Is It All Relative? Relative Pitch and L2 Lexical Tone Perception/Tone Language Comprehension by Adult Tone and Non-Tone Language Speakers

Sloane C. von Wertz

The Graduate Center, City University of New York

How does access to this work benefit you? Let us know!

Follow this and additional works at: https://academicworks.cuny.edu/gc_etds

Part of the East Asian Languages and Societies Commons, Language and Literacy Education Commons, and the Music Commons

Recommended Citation

von Wertz, Sloane C., "Is It All Relative? Relative Pitch and L2 Lexical Tone Perception/Tone Language Comprehension by Adult Tone and Non-Tone Language Speakers" (2017). CUNY Academic Works.
https://academicworks.cuny.edu/gc_etds/2247

This Dissertation is brought to you by CUNY Academic Works. It has been accepted for inclusion in All Dissertations, Theses, and Capstone Projects by an authorized administrator of CUNY Academic Works. For more information, please contact deposit@gc.cuny.edu.
IS IT ALL RELATIVE?
RELATIVE PITCH AND L2 LEXICAL TONE PERCEPTION/TONE LANGUAGE COMPREHENSION BY ADULT TONE AND NON-TONE LANGUAGE SPEAKERS

by

SLOANE CELESTE VON WERTZ

A dissertation submitted to the Graduate Faculty in Linguistics in partial fulfillment of the requirements for the degree of Doctor of Philosophy, The City University of New York

2017
Is It All Relative?

Relative Pitch and L2 Lexical Tone Perception/Tone Language Comprehension by Adult Tone and Non-Tone Language Speakers

by

Sloane Celeste von Wertz

This manuscript has been read and accepted for the Graduate Faculty in Linguistics in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy.

________________________________________
Date                                   Gita Martohardjono
                                         Chair of Examining Committee

________________________________________
Date                                   Gita Martohardjono
                                         Executive Officer

Supervisory Committee:

Gita Martohardjono

Andrew Rosenberg

Joseph Straus

THE CITY UNIVERSITY OF NEW YORK
ABSTRACT

Is It All Relative?

Relative Pitch and L2 Lexical Tone Perception/Tone Language Comprehension by Adult Tone and Non-Tone Language Speakers

by

Sloane Celeste von Wertz

Advisor: Professor Gita Martohardjono

Languages generally use musical pitch variation of the voice as part of their sound systems (Maddieson, 2011)—pitch variations that can be somewhat reminiscent of music. Music ability and/or training may influence language processing (e.g., Bidelman et al, 2011; Delogue et al, 2010). In particular, studies have concluded that there may be a relationship between absolute pitch and tone language acquisition (e.g., Lee, Lee, and Shi, 2011; Tillmann et al., 2011; Pfordresher and Brown, 2009). Other research has shown that fundamental frequency (F0) and F0 slope are crucial elements upon which native tone language speakers rely in tone perception (Guion and Pederson, 2007). With the given observations in mind, we could infer that an important tool in tone language processing and/or acquisition would be the ability to identify the relationship between notes (relative pitch ability). This study endeavors to explore the possible relationship between relative pitch aptitude or ability and adult L2 lexical tone perception/tone language comprehension. This study tested native Thai, Mandarin Chinese, and English-only speakers to determine each group’s relative pitch proficiency and the possible correlation
between this proficiency and lexical tone perception and tone language comprehension. The results of this study reveal that tone language speakers are more proficient at relative pitch perception. In addition, Thai speakers are more adept at perceiving the four lexical tones in Mandarin Chinese as well as perceiving Mandarin Chinese words than English-only speakers. Finally, this study found a moderate positive correlation between relative pitch proficiency and lexical tone perception for the English-only speakers but not for the Thai speakers. These findings lead to the conclusion that relative pitch proficiency may be relevant to non-tone language speakers endeavoring to learn a tone language.

Key words: absolute pitch, lexical tone, relative pitch, second language acquisition
ACKNOWLEDGMENTS

There are so many individuals to whom I am indebted for helping me to arrive at the completion of this research. I begin by thanking Gita Martohardjono, my advisor to whom I am eternally grateful. There really are no words that allow me to express fully how much I appreciate Gita for believing in my vision with this research topic, for patiently guiding me through this research, for encouraging me through the rough patches, for helping me to stay focused (it is no secret that I would have followed many odd tangents), and for supporting me from the genesis to the culmination of this work. I also thank my other committee members, Andrew Rosenberg and Joseph Straus. To Andrew, I am grateful for his guidance with experimental design and the fresh ideas to keep participants interested and on task. To Joe, I am grateful for guiding me through aspects of music that I never realized that I did not know, especially cultural-specific features. Indeed, I could not have asked for a better committee than the one that I had! The direction provided by my committee has been invaluable and will be forever treasured.

