-279.

Belotel-Grenié, A., & Grenié, M. (2004). The creaky voice phonation and the organization of Chinese discourse. Paper presented at the *International Symposium on Tonal Aspects of Languages: With Emphasis on Tone Languages, Beijing,*

Bent, T. (2005). *Perception and Production of Non-Native Prosodic Categories,* Doctoral dissertation, Northwestern University.

Bent, T., Bradlow, A. R., & Wright, B. A. (2006). The influence of linguistic experience on the cognitive processing of pitch in speech and nonspeech sounds. *Journal of Experimental Psychology: Human Perception and Performance*, 32(1), 97-103.

Best, C. T. (1995). A direct realist view on cross-language speech perception. In W. Strange (Ed.), *Speech Perception and Linguistic Experience* (pp. 171-204). Timonium, MD: York Press.

Best, C. T., & Avery, R. A. (1999). Left-hemisphere advantage for click consonants is determined by linguistic significance and experience. *Psychological Science*, 10(1), 65-70.

Best, C. T., McRoberts, G. W., & Goodell, E. (2001). Discrimination of non-native consonant contrasts varying in perceptual assimilation to the listener's native phonological system. *The Journal of the Acoustical Society of America, 109*, 775-794.

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

Binder, J. R., Frost, J. A., Hammeke, T. A., Cox, R. W., Rao, S. M., & Prieto, T. (1997). Human brain language areas identified by functional magnetic resonance imaging. *The Journal of Neuroscience*, 17(1), 353-362.

Blicher, D. L., Diehl, R. L., & Cohen, L. B. (1990). Effects of syllable duration on the perception of the Mandarin Tone2/Tone3 distinction: Evidence of auditory enhancement. *Journal of Phonetics,* 18 (1), 37-49.

Bloch, B. (1950). Studies in colloquial Japanese IV phonemics. *Language,* 26(1), 86-125.

Boersma, P., & Weenink, D. (2009). *Praat: Doing Phonetics by Computer (Version 5.1.05) [Computer Program].*

Boemio, A., Fromm, S., Braun, A., & Poeppel, D. (2005). Hierarchical and asymmetric temporal sensitivity in human auditory cortices. *Nature Neuroscience*, 8(3), 389-395.

Boomershine, A., Hall, K. C., Hume, E., & Johnson, K. (2008). The influence of allophony vs. contrast on perception: The case of Spanish and English. *Contrast in Phonology: Perception and Acquisition,* 145-171.

Brådvik, B., Dravins, C., Holtås, S., Rosén, I., Ryding, E., & Ingvar, D. (1991). Disturbances of speech prosody following right hemisphere infarcts. *Acta Neurologica Scandinavica*, 84(2), 114-126.

Brancucci, A., D'Anselmo, A., Martello, F., & Tommasi, L. (2008). Left hemisphere specialization for duration discrimination of musical and speech sounds. *Neuropsychologia, 46*(7), 2013-2019

Brown-Schmidt, S., & Canseco-Gonzalez, E. (2004). Who do you love, your mother or your horse? an event-related brain potential analysis of tone processing in mandarin chinese. *Journal of Psycholinguistic Research*, 33(2), 103-135.

Bundgaard-Nielsen, R. L., Best, C. T., & Tyler, M. D. (2011). Vocabulary size matters: The assimilation of second-language Australian English vowels to first-language Japanese vowel categories. *Applied Psycholinguistics, 32*(1), 51-67.

Burnham, D., Francis, E., Webster, D., Luksaneeyanawin, S., Attapaiboon, C., & Lacerda, F. (1996). Perception of lexical tone across languages: Evidence for a linguistic mode of processing, *The 4th International Conference on Spoken Language Processing.* Philadelphia, PA.

Cebrian, J. (2006). Experience and the use of non-native duration in L2 vowel categorization. *Journal of Phonetics, 34*(3), 372-387.

Chuang, C. K., & Hiki, S. (1972). Acoustical features and perceptual cues of the four tones of standard colloquial Chinese. *The Journal of the Acoustical Society of America,* 52, 146.

Chao, Y. R. (1948). *Mandarin primer: An intensive course in spoken Chinese,* Harvard University Press.

Chen, M. Y. (2000). *Tone sandhi: Patterns across Chinese dialects.* Cambridge Univ Press.

Clark, J. E., Yallop, C., & Fletcher, J. (2007). *An introduction to phonetics and phonology.* Wiley-Blackwell.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences.* Lawrence Erlbaum

Cutler, A., & Chen, H. C. (1997). Lexical tone in Cantonese spoken-word processing. *Attention, Perception, & Psychophysics*, 59(2), 165-179.

Cutler, A., & Clifton, C. (1999). Comprehending spoken language: A blueprint of the listener. *The Neurocognition of Language*, 123-166.

Cutler, A., Dahan, D., & Van Donselaar, W. (1997). Prosody in the comprehension of spoken language: A literature review. *Language and Speech, 40*(2), 141-201.

Dehaene-Lambertz, G. (1997). Electrophysiological correlates of categorical phoneme perception in adults. *Neuroreport,* 8(4), 919.

Dehaene-Lambertz, G., Dupoux, E., & Gout, A. (2000). Electrophysiological correlates of phonological processing: A cross-linguistic study. Journal of Cognitive *Neuroscience*, 12(4), 635-647.

Davison, D. S. (1991). An acoustic study of so-called creaky voice in Tianjin Mandarin. *UCLA Working Papers in Phonetics, 78*, 50-57.

Ding, H., Jokisch, O., & Hoffmann, R. (2010). *Perception and production of mandarin tones by German speakers.* Paper presented at the Speech Prosody 2010-Fifth International Conference.

Duanmu, S. (2007). *The phonology of standard Chinese* Oxford University Press, USA.

Fedorenko, E., Behr, M. K., & Kanwisher, N. (2011). Functional specificity for high-level linguistic processing in the human brain. *Proceedings of the National Academy of Sciences,* 108(39), 16428-16433.

Flege, J. (2007). Language contact in bilingualism: Phonetic system interactions. *Laboratory Phonology, 9,* 353-382.

Flege, J. (1995). Second-language Speech Learning: Theory, Findings, and Problems. In W. Strange (Ed.), *Speech Perception and Linguistic Experience: Issues in Cross-language research* (pp.) 229-273. Timonium, MD: York Press.

Flege, J. (1987). The production of 'new' and 'similar' phones in a foreign language: Evidence for the effect of equivalence classification. *Journal of Phonetics*, 15, 47-65.

Flege, J. E., Munro, M. J., & MacKay, I. R. A. (1995). Effects of age of second-language learning on the production of English consonants. *Speech Communication*, 16(1), 1-26.

Flege, J. E., Schirru, C., & MacKay, I. R. A. (2003). Interaction between the native and second language phonetic subsystems. *Speech Communication, 40*(4), 467-492.

