Effects of Musical Experience and Linguistic Experience on Categorization of Melodic and Lexical Tones

by

Daniel Chang

B.A. (Linguistics), Simon Fraser University, 2012

Thesis Submitted in Partial Fulfillment of the
Requirements for the Degree of
Master of Arts

in the
Department of Linguistics
Faculty of Arts and Social Sciences

SIMON FRASER UNIVERSITY

Summer 2014

Approval

Name: **Daniel Chang**

Degree: **Master of Arts (Linguistics)**

Title of Thesis: ***Effects of Musical Experience and Linguistic Experience on Categorization of Melodic and Lexical Tones***

Examining Committee: **Chair:** Dr. Maite Taboada
Associate Professor

Dr. Yue Wang
Senior Supervisor
Associate Professor

Dr. Nancy Hedberg
Co-Supervisor
Professor

Dr. Yang Zhang
External Examiner
Associate Professor, Department of
Speech-Language-Hearing Sciences
University of Minnesota

Date Defended/Approved: August 13, 2014

Ethics Statement

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.

Abstract

Tone language and music both use pitch to convey categorical information. Previous research has shown shared processing of linguistic and musical tones in that pitch experience in one domain may affect the perception of the other. However, how such transfer facilitates tonal categorization is less clear.

The present study investigates the categorical perception of linguistic and music tones by three groups of listeners differing in their tone language and musical training experiences (native Mandarin non-musicians, native English musicians, and native English non-musicians). Linguistic tonal continua were created from Mandarin level to rising tones, and from level to falling tones. Melodic music continua were created by varying the note D under the context of C and E (CDE to CEE, and CCE to CDE). The participants performed a discrimination task and an identification task for the Mandarin tones and the music melodies.

The results show facilitative effects of musical experience on Mandarin tone. First, the English musicians' tone discrimination outperformed the native Mandarin non-musicians, indicating musicians' enhanced sensitivity to pitch variations. Moreover, the English musicians identified the Mandarin tones in a more categorical manner than English non-musicians, showing musicians' greater ability in categorizing pitch information. On the other hand, experience with linguistic tone is also shown to affect melodic perception. Specifically, the Mandarin non-musicians' experience in linguistic tone categorization appears to facilitate their perception in music, as their identification of melodic CDE-to-CEE exemplars reveals a more categorical pattern than the English non-musicians. However, the Mandarin non-musicians' experience in tone categorization may have decreased their sensitivity to subtle pitch differences, as their discrimination of melodic pitch differences was the poorest among the three listener groups. Taken together, these results suggest bi-directional transfer between language and music in perception and categorization processes, pointing to an integration of domain-general and domain-specific pitch processing as a function of experiences.

Keywords: Music Perception; Categorical Perception; Effects of Interference; Mandarin Tones; Musical Training; Perceptual Transfer

To the special friends

Whom I have met in life...

Acknowledgements

I am not good at writing the acknowledgement section because thousands of words and sentences would not be enough for me to express my profound gratitude to all the people that I have met at SFU. I am sure if you are reading this section, you are 99.9% of the ones that I want to say thank you to. Therefore, I don't want to be pretentious, nor do I want to be sentimental; thus, I decided to end this section very quickly with a quote from Yogi Berra. This best represents what I've been feeling throughout my undergraduate career and graduate career at SFU.

"*Baseball is ninety percent mental and the other half is physical*" – Yogi Berra.

By the way, hello to those people that I want to say "thank you". I will probably keep bothering you for another 4-5 years or forever at SFU.

Thank you, Yue, Nancy, and Yang for being so supportive in the last stage of my thesis. Your guidance, support, and encouragement were more important than anything else. I also want to thank my family members, all the previous and current lab members, previous and current undergraduate/graduate students in the department, all the administrative staff in the department, all the professors/instructors in the department (plus Mr. Ian Bercovitz in the Statistics Department), and all the masters at the *LingYen Mountain Temple* (because they served me food and encouraged me a lot while I wrote my thesis). Your continuous encouragement has helped me to reach this point in my life time.

Table of Contents

List of Tables

List of Figures

Chapter 1. Introduction

The interaction between music and language has been studied extensively. The underlying assumption is that the physical realizations of musical and linguistic information share similar acoustic properties, such as fundamental frequency (pitch). For instance, tone languages use pitch height and contour shape for differentiating tonemic categories (Chen, 1984). Likewise, music uses such pitch information to distinguish different levels of musical tones.

The present study investigates the interaction between musical experience and tone language experience on the categorical perception of lexical tones and melodic tones. This chapter reviews the relevant literature: Section 1.1 first discusses the theoretical background in terms of the relationship between music and language. Section 1.2 reviews the empirical findings on the effects of musical and linguistic experiences on perception of melodic and lexical tones. Section 1.3 describes the present research.

1.1. Theoretical Framework

The theoretical foundation of the relationship between language and music has been thoroughly discussed in Patel's (2008) book, *Music, Language and the Brain*, in which he argues that music and language may utilize a shared processing mechanism in the human brain. Because the fundamental elements of music and languages are sounds, the brain may process language in the same way it processes music information. Specifically, Patel (2008) proposes a *Shared Sound Category Learning Mechanism Hypothesis (SSCLMH)*, arguing that despite the different surface-level representations in language and music, the underlying general processing mechanism may develop similar representations of sound categories in language and music. That is, when humans perceive pitch information, either linguistic or music, the auditory system in the brain starts to process the sounds in a domain-general manner. During the sound-category development, the brain will classify the received sounds into categories, which may later become domain-specific.

SSCLMH aligns well with another theoretically grounded work by Zatorre and Gandour (2008), which argues for an integrated network in pitch processing (namely the integrated account of pitch processing) on the basis of two opposing models involving domain-specific and domain-general processes. Specifically, the domain-specific model supports that pitch in language is processed exclusively by a speech-specific mechanism in the brain. On the other hand, the domain-general model predicts that there is a shared processing mechanism of pitch, which is responsible for handling any kinds of pitch information. According to Zatorre and Gandour (2008) these two models do not have to be mutually exclusive, in that some aspects of pitch information may be processed by the domain-general mechanism (e.g. the onset of categorical formation), while other aspects may be domain-specific (e.g. the end product of the sound-category formation).

In the book, *Principles of Perceptual Learning and Development*, Eleanor Gibson (1969) used the cognitive perspective to propose the effect of perceptual learning and argued that experience in one cognitive domain may trigger an improved performance in the other domain. This theory has been extended by subsequent research on different sensory modalities, such as speech and non-speech signals (Banai & Amitay, 2012), categorization ability (Sloutsky, 2003 Smits, Sereno & Jongman, 2006), or other natural acoustic signals (Amitay, Zhang & Moore, 2012). These all point to the direction that there may be a bottom-up perceptual transfer, which is subject to human's naïve perceptual experience (Ahissar & Hochstien, 1997; Bradley, 2012; Moore, Amitay, & Hawkey, 2003). Thus, it is conceivable that pitch experiences in language and music may be mutually transferrable.

Taken together, both neurological perspectives and the cognitive perspective have provided a well-established basis for empirical research examining the interplay of linguistic and musical experience on the perception of musical tones and lexical tones. The next section presents a review on the empirical findings with respect to the influence of musical training experience on language processing, language effects on music processing, and the issue of categorical perception in speech and music.

1.2. Empirical Findings

This section is divided into four sub-sections. Section 1.2.1 discusses some recent findings on the effects of musical experience on language learning or processing. Section 1.2.2 reviews the literature on the effects of language experience on music perception. Section 1.2.3 focuses on the categorical perception of lexical tones by different populations. Section 1.2.4 summarizes the overall findings.

1.2.1. Effects of Musical Experience on Tone Language Processing

There has been a growing interest in how the connection of music and language influence pitch processing (e.g. Milovanov & Tervaniemi, 2011; Milovanov et al., 2010; Moreno et al., 2009; Morrison & Demorest, 2009). Because music and language share several properties, one line of research has focused on how musical training influences language processing and learning (i.e., the "music-to-language" effects). Overall, consistent evidence has shown a positive effect of musical training on language abilities, such as enhanced verbal memory (Franklin et al., 2008; Ho, Cheung & Chan, 2003), intelligence quotient (Schellenberg, 2004), lexical stress processing (Kolinsky et al., 2009), speech segmentation (Francois, Chobert, Besson & Schon, 2012) and foreign language learning (Alexander, Wong & Bradlow, 2005; Cooper, 2010; Cooper & Wang, 2012; Delogu, Lampis, & Belardinelli, 2010; Wong & Perrachione, 2007).

In a study examining the discrimination of Mandarin tones by Italian listeners whose native language is non-tonal, Delogu, Lampis, and Belardinelli (2006) show that listeners' performance on a music intelligence test positively correlates with the correct perception of Mandarin tonal contrasts; that is, the Italian listeners who score higher on the music test achieved greater accuracy in the tone discrimination task. The authors thus conclude that music skills can be transferrable to linguistic tone skills, motivating subsequent research on how musical experience may facilitate linguistic tone learning.

Wong and Perrachione (2007) conducted a perceptual training study to examine how English listeners with no linguistic tone background use their music pitch experience to acquire Mandarin tone words. In the study, they trained native English listeners to associate pseudo-syllables with the Mandarin tonal pitch patterns. Their results indicate that the listener's pitch awareness is positively associated with their performance on the pseudo word identification task. That is, those English listeners who were more sensitive to pitch showed greater improvements after training. Furthermore, they found that the listeners with more years of musical experience performed better on the tone learning. The authors suggest that the observed positive behavioural learning patterns may be a result of increased cortical auditory representations and pitch awareness related to music experience.

Following Wong and Perrachione (2007), Lee and Hung (2008) further explored how musical training influences non-tone speakers' learning of real Mandarin words. In the study, they examined the effects of musical training on Mandarin tone identification by 36 native English-speaking musicians and 36 native English-speaking non-musicians. Their results showed that musicians were more successful at identifying Mandarin tones produced by multiple speakers than non-musicians. This suggests that musical training reinforces a listener's pitch awareness, which in turn facilitates the perception of linguistic tones in a non-native language.

Moreover, other researchers have investigated the interaction of linguistic and music tone experience on the perception and learning of non-native tones. Cooper and Wang (2012) examined the perceptual learning of Cantonese tone words by four groups of listeners differing in their backgrounds with linguistic and music tones: native Thai musicians, native Thai non-musicians and native English musicians and non-musicians without any linguistic tone experience. Their training results showed that the Thai non-musicians indeed obtained higher Cantonese word accuracy after training compared to English non-musicians, indicating the positive effects of tone language experience. On the other hand, after training, English musicians' tone performance is significantly higher than the English non-musicians, showing that musical experience facilitates linguistic tone learning. However, the results do not show an additive effect of music and tone experiences, meaning that for the Thai musicians, having both a tone language background (Thai) and musical training does not seem more advantageous for learning the Cantonese tone words. These results imply that long-term exposure to either linguistic pitch experience or music pitch experience indeed benefits the learning of linguistic tones in a non-native language.

Consistently, at the neurological level, research using electroencephalography (EEG) has shown that musical training indeed facilitates pitch sensitivity (Fujioka et al., 2004; Kraus & Chandrasekaran, 2010; Magne, Schon, & Besson, 2006; Marques et al., 2007; Skoe & Kraus, 2012) and lexical tonal variations (Marie et al., 2011). For instance, Marie et al. (2011) used a four-interval Mandarin discrimination task (4I2AFC) to examine whether musicians are better at detecting a tonal variation as compared to non-musicians. At the same time event-related potentials (ERPs) are recorded during the discrimination task. On the behavioral side, results have supported that musicians can more accurately detect tonal deviations than non-musicians. On the neurological side, the "different" stimulus sequences elicited more enhanced ERP (N2/N3) components around 100 milliseconds post stimulation than the "same" stimulus sequences in the discrimination task; this reflected early discrimination sensitivity for musicians compared to non-musicians. Additionally, a late ERP component (P3b) showed a larger amplitude difference in musicians than in non-musicians. The authors then conclude that there is a music-to-language perceptual transfer effect at both the behavioral level and neurological level.

Taken together, these studies consistently show that having extensive musical training enhances linguistic tone processing and facilitates non-native tone acquisition at the behavioral level and the neurological level. However, research on the influence of linguistic tone on music tone processing has been controversial and less consistently reported in the literature. The following sub-section discusses the general findings from the language-to-music stream and provides some insights to why the findings are inconsistent across.

1.2.2. Effects of Language on Music Tone Processing

While research in the direction from music to language has consistently reported that there is a positive transfer effect, the reverse direction (e.g. language to non-speech tones and language to music) has caught some researchers' attention as well.

1.2.2.1 Effects of Linguistic Experience on Non-speech Perception

Since linguistic and music tone processing may both involve domain-general processing of the acoustic properties of pitch, one line of research has focused on the neural representations of non-speech pitch as a result of linguistic pitch experience (Bidelman et al., 2011; Chandrasekaran et al., 2007; Gandour et al., 1988; Klein, Zatorre, Milner, Zhao et al., 2001; Krishnan et al., 2010). For instance, Gandour et al. (1998) made use of a neuro-imaging technique (*Positron Emission Tomography* PET) to investigate native tone (Thai) and non-tone (English) listeners' processing patterns of Thai Tones and non-speech pitch with low-pass filtered tones. Results showed that Thai listeners and English listeners exhibited no behavioural differences in the discrimination task. However, the Thai listeners but not English listeners showed more robust neural activations in the left frontal operculum (Broca's area) while listening to the speech and non-speech tones. This suggests that native tone language speakers are able to utilize and activate the neural linguistic network (e.g. Broca's area) to encode the lower-level non-speech pitch as well as the higher-level linguistic pitch information. These results motivate further research on how linguistic pitch experience influences the underlying neural auditory processing mechanism.

Recent behavioral research has also documented a strong influence of linguistic experience on non-speech pitch perceptual performance (Bent, Bradlow, & Wright, 2006;

Mattock & Burnham, 2006). Bent et al. (2006) compared Mandarin speakers and English speakers' ability to discriminate Mandarin tones and non-speech tones with either flat or contour tones (containing rising or falling pitch sequences). They have found that Mandarin listeners are better at identifying their own native tones, but their non-speech tone discrimination does not differ from the English listeners. In addition, an interesting finding from Bent et al. (2006) is that Mandarin speakers are more likely to mis-identify the non-speech tones (especially the falling pitch contour) than English speakers because Mandarin speakers may have employed different listening strategies in identifying the non-speech tones. They explain that the differences may be attributed to the factor of listeners' pitch range threshold. It is because, for example, the Mandarin speakers may be more likely to have a stricter criterion for a falling pitch to be classified as a falling sound. This has pointed to the implications that tone speakers' auditory processing may be affected by task and stimulus conditions, or language-specific tone category.

Overall, the effect of language experience influences the perception of non-speech tones, as some research has pointed that tone language speakers are able to utilize their own native pitch experience in the processing of filtered non-speech tones (e.g. Gandour et al., 1998) or identifying pure tones (e.g. Bent et al., 2006). It has been unclear whether tone language speakers can process music tones, different from non-speech tones, by the same auditory mechanism as they process non-speech tones. The next section will review studies that address the influence of language on music processing.

1.2.2.2 Effects of Language on Music Performance

The neural representation of music by tone language non-musicians seems to be the same as that of non-tone language musicians (e.g. Bidelman, Gandour, & Krishnan, 2011). In an attempt to investigate the effect of long-term pitch experience on music processing, Bidelman et al. (2011) recruited 11 English-speaking musicians, 11 English-speaking non-musicians, and 11 native speakers of Mandarin non-musicians to perform two tasks: (1) a two-alternative forced choice discrimination task (2AFC Discrimination) on two sequential music intervals, and (2) brainstem responses using FFR (*frequency-following responses*) recorded while four music triad arpeggios are played to the participants. Results indicate that native Chinese speakers and native English musicians show stronger brainstem encoding. Nevertheless, the native Chinese speakers do not translate such neurological superiority to behavioral music discrimination. These findings suggest that a long-term experience in one domain (e.g. language) may influence the subcortical processing of the acoustic information available in the music domain, yet such enhancement does not necessarily imply a behavioral benefit for music perception in tone language speakers.

Similarly, a study by Peretz et al. (2011) showed an interference effect of tone language experience on the discrimination of synthesized piano tones. In their study, the researchers asked speakers with varied tone language backgrounds (Mandarin, Cantonese, and Vietnamese) and speakers of non-tone languages (French or English) to perform a same-and-different pitch discrimination task. In the task, the participants were asked to tell whether they could detect a pitch change in the target sequence compared to the primed sequence (each sequence contains 5 successive tones, with the fourth tone being altered). Results showed that there is no clear benefit for the tone language speakers to discriminate the sequences in which the non-tone speakers outperformed tone-speakers in the discrimination task. They attribute this finding to the effect of interference, in which the tone language speakers' perception of music sequences may be interfered with their own native language experiences, particularly when they perceive a downward pitch sequence. The authors then conclude that the detection of pitch deviation may be speech specific for tone language speakers if the pitch change threshold is small.

While the findings reported by Bidelman et al. (2011) do not report a perceptual music benefit for tone language non-musicians, other studies reported that professional tone language musicians show superior musical abilities (Bahr, Christensen & Bahr, 2005; Deutsch, Dooley, Henthorn & Head, 2009; Deutsch, Henthorn & Dolson, 2004; Deutsch, Henthorn, Marvin & Xu, 2006; Levitin & Rogers, 2005; Schellenberg & Trehub, 2008; Wilson et al., 2009; Wong et al., 2012). For professionally trained musicians, Deutsch et al. (2006) examined a group of musicians, who were also fluent native tone language speakers, in an absolute-pitch music task. The results have revealed that these musicians performed better on the absolute-pitch music task than those native tone language speakers who cannot fluently speak the tone language, suggesting that speaking a tone language may provide a benefit in acquiring the ability of absolute pitch for music. Additional results also indicate that those who have learned music earlier tend to perform better on the task, but earlier onset of musical training, coupled with a tone language background does not give rise to an additional, better absolute-pitch score. They attribute the results to the fact that tone language speakers may possess absolute pitch from their tone language experience, and this can transfer to the acquisition of absolute pitch while they are learning music.

Although musicians with a tone language background have enhanced music skills, non-musically-trained tone language speakers have been found to have enhanced musical skills as well (Hove et al., 2010; Krishnan et al., 2010; Pfordresher & Brown, 2009).This is contradictory to Bidelman et al. (2011)'s findings. In Pfordresher and Brown's study (2009), they examined whether experience with a tone language facilitates better pitch discrimination and imitation in music. To investigate this, they asked musically-untrained tone language and non-tone language listeners to produce and perceive musical tones. In the production part, there were three levels: note sequence, interval sequence, and melodic sequence. In the perception part, there were two levels: note discrimination and interval discrimination. The results showed that tone language listeners have the advantage of producing interval sequence and discrimination music intervals in comparison to English listeners, suggesting that speaking a tone language facilitates the processing of musical intervals, as those tone language speakers are more aware of the direction of pitch changes in their native language.

Moreover, some of the research findings also indicate that speaking a tone language only facilitates low-level processing of tones, such as discriminating music tones (Alexander, Bradlow, Ashley & Wong, 2011; Giuliano et al., 2011; Stevens, Keller, & Tyler, 2004), but not high-level processing, such as identifying or matching music tones (Alexander et al., 2011; Bidelman et al., 2011). For instance, examining melodic discrimination and identification by native Mandarin (tone) speakers and native English speakers (all non-musicians), Alexander et al. (2011), found that the native Mandarin speakers are capable of discriminating unfamiliar music tones but not identifying them. On the contrary, the native English speakers perform better at matching the music tones with melodic sequential graphics than the native Mandarin speakers. The authors attribute this result to the effect of tone-category interference, such that when a native tone language speaker perceives a musical tone and is asked to identify the tone, the interference effect will occur because the long-term established lexical tones do not match the perceived musical categories, suggesting that long-term experience with a tone language may inhibit the top-down processing of the melodic identification but reinforce the bottom-up processing of music tones.

From the review above, it can be seen that findings on the role of tone language in music processing are not consistent. Some researchers argue that there may be methodological concerns in the literature (Bidelman et al., 2013). First, the definitions of non-speech vary across studies, as the term “non-speech” can be interpreted as music (e.g. Bidelman et al., 2011), or pure tones (e.g. Bent et al., 2006). Inconsistent definitions of non-speech tones and music tones may affect the interpretations of the findings in the language-to-music studies. Because music tone perception has been considered meaningful and generative at a higher cognitive domain, it may be processed by a domain-specific strategy (Patel, 2008; Stevens, Keller & Tyler, 2004). Furthermore, non-speech pitch processing is considered to involve basic acoustic properties that are governed by a domain-general processing mechanism. Secondly, the existing language-to-music studies have tested non-unified tone-language speakers (e.g. taking a broad range of tone language speakers as participants) to make a claim that speaking tone languages facilitates music perception (e.g. Bidelman et al., 2013; Pfordresher & Brown, 2009).

In sum, the previous language-to-music research suggests that tone language experience may affect the processing and perception of music pitch in a positive perceptual transfer at the lower-level acoustic domain (Gandour et al., 1998; Pfordresher & Brown, 2009), or the effects may be negative due to speech-specific top down influence (Peretz., et al., 2011). These findings further motivate the examination of categorical perception due to the notion that categorical perception is a domain-specific phenomenon that governs the processing of sound categories (Husain et al., 2006; Liberman et al., 1957). Therefore, the next section reviews the issue of categorical perception in linguistic tones and provides a connection with non-speech, music, and language perception.

1.2.3. Categorical Perception of Lexical Tones

1.2.3.1 Overview of Categorical Perception

Whether or not a listener is capable of perceiving differences between speech sounds in his/her native language has been studied in the context of categorical perception. Categorical Perception (CP) is a sensory perceptual phenomenon that shows the point where individuals are able to successfully differentiae or label one category from another in their native language. In paradigm such as this, it is assumed that if native listeners perceive speech sounds in a categorical manner, they will find it hard to perceive differences for within-category stimuli, whereas it will be easier for the listeners to perceive differences for between-category stimuli. As a result, in a typical CP experiment, two types of tasks have been used: (1) a discrimination task, and (2) an identification task.

In order to examine the results of CP for speech, a synthesized stimulus continuum is typically created (Klatt, 1973; Pisoni & Lazarus, 1974). Typical CP results always indicate a strong relationship between identification responses and discrimination performance. For example, if speech stimuli are identified as belonging to the same category, then it will be hard to discriminate them. However, if speech stimuli belong to different categories, then listeners might find it easier to discriminate. Based on this idea, it has been suggested that discrimination can be predicted from the identification responses (e.g. Bimler & Kirkland, 2001; Liberman et al., 1957; Repp, 1984; Repp & Liberman, 1987; Pisoni, 1977; Wang, 1976). In sum, with the understanding from categorical perception, we will be able to investigate whether individuals exploit language-specific mechanism in the processing of the speech sounds that are existent or non-existent in their native language, and how musical training and linguistic tone experience affect the formation of tone categories.

1.2.3.2 Categorical Perception of Lexical Tones

A natural consequence of categorical perception comes from extensive experience of our native language, and it has been argued that the assumption for CP is that listeners tend to exploit a language-specific strategy in processing their own native phonemic categories (e.g. Bidelman, Moreno & Alain, 2013; Liberman et al., 1957; Liberman et al., 1993; Liberman et al., 1961). Recently, more and more researchers have attempted to uncover whether tone language listeners can make use of the same speech-specific strategy they use to process lexical tones. Earlier work on lexical tone CP found that Mandarin Chinese listeners' perception of level tone and rising tone is categorical when they were asked to perceive the tone continuum varied from the level tone to the rising tone (Wang, 1976). However, for native English speakers, the pattern seems to be more continuous, because the two lexical tones do not exist in English. This has provided support for the evidence that Mandarin listeners have made use of a language-specific mechanism in processing their native tones, whereas native English listeners' processing of lexical tones has been general (Burnham & Jones, 2002; Chan et al., 1975; Peng et al., 2010; Xi et al., 2010; Xu, Gandour & Francis, 2006; Wang, 1976).

Subsequent research on Thai tones by Thai listeners had yielded different results, therefore motivating further research on non-speech tones (Abramson, 1979; Francis, Ciocca & Ng, 2003). Abramson (1979) revealed that Thai listeners' perception of their level tones is not categorical, because the Thai listeners do not exhibit a discrimination peak at the boundary location along the continuum. Abramson (1979) suggested that complexity of tones or syllable types might have influenced the listeners' auditory sensitivity. This is because in Chan et al. (1975) and Wang (1976)'s study, these authors used a single syllable /yi/ with a contour tone (rising) in Mandarin, whereas Abramson (1979)'s study used a complex syllable /khaa/ with three level tones.

A more recent study by Sun and Huang (2012) attempted to replicate Wang (1976)'s Mandarin research and Abramson (1979)'s Thai research by using a different tone language as the continuum material. They examined the categorical perception of Hokkien tones[1] by native Taiwanese speakers and native English speakers. Results show that Hokkien contour tones are perceived categorically by Taiwanese speakers but not the level tones, replicating the findings from Wang (1976) for contour tones and Abramson (1979) for level tones. This implies that tone language speakers have become accustomed to using a language-specific strategy for detecting the dynamic changes of contour tones. Compared to tone speakers, non-tone speakers never learned to use such dynamic acoustic cues in detecting pitch differences; therefore, they might be more proficient in using pitch height (stable acoustic change) as the potential indicator of discriminating the perceived tones, which appears to be a more domain general technique (Pisoni, 1973).

These findings have motivated subsequent research on categorical perception of non-speech tones, in an attempt to uncover whether a listener's auditory processing is governed by their language-specific strategy or a more general mechanism (e.g. Peng et al., 2010; Xu et al., 2006; Zhang, Xi, Wu, Shu & Li, 2012).