I would like to thank those individuals who helped me to begin this academic journey. I thank Linda Grasso for taking a moment one day and sharing with me the joys of the scholarly endeavor and instilling in me the idea that my undertaking would allow me to utilize my love for both language and music. I thank Christina Tortora for lighting the linguistics fire in my heart. How vividly I remember sitting in her Dialects of English Syntax class and thinking, “I want to become a linguist!” I thank all of the faculty and classmates with whom I had the pleasure of studying in the department of Linguistics at the Graduate Center. I especially thank Dianne Bradley and Robert Fiengo who co-taught Introduction to Theoretical Linguistics which helped
to facilitate my transition into linguistics. I would also like to thank the Second Language Acquisition Lab staff for allowing me the opportunity to begin to test my ideas early in my program and for helping me in each endeavor. I thank Juliette Blevins for teaching me about linguistic tone. Thank you Nishi Bissoondial for always having that calming spirit and smile whenever I walked into the department.

I am most grateful to the Magnet Scholars Program and the Dean K. Harrison Scholarship at the Graduate Center. Without the funding provided by these two significant programs, I would never have been able to complete my research nor the program. I thank the Parsons Family to whom I am eternally grateful for funding the last three years of my research. Indeed, the most generous scholarship in memory of the Parsons’ matriarch and patriarch helped me to finish this program. Thank you to Sherry Greene for tirelessly following-up with the various resources to help ensure that the funding was in place.

Without the help of many collaborators, I would not have been able to collect and analyze data for this project. I thank Lie Liu, Hai Ning, Linda Chang, and several other anonymous individuals for recording so many versions of various tokens. I am so very grateful to Donna Grant, Blerina Likollari, and Polina Belimova for helping with participant recruitment.

Outside of academics, I thank Marina Rousseau for supporting me and guiding me in so many wonderful ways. I thank Bernice Cohen for the meals and the motherly listening ear. I thank Catherine Perkins as well as the members of the New Hope Lutheran Choir who walked with me through the first portion of this journey and encouraged me in so many ways along the way. I thank Linda Assini for joining me on the last leg of this journey and being a cheerleader whose encouraging words helped me to cross the finish line.
I thank my family for all of the tireless support and encouragement. I especially thank my dearest son, Derrick, for tolerating my craziness as I strove to reach this milestone. (For the record, you are still my sunshine and have always been my inspiration.) I thank my brother, Trevor, for subsidizing my internet access throughout my program. To my mother and father, I am so very grateful for teaching me the importance of education and for mainstreaming me in a way that taught me how to be persistent. Although I have struggled with vision issues my entire life, the persistent spirit that I learned as a child helped me to become who I am and has brought me to where I am today.

There are so many friends whom I wish to thank for helping me along the way. I begin by thanking my York College family and my College of Staten Island family for the support. I also wish to thank the following individuals who helped me in their own special way: Cady-Ann Parris David, Carolyn Kirkpatrick, Eva Roca, Eva and Arek Szymanski, Fenix Arias, Fredricka Grasso, Linda Perry, Marcia Keizs, Monica and Charles Dyer, Patricia Milanes, Olga Dais, Sharon Beharry-Singh, Valerie Anderson, Wendy Pearson, and Zofia and Wojtek Stankiewicz. I hope that anyone whom I may have somehow mistakenly forgotten to include will understand that I am so grateful to him or her as well for enhancing my life, particularly during this academic endeavor.