Fox, R. A., Flege, J. E., & Munro, M. J. (1995). The perception of English and Spanish vowels by native English and Spanish listeners: A multidimensional scaling analysis. *The Journal of the Acoustical Society of America*, 97, 2540-2551.

Francis, A.L., Ciocca, V., Ma, L., & Fenn, K. (2008). Linguistic experience and the acquisition of lexical tones. *Journal of Phonetics,* 36, 268-294.

Francis, A.L., Ciocca, V.C., & Ng, B.K.C. (2003). On the (non)categorical perception of lexical tones. *Perception & Psychophysics,* 65(6), 1029-1044.

Galaburda, A. M. (1995). Anatomic basis of cerebral dominance. *Brain Asymmetry* , 51-73.

Gandour, J. (1983). Tone perception in far eastern languages. *Journal of Phonetics, 11*(2), 149-175.

Gandour, J. (2006). Tone: Neurophonetics. *Encyclopedia of Language and Linguistics, 12*, 751-760.

Gandour, J., Wong, D., & Hutchins, G. (1998). Pitch processing in the human brain is influenced by language experience. *Neuroreport*, 9(9), 2115.

Gandour, J., Potisuk, S., Ponglorpisit, S., & Dechongkit, S. (1991). Inter- and intra-speaker variability in fundamental frequency of Thai tones. *Speech Communication, 10*(4), 355-372.

Gandour, J.T., Potisuk, S., Dechongkit, S., & Suvit, P. (1992). Tonal coarticulation in Thai disyllabic utterances: a preliminary study. *Linguistics of the Tibeto-Burman Area,* 15 (1), 93-110.

Gandour, J., Tong, Y., Wong, D., Talavage, T., Dzemidzic, M., Xu, Y., & Lowe, M. (2004). Hemispheric roles in the perception of speech prosody. *NeuroImage, 23*(1), 344-357.

Gandour, J., Wong, D., Hsieh, L., Weinzapfel, B., Lancker, D. V., & Hutchins, G. D. (2000). A crosslinguistic PET study of tone perception. *Journal of Cognitive Neuroscience, 12*(1), 207-222.

Gerratt, B. R., & Kreiman, J. (2001). Toward a taxonomy of nonmodal phonation. *Journal of Phonetics, 29*(4), 365-381.

Geschwind, N. (1979). Specialization of the human brain. *Scientific American,* 241(3),180–199.

Gilichinskaya, Y. D., & Strange, W. (2010). Perceptual assimilation of American English vowels by inexperienced Russian listeners. 128, EL80–EL85.

Goldstein, L., & Fowler, C. A. (2003). Articulatory phonology: A phonology for public language use. *Phonetics and Phonology in Language Comprehension and Production: Differences and Similarities,* 159-207.

Golestani, N., & Zatorre, R. J. (2004). Learning new sounds of speech: Reallocation of neural substrates. *NeuroImage, 21*(2), 494-506.

Grabe, E., Rosner, B. S., García-Albea, J. E., & Zhou, X. (2003). Perception of english intonation by English, Spanish, and Chinese listeners. *Language and Speech*, 46(4), 375-401.

Grosjean, F. (2001). The bilingual's language modes. *One Mind, Two Languages: Bilingual Language Processing,* 1-22.

Guion, S. G., & Pederson, E. (2007). Investigating the role of attention in phonetic learning. O.S. Bohn, & M. Munro (Eds.), *Second language speech learning: The role of language experience in speech perception and production: In honour of James E. Flege, (*pp. 57–77), John Benjamins, Amsterdam.

Hallé, P. A., & Best, C. T. (2007). Dental-to-velar perceptual assimilation: A cross-linguistic study of the perception of dental stop /l/clusters. *The Journal of the Acoustical Society of America, 121*(5 PT1), 2899-2914.

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*(3), 395-421.

Hallé, P. A., Segui, J., Frauenfelder, U., & Meunier, C. (1998). Processing of illegal consonant clusters: A case of perceptual assimilation? *Journal of Experimental Psychology: Human Perception and Performance, 24*(2), 592.

Hillenbrand, J. M., Clark, M. J., & Nearey, T. M. (2001). Effects of consonant environment on vowel formant patterns. *The Journal of the Acoustical Society of America,* 109(2), 748-763.

Halliday, M. A. K., & Greaves, W. S. (2008). *Intonation in the grammar of English.* Equinox Publishing.

Ho, P. K. (2010). An ERP study on the effect of tone features on lexical tone lateralization in Cantonese. Master's thesis. The Chinese University of Hong Kong.

Howie, J. (1976). An acoustic study of Mandarin tones and vowels. *London: CUP,*

Hsieh, L., Gandour, J., Wong, D., & Hutchins, G. D. (2001). Functional heterogeneity of inferior frontal gyrus is shaped by linguistic experience. *Brain and Language, 76*(3), 227-252.

Huang, J., & Holt, L. L. (2009). General perceptual contributions to lexical tone normalization. *The Journal of the Acoustical Society of America,* 125, 3983-3994.

Huang, T. (2001). The interplay of perception and phonology in tone 3 sandhi in Chinese putonghua. *Working papers in linguistics-Ohio state university department of linguistics,* 23-42.

Huang, T., & Johnson, K. (2010). Language specificity in speech perception: Perception of Mandarin tones by native and nonnative listeners. *Phonetica, 67*(4), 243-267.

Hume, E. and Johnson, K. (2003) The Impact of Partial Phonological Contrast on Speech Perception. *15th International Congress of Phonetic Sciences*, 2385-2388.

Humstone, H. J. (1919). Memory Span Tests. *Journal of Clinical Psychology*, 12, 196-200.

International Phonetic Association. (1999). *Handbook of the international phonetic association: A guide to the use of the International Phonetic Alphabet* Cambridge University Press.

Ishi, C. T., Sakakibara, K. I., Ishiguro, H., & Hagita, N. (2008). A method for automatic detection of vocal fry. *Audio, Speech, and Language Processing, IEEE Transactions on, 16*(1), 47-56.

Iverson, P., Kuhl, P. K., Akahane-Yamada, R., Diesch, E., Tohkura, Y., Kettermann, A., & Siebert, C. (2003). A perceptual interference account of acquisition difficulties for non-native phonemes. *Cognition, 87*(1), B47-B57.

Ivry, R. B., & Lebby, P. C. (1993). Hemispheric differences in auditory perception are similar to those found in visual perception. *Psychological Science,* 4(1), 41.

Ivry, R. B., & Robertson, L. C. (1998). *The two sides of perception.* The MIT Press.

Jacquemot, C., Pallier, C., LeBihan, D., Dehaene, S., & Dupoux, E. (2003). Phonological grammar shapes the auditory cortex: A functional magnetic resonance imaging study. *The Journal of Neuroscience*, 23(29), 9541.