1.2.3.3 Categorical Perception of Non-speech Tones

Recent research on the issue of lexical tone CP has added another non-speech condition to examine whether non-speech tones give rise to the same categorical effect. Because non-speech tones bear little resemblance to lexical tones, research in this regard has always hypothesized that a more domain-general strategy might be applied to perceive the lower-level non-speech tones (e.g. Peng et al., 2010; Xu et al., 2006).

For instance, Xu et al. (2006) studied 15 native Mandarin listeners and 15 native English listeners to examine the effect of language experience on categorical perception of Mandarin tones and non-speech tones. To investigate whether the auditory processing of CP involves domain-specific mechanism or domain-general mechanism, they used

[1] Hokkien (Taiwanese Southern Min) contains seven different lexical tones: two level tones, three contour tones, and two entering tones

harmonic non-speech tones which share the same pitch, intensity, and duration as in the speech materials (Mandarin tones). Results indicated that the overall scores for the non-speech tone discrimination are better than the speech tone discrimination. They attribute this to the possibility that a more general auditory mechanism might have been applied to process these non-speech tones, which are less complex than speech tones. Therefore, the processing of speech tones may require more cognitive demand than the processing of non-speech tones. Secondly, they also found that Mandarin speakers exhibit CP for both speech and non-speech tones, suggesting a possible transfer link between the experience of native tones and the processing effect of non-speech tones. Similarly, Peng et al. (2010) reported similar findings regarding the categorical perception of non-speech tones. In their study, in the non-speech condition, they found that all groups of listeners (Cantonese, Mandarin, and German) attached the greatest importance to the falling-end in the falling continuum, suggesting that for the non-speech condition, a domain-general auditory mechanism might have been utilized in processing the falling non-speech tones.

Taken together, it can be seen that categorical perception of lexical tones is governed by a more speech specific strategy by tone language speakers (e.g. Sun and Huang, 2012), whereas the categorical perception of non-speech tones is governed by a more domain-general processing strategy as it is considered a lower-level acoustic domain (e.g. Peng et al., 2010). This has motivated further research on categorical perception of lexical tones by musicians who do not speak any tone languages, because previous research has suggested that musical training facilitates tone word learning (e.g. Cooper and Wang, 2012). If this is the case, then it might be expected that musicians would be able to use a music-specific strategy in classifying the tones in Mandarin, resulting in a positive perceptual transfer (e.g. domain-general processing) from music to language.

1.2.4. Summary

While previous research has shown categorical perception in linguistic tones (e.g. Wang, 1976) as well as in non-speech tones (e.g. Xu et al., 2006), very limited studies so far have attempted to examine the categorization of musical tones, particularly the categorical perception of music by tone language speakers. Relevant studies regarding music categorization have focused on musical instrument classification (Hannon, 2009; Lewis & Wieczorkowska, 2007), melodic emotional categorization (Filipic & Bigand, 2005), music triads categorization (Blechner, 1977; Burns & Ward, 1978) or timbre categorization (Loureiro, de Paula & Yehia, 2004).

The importance for examination of CP in music comes from the research findings that tone language speakers generally have enhanced music perception (e.g. Deutsch et al., 2009; Pfordresher & Brown, 2009). If linguistic tone experience facilitates such advantage, then it is possible that tone language speakers are sensitive to musical tones, and very likely that their experience may affect their categorization of musical pitches.

Secondly, examinations of musicians' CP of lexical tones come from the findings that musical training benefits tone language learning (e.g. Cooper and Wang, 2012) and identification (Lee and Hung, 2008). If musical training does facilitate tone language acquisition, then it will be expected that it may facilitate categorical perception of lexical tones, and it is likely that musicians will use their musical experience in the same way Mandarin speakers use their linguistic experience in lexical tone categorization.

1.3. The Present Study

The present study examines tone language speakers' and non-tone language speaking musicians' ability to identify and discriminate linguistic and music tonal pitch in order to investigate the effects of linguistic and musical experience on lexical and melodic pitch categorization.

Based on the review in Section 1.2, the previous research has generally reported facilitative effects of musical experience on language learning (e.g. Lee & Hung, 2008). However, it has not been clear how musical training affects the categorization of lexical tones. On the other hand, research has not been consistent about the role of tone language background in the categorization of melodic pitch (e.g., Bidelman et al., 2013; Peretz et al., 2011). The present study aims at contributing to the existent findings by investigating the interaction between musical experience and linguistic experience on tone categorization in Mandarin and music. This is because the interplay between language and music experiences may be bi-directional. This then allows us to identify whether pitch proficiency in one cognitive domain affects the perception of another, and therefore to further explore the underlying mechanisms for pitch processing – whether it is domain-general, domain-specific, or integration of both (Zatorre and Gandour, 2008; Patel, 2008).

Moreover, the present study also intends to contribute to the understanding of how the transfer effect works in relation to the categorical perception of pitch. For instance, Gibson's Perceptual Learning Theory (1969) argues that the expertise in one area may lead to an improved performance of another cognitive area through practice or experience. Furthermore, from Sloutsky (2003)'s discussion on the development of categorization, she has argued that the development of categorization is grounded in perceptual learning mechanism (e.g. Gibson, 1969). That is, human beings are able to categorize perceptual auditory inputs by extracting the regularities, and further compare them with their own experience. Hence, if this is the case, lexical pitch experience may enhance the categorization of music pitch, whereas music pitch experience may enhance the categorization of lexical pitch. Since research with respect to categorical perception has found that tone language speakers are used to categorizing their own native linguistic tones (Peng et al., 2010; Xu et al., 2006), it is possible, in conjunction with Sloutsky (2003)'s prediction, that musical experience may strengthen the ability to categorize linguistic tones. Similarly, having a linguistic tone experience may provide enhanced benefits for categorizing melodic tones. Therefore, these arguments have motivated the two research questions addressed in this study, as stated below in Section 1.3.1.

1.3.1. Research Questions

The present study aims at answering the following questions which have not been fully addressed previously:

(1) **Effects of musical training on lexical tone perception**: Does musical training facilitate categorical perception of lexical tones?

(2) **Effects of tone language experience on music perception:** Does speaking a tone language facilitate categorization of tones in music?

1.3.2. Overview of Research Design

There are two experimental conditions in this study: linguistic tone categorical perception and music tone categorical perception. Participants include (1) native Mandarin-speaking non-musicians (MN) (2) native English-speaking musicians (EM), and (3) native English-speaking non-musicians (EN). The stimuli include between-category Mandarin tones (Level, Rising, and Falling), between-category musical tones (C4, D4, E4), as well as within-category linguistic and music tone exemplars that were derived from these tones by creating step-wise pitch changes from one tone category to another (e.g. for Mandarin, the within-category stimuli are from Level to Rising and Level to Falling; for Music, the within-category stimuli are from C4 to E4). The participants are asked to perform an AX discrimination task on Mandarin tones as well as the musical tones, followed by an identification task on both. The AX discrimination task (DI) requires the participants to judge whether a pair of two tones sound the same or different. The identification task (ID) is a two-alternative forced choice identification task for which participants are to identify the target linguistic or music tone. In terms of data analysis, accuracy will be computed for discrimination, whereas the sharpness of the identification curve and the boundary location will be computed for identification using logistic regression analysis by the *Generalized Linear Model*.

1.3.3. Hypotheses

(1) **Effects of musical training on lexical tone perception**

Based on previous studies showing a facilitative effect of musical training on non-native tone perception and learning (Cooper & Wang, 2012; Wong et al., 2007), we expect that English musicians will be able to establish new Mandarin tone categories, better than English non-musicians. That is, in terms of between-category tone perception, native English musicians will exhibit more distinct Mandarin tone categorical boundaries than the English non-musicians due to the musicians' trained ability to categorize music tones. In terms of within-category tone discrimination, the native English musicians who are sensitive to subtle pitch differences due to musical training will outperform the native English non-musicians.

(2) **Effects of tone language experience on music perception**

Based on the previous findings of the effects of tone language experience on musical pitch perception (Bidelman et al., 2013; Pfordresher & Brown, 2009; Tervaniemi et al., 2005), the native Mandarin speakers' categorical perception of music tones are expected to be influenced by their native tone language experience. Therefore, the hypotheses are left open.

First, Mandarin non-musicians who utilize a more domain-general mechanism in processing pitch in music (Gibson, 1969; Patel, 2008), are expected to excel at discrimination of the within-category musical tones due to their overall fine-tuned sensitivity to pitch. Alternatively, if they utilize a more speech-specific mechanism in processing the musical tones (Patel, 2008; Peretz et al., 2011), they will not be able to discriminate within-category music tone differences as efficiently as their experience with Mandarin tone categories may decrease their sensitivity to subtle acoustic differences.

Second, regarding between-category perception, research has shown that native Mandarin speakers demonstrate categorical perception in their native tones (Peng et al., 2010; Xu et al., 2006). If domain-general mechanisms are involved in this categorical perception of linguistic tones, then native Mandarin non-musicians may also exhibit distinct categorical boundaries for music tones as the general ability in categorization may be transferred from language to music. Alternatively, if a top-down speech-specific mechanism is used (Tervaniemi et al., 2009), then they will not be able to categorize between-category pitch differences in music as their established Mandarin tone categories may interfere with the categorization of music tones (if the category boundaries of the two domains do not overlap).

Chapter 2. Methodology

In this chapter, the methodological aspect of the present study is discussed. Section 2.1 details the criteria for the participants and the characteristics of the selected participants. Section 2.2 describes the stimuli used for the current study, regarding tone production, recording, and measurement, as well as the creation of tone continua. Section 2.3 gives an overview of the tasks used in the categorical perception experiments – discrimination and identification. Section 2.4 lists the variables for the current study, and Section 2.5 introduces the independent variables and dependent variables in the study. Finally, section 2.6 describes the data analysis methods used regarding the categorical perception data

2.1. Participants

2.1.1. Definition of Native Language

Because the current study was conducted in the multi-lingual city of Vancouver, British Columbia, Canada, the definition of native language must be provided. In this study a native speaker of Mandarin is defined as a person who was born in a country where the first, primary language is Mandarin Chinese (e.g. China or Taiwan), grew up at least to the age of 15 in his or her home country, and came to Canada after the age of 15. This age is the cut-off age in terms of physiological and native language development, and thus used in the field of second language (L2) acquisition research examining the L2 proficiency between late bilinguals and early bilinguals (Flege et al., 2003).

Likewise, native speakers of English must be defined as well. A native speaker of English is defined as a person who was born in Canada, and has not learned a second tone language (e.g. Mandarin, Cantonese, or Thai) in his or her lifetime. If a native English speaker was born to a tone language family, the speaker should not be able to comprehend or speak the tone language.

2.1.2. Definition of Musical Training

The definition of non-musicians seems to vary across studies. Non-musicians should have no more than three years of piano musical training, and should not have learned or played a musical instrument in the past five years (Bidelman et al., 2011; Chandrasekaran et al., 2007; Cooper & Wang, 2012; Hove et al., 2010).

The definition of musicians, however, is considered to be with five or more years of western-style instrument (e.g. Piano preferably) continuously, and they should report that they have played the instrument in the past five years on a regular basis, as of the time of the experiment (Bidelman et al., 2013; Cooper & Wang, 2012; Wong et al., 2007).

2.1.3. Participant Characteristics

A total of 55 participants were recruited at Simon Fraser University, including 17 Mandarin-speaking non-musicians (MN), 19 English-speaking musicians (EM), and 19 English-speaking non-musicians (EN). A summary of their information is provided in Table 1.

The Mandarin-speaking non-musicians (12 male, 5 female; mean age: 26) were all born and raised in China or Taiwan with an average of 1.04 years of musical experience. In addition, they self-reported that they had no other tone language experience or had minimal proficiency in a 2nd tone language (e.g. Cantonese or Taiwanese).

The English-speaking non-musicians (9 male, 10 female; mean age: 28) were all born and raised in Canada with an average of 1.13 years of musical experience. One English-speaking participant was raised in Montreal, Quebec, Canada, and therefore both French and English are the dominant languages of the participant. The remaining participants only spoke English as their native language and had no previous knowledge of a tone language.

The English-speaking musicians were 7 males and 12 females (mean age: 25), and were all born and raised in Canada with an average 10.94 years of musical experience. Thirteen of them reported that piano has been the major instrument during their musical training, although they may have learned multiple instruments at the same time. The remaining participants, however, reported that violin, clarinet, guitar, percussion, or flute was the major musical instrument.

Table 1 - Summary of the Participants' Age and Average Musical Training

Group	Number (M, F)	Age (Min, Max)	Average years of musical training (Standard Deviation)
Mandarin non-musicians	17 (12,5)	26 (20, 34)	1.04 (1.67)
English musicians	19 (7,12)	25 (19, 38)	10.94 (4.69)
English non-musicians	19 (9, 10)	28 (21, 64)	1.13 (1.89)

2.2. Mandarin Tone Stimuli

2.2.1. Recording

The purpose of the recording is to (1) establish a solid acoustic basis for the Mandarin-tone stimuli used for the study, such as fundamental frequency, duration, and amplitude and (2) to provide the stimuli used for the categorical perception experiment.

2.2.1.1 Speakers

Six native Mandarin speakers (three female speakers and three male speakers) from Simon Fraser University were recruited to repeat 32 monosyllabic real Mandarin words (See Appendix A) three times. The recording session took 1 hour to complete and each speaker was compensated CAD$10 for their participation.

The native Mandarin speakers were selected based on the following criteria: (1) They must be born in China or Taiwan (2) The length of residence in Canada must be less than 4 years (3) They must not speak another tone dialect (e.g. Cantonese or Shanghainese) (4) They must not have any speech or hearing impairments.

2.2.1.2 Setting and Equipment

The recording took place in a sound-attenuated recording booth at the Language and Brain Lab, Simon Fraser University. Before the recording session began, the participants were asked to read over the consent form. They were also required to pass a Mandarin quiz to ensure that each Mandarin word was pronounced correctly. The research assistants then instructed them to sit in the recording booth and reminded them not to make noise when the recording began.

The 32 words were presented using *Microsoft PowerPoint*, and the first few slides were detailed instructions and practice sections. The purpose of this was to ensure that each participant followed the procedure, the microphone clearly caught the production, and they did not make any noise during the practice section.

During recording, the research assistants had a recording checklist, in which the order of presented Mandarin words was listed on the checklist, to ensure proper productions of each word. If the participant accidentally mis-pronounced a word/or pronounced the word with noise, then the researchers would be able to track down these words and ask them to repeat the word at the end of the recording section. The entire PowerPoint presentation was repeated three times from the very beginning to the end.

2.2.2. Editing

As mentioned in Section 2.2.1, each participant was asked to read 32 Mandarin words three times. This resulted in a long sound file after recording each participant, which was then automatically segmented into individual word files (250ms each) using a script in Praat (Version 5.3.80) (Boersma & Weenink, 2014).

The amplitude for each sound file was normalized to 65 dB using a script implemented in Praat. This is the amplitude commonly found in the literature of lexical tone categorization by native Mandarin speakers (Zheng et al., 2012; Peng et al., 2010).

2.2.3. Goodness Rating

As mentioned above, the 32 words with three repetitions per speaker were evaluated by two phonetically-trained native Mandarin-speaking students (2 evaluators x 3 speakers = 6 evaluators in total). The objectives for performing the goodness rating are to (1) select the best exemplars for developing the tone continua and (2) ensure that the chosen stimulus is not deviant from Standard Mandarin. Each evaluator was asked to pay attention to the pitch contour of the tone words and rated each word on a scale of five, with one as the worst production and five as the best production of the word. The results from the goodness rating indicate a ceiling score for each word; therefore, the present study decided to go with the most commonly used syllable /yi/ in previous studies (Peng et al., 2010; Xu et al., 2006; Wang, 1976; Zheng et al., 2012).

2.2.4. Mandarin Tone Continuum Creation and Development

Based on the goodness-rating results gathered from the six native Mandarin speakers, three tone words, produced by a male speaker, were selected for the creation of the tone continua used in categorical perception. The selection was based on the onset and offset of fundamental frequency, so that when the sound is normalized the tones will not be too deviant from the naturally-produced tones. For instance, if the average fundamental frequency for a male speaker for the syllable /yi/ is 150 Hz, then we chose the tone words in which the onset and/or offset are not noticeably deviant from 150 Hz. Therefore, the syllable used was /yi1/ with the high flat level tone (一) meaning “one”, with the rising tone /yi2/ (移) meaning “move”, and with the falling tone /yi4/ (益), meaning “benefit”. Figure 1 presents the pitch contour for the best produced tones by the male Mandarin speaker.

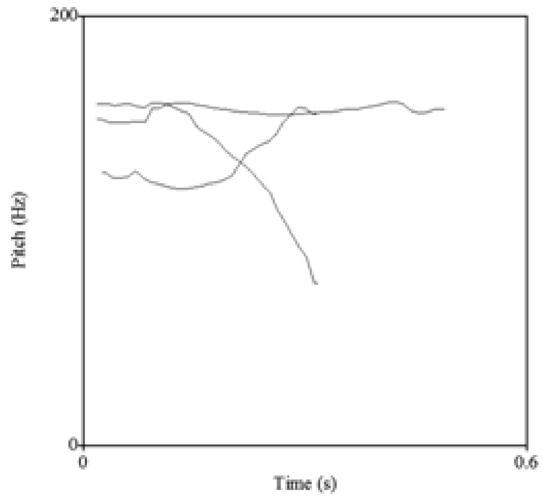


Figure 1 – The three best exemplars of Mandarin tones
1, 2 and 4 produced by a male Mandarin speaker

2.2.4.1 Tone Continua Overview

There are two types of lexical tone continua: level-to-rising contour continuum and level-to-falling contour continuum. Following and adapting a previously established paradigm, the PSOLA (*Pitch Synchronous Overlap and Add*) method was used (Peng et al., 2010; Xu et al., 2006; Zheng et al., 2012), and two Mandarin tone continua for each type were created, resulting in a total of four continua: level-to-rising, rising-to-level, level-to-falling, and falling-to-level. The continua are illustrated in Figure 2.

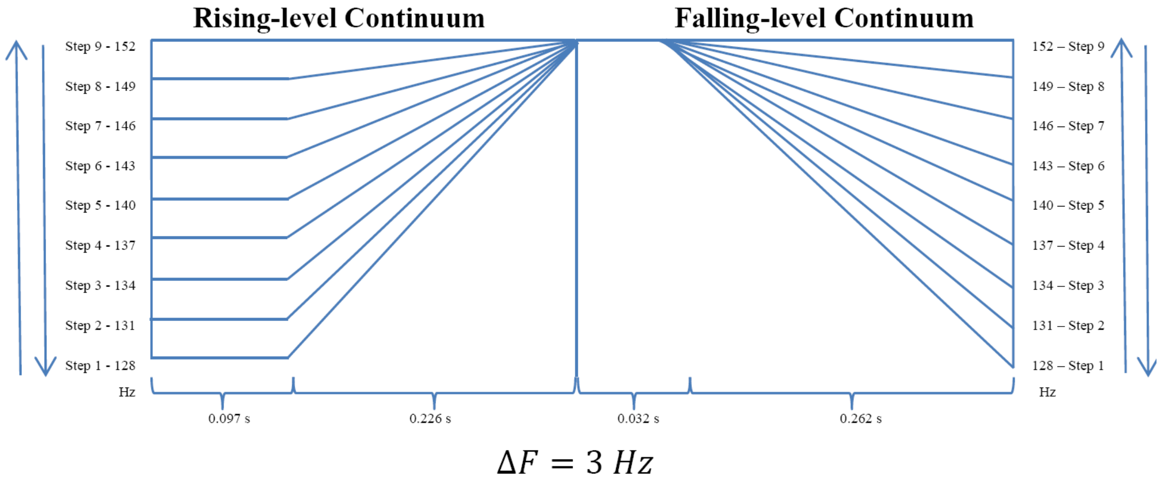


Figure 2 – Mandarin Tone Continua Examples

2.2.4.2 Acoustic Details for the Continua

As shown in Figure 2, the fundamental frequency (F0) for the level tone as well as that for the offset of the rising tone and the onset of the falling tone was fixated to 152 Hz. This value was determined based on the male Mandarin speaker's average onset and offset F0 of the /yi1/ productions, the offset F0 of the /yi2/ productions, and the onset F0 of the /yi4/ productions. Likewise, a fixed value (128 Hz) was used as the onset F0 for the rising tone and the offset F0 for the falling tone, based on the averaged F0 at these points by the above-described male speaker. The duration of all the continua was fixed to 323ms based on the average length of /yi1/, /yi2/, and /yi4/ produced by the male Mandarin speaker as well. Additionally, since Tone 2 (Rising Tone) and Tone 4 (Falling Tone) each involve a turning point in the pitch contour (Moore & Jongman, 1997; Nordenhake & Svantesson, 1983), measurements were conducted to determine the time when the tone starts to rise for Tone 2, and when the tone starts to decline for Tone 4 (See Appendix G for the rising tone example). The measurements show that the speaker's turning point is at 30% of the entire duration for Tone 2, 10% of the entire duration for Tone 4. Therefore, the first 97 ms for Tone 2 continuum remained stable, whereas the remaining 226 ms was considered the rising part. Similarly, the first 32 ms for Tone 4 remained stable, and the remaining 262 ms was considered the falling part. Refer to Appendix H for the acoustic details of the tone productions.

2.2.4.3 Development of Continua

To derive a continuum using PSOLA, the onset of the Tone 2 was varied to obtain a rising-level continuum. Furthermore, by varying the offset of Tone 4, a falling-level continuum was obtained. For Tone 2, with the first 97 ms fixed at 152 Hz, and the remaining 226 ms was varied accordingly by 3 Hz increments, resulted in 9 steps in total. For Tone 4, with the first 32 ms fixed at 152 Hz, and the remaining 262 was varied accordingly by 3 Hz, resulted in 9 steps in total as well.

Two additional continua were created from the opposite direction. Once a level-to-rising continuum was created, a continuum that goes from rising-to-level was also created by applying the same method described above. Additionally, once a level-to-falling continuum was created, a continuum from falling-to-level was created as well.

This particular increment of 3 Hz was chosen because Wang (1976) has claimed that 3 Hz is the best value for determining a psychoacoustic boundary for two different tonal categories, and this frequency was then later supported by Liu (2013)[2] as the just noticeable difference (JND) for Mandarin tones. Lastly, the intensity of the stimuli was normalized to 65 dB (Zheng et al., 2012; Peng et al., 2010) to ensure that all stimuli have the same volume.

After the 36 tokens (9 steps x 4 continua) were created, the end-tokens (e.g. a typical flat tone (or a level tone), a synthesized rising tone, a typical rising tone, a synthesized flat tone, a synthesized falling tone, a typical falling tone, and a synthesized flat level tone) were evaluated by two native Mandarin phoneticians in order to ensure the typicality of these tones in Mandarin. All the end tones were correctly identified and judged as “typical” by these evaluators.

[2] In Liu (2013)’s study, he found that the range of JND in Mandarin tones is 3Hz to 14 Hz. 3 Hz is chosen because it corresponds to Wang (1976)’s value.

2.3. Music Tone Continuum

For the musical stimuli the researcher created a three note arpeggio using *Garageband* (see Bidelman et al., 2011b). To control for familiarity of the arpeggios, a universally familiar arpeggios was used, such as C4 (261.6Hz), D4 (293.6 Hz), and E4 (329.6 Hz). The within-category (mis-tuned D4) sequences was created from 261.6 Hz C4 to 329.6 Hz E4, each differed in 4 Hz. 4 Hz was chosen because Harrison (1996) has reported that this is the just noticeable difference (JND) for individuals in terms of tone perception. Each note (C, D, E) is 100 ms in duration, and therefore one arpeggio is 300 ms in duration (Bidelman et al., 2011b), which is comparable to the tones used for Mandarin. A detailed summary is provided in Figure 3 below.

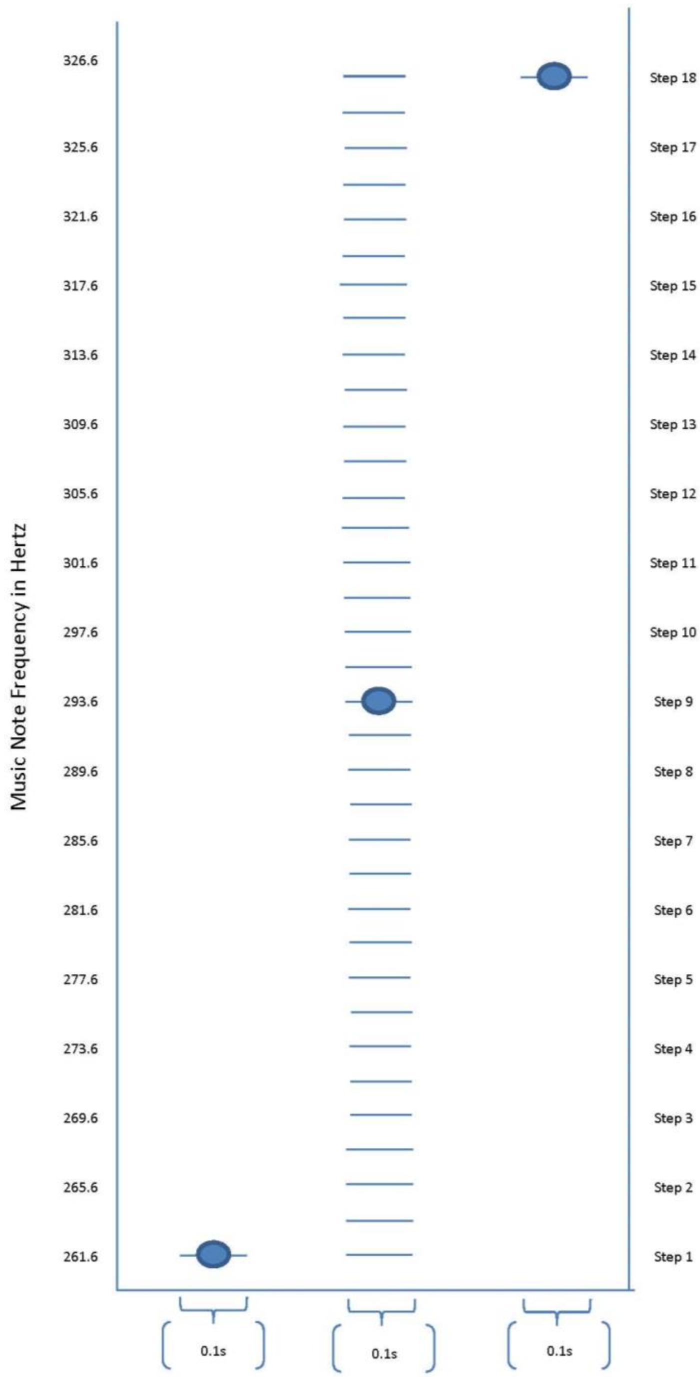


Figure 3 – Music Tone Continuum – the note C starts at 261.6 Hz, the note D starts at 293.6 Hz, and the note E starts at 326.6 Hz. The CCE-to-CDE continuum was created by varying the fundamental frequency of the note D down, whereas the CDE-to-CEE continuum was created by varying the fundamental frequency of the note D up.