Finally, and most importantly, I thank God for this opportunity, for the magnificence of humanity, and for the privilege to study the marvelous human feature—language.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Introduction</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>Literature Review</td>
<td>7</td>
</tr>
<tr>
<td>2.1</td>
<td>Language and Music</td>
<td>7</td>
</tr>
<tr>
<td>2.1.1</td>
<td>The Development of Language and Music</td>
<td>8</td>
</tr>
<tr>
<td>2.1.2</td>
<td>The Biological Mechanisms Responsible for Language and Music</td>
<td>14</td>
</tr>
<tr>
<td>2.2</td>
<td>Pitch, Fundamental Frequency, and Lexical Tone</td>
<td>16</td>
</tr>
<tr>
<td>2.2.1</td>
<td>Pitch</td>
<td>16</td>
</tr>
<tr>
<td>2.2.2</td>
<td>Fundamental Frequency</td>
<td>17</td>
</tr>
<tr>
<td>2.2.3</td>
<td>Lexical Tone</td>
<td>19</td>
</tr>
<tr>
<td>2.3</td>
<td>Lexical Tone and Adult L2 Acquisition</td>
<td>26</td>
</tr>
<tr>
<td>2.4</td>
<td>Language Acquisition and Musicians/Non-Musicians</td>
<td>30</td>
</tr>
<tr>
<td>2.4.1</td>
<td>L1 and Musicians/Non-Musicians</td>
<td>30</td>
</tr>
<tr>
<td>2.4.2</td>
<td>Musicians/Non-Musicians and L2 Acquisition</td>
<td>31</td>
</tr>
<tr>
<td>2.4.3</td>
<td>Musicians/Non-Musicians and Tone Language L2 Acquisition</td>
<td>33</td>
</tr>
<tr>
<td>2.5</td>
<td>In Summary</td>
<td>37</td>
</tr>
<tr>
<td>3</td>
<td>Methods and Materials</td>
<td>40</td>
</tr>
<tr>
<td>3.1</td>
<td>Hypothesis and Research Questions</td>
<td>40</td>
</tr>
<tr>
<td>3.2</td>
<td>Participants</td>
<td>42</td>
</tr>
<tr>
<td>3.3</td>
<td>Methodology</td>
<td>43</td>
</tr>
</tbody>
</table>
3.3.1 Pre-Experiment Training 43
3.3.2 Experiment One 44
3.3.3 Experiment Two 49
3.3.4 Experiment Three 50
3.3.5 Experiment Four 53
3.4 Hypothesized Results 54

Chapter 4 Results and Discussion 56

4.1 Results 56

4.1.1 Descriptive Statistics 56
4.1.2 Test Reliability: Assessment Tasks 64
4.1.3 Inferential Statistics 64
4.1.4 Results: Research Question One 65
4.1.5 Results: Research Question Two 67
4.1.6 Results: Research Question Three 68
4.1.7 Results: Research Question Four 70

4.2 Discussion 73

4.2.1 Relative Pitch Proficiency in Tone Language and Non-Tone Language Speakers 74
4.2.2 Relative Pitch Proficiency and Lexical Tone Perception 78
4.2.3 Lexical Tone Perception of Tone Language Speakers and Non-Tone Language Speakers 81
4.2.4 A Final Word about Music Training and Lexical Tone Perception 88
4.3 Limitations and Further Research 88
  4.3.1 Limitations 88
  4.3.2 Further Research 89

Chapter 5 Conclusion 91

Appendices 92

References 124
LIST OF FIGURES

Figure 1: Music and Language Development Timeline
(Adapted from http://www.frontiersin.org/files/Articles/31563/fpsyg-03-00327-r2/image_m/fpsyg-03-00327-g002.jpg;
Retrieved December 5, 2014) 13

Figure 2: Musical Notes and their General Corresponding Frequencies
((Based on Suits, 2015). 18

Figure 3: Mandarin Lexical Tones F0 contours
(Adapted from Xu, 1997) 23

Figure 3: Samples Tokens For Experiment 3

Figure 4: Praat-Generated Waveforms and Spectrograms for
Experiment 1 Part A Stimuli 46

Figure 5: Praat-Generated Waveforms and Spectrograms for
Experiment 1 Part B Stimuli 47