Kaan, E., Wayland, R., Bao, M., & Barkley, C. M. (2007). Effects of native language and training on lexical tone perception: An event-related potential study. *Brain Research, 1148*, 113-122.

Kazanina, N., Phillips, C., & Idsardi, W. (2006). The influence of meaning on the perception of speech sounds. *Proceedings of the National Academy of Sciences*, 103(30), 11381.

Ke, C. (1992). Dichotic listening with Chinese and English tasks. *Journal of Psycholinguistic Research, 21*(6), 463-471.

Keating, P., & Kuo, G. (2010). Comparison of speaking fundamental frequency in English and Mandarin. *UCLA Working Papers in Phonetics*, 108, 164-187

Khouw, E., & Ciocca, V. (2007). Perceptual correlates of Cantonese tones. *Journal of Phonetics*, 35, 104-117.

Kimura, D. (1961). Cerebral dominance and the perception of verbal stimuli. *Canadian Journal of Psychology*, 15(3), 166.

Kimura, D. (1964). Left-right differences in the perception of melodies. *Quarterly Journal of Experimental Psychology*, 16(4), 355-358.

King, R. D. (1967). Functional load and sound change. *Language,* 831-852.

Kiriloff, C. (1969). On the auditory perception of tones in mandarin. *Phonetica*, 20(2-4), 63-67.

Kondaurova, M. V., & Francis, A. L. (2010). The role of selective attention in the acquisition of English tense and lax vowels by native Spanish listeners: Comparison of three training methods. *Journal of Phonetics, 38*(4), 569-587.

Ladd, D. R. (2008). *Intonational phonology.* Cambridge University Press.

Ladd, D. R., Silverman, K. E. A., Tolkmitt, F., Bergmann, G., & Scherer, K. R. (1985). Evidence for the independent function of intonation contour type, voice quality, and F0 range in signaling speaker affect. *Journal of the Acoustical Society of America, 78*(2), 435-444.

Ladefoged, P. (2001). *A course in phonetics.* Orlando: Harcourt College Publishers.

Lee, Y. S., Vakoch, D. A., & Wurm, L. H. (1996). Tone perception in Cantonese and Mandarin: A cross-linguistic comparison. *Journal of Psycholinguistic Research, 25*(5), 527-542.

Leung, A. S. (2008). Tonal assimilation patterns of Cantonese L2 speakers of Mandarin in the perception and production of Mandarin tones. Paper presented at the *Proceedings of the 2008 CLA Annual Conference.*

Levelt, W. J. M. (1999). Producing spoken language: A blueprint of the speaker. The *Neurocognition of Language*, 83-122.

Levy, E. S. (2009). Language experience and consonantal context effects on perceptual assimilation of French vowels by American English learners of French. *The Journal of the Acoustical Society of America, 125*, 1138-1152.

Li, C., & Thompson, S. A. (1977). The acquisition of tone in Mandarin-speaking children. *Journal of Child Language,* 4,185-99.

Liang, J., & Van Heuven, V. J. (2004). Evidence for separate tonal and segmental tiers in the lexical specification of words: A case study of a brain-damaged chinese speaker. *Brain and Language*, 91(3), 282-293.

Lieberman, P. (1963). Some effects of semantic and grammatical context on the production and perception of speech*. Language and Speech*, 6(3), 172-187.

Lin, H. (1996). *Mandarin tonology.* Taipei, Taiwan: Pyramid Press.

Lin, M.C. (1988). 普通话声调的声学性质与知觉征兆, *《中国语文》*, 204 (3), 182.

Lin, H. B., & Repp, B. H. (1989). Cues to the perception of Taiwanese tones. *Language and Speech,* 32(1), 25-44.

MacKay, I. R. A., Meador, D., & Flege, J. E. (2001). The identification of English consonants by native speakers of Italian. *Phonetica*, 58(1-2), 103-125.

Marian, V., Blumenfeld, H. K., & Boukrina, O. V. (2008). Sensitivity to phonological similarity within and across languages. *Journal of Psycholinguistic Research, 37*(3), 141-170.

Maye, J., Werker, J. F., & Gerken, L. A. (2002). Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition,* 82(3), B101-B111.

Millay, K. K., Roeser, R. J., & Godfrey, J. J. (1977). Reliability of performance for dichotic listening using two response modes. *Journal of Speech & Hearing Research*, 510-518.

McCabe, D. P. (2008). The role of covert retrieval in working memory span tasks: Evidence from delayed recall tests. *Journal of Memory and Language*, 58(2), 480-494.

Moore, C. B., & Jongman, A. (1997). Speaker normalization in the perception of mandarin Chinese tones. *The Journal of the Acoustical Society of America*, 102, 1864-1877.

Nishi, K., Strange, W., Akahane-Yamada, R., Kubo, R., & Trent-Brown, S. A. (2008). Acoustic and perceptual similarity of Japanese and American English vowels. *The Journal of the Acoustical Society of America, 124*, 576-588.

Peabody, M., & Seneff, S. (2009). Annotation and features of non-native Mandarin tone quality. Paper presented at the *Tenth Annual Conference of the International Speech Communication Association,* 480-483.

Pegg, J. E., & Werker, J. F. (1997). Adult and infant perception of two English phones. *Journal of the Acoustical Society of America*, 102(6), 3742-3753.

Patel, A. D., & Balaban, E. (2001). Human pitch perception is reflected in the timing of stimulus-related cortical activity. *Nature Neuroscience*, 4(8), 839-844.

Penhune, V. B., Zatorre, R. J., & Evans, A. C. (1998). Cerebellar contributions to motor timing: A PET study of auditory and visual rhythm reproduction. Journal of Cognitive *Neuroscience*, 10(6), 752-765.

Peperkamp, S., Pettinato, M., & Dupoux, E. (2003). Allophonic variation and the acquisition of phoneme categories. In B. Beachley, A. Brown, & F. Conlin (Eds.), *Proceedings of the 27th Annual Boston University Conference on Language Development* (Vol. Volume 2, 650-661). Sommerville, MA: Cascadilla Press.

Pierrehumbert, J., & Hirschberg, J. (1990) The meaning of intonational contours in the interpretation of discourse. In P. Cohen, J. Morgan, and M. Pollack (Eds.), Intentions in Communication. MIT Press, Cambridge, MA.

Poeppel, D. (2003). The analysis of speech in different temporal integration windows: Cerebral lateralization as 'asymmetric sampling in time'. *Speech Communication*, 41(1), 245-255.

Poeppel, D., Guillemin, A., Thompson, J., Fritz, J., Bavelier, D., & Braun, A. R. (2004). Auditory lexical decision, categorical perception, and FM direction discrimination differentially engage left and right auditory cortex. *Neuropsychologia*, 42(2), 183-200.