2.4. Tasks

2.4.1. Discrimination

An AX discrimination task for Mandarin tones and musical tones was prepared. For the Mandarin tone condition, there were 9 same pairs with each step (9 in total) repeating itself (e.g. [Step 1, Step 1], [Step 2, Step 2]) and 16 different pairs with 8 combinations (e.g. [Step 1, Step 2], [Step 2, Step 3]) and the reverse pairs (e.g. [Step 2, Step 1], [Step 3, Step 2]), resulting in 8 x 2 = 16 pairs (e.g. [Step 1, Step2], [Step2, Step3], [Step 2, Step1], [Step 3, Step 2]) per continuum. For the music tone condition there were 18 same pairs with each step (18 in total) repeating itself (e.g. [Step 1, Step 1], [Step 2, Step 2]) and 34 different pairs with 17 combinations (e.g. [Step 1, Step 2], [Step 2, Step 3]) and the reverse pairs (e.g. [Step 2, Step 1], [Step 3, Step 2]), resulting in 17 x 2 = 34 pairs. For both conditions each pair was repeated 3 times, resulting in 75 pairs (Same Pairs: 9 x 3 = 27; Different Pairs: 16 x 3 = 48; Total: 48 + 27 = 75) per continuum for Mandarin Tones and 156 pairs for Music Tones (Same Pairs: 18 x 3 = 54; Different Pairs: 34 x 3 = 102; Total: 54 + 102 = 156). The participants were instructed to judge whether two tones in a pair sounded the same or different. The inter-stimulus interval (ISI) for the discrimination pairs was set at 250 ms, where Pisoni (1976) has found that this is the ideal interval for reaching maximum discrimination scores.

For the Mandarin condition, the rising-level continuum (level to rising, and rising to level) and the falling-level continuum (level to falling, and falling to level) were blocked separately. The presentation of either continuum was counter-balanced, in which some participants received the rising-level continuum first and the falling-level continuum later, or the vice versa. For the music condition, the presentation order for either CCE-to-CDE continuum or CDE-to-CEE continuum was counter-balanced as well, in which some participants received CCE-to-CDE first and CDE-to-CEE later, or the vice versa.

2.4.2. Identification

The participants were asked to perform a 2AFC (2-way-alternative-forced-choice) identification task for linguistic tones as well as music tones.

For Mandarin tones, after a tone is displayed, listeners were to judge whether the tone was a high-level flat tone (Sound 1), or a rising/falling tone (Sound 2). The task for melodic tones is a contextual melody categorization. A three-note arpeggio (CXE) was presented, where X is the varied D note. They were asked to perform a 2AFC identification task and determine whether the sequence sounds like (1) C-D-E (Do-Re-Mi) or C-E-E (Do-Mi-Mi), and (2) C-D-E (Do-Re-Mi) or C-C-E (Do-Do-Mi).

2.4.3. Procedures

As mentioned above, there was a Mandarin tone condition and a music tone condition, and each condition has a discrimination task [DI] and an identification task [ID]. The order of condition presentation was counterbalanced across listeners so that half of the listeners took the Mandarin tone condition first and the other half took the music tone condition first. This was to avoid any potential bias resulting from an order effect such as a prior condition leading to the improvement of another. Also, before each condition the participants received a short 1-2 minute training task in order to be familiarized with the stimuli in each condition. However, given that previous research has suggested that short-term memory may greatly affect the identification function (Caruso & Detterman, 1983; Cowan, 1984; Wickelgren, 1965), the participants always performed the DI task first and then the ID task in order to control for the effect of short-term memory on sound identification. Closer to the end of the experiments a short music aptitude discrimination task, modified from Gordon (1988)'s Music Aptitude Profile, was administered. The task contained five tonally-changed sequences and five tonally-unchanged sequences. The Mandarin DI task took approximately 30 minutes, whereas the Mandarin ID task took approximately 15 minutes. The music DI, however, took 15 minutes, and the music ID task took around 10 minutes. The entire experiment including instructions, practice, and the music aptitude task took about 1.5 hours. Refer to Appendix I for an example.

2.5. Variables

For identification, for each condition each identification condition, the independent variables included Group (English Musicians [EM], English non-musicians [EN] and Mandarin non-musicians [MN]) as the between-subject factor, and Continuum Steps (Mandarin: Step 1 – Step 9; Music: Step 1 – Step 18) and Directions (Mandarin rising tone: level-to-rising and rising-to-level; Mandarin falling tone: level-to-falling and falling-to-level) as the within-subject factors. The dependent variable is binary (either Sound 1 or Sound 2 for Mandarin tone condition; either CDE or CEE; CDE or CCE for music tone condition). For discrimination, the independent variables include Group (EM, EN and MN) as the between-group factor, and Tone Pairs, and Directions as the within-subject factors. The dependent variable is the accuracy for correctly judging the sameness of the pairs.

2.6. Data Analysis

2.6.1. Identification Function

Following and adapting the previous procedures for categorical analyses of tones (Peng et al., 2010; Xu et al., 2006; Zheng et al., 2012), we used *the Generalized Linear Model* to investigate the sharpness of the boundary and the boundary location. *The Generalized Linear Model* can be considered a type of logistic regression because we are interested in how a participant classifies a sound (e.g. Sound 1 or Sound 2) in terms of the boundary and the location, and how the participant's responses fit into a sigmoid curve (or a categorical curve). Therefore, *the Generalized Linear Model* computes the scores for identification based on the following equation in (1):

$$(1)\ \ln\left(\frac{PI}{1-PI}\right) = b0 + b1x$$

PI in (1) is the probability for the identification options provided to the participants (e.g., Sound 1 or Sound 2), and therefore it is set to be 0.5 (50%, because there are two response options). The variable *x* refers to a step number in a continuum, indicating the potential categorical boundary for a continuum. The variable b_1 refers to the slope of the boundary between the two categories in a continuum, considered as the sharpness of the categorical boundary in the continuum (Xu et al., 2006). The variable b_0 refers to the intercepts of the regression equation. Therefore, to compute the categorical boundary X_{cb} based on the identification responses gathered from each listener for each group, the following equation (2) is derived from (1).

$$(2)\ xcb = -\frac{b0}{b1}$$

Three predictor variables GROUP (MN, EM, EN), Direction (Level-to-Rising and Rising-to-Level; Level-to-Falling and Falling-to-Level), and Step (Mandarin: Step 1 to Step 9; Music: Step 1 to Step 18) were submitted to JMP using the Generalized Linear Models to examine the fitness of the categorical curves for each tone continuum and each listener. Using this model, we thus computed the slope (category boundary sharpness) and the category boundary position at the 50% crossover point for each continuum per subject can be computed using this model, and then submitted for subsequent analyses of variance (ANOVAs) to compare group differences.

2.6.2. Discrimination Accuracy

The discrimination accuracy (scores) (P) is computed based on the following equation (3), as used in previous studies (e.g. Peng et al., 2012; Xu et al., 2006).

$$(3)\ P = P(\mathrm{S}|S) * P(S) + P(\mathrm{D}|D) * P(D)$$

In (3), P(S) refers to the percentage of Same Pairs, whereas P(D) refers to the percentage of Different Pairs. Therefore, the scores for of obtaining either P(S) or P(D) are set to be 0.5 because the probability of getting same or different responses is 50%. Furthermore, P("S"|S) refers to the percentage of "Same" responses of all the same pairs, whereas P("D"|D) refers to the percentage of "Different" response among all the different pairs. The reason for using this equation is that this equation pays equal attention to the scores for the same pairs and the scores for the different pairs. For instance, in any given pair (Step 1, Step 2), the equation takes care of how often a participant correctly perceives the same pairs (Step 1, Step 1) and (Step 2, Step 2), and how often the participant correctly perceives the different pairs (Step 1, Step 2) and (Step 2, Step 1). Therefore, a higher score indicates that a participant is more likely to correctly judge a given different pair and a given same pair.

Based on this equation, the discrimination scores for each pair in each continuum could be computed per subject, and these scores were subsequently submitted to JMP 10 (Version 6.2.9200.0) for the analyses of variance to examine group differences.

2.6.3. Relation between Discrimination and Identification

The predicted discrimination scores were derived from the identification responses in each continuum using the formula in (4). In order to claim categorical perception, it is necessary to justify that listeners can identify as well as discriminate the target stimuli. That is, the identification boundary location (indicating detection of category boundary) should correspond to the discrimination peak location (indicating high accuracy in differentiating the between-category pairs). Whether such correspondence exists can be evaluated by examining the extent to which identification boundary location may predict the discrimination scores, especially the prediction of the peak. Therefore, the following equation (4) is used to compute the predicted discrimination accuracy (**P'**) (Liberman et al., 1957; Pollack and Pisoni, 1971; Xu et al., 2006)

$$(4)\ P' = \frac{1 + (PA - PB)^2}{2}$$

The predicted discrimination score for a pair (e.g. Pair (Step 1, Step 2) is calculated based on the actual identification responses in a continuum. For example, to derive the predicted score for the pair (Step 1 and Step 2), P_A will refer to the number of Sound 1 responses when the participant perceive Step 1, and P_B will refer to the number of Sound 1 responses when the participant perceive Step 2. Imagine that a participant makes no Sound 1 responses among the three repetitions (0/3) for Step 1, but the participant makes a Sound 1 response among the three repetitions (1/3) for Step 2. Then, the predicted score for the pair (Step 1, Step 2) will be 0.56 (0.56 = 1+(0-1/3)^2/2) . This then indicates that the predicted score for Pair (Step 1, Step 2), derived from the identification, is 0.56, suggesting that the participant will be expected to perform at an above chance level for the pair (1 & 2) based on the identification responses. Then, a subsequent analysis of variance was used to will examine whether there is a group, pair, or direction difference for the predicted discrimination scores.

Secondly, another definition of observing CP is to examine whether the discrimination peak location corresponds to the identification boundary location. In order to obtain this relationship, a regression analysis was used to determine such correspondence. Based on the discrimination scores, the corresponding predicted categorical position was computed (Schneider, Dogil & Mobius, 2009; Schneider & Lintfert, 2003). For instance, if the discrimination peak was located in the pair (5&6), then the corresponding predicted identification boundary location would be 5.5. Then, a regression analysis is done using the predicted boundary locations (e.g. 5.5) and the obtained discrimination peak locations.

2.6.4. Summary of Analysis

In sum, the identification data was analyzed by the Generalized Linear Model in JMP. *The Generalized Linear Model* then provided the slope information, the intercept, and then the 50% crossover boundary location for each listener. Then, a further analysis of variance was conducted to compare the group differences on the slope information and the boundary location for each participant and continuum. Next, the discrimination accuracy, on the other hand, was computed by the equation in (3). A further analysis of variance was also conducted to compare the group differences on the discrimination accuracy. Predicted discrimination scores, which were derived from the identification data, were computed as well, using the equation in (4). A further ANOVA was conducted to see if the predicted scores have any group differences. Finally, to establish the relationship between the predicted boundary locations and the actual peak boundary locations, a regression analysis was conducted to see whether the two locations (predicted versus obtained) were aligned.

Chapter 3. Results

In this chapter we will report the results of the present study. Section 3.1 presents the categorical perception findings from Mandarin tones, divided by tasks and continua. Section 3.2 reports the findings from the music condition, divided by tasks.

3.1. Mandarin Tones

3.1.1. Discrimination Task

A mixed 3-way analysis of variance (ANOVA) with Pairs (Pair1&2; Pair2&3; Pair3&4; Pair4&5; Pair5&6; Pair6&7; Pair7&8; Pair8&9), Directions (Rising-level continuum: level-to-rising and rising-to-level; Falling-level continuum: level-to-falling and falling-to-level) as repeated measures, and Group (English Musicians [EM], English non-musicians [EN] and Mandarin non-musicians [MN]) as a between-subject factor, was submitted to JMP for analysis. The measurement for the discrimination task is the accuracy, ranging from 0 to 1. A score of 0 indicates that the participant have incorrectly judged for all the same pairs and different pairs, whereas a score 1 indicates perfect discrimination for the pairs.

Rising-level Continuum

Figure 4 shows the discrimination accuracy for all of the three groups across the pairs. It can be seen that English musicians slightly outperformed Mandarin non-musicians and English non-musicians, and the peak is located at the Pair (5&6).

For the rising-level continuum, a marginally significant main effect of Group was obtained [$F(2, 816) = 2.9325$, $p = 0.0569$], indicating that there was a tendency of group differences on the Mandarin discrimination scores across pairs and directions. Post Hoc

Tukey adjusted pairwise comparisons among the groups indicated that English musicians (M = 0.542; SD = 0.0078; p = 0.0569) marginally performed better than Mandarin non-musicians (M = 0.537; SD = 0.0075). Other pairwise group comparison does not yield any significant differences. Furthermore, there is no significant main effect of Direction [F (1, 816) = 0.2207, p>0.05], indicating that the performance in the discrimination task did not differ by the direction of continuum across the groups. However, a significant main effect of Pairs was also found [F (7, 816) = 0.66547303, p<0.0001]. Post hoc Tukey HSD pair-wise comparisons showed that the accuracy for the pair (5&6) (M=0.599; SD = 0.0127; p<0.0001) is significantly higher than and the rest of pairs (M=0.521; p>0.05), indicating the discrimination peak location. For detailed comparisons, refer to Appendix C. Moreover, the interaction between Group x Pairs x Direction was not significant [F (14, 816) = 0.2992, p=0.9941]. There is no significant interaction for Group x Direction [F (2, 816) = 0.4820, p=0.6177], Group x Pairs [F (14, 816) = 0.5893, p=0.8748], and Direction x Pairs [F (7, 816) = 0.8427, p=0.5520].

Because there is a peak location at the pair (5&6), a further 1-way ANOVA using Group as the between subject factor and the discrimination scores at Pair (5&6) as the dependent variable was submitted to JMP to determine whether there is any group difference at the peak. Results showed that there is no significant effect of Group in the peak [F (2, 105) = 0.0950, p=0.9095 (p>0.05)], indicating that the peak scores do not yield any group differences.

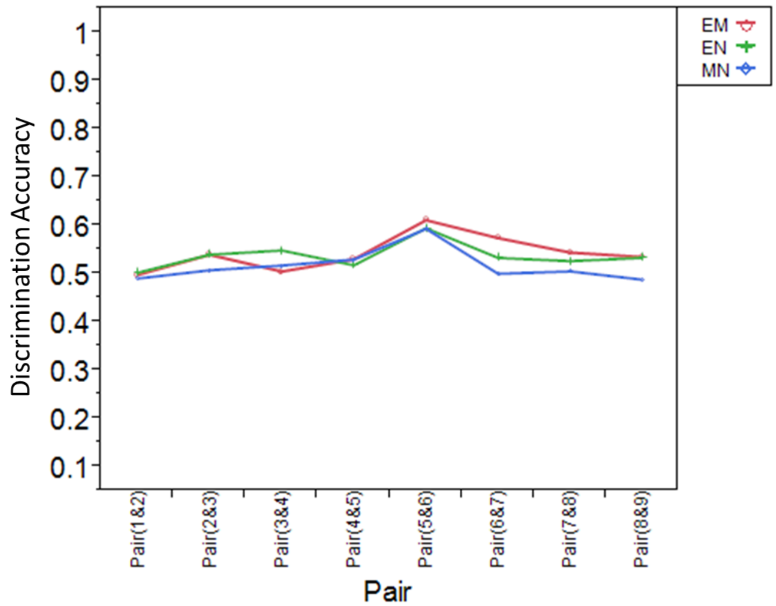


Figure 4 – Average Rising Discrimination Scores By Group. The y-axis is the discrimination score range from 0 to 1, with 1 being perfect discrimination of the pairs, and 0 is poor discrimination of the pairs. This figure has shown that the peak location is at the pair (5&6) for the rising-level continuum.

Falling-level Continuum

Figure 5 shows the discrimination accuracy for all of the three groups across the pairs. It can be seen that English musicians slightly outperformed Mandarin non-musicians and English non-musicians, and the peak is located at the Pair (5&6) as well.

For the falling-level continuum, a significant main effect of Group was obtained [F (2, 816) = 3.9327, p=0.02], indicating that there is a significant overall group difference on the overall Mandarin discrimination scores. Post Hoc Tukey adjusted pairwise comparisons indicated that English musicians (M = 0.552; SD= 0.0079) significantly performed better than Mandarin non-musicians (M = 0.521; SD = 0.0082; p<0.05). Other pairwise group comparison does not yield any significant differences. There is no significant main effect of Direction [F (1, 816) = 0.2207, p=0.3532]. Finally, a significant main effect of Pairs was also found [F (7, 816) = 10.0381, p=0.0001]. Post hoc Tukey HSD pair-wise comparisons showed that there is a significant difference between the pair (5&6) (M = 0.6350; SD = 0.0129; p<0.001) and the rest of pairs. For detailed comparisons, refer to Appendix D.

The interaction between Group x Pairs x Direction is not significant [F (14, 816) = 1.4772, p>0.05]. There is no significant interaction for Group x Direction [F (2, 816) = 0.7251, p=0.1131], or Group x Pairs [F (14, 816) = 0.9926, p=0.4586]. However, there is a significant interaction between Direction x Pairs [F (7, 816) = 0.7251, p=0.0051].

A further 1-way ANOVA using Group as the between subject factor and the discrimination scores at Pair (5&6) as the dependent variable was submitted to JMP to determine whether there is any group difference at the peak. Results showed that there is no significant effect of group difference in the peak [F (2, 105) = 2.5031, p=0.0867].

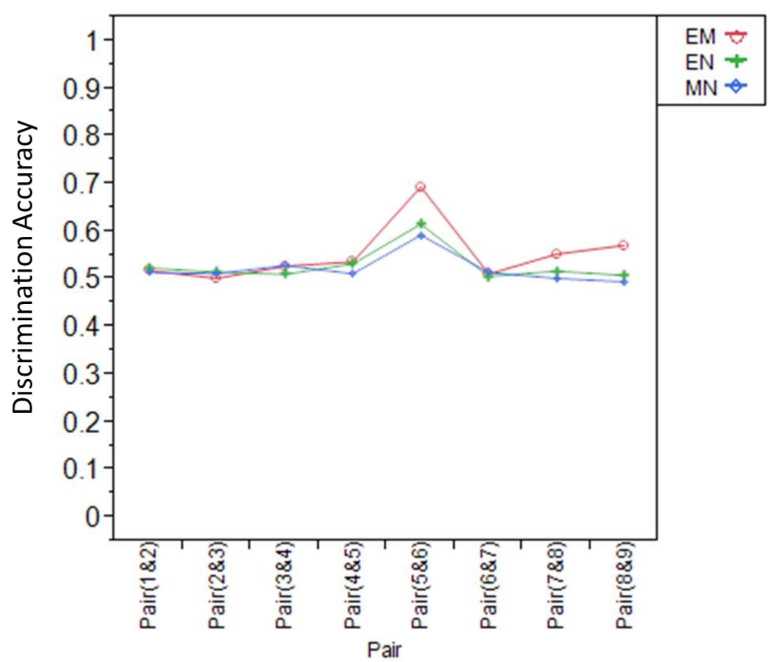


Figure 5 –Average Falling Discrimination Scores By Group. The y-axis shows the discrimination score range from 0 to 1, with 1 being perfect discrimination of the pairs and 0 is poor discrimination of the pairs. This figure has shown that the peak location is at the pair (5&6) for the falling-level continuum.

As shown in Figure 4 and Figure 5, the Mandarin discrimination results indicate that English musicians always outperform English non-musicians, followed by Mandarin non-musicians, in both the rising-level and falling-level continuum. In terms of the peak location, the Pair (5&6) was found to be the peak location for all the groups. A further one-way ANOVA, however, does not show any group differences on the peak location.

3.1.2. Identification Task

The Generalized Linear Model – Logistic Regression

The purpose of using the Generalized Linear Model to analyze the identification is to examine whether the binary identification responses from all groups can fit into the sigmoid shape of the logistic functions. The estimated slopes, intercepts, and categorical boundary positions between different groups were also computed using the Generalized Linear Model. In terms of slopes, a higher slope value indicates a sharper boundary. Moreover, the boundary location ranges from 1 to 9, in which 1 represents the rising end or the falling end, and 9 represents the level end. For instance, a boundary location value of 8.6 means that there are more rising or falling responses, pointing that the location is closing to the level end. A summary of these coefficients are presented in Table 2. These values were used to plot Figure 6 and Figure 7.

The logistic regression in the Generalized Linear Model indicated a three-way interaction between Group, Step Number, and Direction ($p<0.0001$), meaning that the three predictors (Group, Step, and Direction) might affect the fitness of the sigmoid shape. Therefore, further analyses were conducted to examine the sharpness of the categorical (slope) boundary for each subject and the location of the categorical boundary to see if there is any group difference.

Table 2 - A Summary of Intercepts, Slopes, and Category Boundary Positions

Group	Direction	b_0 - intercepts	b_1 - Slopes	$X_{cb} = -b_0/b_1$
EM	Level-to-Rising (1to2)	-4.018	0.716	5.614
EN		-2.202	0.435	5.058
MN		-10.522	2.029	5.186
EM	Rising-to-Level (2to1)	-3.147	0.643	4.888
EN		-1.538	0.343	4.473
MN		-7.467	1.584	4.715
EM	Level -to-Falling (1to4)	-4.460	0.804	5.550
EN		-2.069	0.486	4.261
MN		-2.950	0.670	4.404
EM	Falling-to-Level (4to1)	-2.729	0.451	6.047
EN		-1.487	0.199	7.486
MN		-3.572	0.411	8.688

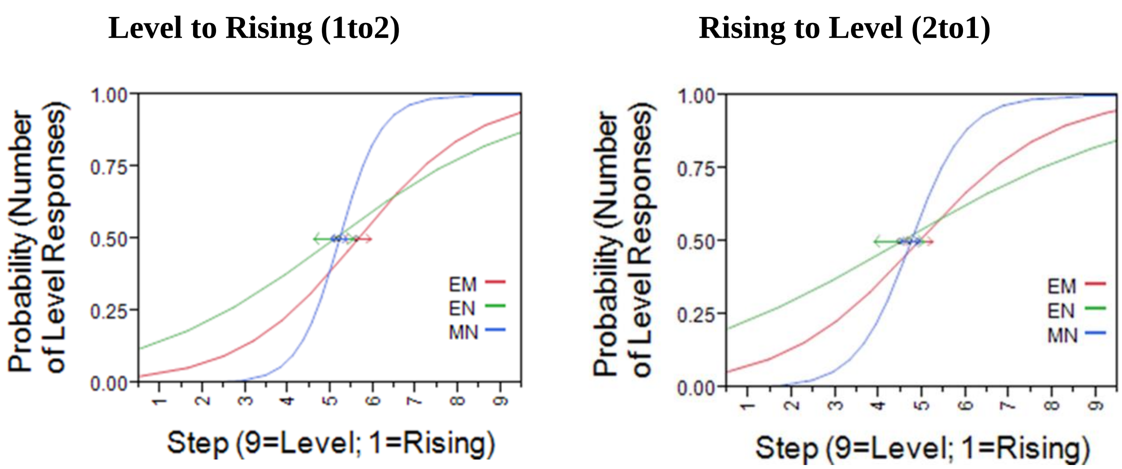


Figure 6 – Logistic Identification Curves for the Rising-level Continuum Per Group – the x axis indicates the boundary range (from 1 to 9), whereas the y axis indicates the percentage of sound 1 (level responses).

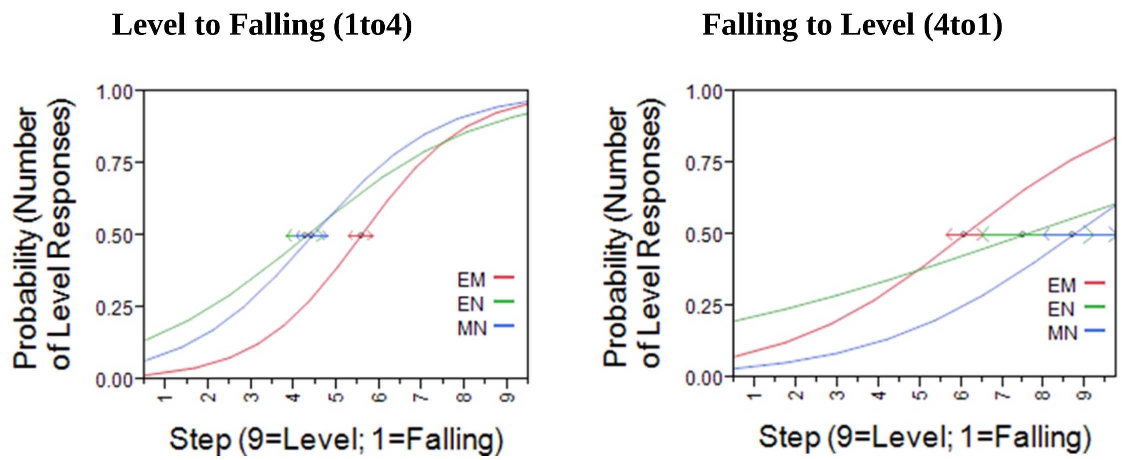


Figure 7 – Logistic Identification Curves for the Falling-Level Continuum Per Group – the x axis indicates the boundary range (from 1 to 9), whereas the y axis indicates the percentage of sound 1 (level responses).