Figure 6: Sample tokens for Experiment 3 52

Figure 7: Groups’ Music Proficiency Means 66

Figure 8: Test Groups’ Means (%) in Tone Identification Tasks 67

Figure 9: Mandarin Chinese Lexical Tones 84

Figure 10: Thai Lexical Tones 84
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table 1:</td>
<td>Tone to Indicate Lexical meaning in Mandarin Chinese</td>
<td>22</td>
</tr>
<tr>
<td>Table 2:</td>
<td>Tone to Indicate Grammatical Function in Ngiti Verbs</td>
<td>24</td>
</tr>
<tr>
<td></td>
<td>(Kutsch Lojenga, 1993)</td>
<td></td>
</tr>
<tr>
<td>Table 3:</td>
<td>Tone to Indicate Grammatical Function in Ngiti Nouns</td>
<td>24</td>
</tr>
<tr>
<td></td>
<td>(Adapted from Booij, 2012)</td>
<td></td>
</tr>
<tr>
<td>Table 4:</td>
<td>The Participant Pool</td>
<td>43</td>
</tr>
<tr>
<td>Table 5:</td>
<td>Stimuli in Experiment One by lexical tone number and speaker gender</td>
<td>48</td>
</tr>
<tr>
<td>Table 6:</td>
<td>Stimuli for Experiment Two</td>
<td>50</td>
</tr>
<tr>
<td>Table 7:</td>
<td>Nonsense Words Lexicon with Mandarin Chinese Lexical Tone</td>
<td>51</td>
</tr>
<tr>
<td>Table 8:</td>
<td>Four Mandarin Color Adjectives</td>
<td>54</td>
</tr>
<tr>
<td>Table 9:</td>
<td>Descriptive Statistics (%) on all four Experiments</td>
<td>58</td>
</tr>
<tr>
<td>Table 10:</td>
<td>Descriptive Statistics (%) on Tone Identification Tasks</td>
<td>59</td>
</tr>
<tr>
<td></td>
<td><em>(Exps One, Three, and Four)</em></td>
<td></td>
</tr>
<tr>
<td>Table 11:</td>
<td>Descriptive Statistics (%) on Music Proficiency</td>
<td>60</td>
</tr>
<tr>
<td>Table 12:</td>
<td>Levene’s Test for Experiments One through Four</td>
<td>61</td>
</tr>
<tr>
<td>Table 13:</td>
<td>One-Way ANOVA Results with Group as Independent Variable for Experiments One</td>
<td>63</td>
</tr>
<tr>
<td></td>
<td>Through four</td>
<td></td>
</tr>
<tr>
<td>Table 14:</td>
<td>Pearson’s Correlation Matrix: By Tasks</td>
<td>64</td>
</tr>
</tbody>
</table>
Table 15: Participant Means (%) in the Music Proficiency Tasks 65
Table 16: Test Groups’ Means (%) in Tone Identification Tasks 66
Table 17: Correlation Matrix: English Speakers’ Music Proficiency and Lexical Tone Identification 69
Table 18: Regression Analysis: Music Proficiency and Tone Identification 72
Table 19: Mandarin Chinese and Thai Lexical Tones 83
LIST OF APPENDICES

Appendix 1: Participant In-take Questionnaire 93
Appendix 2: Script for Pre-Experiment Training 96
Appendix 3: Stimuli Blocks for Pre-Experiment training 99
Appendix 4: Pre-Experiment Training Handout 100
Appendix 5: Experiment 1: Lexical Tone Identification Baseline Stimuli 101
Appendix 6: Experiment 1: Lexical Tone Identification Baseline Script 102
Appendix 7: Experiment 2: Relative Pitch Proficiency Measurement Script 104
Appendix 8: Experiment 3: Monosyllabic Nonsense Words
And Lexical Tone Identification Pre-Training materials 106
Appendix 9: Experiment 3 Monosyllabic Nonsense Words
And Lexical Tone Identification Script 111
Appendix 10: Experiment 3 Monosyllabic Nonsense Words
And Lexical Tone Identification Stimuli 114
Appendix 11: Experiment 4: Mandarin Chinese Color Adjectives Script
With Training Materials 116
Appendix 12: Experiment 4: Mandarin Chinese Color Adjectives
Stimuli 120
In memory of my mother who taught me that the sky is the limit.
“Music is the universal language of mankind.”

—Henry Wadsworth Longfellow

“Language is a process of free creation; its laws and principles are fixed, but the manner in which the principles of generation are used is free and infinitely varied. Even the interpretation and use of words involves a process of free creation.”

—Noam Chomsky
Spoken language and music: Is there a relationship between these two universal features of the human experience? With the numerous different languages alive today, there is a gap in communication via language amongst all of humanity; however, there are consistent cross-cultural sounds that significantly narrow this communication gap. For example, there are consistent cross-cultural sounds that communicate the same emotions from culture to culture, such as crying for sadness or laughter for happiness (Sauter et al., 2009). In addition, research has shown that there is similar cross-cultural emotional response to specific chords and melodies. For example, Bakker and Martin (2015) found that certain music chords appear to communicate specific emotions cross-culturally (e.g., major chords are associated with positive emotions while minor chords are associated with negative emotions). Ludden (2015) pointed out that one may watch a foreign film and may not understand the language or exactly what is transpiring, but can still comprehend the emotion of the characters on screen. In this regard, there is cross-cultural recognition of language pitch, rhythm, and tempo; for example, high-pitched, fast speech can indicate happiness or excitement (e.g., Ludden, 2015).

In addition to the aforementioned comments about language and music, there are three factors that point to the strong probability of a relationship between language and music. First, from an ontological viewpoint, both spoken language and vocal music rely on sounds made by
the same vocal apparatus; humans utilize the larynx for all phonation, whether linguistic or musical. Second, all spoken languages utilize musical features including pitch, rhythm, and tempo. Maddieson (2011) points out, for example, that pitch variation is a feature in general of spoken language. Moreover, just as it takes pitch, rhythm, and tempo to communicate a melody in music, language also has its own melody—prosody. Third, although it has traditionally been argued that language and music are two separate general functions that are lateralized differentially in the brain (i.e., language is governed by the left cerebral hemisphere, whereas music is governed by the right cerebral hemisphere), recent neurophysiological evidence supports the argument that music and language are actually related cognitive abilities which are processed by an overlapping of the two cerebral hemispheres (Sammler et al., 2009; Patel, 2003).