Repp, B. H., & Lin, H. B. (1989). Acoustic properties and perception of stop consonant release transients. *Journal of the Acoustical Society of America, 85*(1), 379-396.

Rose, P. (1987). Considerations in the normalisation of the fundamental frequency of linguistic tone. *Speech Communication, 6*(4), 343-352.

Saito, K., & Lyster, R. (2011). Effects of Form-Focused instruction and corrective feedback on L2 pronunciation development of/ɹ/by Japanese learners of English. *Language Learning,* EarlyView, 1-39.

Schneider, W., Eschmann, A., & Zuccolotto, A. (2002). *E-Prime user's guide*. Pittsburgh, PA: Psychology Software Tools, Inc..

Sergent, J., Ohta, S., & MacDonald, B. (1992). Functional neuroanatomy of face and object processing. *Brain*, 115(1), 15-36.

Sereno, J. A., & Wang, Y. (2007). Behavioral and cortical effects of learning a second language. O.S. Bohn, & M. Munro (Eds.), *Second language speech learning: The role of language experience in speech perception and production: In honour of James E. Flege*, (pp. 239-258.), John Benjamins, Amsterdam.

Shen, X. S. (1989). Toward a register approach in teaching Mandarin tones. *Journal of the Chinese Language Teachers Association,* 24(3), 27-47.

Shen, X. S., & Lin, M. (1991). A perceptual study of Mandarin tones 2 and 3. *Language and Speech, 34*(2), 145-156.

Shipley-Brown, F., Dingwall, W. O., Berlin, C. I., & Yeni-KomshianSandra, G. (1988). Hemispheric processing of affective and linguistic intonation contours in normal subjects. *Brain and Language*, 33(1), 16-26.

Simpson, E. H. (1949). Measurement of diversity. *Nature, 163*(4148), 688-688.

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

Spinelli, E., Segui, J., & Radeau, M. (2001). Phonological priming in spoken word recognition with bisyllabic targets. *Language and Cognitive Processes, 16*(4), 367-392.

SPSS, S. (2001). For windows. SPSS Inc., Chicago, IL.

Stagray, J. R., & 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. (1995). Cross-language studies of speech perception: A historical review. In W. Strange (Eds.), *Speech perception and linguistic experience* (pp. 3–45). Baltimore: York Press.

Strange, W., Akahane-Yamada, R., Kubo, R., Trent, S. A., & Nishi, K. (2001). Effects of consonantal context on perceptual assimilation of American English vowels by Japanese listeners. *The Journal of the Acoustical Society of America, 109*, 1691-1704.

Strange, W., Bohn, O. S., Nishi, K., & Trent, S. A. (2005). Contextual variation in the acoustic and perceptual similarity of north German and American English vowels. *The Journal of the Acoustical Society of America, 118*, 1751-1762.

Strange, W., Bohn, O. S., Trent, S. A., & Nishi, K. (2004). Acoustic and perceptual similarity of north German and American English vowels. *The Journal of the Acoustical Society of America, 115*, 1791-1807.

Surendran, D., & Niyogi, P. (2006). Quantifying the functional load of phonemic oppositions, distinctive features, and suprasegmentals. *Amsterdam studies in the theory and history of linguistic science series 4, 279*, 43.

Tartter, V. (1988). Acoustic and phonetic feature effects in dichotic listening. In K. Hugdahl (ed.), *Handbook of dichotic listening: Theory, methods and research* (pp. 283–321). Chichester, England: Wiley.

Techentin, C., Voyer, D., & Klein, R. M. (2009). Between-and within-ear congruency and laterality effects in an auditory semantic/emotional prosody conflict task. *Brain and Cognition,* 70(2), 201-208.

Tong, Y., Francis, A., & Gandour, J. (2007). Processing dependencies between segmental and suprasegmental features in Mandarin Chinese. *Language & Cognitive Processes*, 23, 689-708.

Tseng, C. (1981). *An acoustic phonetic study on Mandarin Chinese.* Doctoral dissertation, Brown University.

Van Lancker, D. (1980). Cerebral lateralization of pitch cues in the linguistic signal. *Research on Language & Social Interaction, 13*(2), 201-277.

Van Lancker, D., & Fromkin, V. (1978). Cerebral dominance for pitch contrasts in tone language speakers and in musically untrained and trained English speakers. *Journal of Phonetics,* 6(1), 19-23.

Vance, T. J. (1976). An experimental investigation of tone and intonation in Cantonese. *Phonetica, 33*(5), 368-392.

Voyer, D., & Techentin, C. (2009). Dichotic listening with consonant-vowel pairs: The role of place of articulation and stimulus dominance. *Journal of Phonetics,* 37(2), 162-172.

Wang, W. S. Y. (1976). Language change. *Annals of the New York Academy of Sciences,* 280(1), 61-72.

Wang, Y., Behne, D. M., Jongman, A., & Sereno, J. A. (2004). The role of linguistic experience in the hemispheric processing of lexical tone. *Applied Psycholinguistics, 25*(3), 449-466.

Wang, Y., Jongman, A., & Sereno, J. A. (2001). Dichotic perception of Mandarin tones by Chinese and American listeners. *Brain and Language, 78*(3), 332-348.

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

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, M. M., Jongman, A., & Sereno, J. A. (1999). Training American listeners to perceive Mandarin tones. *The Journal of the Acoustical Society of America, 106*, 3649-3658.

Warrington, E. K., & Taylor, A. M. (1978). Two categorical stages of object recognition. *Perception,* 7(6), 695-705.

Wayland, R., & Guion, S. (2003). Perceptual discrimination of Thai tones by naïve and experienced learners of Thai. *Applied Psycholinguistics, 24*(1), 113-130.

Wayland, R. P., & Guion, S. G. (2004). Training English and Chinese listeners to perceive Thai tones: A preliminary report. *Language Learning, 54*(4), 681-712.

Wells, J. C. (2006). *English intonation: An introduction.* Cambridge University Press.

Whalen, D. H., Best, C. T., & Irwin, J. R. (1997). Lexical effects in the perception and production of American English /p/ allophones. *Journal of Phonetics*, 25, 501-528

Wilkinson, L. (1999). Statistical methods in psychology journals: Guidelines and explanations. *American Psychologist*, 54(8), 594.

Wong, P. (2002). Hemispheric specialization of linguistic pitch patterns. *Brain Research Bulletin, 59*(2), 83-95.

Wong, P. C. M., Skoe, E., Russo, N. M., Dees, T., & Kraus, N. (2007). Musical experience shapes human brainstem encoding of linguistic pitch patterns. *Nature Neuroscience*, 10(4), 420-422.

Wu, X., & Lin, H. (2008). Perception of Mandarin tones by Mandarin and English listeners. *Journal of Chinese Language and Computing*, 18(4), 175-187.