Sharpness of Categorical Boundary

The slope (the measurement) for each subject and each continuum was computed comparing the sharpness of the category boundary between the groups. A 2-way ANOVA using Group (EM, EN, MN) as the between subject factor and Direction as the within-subject repeated measures was conducted.

There is a significant main effect of Group for both of the rising and falling continua [Rising: $F(2, 208) = 10.9083$, $p=0.0001$; Falling: $F(2, 208)=6.8298$, $p=0.0016$ Post Hoc Tukey adjusted pairwise comparisons for the rising continua indicated that Mandarin non-musicians (1to2: 2.029; 2to1: 1.584) show a significantly sharper categorical boundary than English non-musicians (1to2: 0.435; 2to1: 0.343, $p<0.0001$). English musicians show a significantly sharper categorical boundary (1to2: 0.716; 2to1: 0.643, $p<0.05$) than English non-musicians Post Hoc Tukey adjusted pairwise comparisons for the falling continua indicated that Mandarin non-musicians (1to4: 0.67; 4to1: 0.411) show a significantly sharper categorical boundary than English non-musicians (1to4: 0.486; 4to1: 0.199, $p<0.05$). English musicians show a significantly sharper categorical boundary (1to4: 0.804; 4to1: 0.451, $p<0.05$) than English non-musicians. Both Mandarin non-musicians and English musicians' sharpness of categorical boundary does not significantly differ. However, there is no significant effect of Direction for both of the rising-level or falling-level continuum [Rising: $F(1, 208) = 3.1087$, $p=0.0808$; Falling: $F(1, 208) = 3.8928$, $p=0.0511$]. There was no significant Group x Direction interaction for both rising and falling continua [Rising: $F(2, 208) = 2.2142$, $p=0.1144$; Falling: $F(2, 208) = 1.3388$, $p=0.2666$].

Figure 8 presents the group comparisons for each continuum. It can be seen that Mandarin non-musicians generally demonstrate a sharper boundary, followed by English musicians. The English non-musicians' boundary is the least sharp.

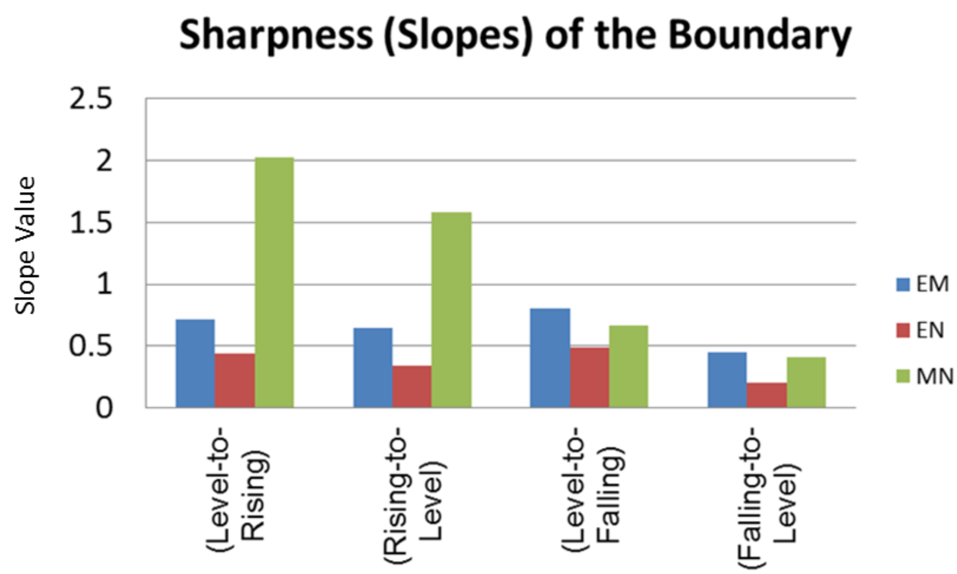


Figure 8 – Average Slope for Each Group in Each Continuum – The bar chart shows that English musicians have a significant sharper boundary than English non-musicians, whereas Mandarin non-musicians have a significantly sharper boundary than English non-musicians for all continua. A higher value indicates that the slope is steeper or sharper, and it is therefore more categorical.

Position of Categorical Boundary

The position of categorical boundary [CB Location] for each subject and each continuum was computed for comparing the boundary location between the groups. A 2-way ANOVA using Group (MN, EM, EN) as the between subject factor and Direction as the within-subject repeated measures was conducted. There is no significant Group x

Direction interaction for either the rising-level or falling-level continuum [Rising: F (2, 95) = 0.0852, p=0.9184; Falling: F (2, 72) = 2.5443, p=0.8056].

For the rising continua, there is no significant main effect of Group [F (2, 95) = 0.7768, p=0.4628]. However, for the falling continua, there is a significant effect of Group on the position of categorical boundary [F (2, 72) = 3.6122, p=0.0320]. Post Hoc Tukey adjusted pairwise comparisons indicates that English musicians' CB location (1to4: 5.55; 4to1: 6.047, p<0.05) occurs significantly towards the level tone (indicating more falling-tone-biased responses), compared to English non-musicians' CB location (1to4: 4.261; 4to1: 7.486, p<0.05), which occurs towards the falling tone (indicating more level-tone-biased responses). However, there is no location difference between English musicians and Mandarin non-musicians, nor is there a difference between English non-musicians and Mandarin non-musicians.

Both of the rising and falling continua exhibited significant main effects of Direction on the position of categorical boundary [Rising: F (2, 95) = 5.1795, p=0.0251; Falling: F (2, 72) = 5.9421, p=0.0173]. Post Hoc Tukey adjusted pairwise comparisons indicated that the CB location for the rising-to-level continuum occurs significantly towards the level tone, compared to the level-to-rising continuum, which occurs significantly towards the rising tone. For the falling continua, Post Hoc Tukey adjusted pairwise comparisons indicated that the CB location for the falling-to- level continuum occurs significantly towards the level tone, compared to the level-to-falling continuum, which occurs significantly towards the falling tone.

Figure 9 and 10 summarize the boundary location for both of the rising and falling continua. For the rising-level continuum, it can be seen that the CB locations are similar among the three groups. For the falling-level continuum, however, it can be seen that the falling-to-level continuum is closer to Step 9, indicating that there are more falling responses. However, the level-to-falling continuum is closer to Step 1, indicating that there are more level responses.

Rising-level Continuum

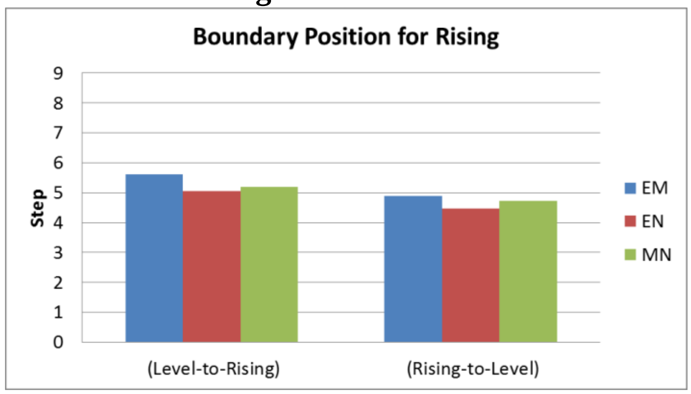


Figure 9 – Average Rising CB Location Per Group

Falling-level Continuum

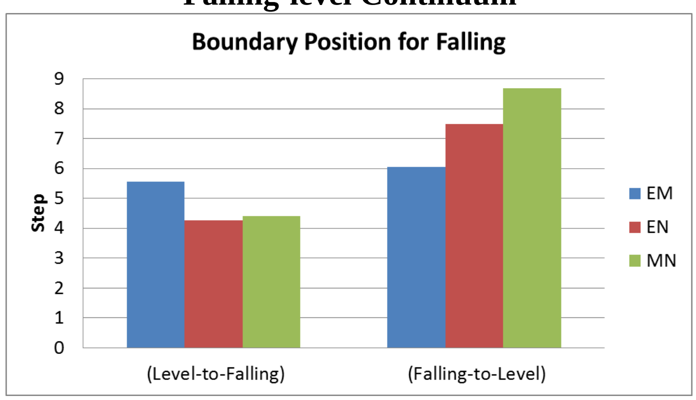


Figure 10 – Average Falling CB Location Per Group

3.1.3. Correlation of discrimination and identification

Further analyses were conducted to examine how well the identification function predicts the discrimination peak. To do so, predicted discrimination scores were derived from the identification functions. Then, a mixed 3-way ANOVA was conducted to examine whether the predicted scores contain any group differences. Secondly, the peak locations were computed from the predicted discrimination scores, and a correlation analysis was conducted to examine whether the predicted peak locations coincide with the obtained discrimination peak locations.

3-way ANOVA for Predicted Discrimination Scores

A mixed 3-way ANOVA with Pairs (Pair1&2; Pair2&3; Pair3&4; Pair4&5; Pair5&6; Pair6&7; Pair7&8; Pair8&9), Directions (Mandarin: level-to-Rising and Rising-to-level; level-to-Falling and Falling-to-level) as repeated measures, and Group (EM, EN, MN) as a between-subject factor, was submitted to JMP for analysis. The dependent variable is the predicted discrimination scores derived from the identification function. Similar to the obtained discrimination scores, a score of 0.5 indicates that the participant is expected to perform at the chance level. A score of 1 indicates that the participant is expected to perform perfectly for discriminating the pairs.

Rising-level Continuum

Figure 11 shows the predicted discrimination accuracy for all of the three groups across the pairs. It can be seen that Mandarin non-musicians slightly outperformed English musicians and English musicians, and the peak is only evident for Mandarin non-musicians at the Pair (5&6).

For the rising-level continuum, No significant main effect of Group was obtained [F (2, 832) = 0.3952, p=6737], indicating that there was not a significant overall group difference on the predicted Mandarin discrimination scores across pairs. There is also no significant main effect of Direction [F (1, 832) = 0.2910, p=0.5897], indicating that the performance in the discrimination task did not differ by the direction of continuum across the groups. However, a significant main effect of Pairs was also found [F (7, 832) = 0.828,

$p=0.0001$]. Post hoc Tukey HSD pair-wise comparisons showed that there is a significant difference between the pair (5&6) ($M = 0.608$; $SD=0.008$; $p<0.001$) and the rest of pairs. For detailed comparisons, refer to Appendix E.

There is a significant three-way interaction between Group x Pairs x Direction [$F(14, 832) = 2.1662$, $p=0.0076$]. A significant 2-way interaction between Group x Pair was also obtained [$F(14, 832) = 2.7421$, $p<0.05$]. However, there is no significant interaction for Group x Direction [$F(2, 832) = 0.7817$, $p>0.05$], or Direction x Pairs [$F(7, 832) = 0.8427$, $p>0.05$]. Post Hoc Tukey adjusted pairwise comparisons for Group x Pair indicated that Mandarin non-musicians significantly outperform English musicians and English non-musicians in the Pair (5&6) ($p<0.05$). In the Pair (6&7) and (7&8), English musicians significantly outperform Mandarin non-musicians ($p<0.05$). Furthermore, for Mandarin non-musicians, the scores in Pair (4&5) and Pair (5&6) are significantly higher than the rest of the pairs ($p<0.05$). For English non-musicians, the score in Pair (4&5) is significantly higher than the Pair (1&2), Pair (2&3), and Pair (6&7) ($p<0.05$). The score in Pair (5&6) is significantly higher than the Pair (1&2), Pair (2&3), Pair (6&7), and Pair (7&8) ($p<0.05$). For English musicians, the score in Pair (4&5) is significantly higher than Pair (1&2), Pair (2&3), and Pair (8&9) ($p<0.05$). The score in Pair (5&6) is also significantly higher than Pair (1&2), Pair (2&3), Pair (3&4), and Pair (8&9) ($p<0.01$). The score in Pair (6&7) is higher than Pair (1&2), Pair (2&3), and Pair (8&9) ($p<0.01$).

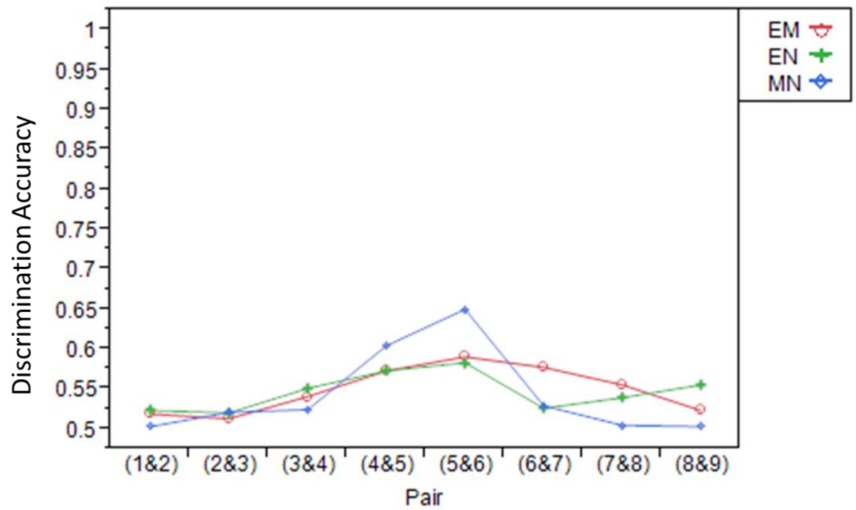


Figure 11 - Average Predicted Rising Discrimination Scores By Groups. The y-axis shows the discrimination score range from 0 to 1, with 1 being perfect discrimination of the pairs and 0 is poor discrimination of the pairs. This figure shows that the peak locations are not consistently at the same pairs.

Falling-level Continuum

Figure 12 shows the predicted discrimination accuracy for all of the three groups across the pairs. It can be seen that the peak is evident at the Pair (6&7) for English musicians. However, the peak is evident at the Pair (5&6) for Mandarin non-musicians and English non-musicians.

For the falling-level continuum, no significant main effect of Group was obtained [$F(2, 832) = 1.1266$, $p=0.3246$], indicating that there was not a significant overall group difference on the Mandarin predicted discrimination scores. There is also no significant main effect of Direction [$F(1, 832) = 3.4474$, $p=0.0637$], indicating that the performance in the discrimination task did not differ by the direction of the continua across the groups. However, a significant main effect of Pairs was also found [$F(7, 832) = 9.1827$, $p=0.0001$]. Post hoc Tukey HSD pair-wise comparisons showed that the pairs (4&5) [$M=0.5458$, $p<0.05$], (5&6) [$M=0.5762$, $p<0.0001$], and (6&7) [$M=0.5564$, $p<0.0001$] are significantly higher than the rest of pairs. For detailed comparisons, refer to Appendix F.

There is no significant three-way interaction between Group x Pairs x Direction [F (14, 832) = 0.8412, p=0.6241]. There is also no significant interaction for Group x Direction [F (2, 832) = 0.3613, p=0.6969]. However, a significant 2-way interaction between Group x Pair was obtained [F (14, 832) = 2.6339, p=0.0009]. Post Hoc Tukey adjusted pairwise comparisons for Group x Pair indicated that at the Pair (6&7), English musicians significantly outperformed English non-musicians, followed by Mandarin non-musicians (p<0.001). Furthermore, for English musicians, the score at the Pair (6&7) is significantly better than the rest of the pairs, except Pair (5&6). There is also a significant Direction x Pair interaction [F (7, 832) = 3.9507, p=0.0003]. A further analysis was not conducted because the Direction x Pair differences are not the focuses of the present study.

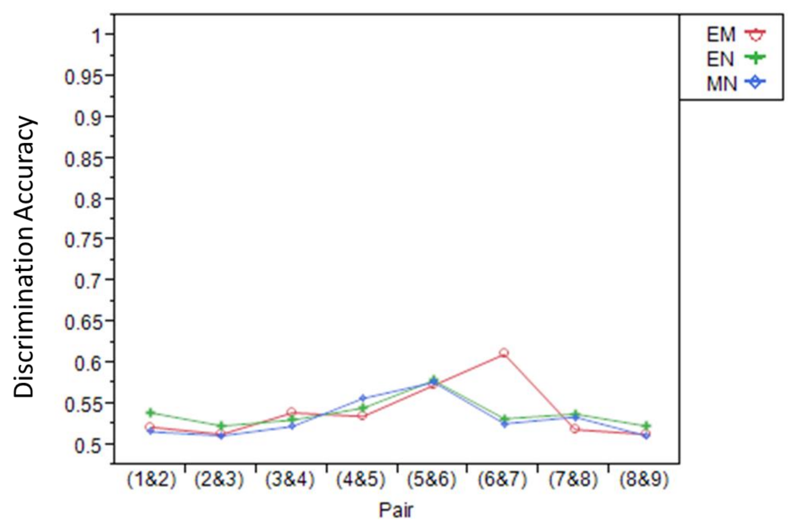


Figure 12 - Average Predicted Falling Discrimination Scores By Pairs Per Group Per Direction The y-axis is the discrimination score range from 0 to 1, with 1 being perfect discrimination of the pairs and 0 is poor discrimination of the pairs. This figure has shown that the peak locations are not consistently at the same pairs.

Correlation between Predicted CB Position and the Obtained Position

One of the requirements for CP is to establish a correlation between the obtained discrimination peak locations and the predicted discrimination peak locations from the identification functions. Therefore, a regression analysis using the predicted categorical boundary (CB) locations (from the predicted discrimination scores at the peak) and the obtained CB locations (from the actual discrimination scores at the peak) was performed to examine whether the boundary locations are aligned.

For both of the rising-level and falling-level continuum, a Pearson Product-moment Correlation Coefficient was computed to investigate the relationship between the predicted discrimination peak and the obtained discrimination peak.

In terms of the rising-level continuum, there is an overall low, non-significant positive correlation between the predicted and the obtained discrimination scores. In regards to the falling-level continuum, the positive correlation between the predicted discrimination peak and the obtained is weak. Only English musicians and Mandarin non-musicians have a significant positive correlation in the falling-to-level continuum (EM: $p<0.05$; EN: $p<0.05$). This has indicated that for English musicians and Mandarin non-musicians, the relationship between the predicted locations and the obtained locations is significantly weak. Table 3 and 4 summarize the results for the correlation. From both tables, it can be seen that the relationship between the predicted peak location and the obtained peak location is weak.

Overall, the Figure 13 summarizes the findings from the discrimination task and the identification task. In terms of the shape of the categorical boundary, it can be seen that Mandarin non-musicians and English musicians both demonstrate similar categorical patterns. However, English non-musicians’ pattern is more continuous.

Table 3 - Summary of Mean Predicted Locations and Mean Obtained Locations for Rising		
Group\Direction	**1to2**	**2to1**
	Level-to-Rising	**Rising-to-Level**
EM	Predicted=6; Obtained=5.13 R^2=0.0025; p=0.88	Predicted=4.82; Obtained=5.85 R^2=0.275; p=0.054
EN	Predicted=6.03; Obtained=5.27 R^2=0.0348; p=0.63	Predicted=5.08; Obtained=5.25 R^2=0.15; p=0.21
MN	Predicted=5.26; OBTAINED=4.38 R^2=0.003; p=0.88	Predicted =4.79; Obtained=4.58 R^2=0.002; p=0.88

Table 4 - Summary of Mean Predicted Locations and Mean Obtained Locations for Falling		
Group\Direction	**1to4**	**4to1**
	Level-to-Falling	**Falling-to-Level**
EM	Predicted =8.19; Obtained=5.13 R^2=0.20; p=0.168	Predicted =5.65; Obtained=5.44 R^2=0.146; p=0.0078*
EN	Predicted =4.81; Obtained=5.27 R^2=0.387; p=0.0736	Predicted =5.55; Obtained=5.65 R^2=0.19; p=0.133
MN	Predicted =4.85; Obtained=4.39 R^2=0.179; p=0.257	Predicted =5.91; Obtained=4.92 R^2=0.02; p=0.0076*

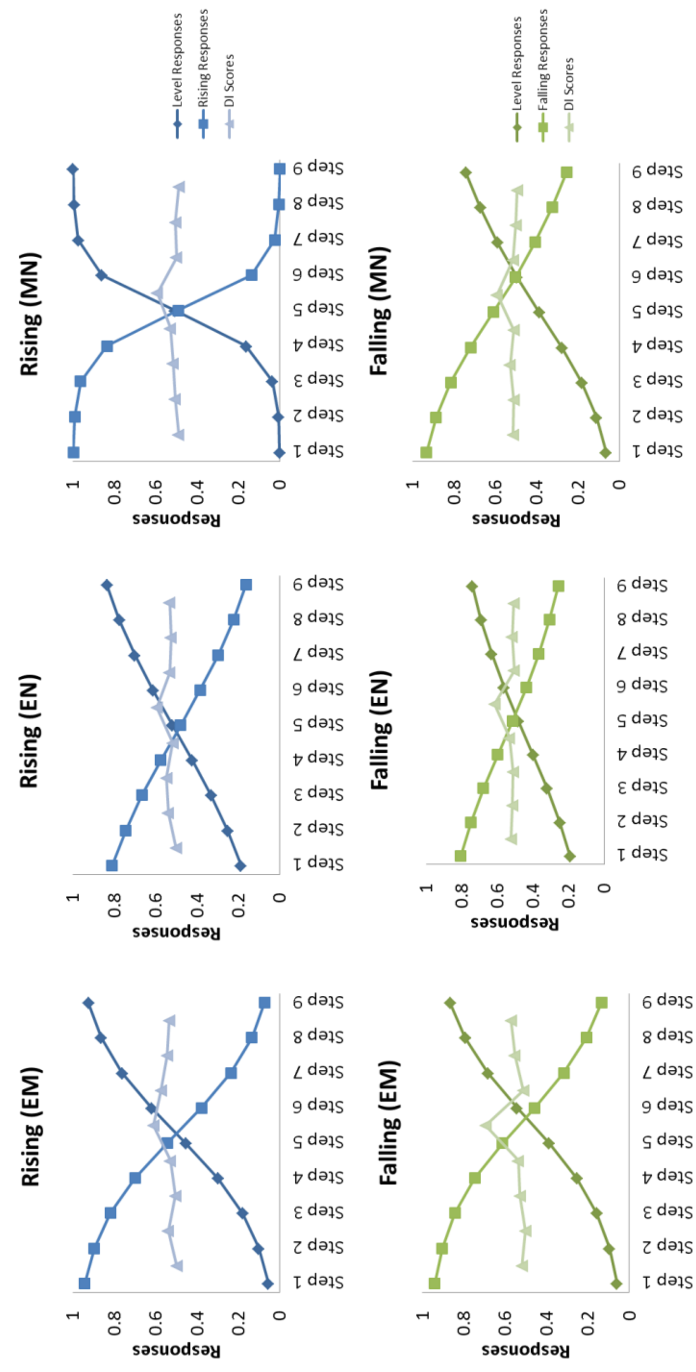


Figure 13 – A summary of identification curves and discrimination curves by Group. It seems that the ID (identification) and DI (discrimination) patterns for English musicians and Mandarin non-musicians are more categorical. However, for English non-musicians, the patterns seem more continuous.

3.2. Music Tones

This section presents the results from the music condition. There are two continua in this condition: (1) CCE to CDE, and (2) CDE to CEE. However, in the discrimination task, the two continua were combined for analysis, so that it is easier to demonstrate the scores for the 17 pairs. For the identification task, the continua must be separated as CCE-to-CDE and CDE-to-CEE in order to calculate the sharpness of the boundary and the boundary locations at the 50% crossover points.

3.2.1. The Discrimination Task

A mixed 2-way analysis of variance (ANOVA) with Pairs (Pair1&2; Pair2&3; Pair3&4; Pair4&5; Pair5&6; Pair6&7; Pair7&8; Pair8&9) as repeated measures, and Group (English Musicians [EM], English non-musicians [EN] and Mandarin non-musicians [MN]) as a between-subject factor, was submitted to JMP for analysis. The measurement for the discrimination task is the accuracy, ranging from 0 to 1. A score of 0 indicates that the participant have incorrectly judged for all the same pairs and different pairs, whereas a score 1 indicates perfect discrimination for the pairs.

CCE-CDE-CEE Continuum

Figure 14 shows the music discrimination accuracy for all of the three groups across the pairs. It can be seen that the peak is not evident for all of the three groups. However, it is apparent that English musicians outperformed English non-musicians, and the Mandarin non-musicians demonstrate the worst performance.

A significant main effect of Group was obtained [$F(2, 866) = 71.6206$, $p = 0.0001$], indicating that there was a significant group difference on the music discrimination scores across pairs. Post Hoc Tukey adjusted pairwise comparisons indicated that English musicians ($M = 0.76$; $SD = 0.0098$) performed significantly better than Mandarin non-musicians ($M = 0.596$; $SD = 0.0101$) and English non-musicians ($M = 0.672$; $SD = 0.0096$, $p<0.0001$). English non-musicians performed significantly better than Mandarin non-musicians ($p<0.0001$). There is no significant main effect of Pair [$F(16, 866) =$

0.3133, p=0.9956], indicating that the performance across the pairs was comparable, with no significant discrimination peak. Figure 15 displays the discrimination scores for each group and each pair. However, the interaction between Group x Pairs was not significant [F (32, 866) = 0.2969, p=1].