Wu, X., Tu, J. Y., & Wang, Y. (2011). Native and nonnative processing of Japanese pitch accent. *Applied Psycholinguistics, 1*(1), 1-19.

Xu, Y. (1997). Contextual tonal variations in Mandarin. *Journal of Phonetics, 25*, 61-84.

Xu, Y., Gandour, J., Talavage, T., Wong, D., Dzemidzic, M., Tong, Y., & Lowe, M. (2006). Activation of the left planum temporale in pitch processing is shaped by language experience. *Human Brain Mapping, 27*(2), 173-183.

Xu, Y., Gandour, J. T., & Francis, A. L. (2006). Effects of language experience and stimulus complexity on the categorical perception of pitch direction. *The Journal of the Acoustical Society of America, 120*, 1063-1074.

Yu, A. C. L. (2007). Understanding near mergers: The case of morphological tone in Cantonese. *Phonology, 24*(01), 187-214.

Yu, K. M. (2010). Laryngealization and features for Chinese tonal recognition. Paper presented at the *Eleventh Annual Conference of the International Speech Communication Association.*

Zatorre, R.J. & Belin, P. (2001) Spectral and temporal processing in human auditory cortex. *Cerebral Cortex,* 11, 946-953.

Zatorre, R. J., Evans, A. C., & Meyer, E. (1994). Neural mechanisms underlying melodic perception and memory for pitch. *The Journal of Neuroscience*, 14(4), 1908-1919.

Zatorre, R. J., & Gandour, J. T. (2008). Neural specializations for speech and pitch: Moving beyond the dichotomies. *Philosophical Transactions of the Royal Society B: Biological Sciences,* 363(1493), 1087-1104.

Zhang, L., Shu, H., Zhou, F., Wang, X., & Li, P. (2010). Common and distinct neural substrates for the perception of speech rhythm and intonation. *Human Brain Mapping,* 31(7), 1106-1116.

Zhao, J., Shu, H., Zhang, L., Wang, X., Gong, Q., & Li, P. (2008). Cortical competition during language discrimination. *NeuroImage*, 43(3), 624-633.

Zheng, H. Y., Minett, J. W., Peng, G., & Wang, W. S. Y. (2010). The impact of tone systems on the categorical perception of lexical tones: An event-related potentials study, Language and cognitive processes, 1-31.

Zsiga, E., & Nitisaroj, R. (2007). Tone features, tone perception, and peak alignment in Thai. *Language and Speech, 50*(3), 343-383.

Zue, V. W. (1976). Some perceptual experiments on the Mandarin tones. *The Journal of the Acoustical Society of America*, 60, S45.

Appendices

Appendix A. Instructions for listeners in tone assimilation experiment (in listeners' native language)

1. General introduction:

In this part, you will listen to syllables presented singly. Please pay attention to the tone only.

First you will indicate which Mandarin tone sounds closest to the tone you heard by pressing the button on the keyboard. And then rate the similarity between the Mandarin tone and the tone you heard. If the two tones sound identical, press 5. If they sound least similar, press 1. You can also press 2, 3, or 4. The larger the number is, the more similar the two tones are. There is no time limit for your response, but please respond as quickly as possible. Press any key to start.

2. Practice session:

Let's have a practice first. Press any key to start.

3. Experiment:

This part is the actual experiment. You will repeat what you did in the practice session. The experiment has two blocks with a break between each block. Press any key to start.

Appendix B. Instructions for listeners in tone identification experiment (in listeners' native language)

(For inexperienced listeners)

1. General introduction:

In this session, you will listen to the four tones in Mandarin: Tone 1 (T1), Tone 2 (T2), Tone 3 (T3) and Tone 4 (T4). Please indicate which tone you heard by pressing the button on the keyboard. The four Mandarin tones are imposed on different syllables. Your response time is **2 seconds**. Press any key to start.

2. Familiarization session:

(1) Let's have a practice first. When your answer is incorrect, the correct answer will appear. Please respond in 2 seconds. Press any key to start.

(2) In next two blocks, you will hear two tones presented in a row, but you only need to identify **the first one**. When your answer is incorrect, the correct answer will appear. Please respond in 2 seconds. Press any key to start.

3. Experiment:

This part is the actual experiment. Please indicate the first tone you heard by pressing the button on the keyboard. You will respond in 2 seconds. Press any key to start.

(For native and experienced listeners)

1. General introduction:

In this session, you will listen to the four tones in Mandarin: Tone 1 (T1), Tone 2 (T2), Tone 3 (T3) and Tone 4 (T4).

You will hear two tones presented in a row, but you only need to identify **the first one** by pressing the button on the keyboard. The four Mandarin tones are imposed on different syllables. Your response time is **2 seconds**. Press any key to start.

2. Practice session:

Let's have a practice first. Press any key to start.

3. Experiment:

This part is the actual experiment. Please indicate the first tone you heard by pressing the button on the keyboard. You will respond in 2 seconds. Press any key to start.

Appendix C. Instructions for listeners in dichotic listening experiment (in listeners' native language)

1. General introduction:

In this part, you will hear two different Mandarin tones in your left and right ears at the same time. These two tones are followed by another tone. Please only indicate the first tone you heard in your left ear and ears. You will have **3 seconds** for each response. Press any key to start.

2. Practice session (response order: first left then right):

Next you will practice this experiment. You will first indicate the first tone you heard in your **LEFT** ear and then the first tone in your RIGHT ear.

Please respond within three seconds for each response. If your answer is incorrect, the correct answer will appear. Press any key to start.

3. Experiment:

This part is the actual experiment. Please select the first tone you heard in each ear by pressing the button on the keyboard. You will first indicate the first tone you heard in your **LEFT** ear and then the first tone in your RIGHT ear. The maximum time for each response is 3 seconds. Press any key to start.

Appendix D. Table of F0 correlates (in T value) for the Mandarin and Thai tones

(1) Level tones

Group	Tone	Speaker	Peak	Valley	Range	Pt1 (0%)	Pt2 (25%)	Pt3 (50%)	Pt4 (75%)	Pt5 (100%)	MeanF0
Mandarin	MT1	Female	4.1	3.8	0.3	3.9	3.8	3.9	3.9	4.1	3.9
			4.2	4.0	0.2	4.0	4.0	4.0	4.1	4.0	4.0
			4.2	4.1	0.1	4.2	4.2	4.2	4.1	4.1	4.2
			4.1	4.0	0.1	4.0	4.1	4.0	4.0	4.0	4.0
			4.1	3.9	0.2	4.1	4.0	3.9	3.9	4.1	4.0
			4.3	4.0	0.3	4.3	4.1	4.0	4.0	4.0	4.1
			4.0	3.9	0.2	4.0	4.0	3.9	4.0	3.9	4.0
			4.2	3.9	0.3	4.2	3.9	3.9	3.9	4.0	4.0
			4.0	3.8	0.2	3.9	3.9	3.9	3.9	4.0	3.9
			4.3	3.9	0.3	3.9	4.2	4.2	4.2	4.2	4.1
			4.2	4.0	0.2	4.1	4.1	4.0	4.0	4.0	4.0
		Male	3.8	3.5	0.3	3.5	3.7	3.8	3.8	3.7	3.7
			3.8	3.4	0.4	3.7	3.7	3.8	3.8	3.4	3.7
			3.9	3.6	0.2	3.7	3.8	3.8	3.9	3.6	3.8
			3.8	3.6	0.2	3.7	3.7	3.7	3.7	3.6	3.7
			3.7	3.4	0.3	3.4	3.5	3.6	3.6	3.6	3.5
			3.6	3.3	0.2	3.4	3.5	3.5	3.6	3.4	3.5
			3.8	3.5	0.3	3.7	3.6	3.7	3.6	3.8	3.7
			3.8	3.3	0.5	3.8	3.3	3.4	3.5	3.3	3.5
			3.6	2.9	0.7	2.9	3.4	3.6	3.6	3.4	3.4
			3.8	3.4	0.4	3.4	3.7	3.8	3.8	3.6	3.7
			3.7	3.4	0.3	3.7	3.6	3.5	3.5	3.4	3.5
Thai	TT1	Female	4.2	3.7	0.5	4.2	3.9	3.8	3.8	3.9	3.9
			4.7	4.0	0.7	4.7	4.2	4.1	4.0	4.0	4.2
			4.5	3.9	0.6	4.5	4.1	4.0	3.9	4.0	4.1
			4.7	4.1	0.6	4.7	4.4	4.3	4.2	4.2	4.4
			4.6	3.9	0.7	4.6	4.1	4.0	3.9	4.1	4.1
			4.4	3.9	0.5	4.4	4.1	4.1	3.9	4.1	4.1
			4.5	3.8	0.7	4.5	4.1	4.0	3.9	3.8	4.1
			3.7	3.5	0.1	3.6	3.6	3.6	3.6	3.7	3.6
			4.2	3.9	0.3	3.9	4.1	4.0	4.0	4.1	4.0
			3.8	3.5	0.3	3.8	3.6	3.6	3.5	3.7	3.6
			4.5	3.9	0.6	4.5	4.2	4.0	3.9	3.9	4.1
		Male	2.4	2.2	0.2	2.4	2.3	2.3	2.3	2.2	2.3
			2.4	2.3	0.1	2.4	2.4	2.3	2.3	2.3	2.3
			2.4	2.0	0.3	2.4	2.2	2.1	2.0	2.0	2.1
			2.4	2.2	0.2	2.4	2.3	2.2	2.3	2.4	2.3
			2.6	2.5	0.1	2.6	2.5	2.5	2.6	2.5	2.5
			2.6	2.3	0.3	2.6	2.4	2.4	2.4	2.3	2.4
			2.6	2.5	0.1	2.5	2.5	2.5	2.5	2.6	2.5
			2.4	2.2	0.2	2.4	2.3	2.3	2.2	2.3	2.3
			2.5	2.3	0.2	2.5	2.4	2.3	2.3	2.3	2.4
			2.1	2.0	0.2	2.1	2.1	2.0	2.0	2.1	2.1
			2.6	2.4	0.2	2.6	2.5	2.4	2.4	2.5	2.5

(2) Rising tones

Group	Tone	Speaker	Peak	Valley	TurnPt (%)	Fall Range	Rise Range	Pt1 (0%)	Pt2 (25%)	Pt3 (50%)	Pt4 (75%)	Pt5 (100%)	Mean F0
Mandarin	MT2	Female	3.9	3.1	15.0	0.2	0.8	3.3	3.1	3.2	3.6	3.9	3.4
			4.3	3.1	32.0	0.3	1.2	3.4	3.2	3.3	3.8	4.2	3.6
			4.1	3.2	6.0	0.0	1.0	3.2	3.2	3.3	3.6	4.0	3.5
			4.0	3.0	17.0	0.1	1.0	3.1	3.1	3.3	3.6	4.0	3.4
			4.1	2.8	19.0	0.6	1.3	3.4	2.9	3.0	3.5	4.1	3.4
			4.0	3.1	17.0	0.3	0.9	3.4	3.1	3.2	3.7	3.9	3.5
			4.0	3.3	38.0	0.3	0.7	3.6	3.4	3.4	3.6	4.0	3.6
			3.9	3.2	9.0	0.1	0.7	3.3	3.2	3.4	3.5	3.8	3.4
			3.9	3.2	34.0	0.4	0.8	3.6	3.2	3.3	3.4	3.9	3.5
			4.3	3.3	6.0	0.3	0.9	3.6	3.6	3.8	4.0	4.3	3.9
			4.0	3.3	38.0	0.5	0.7	3.8	3.4	3.4	3.6	4.0	3.6
		Male	3.3	2.2	23.0	0.1	1.1	2.3	2.2	2.6	3.1	0.0	2.0
			3.3	2.3	17.0	0.2	1.1	2.5	2.3	2.9	3.2	3.3	2.8
			3.3	2.2	13.0	0.1	1.1	2.3	2.2	2.6	3.2	3.3	2.7
			3.3	2.4	19.0	0.3	0.8	2.7	2.5	2.7	3.0	3.3	2.8
			3.7	2.2	26.0	0.1	1.5	2.3	2.2	2.5	3.2	3.7	2.8
			3.2	2.2	43.0	0.3	1.0	2.4	2.2	2.2	2.7	3.2	2.5
			3.4	2.2	30.0	0.4	1.2	2.6	2.2	2.5	3.0	3.4	2.7
			3.2	2.2	4.0	0.0	1.0	2.2	2.3	2.7	3.1	3.1	2.7
			3.3	2.2	40.0	0.4	1.2	2.6	2.3	2.4	2.8	3.3	2.7
			3.7	2.5	13.0	0.0	1.1	2.5	2.6	2.9	3.5	3.6	3.0
			3.3	2.3	26.0	0.5	1.0	2.8	2.3	2.4	2.8	3.3	2.7
Thai	TT4	Female	4.4	3.8	17.0	0.5	0.6	4.3	3.8	3.9	4.2	4.3	4.1
			4.7	4.2	39.0	0.5	0.5	4.7	4.2	4.2	4.5	4.7	4.5
			4.8	3.9	47.0	0.6	0.8	4.5	4.0	4.0	4.2	4.8	4.3
			4.3	3.8	41.0	0.4	0.5	4.2	3.8	3.9	4.2	4.1	4.0
			4.7	4.1	41.0	0.6	0.6	4.7	4.1	4.1	4.5	4.6	4.4
			4.5	4.1	20.0	0.4	0.4	4.5	4.1	4.1	4.5	4.3	4.3
			4.4	3.9	31.0	0.2	0.5	4.1	4.0	4.0	4.4	4.1	4.1
			4.2	3.6	0.0	0.0	0.8	3.6	3.7	3.7	4.0	4.1	3.8
			4.5	3.8	0.0	0.0	0.4	3.8	4.0	4.2	4.5	4.4	4.2
			4.0	3.5	5.0	0.3	0.5	3.9	3.6	3.7	4.0	3.7	3.8
			4.6	4.0	22.0	0.6	0.6	4.6	4.0	4.1	4.4	4.3	4.3
		Male	2.7	2.1	20.0	0.1	0.5	2.2	2.1	2.3	2.6	2.7	2.4
			3.3	2.7	0.0	0.0	0.6	2.7	2.7	2.9	3.1	3.3	2.9
			3.1	2.4	27.0	0.1	0.7	2.4	2.4	2.5	2.8	3.1	2.6
			3.3	2.4	22.0	0.2	0.8	2.6	2.4	2.6	2.7	3.3	2.7
			3.1	2.5	29.0	0.1	0.6	2.6	2.5	2.6	2.8	3.1	2.7
			3.0	2.5	22.0	0.1	0.5	2.5	2.5	2.6	2.8	3.0	2.7
			3.2	2.4	25.0	0.1	0.8	2.5	2.4	2.5	2.8	3.2	2.7
			2.9	2.2	15.0	0.0	0.7	2.3	2.3	2.3	2.7	2.9	2.5
			3.1	2.6	12.0	0.1	0.5	2.7	2.6	2.7	3.1	3.0	2.8
			2.7	2.2	36.0	0.2	0.6	2.3	2.2	2.2	2.4	2.7	2.4
			3.2	2.6	32.0	0.2	0.5	2.8	2.7	2.7	2.8	3.2	2.8