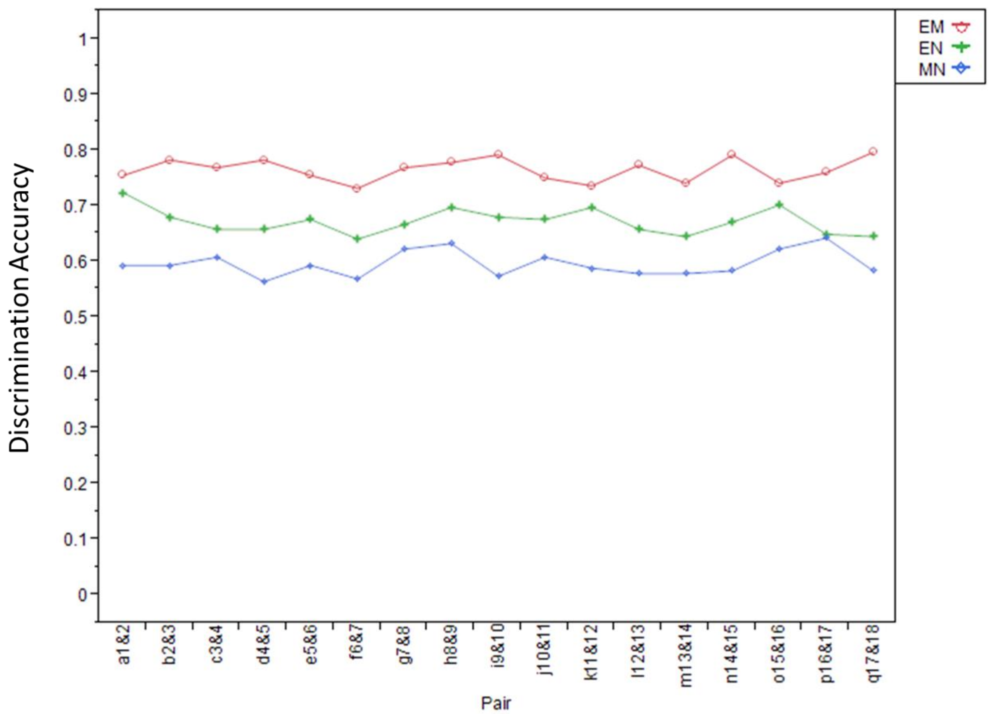


Figure 14 – Average Discrimination Scores for Each Pair per Group. The y-axis is the discrimination scores, ranging from 0 to 1. A score of 1 means perfect discrimination, and a score of 0 means poor discrimination. It can be seen that there are no peaks across the three groups (English musicians, English non-musicians, and Mandarin non-musicians).

3.2.2. The Identification Task

The Generalized Linear Model – Logistic Regression

The estimated slopes, intercepts, and categorical boundary positions between different groups were computed using the Generalized Linear Model. A higher slope value indicates a sharper boundary. Moreover, the boundary location ranges from 1 to 9 for the CCE-to-CDE continuum. For the CDE-to-CEE continuum, the boundary location ranges from 1 to 10.

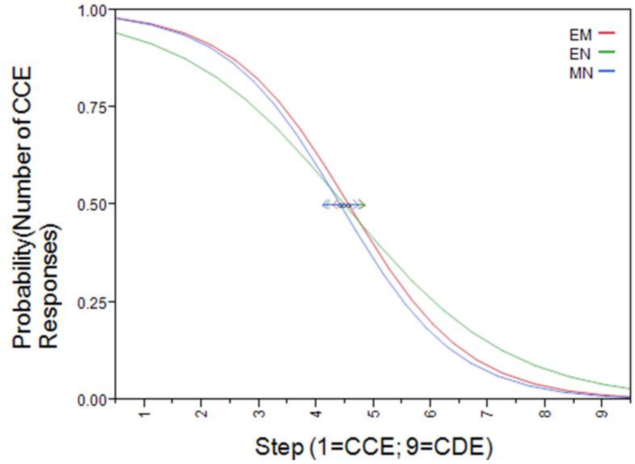
For the CCE-to-CDE continuum, a score of 1 represents that there are more CDE responses (closing to the CCE end). However, a score of 9 indicates that there are more CCE responses (closing to the CDE end) In terms of the CDE-to-CEE continuum, a score of 1 stands for more CEE responses, whereas a score of 10 indicates that there are more CDE responses. A summary of these coefficients is presented in Table 5. These values were used to plot Figure 15.

The logistic regression in the generalized linear model indicated a 2-way interaction between Group and Step Number ($p<0.001$) meaning that Group and Step Number are most likely to be the predictor variables that affect the sigmoid curves. Therefore, as seen with the Mandarin condition, it is worthwhile to examine the sharpness of the boundary and the boundary position for each group.

Table 5 - Summary of Intercepts, Slopes and Categorical Positions

Group	Direction	b_0 - intercepts	b_1 - Slope	$X_{cb} = -b_0/b_1$
EM	**CCE-CDE**	4.419	0.968	4.56
EN		3.149	0.701	4.493
MN		4.37	0.99	4.415
EM	**CDE-CEE**	6.857	0.975	7.03
EN		4.075	0.615	6.621
MN		6.379	0.916	6.967

CCE-CDE Continuum



CDE-CEE Continuum

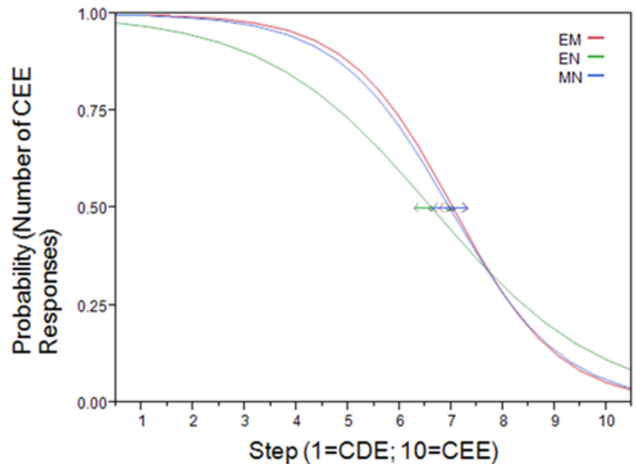


Figure 15 – The Identification Curves for CCE-to-CDE Continuum and CDE-to-CEE Continuum - the x axis indicates the boundary range. The CCE-to-CDE continuum ranges from 1 to 9. A socre of 1 indicates that CCE responses, whereas a score of 9 indicates CDE responses. Moreover, The CDE-to-CEE continuum ranges from 1 to 10. A score of 1 means CDE responses, whereas a score of 10 means CEE. The y axis represents the percentage of CCE or CDE responses in either continuum.

Sharpness of Categorical Boundary

The slope for each subject in each music continuum was computed to compare the sharpness of the category boundary between the groups. A 1-way ANOVA using Group (EM, EN, MN) as the between subject factor was conducted. There was no significant Group effect for the CCE-CDE continuum [F (2, 48) = 2.5362, p=0.0897]. However, there was a significant effect of Group for the CDE-CEE continuum [F (2, 48) = 4.9679, p=0.0109]. For CDE-CEE continuum, post hoc Tukey adjusted pairwise comparisons (HSD) indicated that Mandarin non-musicians (M = 0.916) show a significantly sharper categorical boundary than English non-musicians (M = 0.615, p<0.0110). There is no significant group difference between English musicians and Mandarin non-musicians, nor is there a difference between English non-musicians and English musicians.

Figure 16 shows the group comparison for each of the music continuum. In terms of the sharpness of the boundary, it can be seen that Mandarin non-musicians and English musicians' pattern is similar, whereas the English non-musicians' boundary is least sharp.

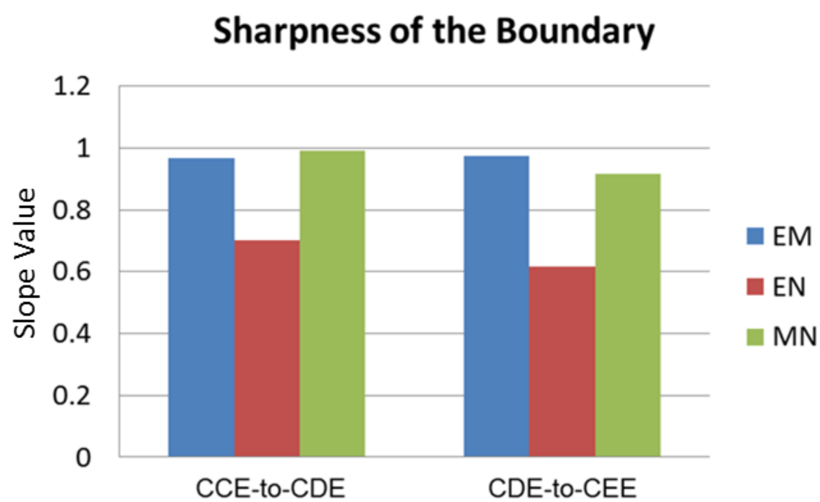


Figure 16 – Average Slope per Group For CCE-CDE And CDE-CEE Continuum- There are three groups: English musicians, English non-musicians, and Mandarin non-musicians. The y-axis indicates the slopes of the logistics curves (e.g. sharpness of the boundary). A higher value means a sharper/steeper boundary. The star indicates that Mandarin non-musicians' boundary is significantly sharper than the English non-musicians in the CDE-to-CEE continuum.

Position of Categorical Boundary

The position of categorical boundary [CB Location] for each subject in each continuum was computed in order to compare the boundary location between the groups. A 1-way ANOVA using Group (EM, EN, MN) as the between subject factor was conducted. For both music continua, there is no significant main effect of Group [CCE-CDE: $F(2, 48) = 0.1962$, $p>0.05$; CDE-CEE: $F(2, 48) = 0.7294$, $p>0.05$].

Figure 17 shows the average boundary position for each group. It can be seen that in each continuum, the group difference is not evident, meaning that the boundary position for each group in each continuum is similar.

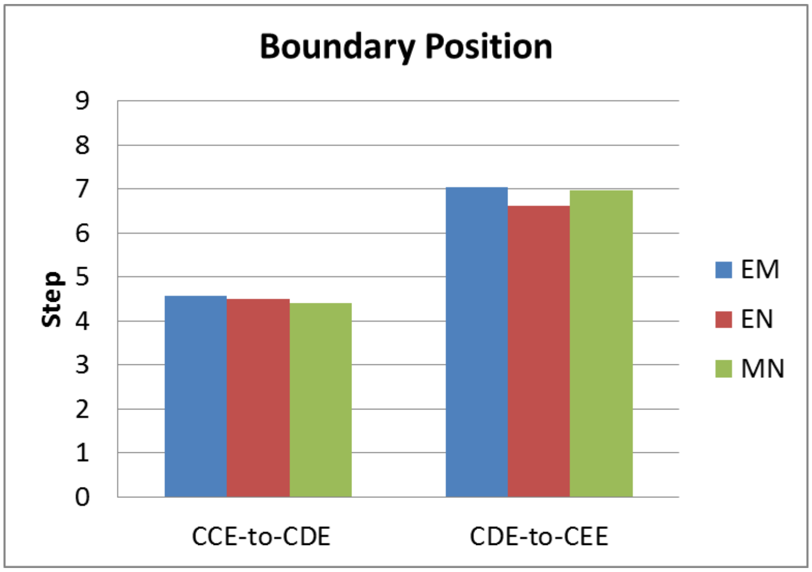


Figure 17 – Average CB Position for each group in the CCE-CDE and CDE-CEE continuum. For CCE-CDE, a higher value indicates that the position is close to CDE end (hense, more CCE responses). For CDE-CEE, a higher value indicates that the position is close to CEE end (hense, more CEE responses).

The correlation analysis was not performed because the music discrimination data does not show peaks for each continuum. Also, the identification data indicates that there are no group differences in terms of the sharpness of the boundary and the categorical boundary locations. Therefore, it is not necessary to correlate the predicted peak locations with the obtained peak locations.

3.3. Summary of the Results

This section presents the overall summary of the results reported in the previous section. The results show that in Mandarin tone discrimination, the perceptual accuracy peak location for all groups was at steps 5 and 6 for both continua. However, in terms of the overall discrimination accuracy, English musicians significantly outperformed Mandarin non-musicians in the falling-level continuum, and they also showed a tendency to be more accurate than the Mandarin non-musicians in the rising-level continuum.

For Mandarin tone identification, the logistic regression analyses indicated that Mandarin non-musicians showed a significantly sharper category boundary (more categorical perception) than English musicians who in turn showed a significantly sharper boundary than English non-musicians. In terms of boundary location along the falling-level continuum, English musicians had more falling tone responses, compared to English non-musicians who had more level tone responses. Group differences were not found for the rising-level continuum.

For music discrimination across steps and continua, English musicians significantly outperformed English non-musicians, followed by Mandarin non-musicians. No discrimination peak was observed for any of the groups. For identification, the logistic regression analysis indicates that the Mandarin non-musicians exhibit a sharper category boundary (more categorical perception) than English non-musicians in the CDE-to-CEE continuum. In terms of boundary position, however, group differences were not found.

Chapter 4. Discussion

The discussion section is separated by the two main research questions in the present study:

(1) Does musical training facilitate categorical perception of lexical tones?

(2) Does speaking a tone language facilitate the categorization of tones in music?

Hence, section 4.1 discusses how musical training experience facilitates the categorical perception of a foreign tone language. Section 4.2 discusses how a lexical tone experience influences music perception, and explores whether native Mandarin non-musicians demonstrate categorical perception of music tones. Section 4.3 discusses the categorical perception findings from native Mandarin listeners. Finally, in section 4.4, current grounded perceptual theories are incorporated to discuss the main findings from the present study.

4.1. Effects of Musical Training on Lexical Tone Perception

The first goal of the present research is to examine whether musical training facilitates categorical perception of lexical tones. A large body of evidence has found that non-tone language musicians' superior lexical tone identification may be resulted from their own musical training experience (Lee & Hung, 2008; Lee & Lee, 2010), and that musical training improves the ability to acquire a second tonal language through long-term exposure and training (Cooper & Wang, 2012; Delogu et al., 2010; Marie, Delogu, Lampis, Belardinelli, & Besson, 2011). If this is the case, then it will be expected that the English musicians might exhibit a categorical boundary of the Mandarin tones, as patterned similarly with the native Mandarin speakers.

To answer the question whether musical experience facilitates categorical perception in Mandarin Chinese (Effects of music on language), two parts of the present findings (1) Discrimination and (2) Identification must be discussed first.

4.1.1. Discrimination

In the rising-level continuum, the Mandarin discrimination results have indicated that the English musicians marginally outperformed the other two groups (Mandarin non-musicians and English musicians). However, in the falling-level continuum, the English musicians significantly outperformed the other two groups. For both continua, it is also apparent that the peak is observable at around Pair (5&6), though the group difference in the peak is not significant.

These results are in line with previous research that musical expertise demonstrates an advantage for detecting pitch-deviant linguistic tones (Chandrasekaran, Krishnan, & Gandour, 2009; Gandour, Wong, & Hutchins, 1998; Magne, Schön, & Besson, 2006; Marie et al., 2011; Sadakata & Sekiyama, 2011; Schön, Magne, & Besson, 2004; Wong, Skoe, Russo, Dees, & Kraus, 2007). These have suggested that English musicians indeed tend to utilize a resource-sharing mechanism in processing the Mandarin tones based on their musical training experience (Patel, 2008; Zattore & Gandour, 2008). Furthermore, compared to English musicians, the English non-musicians outperformed the Mandarin non-musicians in the discrimination task. This has also supported the notion that the within-category discrimination is influenced by lexical tone experience. It is because English non-musicians do not have any lexical tone experience; therefore, it will be easier for the English non-musicians to judge the sameness of the pairs involving subtle acoustic changes, resulting in a more domain-general processing of the within-category pairs. Overall, the present study does support the hypothesis that musical training does facilitate the discrimination of within-category lexical tone differences.

4.1.2. Identification

In the rising-level continuum and the falling-level continuum, Mandarin non-musicians have demonstrated a sharper categorical boundary than English musicians. Interestingly, English musicians have a sharper categorical boundary than English non-musicians. This finding is in line with previous research that has reported a facilitative effect of musical training on tone word identification (Lee & Hung, 2008), or tonal language learning (e.g. Cooper & Wang, 2012).

The presenting findings have also provided more evidence that the English musicians are able to recruit the resources available in the music domain, and then further transfer the pitch information in music to the categorization of the linguistic tone domain (Patel, 2008; Zattore & Gandour, 2008). This might explain why English musicians are better at categorizing novel linguistic tones because of their musical training. Moreover, compared to English musicians, the English non-musicians do not seem to transfer the categorization ability. It is because the English non-musicians do not have any experience with a lexical tone language, nor do they have musical training experience; therefore, the perception of the Mandarin tones seems to be more continuous rather than categorical for English non-musicians (Wang, 1976; Xu et al., 2006). Thus, this also supports the hypothesis that musical training does have a positive impact on lexical tone perception.

Moreover, in terms of the position for categorical boundary, we have found a significant group difference in the falling-level continuum. It was found that English musicians' boundary position is closer to the level-end (closest to Step 9), resulting in more falling responses. Furthermore, English non-musicians' boundary location is closer to the contour-end (closest to Step 1), resulting in more level responses. Consistent with previous research showing that musicians have superior pitch sensitivity (e.g. Magne, Schön, & Besson, 2006; Marie et al., 2011; Schön et al., 2004), the present study suggests that musicians are also able to transfer this lower-level processing ability to a more high-level task. That is, because the musicians have a greater sensitivity to pitch, this has impacted their ability to correctly classify the contour lexical tones (e.g. falling tones) as contour sounds (e.g. falling sounds). Furthermore, the present study has also found that

English non-musicians have more level responses in the falling-level continuum. This is, however, not consistent with Bent et al. (2006)'s findings that English listeners (non-musicians) are more likely to correctly perceive falling contour pitch. A possible explanation might be due to musical training experience. In Schon et al. (2004)'s study, they found that musicians were better at detecting weak F0 changes (or stable F0 changes) than non-musicians; it is possible that musical training has facilitated pitch contour sensitivity for musicians but not for non-musicians. Furthermore, another possibility is that there might be a natural auditory category for the English non-musicians (Xu et al., 2006). The English non-musicians might have classified the falling by utilizing their general auditory knowledge, thereby causing more level responses in the falling-level continuum. However, such claim should await further investigation, as the present study does not include a non-speech condition; therefore, the existence of natural tone categories has remained unclear at this point. Future research may further examine this issue by including a linguistic tone condition, a non-speech tone condition, and a music condition in an attempt to make all the auditory materials comparable in each condition. By this, it is easier to examine how the pitch processing at a higher level domain (e.g. language and music) interacts with the pitch processing at a lower acoustic domain (e.g. non-speech tones).

Taken together, the present findings have found a facilitative effect of musical training on linguistic tone categorization. This implies that not only are musicians able to transfer the music sensitivity to the linguistic tone sensitivity, but they are also able to transfer such sensitivity to linguistic tone categorization.

4.1.3. Categorical Perception of Lexical Tones for English Musicians

In answering whether English musicians possess categorical perception in Mandarin, from Section 4.1.1 and Section 4.1.2, it can be concluded that the criteria for categorical perception has been marginally met: (1) an observed discrimination peak at the Pair (5&6) and (2) a sharper categorical boundary for both rising and falling continua. The third criterion, which is the correlation between predicted peak and obtained peak, has not been met because there has not been a correlation between the predicted peak location and the obtained peak location. This result, however, is consistent with Halle et al. (2004)'s definition of categorical perception of lexical tones[3], in which a *quasi-categorical perception* of tones has been found for English musicians.

Within-category Differences

The present study has shown that English musicians are able to correctly perceive the within-category Mandarin tonal differences as well as correctly classify the tones involving dynamic contour changes. This is actually consistent with several previous lines of research that has found a positive transfer effect from musicality to language learning (Gandour et al., 1998; Koelsch, Schröger, & Tervaniemi, 1999; Lee & Lee, 2010). There are several reasons for this consistency, which can be separated into two levels, such as the pre-attentive neurological level and the attentive behavioral level.

First, musicians' stored musical tonal category representation may possibly influence the linguistic pitch contour perception at the pre-attentive level (Chandrasekaran et al., 2009; Koelsch, Schroger & Tervaniemi, 1999; Moreno, 2009; Moreno & Besson, 2006; Parbery-Clark, Strait, & Kraus, 2011; Parbery-Clark, Tierney, Strait, & Kraus, 2012; Strait, Kraus, Parbery-Clark, & Ashley, 2010; Strait, O'Connell, Parbery-Clark, & Kraus, 2013; Wong & Perrachione, 2007). For instance, in Chandrasekaran et al. (2009)'s study, English musicians and Chinese non-musicians have been found to show larger MMN responses (stronger ERPs) than English non-musicians when they perceive the within-category rising tones in Mandarin. Additionally, a stronger *Frequency-following response* (FFR) has been obtained for musicians when judging the within-category tones. This situation points to the assumption that musicians and tone language speakers may indeed share a similar processing strategy when they were asked

[3] One of the criteria for categorical perception is that the predicted discrimination scores, which derived from the identification functions, should be correlated with the obtained discrimination scores. However, in the present findings, this correlation is not statistically significant for all the groups and all the continuum.

to attend to pitch-related information based on their long-term pitch experience at the early processing stage.

In addition, consistent with Bidelman et al. (2011)'s results, the present finding indicates that musicians possess a more superior music pitch encoding ability at the attentive behavioral level. This is probably because a music-specific (domain-specific) knowledge from musicians has helped them to pay attention to the pitch-related information in language. This implies that a directional transfer effect from the music domain to the language domain may have occurred, meaning that the musicians can first make use of the general pitch resources in the brain and further apply the knowledge to the understanding of the tone information in the linguistic domain. Furthermore, as mentioned earlier in Chandrasekaran et al. (2009)'s study, it might be possible that musicians' early attention to pitch difference has been successfully transferred to the perception of the within-category lexical tone differences. Hence, it can be concluded that combination of the pre-attentive pitch processing and attentive processing can result in a more robust, native-like linguistic tone categorical perception for English musicians.

Between-category Differences & Categorical Perception

In the study of categorical perception (CP), it has always been assumed that the effect of categorical perception is manifested by an observed discrimination peak in a continuum along with a sharper categorical boundary (Hallé, Chang, & Best, 2004; Liberman, Harris, Hoffman, & Griffith, 1957; Wang, 1976; Xu, Gandour, & Francis, 2006). Relating back to the present study, in terms of the between-category differences for Mandarin tones, the English musicians have shown discrimination peak between Step 5 and Step 6 and a sharper boundary for the rising-level and falling-level continuum, which is patterned similarly with the native Mandarin non-musicians (see Figure 13 for references). These findings are also in line with several of previous research that has found that musicians excel at the discrimination task for linguistic tones (Marie et al., 2011; Sadakata & Sekiyama, 2011; Schön et al., 2004) as well as the acquisition of a tonal language (G. M. Bidelman et al., 2011; Cooper & Wang, 2012; Delogu et al., 2010; Magne et al., 2006; Marie et al., 2011; Wayland, Herrera, & Kaan, 2010). Musicians' greater sensitivity to the within-category and between-category differences altogether has suggested that they might recruit the shared-resource center in the brain to process the pitch in language, and their musical training has further strengthen the shared resource in acquiring the Mandarin tones (Patel, 2008; Zattore & Gandour, 2008).

There are several explanations that potentially contribute to this facilitation phenomenon. First of all, as previous mentioned, musical training experience might strengthen the sensitivity to pitch differences, and further strengthen the ability to acquire a tone language. For example, the observed discrimination peak for the deviant linguistic tones and the sharper categorical boundary might be facilitated by the musicians' superior music pitch sensitivity, and this low-level sensitivity has been transferred to higher level categorization ability. For instance, Sadakata & Sekiyama (2011) reported that native Dutch-speaking musicians were able to detect the temporal deviance in a second language [L2] or their native language [L1]. This points to the possibility that musicians have a shared auditory resource center between the domains of music and language, as musicians are able to make use of the resources in the music domain, retrieve the pitch/timing information from the center, and transfer the information into the discrimination of the deviant foreign language tones (Besson, Chobert, & Marie, 2011; Magne et al., 2006; Marie et al., 2011; Moreno et al., 2011; Sadakata & Sekiyama, 2011; Wayland et al., 2010).

Secondly, it was found that the English musicians demonstrated more contour responses (falling) than the level responses with respect to the boundary position in the identification task. That is, the boundary location for the English musicians occurs more towards the level-end (meaning more falling responses). This phenomenon can be explained under the context of long-term pitch experience. In Bent et al. (2006)'s findings, they have found that tone language listeners paid more attention to the falling contour, as it involves a larger pitch range than the rising contour. Since the current English musicians had extensive music pitch training experiences, it is possible that they are better able to make use of the low level processing strategy in music (e.g. music sensitivity), and transfer this ability to another low-level processing in linguistic tones (e.g. linguistic tone sensitivity). That is, their lower-level perception of the falling-tone deviance can be facilitated by their musical training experience (Schon et al., 2004). This might lend support to the facilitation of language learning or language acquisition through extensive musical training because of the transfer effect from the music sensitivity to linguistic tone sensitivity.

In sum, the present study is an extension of these previous findings and suggests that if musical training facilitates language processing (Chandrasekaran et al., 2009) or language learning (Cooper & Wang, 2012), then it is also a possibility that musicians might be able to categorize the Mandarin tones. The present results have provided support to the hypothesis that musicians demonstrated a quasi-CP effect (e.g. Halle et al., 2004) in the linguistic tones, yet such an effect is stronger than the native Mandarin speakers and the native English non-musicians. These findings might be due to the effect of musical training and the effect of pitch direction at both the neurological level and behavioral level.

The present findings motivate future research on the examination of tone language musicians' categorical perception of the native tones. If musical training facilitates categorization in a foreign language, it is possible that the combination of the two types of pitch experience (musical training & Mandarin tone experience) might give rise to an additive performance of categorical perception. Related evidence from previous work (e.g. Cooper and Wang, 2012) does not find any significant additive effects for Thai musicians (L1 Thai + musical training) learning Cantonese tones (L2 Learning), yet it is still unclear whether there is an additive categorization effect for tone musicians (L1 Tone + musical training) perceiving their L1 tones (L1 perception). It is because the tone language musicians have both types of pitch experience, and can be considered expert perceivers of both types of pitch. While the tone language musicians perceive either type of tones, it is unclear which type of pitch experience the tone musicians will utilize in processing the tones. Therefore, It could be hypothesized that there should be an additive effect; if musical training does facilitate sensitivity, then pitch experience from the two domains might be more easily retrieved and manifested by the identification task and the discrimination task (Besson et al., 2011; Magne et al., 2006; Marie et al., 2011; Moreno et al., 2011; Sadakata & Sekiyama, 2011; Wayland et al., 2010).