(3) Falling rising tones

Group	Tone	Speaker	Peak	Valley	TurnPt (%)	Fall Range	Rise Range	Pt1 (0%)	Pt2 (25%)	Pt3 (50%)	Pt4 (75%)	Pt5 (100%)	MeanF0
Mandarin	MT3	Female	3.9	1.0	44.0	2.1	2.4	3.2	3.9	0.0	3.0	3.4	2.7
			3.4	0.3	34.0	3.1	2.9	3.4	0.8	1.5	2.8	3.0	2.3
			3.7	0.5	36.0	3.2	2.8	3.7	2.9	1.8	3.1	3.2	2.9
			3.2	0.4	56.0	2.8	2.5	3.2	1.9	0.4	2.5	2.9	2.2
			3.1	2.0	31.0	1.2	1.1	3.1	2.1	2.4	2.8	2.1	2.5
			5.0	0.0	47.0	3.4	5.0	3.4	2.5	0.0	3.1	4.7	2.7
			3.3	0.1	41.0	3.3	3.3	3.3	1.0	1.4	3.0	3.2	2.4
			3.8	0.5	14.0	2.6	3.3	3.2	1.2	2.2	2.8	2.4	2.4
			3.4	0.0	46.0	3.2	3.1	3.3	2.6	0.4	3.1	3.1	2.5
			3.7	3.0	47.0	0.1	0.7	3.1	3.1	3.0	3.3	3.7	3.2
			3.6	0.5	53.0	3.1	2.8	3.6	0.0	0.7	2.8	3.2	2.1
		Male	2.5	0.0	19.0	0.3	2.4	0.3	0.0	0.0	2.1	2.0	0.9
			2.5	0.3	56.0	2.0	2.2	2.4	1.1	0.0	2.3	2.5	1.7
			2.4	0.1	64.0	2.3	2.1	2.4	0.0	0.0	1.9	2.1	1.3
			2.9	0.1	66.0	2.8	2.3	2.9	1.5	0.0	0.9	2.3	1.5
			2.6	0.1	46.0	1.9	2.5	2.0	0.0	0.5	2.4	2.5	1.5
			2.3	0.0	47.0	1.3	2.3	1.3	0.0	0.2	2.0	2.1	1.1
			2.4	0.1	59.0	2.3	2.3	2.4	0.0	0.0	1.9	2.1	1.3
			2.2	0.5	53.0	0.0	1.6	0.6	0.0	0.5	1.3	1.9	0.9
			2.5	0.1	37.0	2.4	2.2	2.5	2.0	0.0	1.4	2.0	1.6
			2.6	1.3	58.0	0.4	1.3	1.7	2.2	1.4	2.0	2.5	2.0
			2.2	0.0	56.0	2.2	1.9	2.2	0.0	0.0	0.9	1.0	0.8
Thai	TT5	Female	4.5	2.1	36.0	1.5	2.4	3.6	2.4	2.6	3.5	4.5	3.3
			4.5	1.9	52.0	2.3	2.6	4.2	2.9	0.0	3.7	4.5	3.1
			4.4	0.7	41.0	1.3	3.8	2.0	0.8	0.8	3.7	4.4	2.3
			4.6	0.5	38.0	3.7	4.1	4.3	1.0	0.6	4.0	4.6	2.9
			4.7	2.8	48.0	1.5	1.9	4.2	3.1	2.8	4.0	4.6	3.7
			4.4	2.6	44.0	1.3	1.8	3.9	3.1	2.7	3.4	4.4	3.5
			4.2	2.2	33.0	1.7	2.0	3.9	2.9	2.6	3.5	4.2	3.4
			4.2	0.5	49.0	2.9	3.6	3.4	1.0	0.5	3.5	4.2	2.5
			4.7	3.2	36.0	0.5	1.5	3.6	3.2	3.2	4.1	4.6	3.7
			3.9	0.2	56.0	3.7	3.7	3.9	3.2	3.1	3.4	3.9	3.5
			4.9	0.0	44.0	4.2	4.9	4.2	3.0	0.7	3.9	4.9	3.3
		Male	3.7	0.6	56.0	2.9	3.1	0.0	0.0	3.5	2.6	2.5	1.7
			3.1	1.0	38.0	1.2	2.1	2.3	1.7	1.5	2.4	3.1	2.2
			2.8	0.4	70.0	1.7	2.5	2.0	0.0	0.0	0.6	2.8	1.1
			3.6	0.3	66.0	1.9	3.3	2.2	1.7	0.0	2.6	3.6	2.0
			3.4	0.3	21.0	1.9	3.0	2.2	0.0	0.0	2.8	3.4	1.7
			3.0	0.0	59.0	1.9	3.0	2.0	1.5	0.0	2.4	3.0	1.8
			3.2	0.3	66.0	1.8	2.9	2.1	0.0	0.0	0.4	3.2	1.1
			3.0	0.3	54.0	1.3	2.7	0.4	1.5	0.5	1.8	3.0	1.4
			3.1	1.6	49.0	0.7	1.6	2.2	1.8	1.6	2.4	3.1	2.2
			3.1	0.0	28.0	2.0	3.0	2.0	0.2	0.0	2.3	2.7	1.4
			2.9	0.0	36.0	2.0	2.9	2.0	0.7	0.0	2.0	2.9	1.5