4.2. Effects of Tone Language Experience on Music Perception

While the majority of previous research has focused on how musical experience facilitates language processing (Cooper & Wang, 2012; Delogu, Lampis, & Belardinelli, 2010; Lee & Lee, 2010; Patel & Iversen, 2007; Perrachione, Fedornko, Vinke, Gibson, & Dilley, 2011; Sadakata & Sekiyama, 2011; Wong & Perrachione, 2007), another unique goal of the present study is to investigate the other direction - an examination of how linguistic tone experience affects music perception. In this case, the present study makes use of both the music discrimination task and identification task in order to examine the interplay between linguistic performance and music performance. Given that several previous studies have found that Mandarin listeners tend to perceive their own native tones categorically (e.g. Peng et al., 2010; Xu et al., 2006), it is possible that Mandarin non-musicians will transfer this categorization ability to the perception of melodies. That is, Mandarin non-musicians will exhibit a categorical pattern for music melodies due to their linguistic tone experience.

4.2.1. Music Discrimination

With respect to the overall accuracy for the music discrimination task, both English non-musicians and musicians significantly outperformed the Mandarin non-musicians. Contrary to our expectations, it seems that speaking a tone language inhibits the sensitivity of musical pitch. This finding is consistent with Peretz, Nguyen, & Cummings (2011)'s prediction that lexical tone experience impairs downward pitch discrimination and pitch-change detection. One of the major conclusions drawn from Peretz et al. (2011)'s study is that if discrimination involves a small change in the pitch threshold (e.g. 4 Hz music increments in the present study), then a speech-specific top-down strategy may be applied and interfered with the novel pitch discrimination. Reversely, if there is a large pitch change, then a more general mechanism of pitch perception will be applied. Similarly, consistent with Bidelman, Gandour, & Krishnan (2011)'s finding, they have attributed such inability of detecting pitch variation to lack of exposure to musical attributes, which might potentially have affected the tone language speakers' pitch detection ability. That is, to successfully detect pitch iterations, musical training might have a greater impact than linguistic experience.

However, since most of the previous research pointed to the direction of a facilitative effect for tone language experience on the discrimination of musical pitch or intervals (e.g. Giuliano et al., 2011; Pfordresher & Brown, 2009), the reasons for such inconsistency might have been due to experimental design (G. Bidelman, Hutka, & Moreno, 2013), or task conditions/ or requirements (Bent, Bradlow, & Wright, 2006).

In previous research examining the similar issue, researchers used either music intervals (Pfordresher & Brown, 2009) or a melodic sequence (Bradley, 2012) in the discrimination. This might contribute to the differences compared to the present results (e.g. Gerrits & Schouten, 2004). For instance, in Pfordresher & Brown's study (2009), they asked the native tone group to listen to a pair of two musical intervals (four tones in total), with only pitch changes on the last note. Their results have indicated that tone language speakers performed significantly better on the music interval discrimination, which points to the fact that the tone language speakers indeed successfully detect the pitch difference in the last note. This, however, motivates further examination of music-interval discrimination by tone language speakers, as they might fully utilize their lexical tone experience as a more domain general strategy in detecting the small differences in the intervallic melodies. By this, a facilitation effect (or a positive transfer effect) might be expected as tone language speakers make use of the pitch information to differentiate the pitches in music.

Also, the reason for such inconsistency might be due to the stimulus processing (task demand) requirement for musical tones by tone language speakers (Bent et al., 2006; Xu et al., 2006). Because music tones and lexical tones are harmonic, the Mandarin non-musicians' mental lexical tones might have competed with the perceived melodic tones during the discrimination task (Alexander et al., 20011). That is, when native Mandarin non-musicians perceive the music tones, they may make use of their lexical tonal experience in competition with the perceived musical tones. If the categorical information does not match between these two types of tones, then it might thereby cause an effect of interference, or the tone speakers may experience negative competition with the music discrimination pairs (Peretz et al., 2011). This motivates further research by suggesting that if the harmonics feature of music could be removed (e.g. using sine-wave pure tones) but kept the altered fundamental frequency, then an effect of facilitation might be expected, as the tone language speakers might utilize a more domain-general, bottom-up processing strategy to discriminate the non-speech tones (Bent et al., 2006).

Another reason for such inconsistency with previous lines of evidence might be due to the effect of tonal inventory in the language. In Mandarin, there are four lexical tones in total, with a high level flat tone (Tone 1), a rising tone (Tone 2), a dipping tone (Tone 3), and a falling tone (Tone 4). Previous research has suggested an influence of L1 experience on the perception of lexical tones (Francis, Ciocca, Ma, & Fenn, 2008; Qin & Mok, 2012). The present music discrimination task involves uniform, stable level acoustic changes of pitch from the note D to note C and the note D to the note E. This might have caused confusion for the Mandarin speakers, as they only have a high level tone in Mandarin. Qin & Mok (2012) have suggested that an assimilation effect might occur from mapping the high level tone in Mandarin to the three level tones in Cantonese. Therefore, it is possible that detecting the level pitch differences in music was difficult for the Mandarin speakers as Mandarin speakers are not used to using different level tones in distinguishing the meanings. This also motivates future research on testing different groups of tone language speakers (e.g. Thai or Cantonese). These languages make use of pitch height as the linguistic cue for meaning; it might therefore be possible for the native speakers of these languages to transfer the ability into the discrimination of music tones (Miran, Holt & McClelland, 2004).

In sum, the current study found that Mandarin speakers perform the weakest on the music discrimination task. This might be due to the effect of interference (Peretz et al., 2011), as the Mandarin speakers have experienced tonal competition while judging the sameness of the musical tones during the process. There are, however, several factors that may contribute to the finding: (1) type of discrimination task, (2) task demand factor, and (3) the influence of tonal inventory in the language.

4.2.2. Music Identification

With respect to the identification task, there is no group effect on the categorical boundary position. Moreover, in terms of the sharpness of the categorical boundary, there is no effect of Group in the CCE-to-CDE continuum, indicating that the slope of the boundary for each group does not differ. Interestingly, there is a significant effect of group in the CDE-to-CEE continuum, and further analyses indicate that the Mandarin non-musicians' boundary is sharper than the English non-musicians. This interesting finding has provided support that the tone language experience might facilitate the categorization of music tones.

Based on past research looking at music categorical perception by professional musicians and non-musicians (Bigand, 2003; Blechner, 1977; Dowling, 1986; Siegel & Siegel, 1977; Smith, Nelson, Grohskopf, & Appleton, 1994; Ziv & Radin, 2014), we can postulate some reasons in relation to the present research findings.

Firstly, the present study reports that the English musicians do not exhibit a categorical boundary for either of the continuum. This is actually consistent with Blechner (1977)'s research findings. In the study, he asked a group of professional musicians and non-musicians to perform an identification task and a discrimination task with a single music tone, varied by its fundamental frequency (F_0). The results have indicated that the musicians did not exhibit a discrimination peak in the single music tone condition, but a peak was observed in the music chord condition. Blechner (1977) then concluded that categorical perception only exists in musical chords when the three fundamental frequencies interact with one another for musicians. Relating back to the present research, the present study has made use of a similar experimental parameter (from Blechner, 1977), though the varied music tone was placed in the context of C-D-E in isolation. The present results are consistent with Blechner (1977)'s research findings that when music notes are placed under an isolated context, categorical perception will not be prevalent for musicians, and now it can be claimed that this is also the case for tone language non-musicians.

By the same token, the present finding reports that the boundary positions for all the music continua are similar across the three groups, and the sharpness of the CCE-to-CDE continuum does not yield any group differences. This phenomenon can be explained by the acoustic nature of the music tones; that is, the acoustic property of the music pitch might have impacted the categorical perception of music tones. As mentioned earlier, the within-category music tones used in the present study involve uniform, stable acoustic changes, whereas the linguistic tones used in the present study involve more dynamic contour changes. It is possible that the Mandarin non-musicians have experienced troubles comparing the boundary of the perceived music tones with the boundary of their lexical tones, as these tones are technically not similar. On the other hand, in a previous CP study, Abramson (1977) found that native Thai did not perceive the Thai level tones categorically. If the music tones are perceived as level tones by the native Mandarin non-musicians, it is then possible to claim that categorization (or categorical perception of music tones) might not be prevalent because the stable music pitch is similar to the level tones in language, and Mandarin non-musicians are not used to classifying these level tones. The present findings, however, imply some motivations for future research on the categorical perception of music triads by tone language speakers. Blenchner (1977) found that a single varied music tone in isolation does not yield a categorical result, yet a musical chord does. It is possible to predict that a music chord might have yielded a categorical result by tone language speakers, given that tone language speakers can be considered a pitch expert comparable to professional musicians. Taking these together, it could be concluded that categorical perception might be prevalent in the context of tone contour linguistically, and in the context of music chord.

Another factor which might have influenced the present results is the effect of long-term melodic memory on the perceived melodies. Several lines of research have indicated that familiar contour sequences or melodies (assuming that familiar melodies are stored in long term memory) might influence music recognition (Dowling, 1978; Dowling, Bartlett, Halpern, & Andrews, 2008; Ettlinger, Margulis & Wong, 2011; Peretz, Radeau & Arguin, 2004; Smith, Nelson, Grohskopf & Appleton, 1994) or pitch detection (Colombo, Deguchi, Boureux, Sarlo, & Besson, 2011; Silverman, 2010; 2012). Though the present research placed the varied D-note under the context of C-X-E, there is no evidence of categorization for CCE-to-CDE. However, interestingly, there is an effect of group on the categorical boundary of CDE-to-CEE, in which the Mandarin non-musicians have shown a sharper categorical boundary than the English non-musicians. This finding might have provided support to the fact that the familiarity of musical melodies might have played a role in categorization (Dowling, 1978; Dowling et al., 2008), if CDE and CEE are considered as the most familiar melody stored in the Mandarin non-musicians' mind. If this is the case, then it is also possible that Mandarin listeners' lexical tone experience might have strengthened the categorization advantage (e.g. Pfordresher & Brown, 2009). This suggests that there might be some degrees of cross-domain transfer from language tone experience to music perception. This explanation, however, still needs to be further investigated, as it is still not clear whether long-term memory that involves melodic information impacts the categorization of melodic tones. If a popular melody involves subtle acoustic changes (e.g. the Happy Birthday song), then it is expected that tone language listeners may be able to perceive the melody in a more categorical manner.

A similar issue has been found in Xu et al. (2006)'s study in which they have found a non-speech quasi-categorical perception by Chinese listeners and English listeners. They have argued that categorical perception might be influenced by memory processing mechanisms and their involvements with different stimulus types. Based on the *Multistore Model of CP* proposed by Xu et al. (2006), it is possible that the CDE-to-CEE have stored in the Mandarin non-musicians' long term memory. When they perceive this music sequence CEE, they are able to quickly recruit their linguistic tone knowledge, pay attention to the dimension of the perceive melody, and further classify the within-category melodies by a top-down mechanism from the linguistic domain. This process might provide ample explanations for why CDE-to-CEE melodies are perceived categorically but not CCE-to-CDE. However, future investigations might include another set of stimuli, such as non-speech non music noise, because the non-speech non-music noise are considered lower-level, and the processing of noise should involve a more bottom-up processes (this assumes that humans do not have any memory for noise sounds). If memory does play a role in categorization, it is possible that tone language listeners or non-tone language listeners are able to categorize the noise based on their short-term memory because they will have to be trained to perceive the noise prior to the categorization tasks.

Overall, based on the obtained results, there has been no prevalent evidence of categorical perception effect on isolated music tones by the three groups. There are several factors which have been discussed: (1) Stimulus Effect – a music tone in isolation might not yield a categorical finding (Blechner, 1977), (2) the effect of stable pitch height – a varied pitch height along with a continuum might not reveal categorical perception, as evidenced in Thai level tones (Abramson, 1977), and finally (3) the effect of memory on music melodies, along with the effect of tone language experience, might explain why there is a group difference in the continuum from CDE to CEE (Colombo et al., 2011; Dowling et al., 2008).

4.3. Categorical Perception of Mandarin Tones for Mandarin non-musicians

Based on the present findings and previous findings (Hallé et al., 2004; Peng et al., 2010; Sun & Huang, 2012; Wang, 1976; Xu et al., 2006; Zheng et al., 2012), the criteria for categorical perception have been marginally met: (1) an observed discrimination peak at Pair (5&6), (2) a sharper categorical boundary than the English non-musicians, and (3) a correlation between the predicted categorical boundary position and the obtained position.

4.3.1. Discrimination

From the Mandarin tone discrimination task, Mandarin non-musicians performed the lowest among the three groups. As mentioned before, we might attribute this phenomenon to the differences in musical training if we compare the performance with the musicians (e.g. Sadakata & Sekiyama, 2011). However, we might also be able to explain the findings under the effect of interference, as put forth by Peretz et al. (2011). The presence of the lowest performance from the Mandarin non-musicians might suggest that the Mandarin speakers have experienced a speech-specific influence when they perceived the within-category Mandarin tones during the discrimination task (Peretz et al., 2011; Strait et al., 2010). By this account, Mandarin non-musicians may have competition in the lexical tonal memory while perceiving the within-category tones; this might result in the lowest performance on the discrimination task (Peretz et al., 2011).

4.3.2. Identification

It was found on the Mandarin identification task that Mandarin non-musicians revealed a sharper categorical boundary in comparison to English-non-musicians for both rising-level and falling-level continuum. This is consistent with Halle et al. (2004), Xu et al. (2006)'s and Peng et al. (2010)'s findings that Mandarin speakers exhibit classical categorical perception pattern for the Mandarin tone condition. In Xu et al. (2006)'s experiment, they found that the Chinese speakers demonstrated some evidence of CP in the non-speech condition by utilizing both speech and non-speech continua as the stimulus materials. They argued that such CP revelation in the homologous non-speech condition might be explained by the perceivers' use of domain-general auditory mechanism[4].

Relating back to the present study, in terms of categorical boundary position, there is no effect of Group in the rising-level continuum, but there is a group difference in

[4] In Xu et al. (2006)'s term, it has been referred to the "natural auditory sensitivity mechanism" (page 1068).

the falling-level continuum. For the falling-level continuum, the native Mandarin speakers' CB position is not different from the English counterpart as well as the English musicians. This is consistent with Xu et al. (2006)'s explanation for the CP effect revealed by English speakers, as these speakers (English musicians and English non-musicians) might have made use of a domain-general bottom-up mechanism in identifying the falling tones based on their tone experience. This is also consistent with Bent et al. (2006)'s prediction that Mandarin speakers are more likely to mis-identify falling contours. Because the falling tone in Mandarin involves a larger pitch range, the pitch tolerance for the falling contour might be lower than the rising one, thereby causing categorization.

4.3.3. Continuum Direction: Stimulus Naturalness Effect

Interestingly, the present study found an effect of Direction. That is, irrespective of groups, the categorical boundary for the continuum, which was derived from a typical contour tone, yielded more contour responses. Reversely, a continuum which derived from a typical level tone yielded more level responses. We might attribute this finding to the effect of stimulus naturalness (van Hessen & Schouten, 1999).

In van Hessen & Schouten (1999)'s research, they have found that the degree of stimulus naturalness used in the study of categorical perception might influence the obtained categorical results. They have concluded that the more complex (or natural) the speech material is, the harder the tasks will be. Based on this assumption, it is possible that if a stimulus is more natural and complex, then it will be very difficult for the perceivers to focus on a specific acoustic parameter for categorical perception (e.g. within-category differences).

This supports the present findings in that if a continuum is derived from a typical rising tone, tone 2 in Mandarin, for example, then it is expected that the perceivers will show more tone 2 responses instead of tone 1 (level tones). If a continuum is derived from a level tone, then we should expect there to be more tone 1 responses. This notion is also aligned with Xu et al. (2006)'s complexity hypothesis with respect to the categorical perception in linguistic tones. According to the *Complexity Hypothesis* (Xu et al., 2006), complex tones are more difficult to process for tone language speakers because such processing imposes a lot of cognitive effort.

4.3.4. Correlation

Another interesting finding from the present study is the correlation between the obtained boundary positions and the predicted positions. According to the results, the predicted scores did not have any group differences, nor was there a correlation between the predicted peak location and the obtained peak location. First of all, for the predicted scores, it was found that the peak is located at Pair (5&6) for the rising-level continuum, which is consistent with the obtained scores. However, for the falling-level continuum, the peak locations obtained ranged from Pair (4&5), Pair (5&6), and Pair (6&7). This is not consistent with the obtained discrimination scores for the falling-level continuum, as the obtained peak location for the falling is located at Pair (5&6), but the predicted peak location spreads from Step 4 to 7. This has suggested/supported previous research that falling tones involves a larger range of pitch difference (Bent et al., 2006), and therefore may exhibit a wider categorical boundary location (Peng et al., 2010).

Secondly, in terms of the correlation results, the present study does not find any correlation between the predicted boundary location and the obtained boundary locations. In Xu et al. (2006)'s study, they have correlated the two dependent variables (the predicted and obtained discrimination scores) in order to compare the shapes of the discrimination curves. They found that the two discrimination curves have similar patterns, suggesting that the correlation between the predicted curves and the obtained curves should be high. However, the present study's results are inconsistent with these findings. The present findings revealed that each group in each continuum exhibited a lower, not statistically significant, correlation. This is inconsistent with the definitions in categorical perception that the predicted peak scores derived from the identification function might successfully align with the obtained discrimination peaks (e.g. Liberman et al., 1961; Wang, 1976).

A possible reason has been put forth in Gerrits & Schouten (2004)'s research on vowel continua, which argues that categorical perception is simply dependent on the types of discrimination tasks used for the experiment. They have argued that a 2-interval alternative forced choice discrimination task (2IAX) will elicit a more categorical result, and the predicted discrimination scores from identification might have given rise to a more accurate measure. Although the AX discrimination, used in our study, can reduce the effect of memory load, the rate of correct rejection could be very low, as the subjects have to be quite certain of their decision in order to choose the "different" label. Based on Gerrits & Schouten (2004)'s finding, future research might focus on replicating their design by using a tone continuum in order to examine which types of discrimination tasks might trigger a more categorical result for a tone study.

In sum, based on the present findings, the identification and discrimination task have indicated a typical, yet marginal, pattern of categorical perception by the Mandarin non-musicians for the rising & falling tones in Mandarin. Specifically, the Mandarin non-musicians have used a speech-specific strategy in processing the within-category pairs. In addition, the naturalness of a stimulus used in a categorical perception study has been found to influence the position of a categorical boundary. Finally, the types of discrimination tasks used for a categorical perception study might affect the predicted discrimination scores and its correlation with the obtained scores.

Interestingly, based on the results, Mandarin non-musicians do not exhibit a categorical-perception pattern for music tones, as (1) there is no observed discrimination peak across the 17 pairs, (2) the categorical boundary is less steep, and (3) the categorical boundary location does not have any group differences.

4.4. Theoretical Significance

The present study has revealed the following major findings: firstly, musical training indeed facilitates categorical perception in lexical tones. Secondly, speaking a tone language does not facilitate music pitch sensitivity, yet speaking a tone language might facilitate the categorization of CDE-to-CEE melodies. We can place these findings within the framework of the integrated account of domain-general and domain-specific pitch processing by Zatorre & Gandour (2008), *The Shared Sound Category Learning Mechanism Hypothesis*, as put forth by Patel (2008), and the theory of perceptual learning by Gibson (1969).

First, the present study reveals that musical experience enhances the sensitivity to linguistic pitch and categorization of linguistic tones. This has lent support for the integrated account of pitch processing by Zatorre & Gandour (2008). In Zatorre & Gandour's integrated model, they have argued that the pitch processing mechanisms should not be exclusive (domain-general processing versus domain-specific processing). Relating back to the present findings about the effects of music on language, we have found that English musicians are better at detecting the linguistic tone differences as well as categorizing the linguistic tones. We can then argue that experience with music may have enhanced the lower-level sensitivity to pitch differences, and such influence results in a more domain-general processing when discriminating linguistic tones. Likewise, experience with musical pitch categorization may have carried over to linguistic pitch categorization (although the category boundaries do not necessarily overlap due to domain-specific differences). Thus, musical experience has facilitated pitch processing in another higher-level cognitive domain, such as tone language processing.

Similar patterns have been revealed with regard to the influence of linguistic pitch background on music pitch categorization. Compared to English non-musicians, Mandarin non-musicians demonstrate a sharper categorical boundary when they perceive the CDE-to-CEE continuum. This suggests that their experience with linguistic tone categorization has enhanced their music tone categorization, indicating domain-general processing resulting in a positive transfer of experiences from language to music (Zattore & Gandour (2008).

The music discrimination results from the Mandarin non-musicians seem to support *The Shared Sound Category Learning Mechanism Hypothesis,* which clearly distinguishes the process of sound category formation and the process of sound category development. Patel (2008) has argued that the development of a novel sound category (regardless of music or language) might be domain general, as the perceivers need to acquire the new sound category using a broader scope of processing mechanism in making sense of the sound category. However, the process will become more domain-specific once a new sound category has been successfully acquired or learned by the perceiver. Relating back to the present finding that the Mandarin non-musicians' discrimination scores of music-tone pairs is the lowest among the three groups, it is possible that the Mandarin speakers' experience in Mandarin tone categories results in declined sensitivities to subtle acoustic differences, indicating the use of a speech-specific processing strategy to discriminate the music tones. It may be that as the music tonal representations are not well-formed in the Mandarin non-musicians' mind, they resorted to a domain-specific processing strategy that they were familiar with (i.e., Mandarin tone processing) to make sense of the perceived music pitch. The categorical nature of Mandarin tones may have led to Mandarin listeners' decreased sensitivity to subtle within-category pitch differences, thus their poor discrimination of music tones. Therefore, an effect of interference occurs during music discrimination under the influence of top-down domain-specific mechanisms. In contrast, the English non-musicians may have relied on a domain-general lower-level sensory processing to discriminate the tones based on their general experiences with pitch information. Both of the non-musician groups are different from the English musicians who have excelled at the music discrimination task. According to *The Shared Sound Category Learning Mechanism Hypothesis,* because the musicians' musical training experience, the music tones may have been well-formed in their minds. Therefore, the English musicians had the advantage of being able to apply a music-specific top-down strategy (Zattore & Gandour, 2008) to perceive the pitch differences in music.

With respect to the results from the English musicians demonstrating categorical perception of lexical tones, this can be explained within the framework of perceptual learning theory by Gibson (1969), which states that experience in one cognitive domain is more likely to trigger an improved perception in the other domain. Based on the current findings, it is possible that a non-tone speaker's musical training may trigger the learning of lexical tones or formation of lexical tone category in the linguistic domains.

In sum, results of the current study seem to support a shared window of perceptual processing theories: *the integrated account of pitch processing* (Zattore & Gandour, 2008), *The Shared Sound Category Learning Mechanism Hypothesis*, as postulated by Patel (2008) and theory of perceptual learning, put forward by Gibson (1969). The three theories have argued that there might be a transfer effect from one cognitive domain to another domain. First of all, regarding the English musicians' demonstration of categorical perception in the lexical tones, there seems to be a positive transfer effect, which musicianship might be beneficial in terms of lexical tone learning and processing. Secondly, in terms of Mandarin speakers' discrimination of music, there seems to be an interference effect – a negative transfer. Lastly, in terms of Mandarin speakers' categorization of music, there seems to be a positive transfer across the two domains. It is because the Mandarin speakers are better at forming the subtle acoustic items into categories based on their linguistic experience. Taken together, these results suggest bi-directional transfer between language and music in perception and categorization processes, pointing to an integration of domain-general and domain-specific pitch processing as a function of experiences.

Chapter 5. Conclusion and Future Directions

The aims of the present study are to examine effects of linguistic and music experiences on tone language speakers' music tone perception and non-tone speakers' linguistic tone perception. Firstly, results indicate that there is an effect of musical training on Mandarin tone discrimination and categorization. For instance, musical training does facilitate the sensitivity of the within-category Mandarin tones and the categorization of the within-category Mandarin tones, resulting in the effect of Mandarin-tone categorical perception by English musicians. Secondly, there is also an effect of lexical tone experience on music categorization. For example, speaking a tone language (e.g. Mandarin) indeed facilitates the categorization of music tones, thereby causing a positive transfer effect from the lexical tone domain to the music categorization domain. However, Mandarin non-musicians' music discrimination scores are the worst among the three groups; that is, Mandarin non-musicians possess the worst sensitivity to the within-category pitch differences for their native tones and music tones. Therefore, there might be an interference effect from the lexical tone experience to the perception of the within-category pitch.

The current study also provides several future directions for psycholinguistic and neurolinguistic research as well as music cognition research on pitch processing. For instance, while musical training benefits tone language learning, it is not clear whether English musicians, who have learned a second tone language for a number of years, exhibit categorical perception of the lexical tones. If music experience does facilitate tone language learning, it would be hypothesized that those English musicians with a second tone language experience might be more capable of categorizing the lexical tones. This will contribute to the theoretical assumption that the perceptual transfer from one domain to another is not uni-directional but bi-directional.

Furthermore, it is also not clear whether lexical tone training benefits the lexical tone forming process. For example, future research may focus on investigating a tone training (learning) effect on categorical perception of lexical tones by non-tone language musicians. If lexical tone training results in more enhanced categorization of lexical tones for non-tone musicians than non-tone non-musicians, then it could be argued that musical training reinforces the category formation process. It will not be simply the training effect which makes a difference in the tone language learning, but the combination of the two (music experience and training) contributes to such enhanced tone categorization. This is important theoretically because this helps to understand whether there is a category formation process during such cross-domain transfer.