(4) Falling tones

Group	Tone	Speaker	Peak	Valley	Fall Range	Pt1 (0%)	Pt2 (25%)	Pt3 (50%)	Pt4 (75%)	Pt5 (100%)	MeanF0
Mandarin	MT4	Female	4.0	3.1	0.9	4.0	3.9	3.6	3.1	0.0	2.9
			4.1	2.6	1.5	4.1	3.9	3.3	2.7	3.2	3.4
			4.4	2.9	1.5	4.3	4.1	3.5	3.0	4.4	3.9
			4.3	0.6	3.7	4.3	4.0	3.5	2.8	1.0	3.1
			4.3	2.4	1.9	4.3	3.9	3.2	2.7	2.8	3.4
			4.3	2.7	1.6	4.3	4.2	3.6	3.1	2.8	3.6
			4.2	0.7	3.6	4.2	4.1	3.4	2.8	0.9	3.1
			4.4	0.4	4.0	4.4	3.9	3.3	2.9	1.4	3.2
			4.1	2.3	1.8	4.1	4.0	3.4	2.8	2.3	3.3
			4.4	3.0	1.4	4.3	4.3	3.8	3.1	3.4	3.8
			4.6	2.4	2.2	4.6	4.3	3.7	3.1	2.5	3.6
		Male	4.7	1.5	3.2	4.2	4.1	3.8	4.3	4.5	4.2
			4.0	0.7	3.3	4.0	3.9	3.5	2.6	0.7	2.9
			4.5	2.5	2.0	4.5	4.4	4.0	3.3	2.5	3.7
			4.7	0.1	4.5	4.7	4.5	4.0	0.3	0.0	2.7
			4.3	0.1	4.2	4.3	4.0	3.6	0.6	0.0	2.5
			4.3	0.8	3.4	4.3	4.1	3.7	1.7	0.0	2.8
			5.0	0.6	4.4	4.3	4.1	3.7	1.9	1.5	3.1
			4.4	1.5	2.8	4.4	4.1	3.9	4.0	3.0	3.9
			3.8	1.8	2.0	2.9	3.6	3.8	1.8	2.0	2.8
			4.1	2.1	1.9	4.0	4.1	3.8	3.0	2.1	3.4
			4.0	0.1	3.9	4.0	3.9	3.6	0.5	0.6	2.5
Thai	TT2	Female	4.3	0.0	4.1	4.3	2.7	0.0	0.0	0.0	1.4
			4.5	1.1	3.4	4.5	3.7	3.2	2.4	0.0	2.8
			4.2	2.1	2.2	4.2	3.5	2.8	2.3	2.5	3.1
			5.0	0.2	4.2	4.4	3.3	2.3	5.0	3.4	3.7
			4.5	3.0	1.5	4.5	3.6	3.2	3.2	3.2	3.5
			4.4	2.3	2.2	4.4	3.8	3.4	3.1	2.7	3.5
			4.3	0.2	4.1	4.3	3.6	2.8	0.0	4.0	2.9
			4.0	0.6	3.2	3.8	3.4	1.1	1.1	0.0	1.9
			3.9	3.1	0.8	3.9	3.4	3.2	3.2	3.3	3.4
			3.8	0.2	3.6	3.8	3.5	3.1	0.0	0.0	2.1
			4.5	0.1	4.4	4.5	3.6	2.8	2.3	0.0	2.6
		Male	2.0	0.0	1.7	2.0	1.7	0.0	0.3	0.0	0.8
			2.3	0.6	1.7	2.3	1.9	1.0	0.7	0.0	1.2
			5.0	1.5	0.5	2.0	1.7	0.0	4.9	0.0	1.7
			2.3	0.2	2.1	2.3	1.8	1.0	0.4	1.3	1.4
			4.7	1.4	0.6	2.0	1.7	1.4	4.5	3.7	2.7
			2.3	0.4	1.6	2.3	2.0	1.5	0.7	0.0	1.3
			3.0	0.5	1.4	2.1	1.3	0.0	0.0	0.0	0.7
			1.9	0.2	1.7	1.9	1.7	0.8	0.2	0.4	1.0
			2.4	1.6	0.7	2.4	2.0	1.7	1.7	1.8	1.9
			2.0	0.0	2.0	2.0	1.9	1.3	0.0	0.0	1.0
			2.3	0.3	1.8	2.1	1.8	0.9	0.3	0.0	1.0

(to be continued)

Thai	TT3	Female	4.6	1.2	3.4	4.6	4.5	4.2	1.2	1.5	3.2
			5.0	1.1	3.8	5.0	4.8	4.6	1.4	3.1	3.8
			4.9	1.3	3.5	4.9	4.7	4.6	1.9	1.4	3.5
			4.8	1.2	3.7	4.8	4.7	4.7	4.0	1.2	3.9
			4.9	3.2	1.7	4.9	4.7	4.4	3.3	3.2	4.1
			4.9	3.2	1.6	4.7	4.9	4.5	3.3	3.3	4.1
			4.8	1.2	3.5	4.6	4.8	4.7	3.8	1.3	3.8
			4.3	1.1	2.7	3.8	4.3	4.2	1.4	1.2	3.0
			5.0	3.4	1.4	4.8	5.0	4.8	3.6	3.4	4.3
			4.2	3.4	0.4	3.8	4.1	4.2	3.6	3.4	3.8
			4.9	1.3	3.7	4.9	4.7	4.5	1.5	1.3	3.4
		Male	2.7	0.3	2.3	2.7	2.7	2.5	0.9	0.0	1.8
			2.9	0.8	0.8	0.8	0.9	2.9	2.4	0.0	1.4
			2.9	0.3	2.6	2.9	2.9	2.4	2.0	0.0	2.0
			2.9	1.9	0.9	2.9	2.9	2.8	2.3	1.9	2.6
			3.0	0.6	2.4	3.0	2.9	2.7	1.3	0.0	2.0
			2.9	1.7	1.2	2.9	2.9	2.7	2.1	1.7	2.5
			2.9	0.2	2.6	2.8	2.9	2.5	2.0	0.0	2.0
			2.6	1.7	0.6	2.3	2.6	2.5	2.3	1.7	2.3
			3.0	1.8	1.2	3.0	2.9	2.7	2.4	1.8	2.6
			2.8	1.9	0.8	2.7	2.8	2.6	2.3	1.9	2.5
			2.9	1.9	1.0	2.9	2.9	2.7	2.3	1.9	2.5