Finally, the current study has also pointed to some theoretical implications for the integrated account of pitch processing (Zattore & Gandour, 2008) *The Shared Sound Category Learning Mechanism Hypothesis* (Patel, 2008) and the theory of perceptual learning (Gibson, 1969) between language and music. The current results have implied that transfer effects across two cognitive domains might be explained by an integration of domain-general and domain-specific pitch processing mechanisms.

References

Abramson, A. S. (1977). Noncategorical perception of tone categories in Thai. *The Journal of the Acoustical Society of America*, *61*(S1), S66–S66.

Ahissar, M. & Hochstein, S. (1997). Task difficulty and the specificity of perceptual learning. *Nature*, *387*(6631), 401–406.

Alexander, J. A., Bradlow, A. R., Ashley, R. D. & Wong, P. (2011). Music-melody perception in tone-language and non-tone-language speakers. *Psycholinguistic representation of tone. Hong Kong. http://www. sfu. ca/~jaa37/index files/Alexander et al*.

Alexander, J. A., Wong, P. C. M. & Bradlow, A. R. (2005). Lexical tone perception in musicians and non-musicians. In *Proceedings of Interspeech* (pp. 397–400).

Amitay, S., Zhang, Y.-X. & Moore, D. R. (2012). Asymmetric transfer of auditory perceptual learning. *Frontiers in psychology*, *3*.

Bahr, N., Christensen, C. A. & Bahr, M. (2005). Diversity of accuracy profiles for absolute pitch recognition. *Psychology of Music*, *33*(1), 58–93.

Banai, K. & Amitay, S. (2012). Stimulus uncertainty in auditory perceptual learning. *Vision research*, *61*, 83–88.

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.

Besson, M., Chobert, J. & Marie, C. (2011). Language and Music in the Musician Brain. *Language and Linguistics Compass*, *5*(9), 617–634.

Bidelman, G., Hutka, S. & Moreno, S. (2013). Tone Language Speakers and Musicians Share Enhanced Perceptual and Cognitive Abilities for Musical Pitch: Evidence for Bidirectionality between the Domains of Language and Music. *PloS one*, *8*(4), e60676.

Bidelman, G. M., Gandour, J. T. & Krishnan, A. (2011a). Musicians demonstrate experience-dependent brainstem enhancement of musical scale features within continuously gliding pitch. *Neuroscience letters*, *503*(3), 203–207.

Bidelman, G. M., Gandour, J. T. & Krishnan, A. (2011b). Cross-domain effects of music and language experience on the representation of pitch in the human auditory brainstem. *Journal of Cognitive Neuroscience*, *23*(2), 425–434.

Bidelman, G. M., Moreno, S. & Alain, C. (2013). Tracing the emergence of categorical speech perception in the human auditory system. *Neuroimage*, *79*, 201–212.

Bigand, E. (2003). More about the musical expertise of musically untrained listeners. *Annals of the New York Academy of Sciences*, *999*(1), 304–312.

Bimler, D. & Kirkland, J. (2001). Categorical perception of facial expressions of emotion: Evidence from multidimensional scaling. *Cognition & Emotion*, *15*(5), 633–658.

Blechner, M. J. (1977). *Musical skill and the categorical perception of harmonic mode*.

Boersma, Paul & Weenink, David (2014). Praat: doing phonetics by computer [Computer program]. Version 5.3.83, retrieved 16 August 2014 from http://www.praat.org/

Bradley, E. D. (2012a). Tone language experience enhances sensitivity to melodic contour. In *LSA Annual Meeting Extended Abstracts*.

Bradley, E. D. (2012b). *Crosslinguistic Perception of Pitch in Language and Music*.

Burnham, D. & Jones, C. (2002). Categorical perception of lexical tone by tonal and non-tonal language speakers. In *Proceedings of the 9th Australian International Conference on Speech Science & Technology* (pp. 515–520).

Burns, E. M. & Ward, W. D. (1978). Categorical perception—phenomenon or epiphenomenon: Evidence from experiments in the perception of melodic musical intervals. *The Journal of the Acoustical Society of America*, *63*(2), 456–468.

Caruso, D. R. & Detterman, D. K. (1983). Individual differences in short-term memory and their relationship to item identification. *The American journal of psychology*, 527–537.

Chan, S. W., Chuang, C.-K. & Wang, W. S.-Y. (1975). Cross-linguistic study of categorical perception for lexical tone. *The Journal of the Acoustical Society of America*, *58*(S1), 119–119.

Chandrasekaran, B., Krishnan, A. & Gandour, J. T. (2007). Mismatch negativity to pitch contours is influenced by language experience. *Brain research*, *1128*(1), 148–156.

Chandrasekaran, B., Krishnan, A. & Gandour, J. T. (2009). Relative influence of musical and linguistic experience on early cortical processing of pitch contours. *Brain and* *language*, *108*(1), 1–9.

Chen, M. Y. (1984). The interfacing of language and music. *Language \& Communication*, *4*(3), 159–194.

Colombo, L., Deguchi, C., Boureux, M., Sarlo, M. & Besson, M. (2011). Detection of pitch violations depends upon the familiarity of intonational contour of sentences. *cortex*, *47*(5), 557–568.

Cooper, A. K. (2010). *Effects of linguistic and musical experience on Cantonese tone word learning*. Burnaby.

Cooper, A. & Wang, Y. (2012). The influence of linguistic and musical experience on Cantonese word learning. *The Journal of the Acoustical Society of America*, *131*(6), 4756–4769.

Cowan, N. (1984). On short and long auditory stores. *Psychological bulletin*, *96*(2), 341–370.

Delogu, F., Lampis, G. & Belardinelli, M. O. (2006). Music-to-language transfer effect: may melodic ability improve learning of tonal languages by native nontonal speakers? *Cognitive Processing*, *7*(3), 203–207.

Delogu, F., Lampis, G. & Belardinelli, M. O. (2010). From melody to lexical tone: musical ability enhances specific aspects of foreign language perception. *European Journal of Cognitive Psychology*, *22*(1), 46–61.

Deutsch, D. (1975). The organization of short-term memory for a single acoustic attribute. In *Short-term memory* (pp. 107–151). New York: Academic Press.

Deutsch, D., Dooley, K., Henthorn, T. & Head, B. (2009). Absolute pitch among students in an American music conservatory: Association with tone language fluency. *The Journal of the Acoustical Society of America*, *125*(4), 2398–2403.

Deutsch, D., Henthorn, T. & Dolson, M. (2004). Absolute pitch, speech, and tone language: Some experiments and a proposed framework. *Music perception*, *21*(3), 339–356.

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

Dowling, W. J. (1978). Scale and contour: Two components of a theory of memory for melodies. *Psychological Review*, *85*(4), 341–354.

Dowling, W. J. (1986). Context effects on melody recognition: Scale-step versus interval representations. *Music Perception*, *3*(3), 281–296.

Dowling, W. J., Bartlett, J. C., Halpern, A. R. & Andrews, M. W. (2008). Melody recognition at fast and slow tempos: Effects of age, experience, and familiarity. *Perception & psychophysics*, *70*(3), 496–502.

Ettlinger, M., Margulis, E. H. & Wong, P. C. M. (2011). Implicit memory in music and language. *Frontiers in Psychology*, *2*(211), 1–10.

Filipic, S. & Bigand, E. (2005). Key processing precedes emotional categorization of western music. *Annals of the New York Academy of Sciences*, *1060*(1), 443–445.

Flege, J. E., Schirru, C. & MacKay, I. R. A. (2003). Interaction between the native and second language phonetic subsystems. *Speech communication*, *40*(4), 467–491.

Francis, A. L., Ciocca, V., Ma, L. & Fenn, K. (2008). Perceptual learning of Cantonese lexical tones by tone and non-tone language speakers. *Journal of Phonetics*, *36*(2), 268–294.

Francis, A. L., Ciocca, V. & Ng, B. K. C. (2003). On the (non) categorical perception of lexical tones. *Perception & psychophysics*, *65*(7), 1029–1044.

François, C., Chobert, J., Besson, M. & Schön, D. (2012). *Music training for the development of speech segmentation*. *Cerebral Cortex*. Oxford Univ Press.

Franklin, M. S., Moore, K. S., Yip, C.-Y., Jonides, J., Rattray, K. & Moher, J. (2008). The effects of musical training on verbal memory. *Psychology of Music*, *36*(3), 353–365.

Fujioka, T., Trainor, L. J., Ross, B., Kakigi, R. & Pantev, C. (2004). Musical training enhances automatic encoding of melodic contour and interval structure. *Journal of cognitive neuroscience*, *16*(6), 1010–1021.

Gandour, J. (1983). Tone perception in Far Eastern languages. *Journal of Phonetics*, *11*(2), 149–175.

Gandour, J. T. & Harshman, R. A. (1978). Crosslanguage differences in tone perception: A multidimensional scaling investigation. *Language and Speech*, *21*(1), 1–33.

Gandour, J., Wong, D. & Hutchins, G. (1998). Pitch processing in the human brain is influenced by language experience. *Neuroreport*, *9*(9), 2115–2119.

Gerrits, E. & Schouten, M. (2004). Categorical perception depends on the discrimination task. *Perception & Psychophysics*, *66*(3), 363–376.

Gibson, E. (1969). *Principles of perceptual learning and development*. Appleton-Century-Crofts New York.

Giuliano, R. J., Pfordresher, P. Q., Stanley, E. M., Narayana, S. & Wicha, N. Y. Y. (2011). Native experience with a tone language enhances pitch discrimination and the timing of neural responses to pitch change. *Frontiers in Psychology*, *2*(146), 1–10.

Gordon, E. E. (1988). *Musical aptitude profile*. Chicago: Riverside Publishing.

Gottfried, T. L. & Riester, D. (2000). Relation of pitch glide perception and Mandarin tone identification. *Journal of the Acoustical Society of America*, *108*(5), 2604.

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.

Hannon, E. E. (2009). Perceiving speech rhythm in music: Listeners classify instrumental songs according to language of origin. *Cognition*, *111*(3), 403–409.

Harrison, P. (1996). An experiment with tone. *UCL Working Papers in Linguistics 8*, 575–593.

Ho, Y.-C., Cheung, M.-C. & Chan, A. S. (2003). Music training improves verbal but not visual memory: cross-sectional and longitudinal explorations in children. *Neuropsychology*, *17*(3), 439.

Hove, M. J., Sutherland, M. E. & Krumhansl, C. L. (2010). Ethnicity effects in relative pitch. *Psychonomic bulletin & review*, *17*(3), 310–316.

Husain, F. T., Fromm, S. J., Pursley, R. H., Hosey, L. A., Braun, A. R. & Horwitz, B. (2006). Neural bases of categorization of simple speech and nonspeech sounds. *Human brain mapping*, *27*(8), 636–651.

Jamieson, D. G. (1987). Studies of possible psychoacoustic factors underlying speech perception. In *The psychophysics of speech perception* (pp. 220–230). Springer.

Klatt, D. H. (1973). Discrimination of fundamental frequency contours in synthetic speech: implications for models of pitch perception. *The Journal of the Acoustical Society of America*, *53*(1), 8–16.

Klein, D., Zatorre, R. J., Milner, B. & Zhao, V. (2001). A cross-linguistic PET study of tone perception in Mandarin Chinese and English speakers. *Neuroimage*, *13*(4), 646–653.

Koelsch, S., Schröger, E. & Tervaniemi, M. (1999). Superior pre-attentive auditory processing in musicians. *Neuroreport*, *10*(6), 1309–1313.

Kolinsky, R., Cuvelier, H., Goetry, V., Peretz, I. & Morais, J. (2009). Music training facilitates lexical stress processing. *Music Perception*, *26*(3), 235–246.

Kraus, N. & Chandrasekaran, B. (2010). Music training for the development of auditory skills. *Nature Reviews Neuroscience*, *11*(8), 599–605.

Krishnan, A., Gandour, J. T. & Bidelman, G. M. (2010). The effects of tone language experience on pitch processing in the brainstem. *Journal of neurolinguistics*, *23*(1), 81–95.

Lee, C. Y. & Hung, T. H. (2008). Identification of Mandarin tones by English-speaking musicians and nonmusicians. *The Journal of the Acoustical Society of America*, *124*(5), 3235–3248.

Lee, C. Y. & Lee, Y. F. (2010). Perception of musical pitch and lexical tones by Mandarin-speaking musicians. *The Journal of the Acoustical Society of America*, *127*(1), 481–490.

Levitin, D. J. & Rogers, S. E. (2005). Absolute pitch: perception, coding, and controversies. *Trends in Cognitive Sciences, 9*(1), 26–33.

Lewis, R. A. & Wieczorkowska, A. (2007). Categorization of musical instrument sounds based on numerical parameters. In *ICCS 2007* (pp. 87–93). Springer.

Liberman, A. M. (1993). Some assumptions about speech and how they changed. *Haskins Laboratories Status Report on Speech Research*, *113,* 1–32.

Liberman, A. M., Harris, K. S., Hoffman, H. S. & Griffith, B. C. (1957). The discrimination of speech sounds within and across phoneme boundaries. *Journal of Experimental Psychology; Journal of Experimental Psychology*, *54*(5), 358.

Liberman, A. M., Harris, K. S., Kinney, J. A. & Lane, H. (1961). The discrimination of relative onset-time of the components of certain speech and nonspeech patterns. *Journal of experimental psychology*, *61*(5), 379–385.

Liu, C. (2013). Just noticeable difference of tone pitch contour change for English-and Chinese-native listeners. *The Journal of the Acoustical Society of America*, *134*(4), 3011–3020.

Loureiro, M. A., de Paula, H. B. & Yehia, H. C. (2004). Timbre Classification Of A Single Musical Instrument. In *ISMIR.*

Magne, C., Schön, D. & Besson, M. (2006). Musician children detect pitch violations in both music and language better than nonmusician children: behavioral and electrophysiological approaches. *Journal of Cognitive Neuroscience*, *18*(2), 199–211.

Marie, C., Delogu, F., Lampis, G., Belardinelli, M. O. & Besson, M. (2011). Influence of musical expertise on segmental and tonal processing in Mandarin Chinese. *Journal of cognitive neuroscience*, *23*(10), 2701–2715.

Marques, C., Moreno, S., & Besson, M. (2007). Musicians detect pitch violation in a foreign language better than nonmusicians: Behavioral and electrophysiological evidence. *Journal of Cognitive Neuroscience*, *19*(9), 1453–1463.

Mattock, K. & Burnham, D. (2006). Chinese and English infants' tone perception: Evidence for perceptual reorganization. *Infancy*, *10*(3), 241–265.

Milovanov, R., Pietilä, P., Tervaniemi, M. & Esquef, P. A. (2010). Foreign language pronunciation skills and musical aptitude: A study of Finnish adults with higher education. *Learning and Individual Differences*, *20*(1), 56–60.

Milovanov, R. & Tervaniemi, M. (2011). The interplay between musical and linguistic aptitudes: a review. *Frontiers in psychology*, *2*, 1–6.

Mirman, D., Holt, L. L. & McClelland, J. L. (2004). Categorization and discrimination of nonspeech sounds: Differences between steady-state and rapidly-changing acoustic cues. *The Journal of the Acoustical Society of America*, *116*(2), 1198–1207.

Moore, C. B. & Jongman, A. (1997). Speaker normalization in the perception of Mandarin Chinese tones. *The Journal of the Acoustical Society of America*, *102*(3), 1864–1877.

Moore, D. R., Amitay, S. & Hawkey, D. J. (2003). Auditory perceptual learning. *Learning & Memory*, *10*(2), 83–85.

Moreno, S. (2009). Can music influence language and cognition? *Contemporary Music Review*, *28*(3), 329–345.

Moreno, S. & Besson, M. (2006). Musical training and language-related brain electrical activity in children. *Psychophysiology*, *43*(3), 287–291.

Moreno, S., Bialystok, E., Barac, R., Schellenberg, E. G., Cepeda, N. J. & Chau, T. (2011). Short-term music training enhances verbal intelligence and executive function. *Psychological science*, *22*(11), 1425–1433.

Moreno, S., Marques, C., Santos, A., Santos, M., & Besson, M. (2009). Musical training influences linguistic abilities in 8-year-old children: more evidence for brain plasticity. *Cerebral Cortex*, *19*(3), 712–723.

Morrison, S. J. & Demorest, S. M. (2009). Cultural constraints on music perception and cognition. In J.Y. Chiao (Ed.), *Progress in brain research* (Vol. 178, pp. 67–77). Elsevier.

Nordenhake, M. & Svantesson, J. O. (1983). Duration of standard Chinese word tones in different sentence environments.

Parbery-Clark, A., Strait, D. & Kraus, N. (2011). Context-dependent encoding in the auditory brainstem subserves enhanced speech-in-noise perception in musicians. *Neuropsychologia*, *49*(12), 3338–3345.

Parbery-Clark, A., Tierney, A., Strait, D. L. & Kraus, N. (2012). Musicians have fine-tuned neural distinction of speech syllables. *Neuroscience*, *219*, 111–119.

Patel, A. D. (2008). *Music, language, and the brain*. Oxford University Press, USA.

Patel, A. D. (2012). Language, music, and the brain: a resource-sharing framework. In Michael A. Arbib (Ed.), *Language, Brain, and Music* (Vol. 10, pp. 204–223). Oxford University Press Oxford, UK.

Patel, A. D., Foxton, J. M. & Griffiths, T. D. (2005). Musically tone-deaf individuals have difficulty discriminating intonation contours extracted from speech. *Brain and cognition*, *59*(3), 310–313.

Patel, A. D. & Iversen, J. R. (2007). The linguistic benefits of musical abilities. *Trends in cognitive sciences*, *11*(9), 369–372.

Patel, A. D., Wong, M., Foxton, J., Lochy, A. & Peretz, I. (2008). Speech intonation perception deficits in musical tone deafness (congenital amusia). *Music Perception: An Interdisciplinary Journal*, *25*(4), 357–368.

Peng, G., Zheng, H. Y., Gong, T., Yang, R. X., Kong, J. P. & Wang, W. S. Y. (2010). The influence of language experience on categorical perception of pitch contours. *Journal of Phonetics*, *38*(4), 616–624.

Peretz, I. & Coltheart, M. (2003). Modularity of music processing. *Nature neuroscience*, *6*(7), 688–691.

Peretz, I., Nguyen, S. & Cummings, S. (2011). Tone language fluency impairs pitch discrimination. *Frontiers in psychology*, *2*(145), 1–5.

Peretz, I., Radeau, M. & Arguin, M. (2004). Two-way interactions between music and language: evidence from priming recognition of tune and lyrics in familiar songs. *Memory & cognition*, *32*(1), 142–152.

Perrachione, T., Fedornko, E., Vinke, L., Gibson, E. & Dilley, L. (2011). Pitch Processing is Shared betwen Language and Music.

Perwez, N. (2011). Categorical Perception of Tones in Vietnamese. In *ICPhs XVII*. Hong Kong.

Pfordresher, P. Q. & Brown, S. (2009). Enhanced production and perception of musical pitch in tone language speakers. *Attention, Perception, \& Psychophysics*, *71*(6), 1385–1398.

Pisoni, D. B. (1973). Auditory and phonetic memory codes in the discrimination of consonants and vowels. *Attention, Perception, & Psychophysics*, *13*(2), 253–260.

Pisoni, D. B. (1977). Identification and discrimination of the relative onset time of two component tones: implications for voicing perception in stops. *The Journal of the Acoustical Society of America*, *61*(5), 1352–1361.

Pisoni, D. B. & Lazarus, J. H. (1974). Categorical and noncategorical modes of speech perception along the voicing continuum. *The Journal of the Acoustical Society of America*, *55*(2), 328–333.

Pollack, I., & Pisoni, D. (1971). On the comparison between identification and discrimination tests in speech perception. *Psychonomic Science*, 24(6), 299-300.

Qin, Z. & Mok, P. (2012). The perception of speech and non-speech tones by tone and non-tone language listeners. In *Speech Prosody 2012*.

Repp, B. H. (1984). Categorical perception: Issues, methods, findings. *Speech and language: Advances in basic research and practice*, *10*, 243–335.

Repp, B. H. & Liberman, A. M. (1987). Phonetic category boundaries are flexible. In *Categorical perception: The groundwork of cognition* (pp. 89–112). Cambridge University Press Cambridge.

Sadakata, M. & Sekiyama, K. (2011). Enhanced perception of various linguistic features by musicians: A cross-linguistic study. *Acta psychologica*, *138*(1), 1–10.

Schellenberg, E. G. (2004). Music lessons enhance IQ. *Psychological Science*, *15*(8), 511–514.

Schellenberg, E. G. & Trehub, S. E. (2003). Good pitch memory is widespread. *Psychological Science*, *14*(3), 262–266.

Schellenberg, E. G. & Trehub, S. E. (2008). Is there an Asian advantage for pitch memory? *Music Perception*, *25*(3), 241–252.

Schneider, K., Dogil, G. & Möbius, B. (2009). German boundary tones show categorical perception and a perceptual magnet effect when presented in different contexts. In *INTERSPEECH* (pp. 2519–2522).

Schneider, K. & Lintfert, B. (2003). Categorical perception of boundary tones in German. In *Proceedings of the 15th International Conference of the Phonetic Sciences* (pp. 631–634).

Schön, D., Magne, C. & Besson, M. (2004). The music of speech: Music training facilitates pitch processing in both music and language. *Psychophysiology*, *41*(3), 341–349.

Siegel, J. A. & Siegel, W. (1977). Categorical perception of tonal intervals: Musicians can't tell sharp from flat. *Perception & Psychophysics*, *21*(5), 399–407.

Silverman, M. J. (2010). The effect of pitch, rhythm, and familiarity on working memory and anxiety as measured by digit recall performance. *Journal of music therapy*, *47*(1), 70–83.

Silverman, M. J. (2012). Effects of melodic complexity and rhythm on working memory as measured by digit recall performance. *Music and Medicine*, *4*(1), 22–27.

Skoe, E. & Kraus, N. (2012). A little goes a long way: how the adult brain is shaped by musical training in childhood. *The Journal of Neuroscience*, *32*(34), 11507–11510.

Sloutsky, V. M. (2003). The role of similarity in the development of categorization. *Trends in Cognitive Sciences*, *7*(6), 246-251.

Smith, J., Nelson, D. G., Grohskopf, L. A. & Appleton, T. (1994). What child is this? What interval was that? Familiar tunes and music perception in novice listeners. *Cognition*, *52*(1), 23–54.

Smits, R., Sereno, J. & Jongman, A. (2006). Categorization of sounds. *JOURNAL OF EXPERIMENTAL PSYCHOLOGY HUMAN PERCEPTION AND PERFORMANCE*, *32*(3), 733–754.

Stevens, C. J., Keller, P. E. & Tyler, M. D. (2013). Tonal language background and detecting pitch contour in spoken and musical items. *Psychology of Music*, *41*(1), 59–74.

Stevens, C., Keller, P. & Tyler, M. (2004). Language tonality and its effects on the perception of contour in short spoken and musical items. In *Proceedings of the 8th lnternational Conference on Music Perception & Cognition. Adelaide, Australia*.

Strait, D. L., Kraus, N., Parbery-Clark, A. & Ashley, R. (2010). Musical experience shapes top-down auditory mechanisms: evidence from masking and auditory attention performance. *Hearing research*, *261*(1), 22–29.

Strait, D. L., O'Connell, S., Parbery-Clark, A. & Kraus, N. (2013). Musicians' enhanced neural differentiation of speech sounds arises early in life: developmental evidence from ages 3 to 30. *Cerebral Cortex*, bht103.

Sun, K. C. & Huang, T. (2012). A cross-linguistic study of Taiwanese tone perception by Taiwanese and English listeners. *Journal of East Asian Linguistics*, *21*, 305–3217.

Tervaniemi, M., Just, V., Koelsch, S., Widmann, A. & Schröger, E. (2005). Pitch discrimination accuracy in musicians vs nonmusicians: an event-related potential and behavioral study. *Experimental Brain Research*, *161*(1), 1–10.

Tervaniemi, M., Kruck, S., De Baene, W., Schröger, E., Alter, K. & Friederici, A. (2009). Top-down modulation of auditory processing: effects of sound context, musical expertise and attentional focus. *European Journal of Neuroscience*, *30*(8), 1636–1642.

Trehub, S. E. & Hannon, E. E. (2006). Infant music perception: domain-general or domain-specific mechanisms? *Cognition*, *100*(1), 73–99.

Van Hessen, A. J. & Schouten, M. (1999). Categorical perception as a function of stimulus quality. *Phonetica*, *56*(1-2), 56–72.

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

Wayland, R., Herrera, E. & Kaan, E. (2010). Effects of musical experience and training on pitch contour perception. *Journal of Phonetics*, *38*(4), 654–662.

Wickelgren, W. A. (1965). Acoustic similarity and retroactive interference in short-term memory. *Journal of Verbal Learning and Verbal Behavior*, *4*(1), 53–61.

Wilson, S. J., Lusher, D., Wan, C. Y., Dudgeon, P. & Reutens, D. C. (2009). The neurocognitive components of pitch processing: insights from absolute pitch. *Cerebral Cortex*, *19*(3), 724–732.

Wong, P. C. M., Ciocca, V., Chan, A. H. D., Ha, L. Y. Y., Tan, L. H. & Peretz, I. (2012). Effects of Culture on Musical Pitch Perception. *PloS one*, *7*(4), e33424.

Wong, P. C. M. & Perrachione, T. K. (2007). Learning pitch patterns in lexical identification by native English-speaking adults. *Applied Psycholinguistics*, *28*(4), 565–585.

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.

Wright, A. A. (1999). Auditory list memory and interference processes in monkeys. *Journal of Experimental Psychology Animal Behavior Processes, 25*(3), 284–296.

Xi, J., Zhang, L., Shu, H., Zhang, Y. & Li, P. (2010). Categorical perception of lexical tones in Chinese revealed by mismatch negativity. *Neuroscience*, *170*(1), 223–231.

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*(2), 1063–1074.

Yang, J. & Liu, C. (2012). Categorical Perception of Lexical Tone in 6 to 8-year-old Monolingual and Bilingual Children. *International Journal of Asian Language Processing*, *22*(2), 49–62.

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., Xi, J., Wu, H., Shu, H. & Li, P. (2012). Electrophysiological evidence of categorical perception of Chinese lexical tones in attentive condition. *Neuroreport*, *23*(1), 35–39.

Zheng, H.-Y., Minett, J. W., Peng, G. & Wang, W. S. (2012). The impact of tone systems on the categorical perception of lexical tones: An event-related potentials study. *Language and Cognitive Processes, 27*(2), 184–209.

Ziv, N. & Radin, S. (2014). Absolute and relative pitch: Global versus local processing of chords. *Advances in Cognitive Psychology*, *10*(1), 15–25.

Appendix A.

Mandarin Word Lists for Recording

土	物	批	罵
五	夫	移	意
無	附	屋	以
一	賭	皮	普
媽	闢	瀑	馬
府	禿	葡	痞
麻	圖	撲	度
督	兔	服	讀

Appendix B.

Participants Musical Training Background

Subject Number	Duration	Onset	Major Type of Instruments	Additional Instruments
Subject 016	10	N/A	Piano	
Subject 027	9	7	Piano	Flute/Guitar
Subject 009	8	8	Piano	
Subject 019	10	10	Piano	Clarinet
Subject 020	16	12	Guitar	
Subject 029	7	12	Clarinet	Guitar
Subject 030	8	17	Flute	
Subject 034	6	11	Piano	Cello
Subject 035	13	8	Violin	
Subject 039	10	13	Piano	Choir (Vocal)
Subject 040	8	14	Piano	Guitar
Subject 036	10	6	Piano	
Subject 042	25	5	Piano,	Turntables/Viola/Percussion/Electronic Music Composition
Subject 038	8	11	Saxophone	Bagpipes/Snare Drums/Piano
Subject 044	5	12	Piano	Clarinet
Subject 043	16	12	Bass	Acoustic Guitar/Trumpet
Subject 045	10	N/A	Piano	Flute
Subject 047	14	4	Piano	Guitar/Trumpet

Appendix C.

Comparisons for Rising Pairs in Mandarin Discrimination

Level	- Level	Difference	Std Err Dif	Lower CL	Upper CL	p-Value
Pair(5&6)	**Pair(1&2)**	**0.1028241**	**0.0180465**	**0.047984**	**0.1576638**	**<.0001***
Pair(5&6)	**Pair(8&9)**	**0.0810617**	**0.0180465**	**0.026222**	**0.1359014**	**0.0002***
Pair(5&6)	**Pair(3&4)**	**0.0761715**	**0.0180465**	**0.021332**	**0.1310112**	**0.0007***
Pair(5&6)	**Pair(7&8)**	**0.0744491**	**0.0180465**	**0.019609**	**0.1292888**	**0.0011***
Pair(5&6)	**Pair(4&5)**	**0.0738327**	**0.0180465**	**0.018993**	**0.1286723**	**0.0012***
Pair(5&6)	**Pair(2&3)**	**0.0707894**	**0.0180465**	**0.015950**	**0.1256291**	**0.0024***
Pair(5&6)	**Pair(6&7)**	**0.0638592**	**0.0180465**	**0.009020**	**0.1186989**	**0.0100***
Pair(6&7)	Pair(1&2)	0.0389649	0.0180465	-0.015875	0.0938046	0.3779
Pair(2&3)	Pair(1&2)	0.0320347	0.0180465	-0.022805	0.0868744	0.6372
Pair(4&5)	Pair(1&2)	0.0289914	0.0180465	-0.025848	0.0838311	0.7465
Pair(7&8)	Pair(1&2)	0.0283750	0.0180465	-0.026465	0.0832147	0.7669
Pair(3&4)	Pair(1&2)	0.0266526	0.0180465	-0.028187	0.0814923	0.8197
Pair(8&9)	Pair(1&2)	0.0217624	0.0180465	-0.033077	0.0766021	0.9302
Pair(6&7)	Pair(8&9)	0.0172025	0.0180465	-0.037637	0.0720422	0.9805
Pair(6&7)	Pair(3&4)	0.0123123	0.0180465	-0.042527	0.0671520	0.9974
Pair(6&7)	Pair(7&8)	0.0105899	0.0180465	-0.044250	0.0654296	0.9990
Pair(2&3)	Pair(8&9)	0.0102723	0.0180465	-0.044567	0.0651120	0.9992
Pair(6&7)	Pair(4&5)	0.0099735	0.0180465	-0.044866	0.0648132	0.9993
Pair(4&5)	Pair(8&9)	0.0072290	0.0180465	-0.047611	0.0620687	0.9999
Pair(6&7)	Pair(2&3)	0.0069302	0.0180465	-0.047909	0.0617699	0.9999
Pair(7&8)	Pair(8&9)	0.0066126	0.0180465	-0.048227	0.0614523	1.0000
Pair(2&3)	Pair(3&4)	0.0053821	0.0180465	-0.049458	0.0602218	1.0000
Pair(3&4)	Pair(8&9)	0.0048902	0.0180465	-0.049949	0.0597299	1.0000
Pair(2&3)	Pair(7&8)	0.0036597	0.0180465	-0.051180	0.0584994	1.0000
Pair(2&3)	Pair(4&5)	0.0030433	0.0180465	-0.051796	0.0578830	1.0000
Pair(4&5)	Pair(3&4)	0.0023388	0.0180465	-0.052501	0.0571785	1.0000
Pair(7&8)	Pair(3&4)	0.0017224	0.0180465	-0.053117	0.0565621	1.0000
Pair(4&5)	Pair(7&8)	0.0006164	0.0180465	-0.054223	0.0554561	.

Appendix D.

Comparisons for Falling Pairs in Mandarin Discrimination

Level	- Level	Difference	Std Err Dif	Lower CL	Upper CL	p-Value
Pair(5&6)	**Pair(2&3)**	**0.1246296**	**0.0183191**	**0.068961**	**0.1802977**	**<.0001***
Pair(5&6)	**Pair(6&7)**	**0.1236502**	**0.0183191**	**0.067982**	**0.1793183**	**<.0001***
Pair(5&6)	**Pair(1&2)**	**0.1154876**	**0.0183191**	**0.059819**	**0.1711557**	**<.0001***
Pair(5&6)	**Pair(3&4)**	**0.1118852**	**0.0183191**	**0.056217**	**0.1675533**	**<.0001***
Pair(5&6)	**Pair(7&8)**	**0.1101913**	**0.0183191**	**0.054523**	**0.1658595**	**<.0001***
Pair(5&6)	**Pair(8&9)**	**0.1093936**	**0.0183191**	**0.053726**	**0.1650618**	**<.0001***
Pair(5&6)	**Pair(4&5)**	**0.1072077**	**0.0183191**	**0.051540**	**0.1628758**	**<.0001***
Pair(4&5)	Pair(2&3)	0.0174219	0.0183191	-0.038246	0.0730900	0.9808
Pair(4&5)	Pair(6&7)	0.0164425	0.0183191	-0.039226	0.0721106	0.9863
Pair(8&9)	Pair(2&3)	0.0152360	0.0183191	-0.040432	0.0709041	0.9913
Pair(7&8)	Pair(2&3)	0.0144383	0.0183191	-0.041230	0.0701064	0.9937
Pair(8&9)	Pair(6&7)	0.0142566	0.0183191	-0.041412	0.0699247	0.9942
Pair(7&8)	Pair(6&7)	0.0134589	0.0183191	-0.042209	0.0691270	0.9959
Pair(3&4)	Pair(2&3)	0.0127444	0.0183191	-0.042924	0.0684125	0.9971
Pair(3&4)	Pair(6&7)	0.0117650	0.0183191	-0.043903	0.0674331	0.9983
Pair(1&2)	Pair(2&3)	0.0091421	0.0183191	-0.046526	0.0648102	0.9997
Pair(4&5)	Pair(1&2)	0.0082798	0.0183191	-0.047388	0.0639479	0.9998
Pair(1&2)	Pair(6&7)	0.0081627	0.0183191	-0.047505	0.0638308	0.9998
Pair(8&9)	Pair(1&2)	0.0060939	0.0183191	-0.049574	0.0617620	1.0000
Pair(7&8)	Pair(1&2)	0.0052962	0.0183191	-0.050372	0.0609643	1.0000
Pair(4&5)	Pair(3&4)	0.0046775	0.0183191	-0.050991	0.0603456	1.0000
Pair(3&4)	Pair(1&2)	0.0036024	0.0183191	-0.052066	0.0592705	1.0000
Pair(4&5)	Pair(7&8)	0.0029836	0.0183191	-0.052684	0.0586517	1.0000
Pair(8&9)	Pair(3&4)	0.0024916	0.0183191	-0.053177	0.0581597	1.0000
Pair(4&5)	Pair(8&9)	0.0021859	0.0183191	-0.053482	0.0578540	1.0000
Pair(7&8)	Pair(3&4)	0.0016939	0.0183191	-0.053974	0.0573620	1.0000
Pair(6&7)	Pair(2&3)	0.0009794	0.0183191	-0.054689	0.0566475	.
Pair(8&9)	Pair(7&8)	0.0007977	0.0183191	-0.054870	0.0564658	.

Appendix E.

Comparisons for the Predicted Rising Pairs

Level	- Level	Difference	Std Err Dif	Lower CL	Upper CL	p-Value
(5&6)	**(1&2)**	**0.0928793**	**0.0118243**	**0.069670**	**0.1160882**	**<.0001***
(5&6)	**(2&3)**	**0.0898119**	**0.0118243**	**0.066603**	**0.1130209**	**<.0001***
(5&6)	**(8&9)**	**0.0806960**	**0.0118243**	**0.057487**	**0.1039050**	**<.0001***
(5&6)	**(7&8)**	**0.0747907**	**0.0118243**	**0.051582**	**0.0979997**	**<.0001***
(5&6)	**(3&4)**	**0.0692294**	**0.0118243**	**0.046020**	**0.0924384**	**<.0001***
(4&5)	**(1&2)**	**0.0683695**	**0.0118243**	**0.045160**	**0.0915784**	**<.0001***
(4&5)	**(2&3)**	**0.0653021**	**0.0118243**	**0.042093**	**0.0885111**	**<.0001***
(5&6)	**(6&7)**	**0.0636968**	**0.0118243**	**0.040488**	**0.0869058**	**<.0001***
(4&5)	**(8&9)**	**0.0561862**	**0.0118243**	**0.032977**	**0.0793952**	**<.0001***
(4&5)	**(7&8)**	**0.0502809**	**0.0118243**	**0.027072**	**0.0734899**	**<.0001***
(4&5)	**(3&4)**	**0.0447196**	**0.0118243**	**0.021511**	**0.0679286**	**0.0002***
(4&5)	**(6&7)**	**0.0391870**	**0.0118243**	**0.015978**	**0.0623960**	**0.0010***
(6&7)	**(1&2)**	**0.0291824**	**0.0118243**	**0.005973**	**0.0523914**	**0.0138***
(6&7)	**(2&3)**	**0.0261151**	**0.0118243**	**0.002906**	**0.0493241**	**0.0275***
(5&6)	**(4&5)**	**0.0245098**	**0.0118243**	**0.001301**	**0.0477188**	**0.0385***
(3&4)	**(1&2)**	**0.0236498**	**0.0118243**	**0.000441**	**0.0468588**	**0.0458***
(3&4)	(2&3)	0.0205825	0.0118243	-0.002626	0.0437915	0.0821
(7&8)	(1&2)	0.0180885	0.0118243	-0.005120	0.0412975	0.1265
(6&7)	(8&9)	0.0169992	0.0118243	-0.006210	0.0402082	0.1509
(7&8)	(2&3)	0.0150212	0.0118243	-0.008188	0.0382302	0.2043
(8&9)	(1&2)	0.0121832	0.0118243	-0.011026	0.0353922	0.3031
(3&4)	(8&9)	0.0114666	0.0118243	-0.011742	0.0346756	0.3325
(6&7)	(7&8)	0.0110939	0.0118243	-0.012115	0.0343029	0.3484
(8&9)	(2&3)	0.0091159	0.0118243	-0.014093	0.0323249	0.4410
(7&8)	(8&9)	0.0059053	0.0118243	-0.017304	0.0291143	0.6176
(3&4)	(7&8)	0.0055613	0.0118243	-0.017648	0.0287703	0.6382
(6&7)	(3&4)	0.0055326	0.0118243	-0.017676	0.0287416	0.6400
(2&3)	(1&2)	0.0030673	0.0118243	-0.020142	0.0262763	0.7954

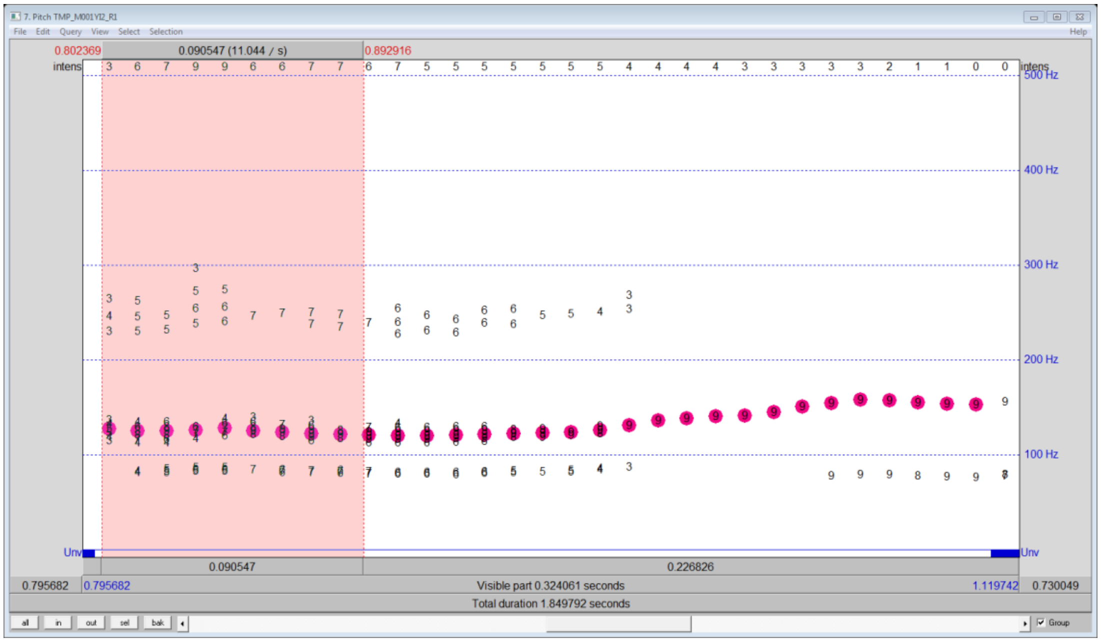
Appendix F.

Comparisons for the Predicted Falling Pairs

Level	- Level	Difference	Std Err Dif	Lower CL	Upper CL	p-Value
(5&6)	**(1&2)**	**0.0928793**	**0.0118243**	**0.069670**	**0.1160882**	**<.0001***
(5&6)	**(2&3)**	**0.0898119**	**0.0118243**	**0.066603**	**0.1130209**	**<.0001***
(5&6)	**(8&9)**	**0.0806960**	**0.0118243**	**0.057487**	**0.1039050**	**<.0001***
(5&6)	**(7&8)**	**0.0747907**	**0.0118243**	**0.051582**	**0.0979997**	**<.0001***
(5&6)	**(3&4)**	**0.0692294**	**0.0118243**	**0.046020**	**0.0924384**	**<.0001***
(4&5)	**(1&2)**	**0.0683695**	**0.0118243**	**0.045160**	**0.0915784**	**<.0001***
(4&5)	**(2&3)**	**0.0653021**	**0.0118243**	**0.042093**	**0.0885111**	**<.0001***
(5&6)	**(6&7)**	**0.0636968**	**0.0118243**	**0.040488**	**0.0869058**	**<.0001***
(4&5)	**(8&9)**	**0.0561862**	**0.0118243**	**0.032977**	**0.0793952**	**<.0001***
(4&5)	**(7&8)**	**0.0502809**	**0.0118243**	**0.027072**	**0.0734899**	**<.0001***
(4&5)	**(3&4)**	**0.0447196**	**0.0118243**	**0.021511**	**0.0679286**	**0.0002***
(4&5)	**(6&7)**	**0.0391870**	**0.0118243**	**0.015978**	**0.0623960**	**0.0010***
(6&7)	**(1&2)**	**0.0291824**	**0.0118243**	**0.005973**	**0.0523914**	**0.0138***
(6&7)	**(2&3)**	**0.0261151**	**0.0118243**	**0.002906**	**0.0493241**	**0.0275***
(5&6)	**(4&5)**	**0.0245098**	**0.0118243**	**0.001301**	**0.0477188**	**0.0385***
(3&4)	**(1&2)**	**0.0236498**	**0.0118243**	**0.000441**	**0.0468588**	**0.0458***
(3&4)	(2&3)	0.0205825	0.0118243	-0.002626	0.0437915	0.0821
(7&8)	(1&2)	0.0180885	0.0118243	-0.005120	0.0412975	0.1265
(6&7)	(8&9)	0.0169992	0.0118243	-0.006210	0.0402082	0.1509
(7&8)	(2&3)	0.0150212	0.0118243	-0.008188	0.0382302	0.2043
(8&9)	(1&2)	0.0121832	0.0118243	-0.011026	0.0353922	0.3031
(3&4)	(8&9)	0.0114666	0.0118243	-0.011742	0.0346756	0.3325
(6&7)	(7&8)	0.0110939	0.0118243	-0.012115	0.0343029	0.3484
(8&9)	(2&3)	0.0091159	0.0118243	-0.014093	0.0323249	0.4410
(7&8)	(8&9)	0.0059053	0.0118243	-0.017304	0.0291143	0.6176
(3&4)	(7&8)	0.0055613	0.0118243	-0.017648	0.0287703	0.6382
(6&7)	(3&4)	0.0055326	0.0118243	-0.017676	0.0287416	0.6400
(2&3)	(1&2)	0.0030673	0.0118243	-0.020142	0.0262763	0.7954

Appendix G.

Measurement of The Turning Point



Appendix H.

Acoustic Details for the Tones

Syllable	Repetition	Onset F0 (Hz)	Coda F0 (Hz)	$\Delta F0$	Duration (s)	Turning Point Duration	
Yi1	1	150.5	153	-2.5	0.301		
Yi2	1	126.9	152.2	-25.3	0.3	0.090547	30.2%
Yi4	2	159	128.3	-7.9	0.369	0.030048	8.1%

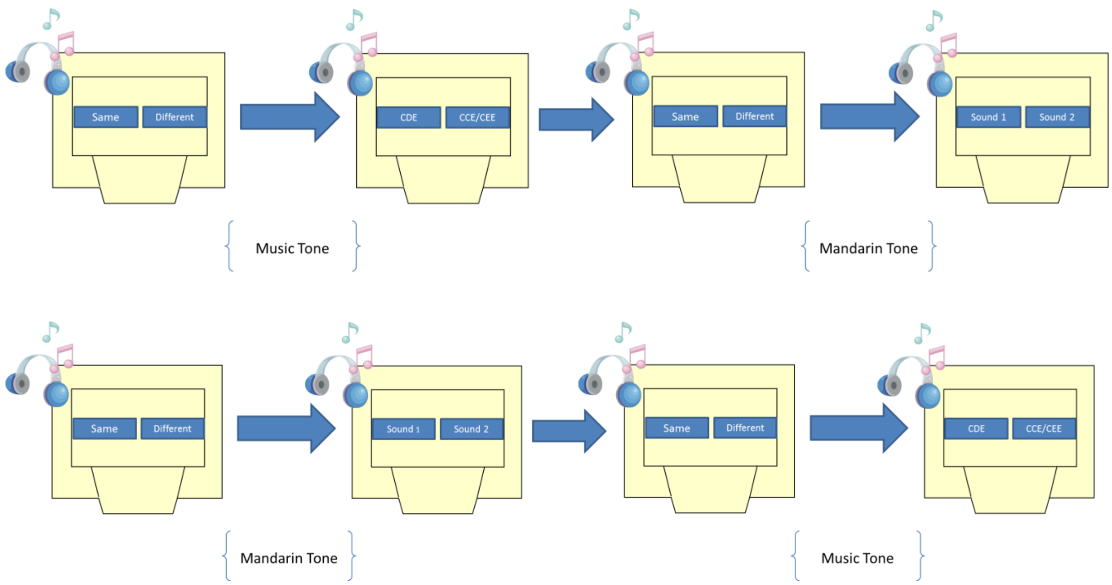
Average Duration for Tone 1, 2, 4 = (0.301+0.3+0.369)/3 = 0.323 s

Tone 2 Turning Point = 323 ms x 30% = 97 ms <- turning point

Tone 4 Turning Point = 323 ms x 10% = 32 ms <- turning point

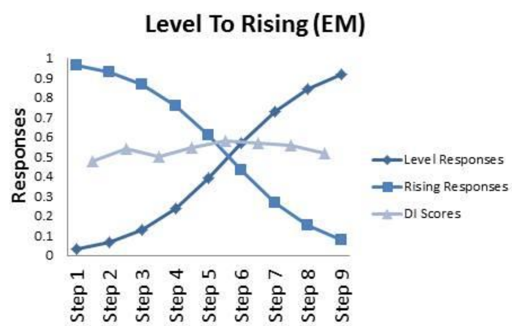
Appendix I.

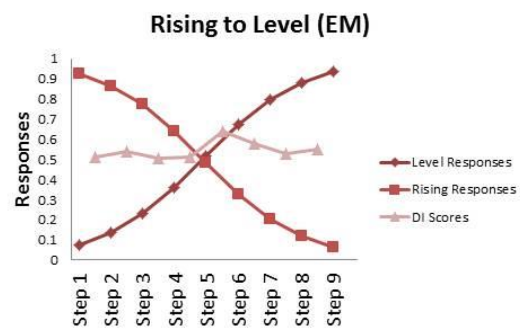
Experiment Flow Chart

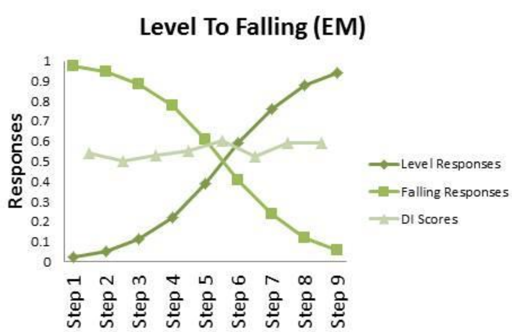


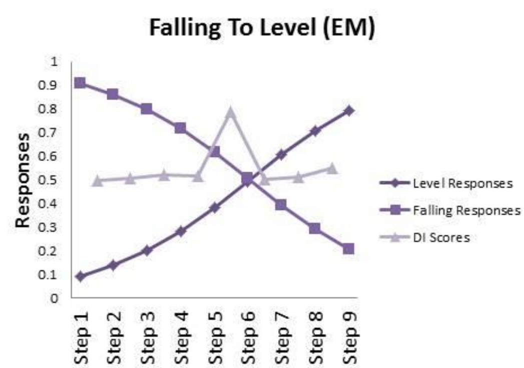
Appendix J.

Combined ID and DI Curves for English Musicians



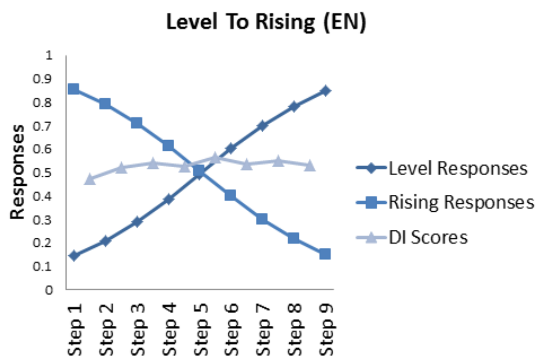


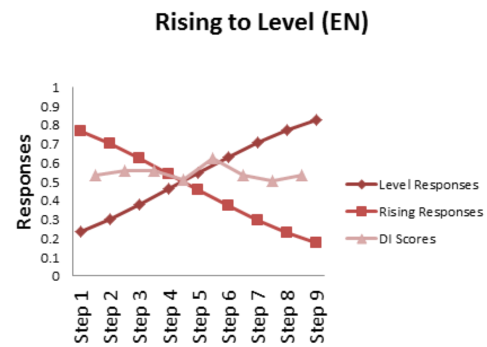


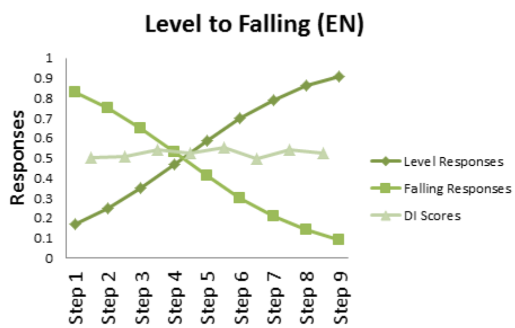


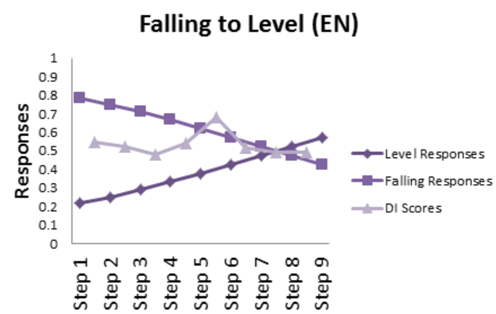
Appendix K.

Combined ID and DI Curves for English Non-musicians









Appendix L.

Combined ID and DI Curves for Mandarin Non-musicians