Since the given factors indicate that there may be a possible relationship between spoken language and music in general, it is logical to explore if there is a relationship between a specific music competency such as pitch discrimination and a specific language area such as second language (L2) acquisition. It is widely recognized that second language learners struggle with learning a second language, especially adult learners. For the adult learner, not only must he or she contend with age effects (i.e., the older an L2 learner, the more difficult it likely is for him or her to acquire native-like proficiency in an L2 [e.g., Flege, Yeni-Komshian, and Lie 1999]), but also he or she must face other hindrances such as first language (L1) interference (e.g., Lado, 1957). Further, the difficulty in acquiring native-like proficiency in an L2 is even more remarkable if the L1 is a non-tone language and the L2 is a tone language.

There several reasons why the added task of recognizing lexical tone proves difficult for L2 learners. The first and most widely recognized reason is based on the differences in the use
of pitch variance for linguistic purposes. Specifically, for the non-tone language speaker learning a tone language, pitch variance in the L1 is oftentimes used as a pragmatic cue; however, in the L2, pitch variance is used as a cue other than (e.g., semantic or grammatical)—or in addition to—a pragmatic one. Therefore, not only must the learner face the general challenges that are associated with acquiring the primary segmental information, but also he or she has the additional task of acquiring the crucial suprasegmental information. Hence, the non-tone language speaker attempting to learn a tone language has to learn to use pitch variation (linguistic tone) for semantic or grammatical information in addition to pragmatic information. Truly, this one reason illustrates how training for the non-tone language speaker learning a tone language takes additional time and effort (Wang et al., 2003). Second, brain processing may be somewhat problematic for the L2 learner. In a brain scan study involving non-tone language speakers listening to Mandarin Chinese, Klein et al. (2001) found that non-tone language speakers process the tones of the words in the right hemisphere of the brain. This processing occurs shortly before processing the semantics in the left hemisphere of the brain (e.g., Devlin, Matthews, and Rushworth, 2003) which may lead to temporary confusion for the learner. This temporary confusion not only slows processing but also hinders progress towards ultimate attainment. Third, it appears that non-tone language speakers process F0 (fundamental frequency—a primary aspect of pitch perception) differently than tone-language speakers (Gandour and Harshman, 1978; Lee, Lee, and Shi, 2011). Research has shown that speakers of non-tone languages are more sensitive to pitch height than direction whereas speakers of tone languages place more emphasis on pitch direction when it comes to lexical tone perception (Qin and Mok, 2011).
These three general observations clearly show the primary issues faced by the non-tone language speaker learning a tone language. However, what is most significant for the purposes of this research is the processing of F0. Since non-tone language speakers are more sensitive to pitch height than direction, it is appears that in order for a non-tone language speaker to begin to comprehend lexical tone, the phonological environment is most crucial in aiding him or her to perceive the lexical tone. Because of the importance of the phonological environment (or context), there must be a mechanism that ultimately enables the non-tone language speaker endeavoring to learn a tone language to distinguish the pitches/pitch contours in context (and the relationships thereof) for there to be viable communication. This mechanism may be music-related, specifically a competency based on pitch assessment.

There are two general music skills when it comes to pitch assessment: absolute pitch and relative pitch. Interestingly, absolute pitch is the ability to name a musical note without hearing a reference note (Bachem, 1955; Profita and Bidder, 1988 [cited in Deutsch, Dooley Henthorn, and Head 2009]). Many tone-language speakers (although still a small minority relative to the overall population) who are musicians have absolute pitch (e.g., Lee, Lee, and Shi, 2011; Deutsch et al, 2006). Deutsch, Henthorn, Marvin, and Xu (2006), for example, found that Chinese speakers are more likely to have absolute pitch than English speakers. Relative pitch (an ability that is more common than absolute pitch) is the ability to recognize the relationships and intervals between notes. Research has shown that tone language speakers are more accurate at imitating musical pitch and discriminating intervals than speakers of non-tone languages (e.g., Pfmdrescher and Brown, 2009). Absolute pitch, an ability that is impossible to acquire as an adult, is inherent. However, although age effects
may possibly impact relative pitch proficiency, relative pitch ability is “virtually universal, or could be—given adequate training” (Straus, 2014). Hence, relative pitch, unlike absolute pitch, can be taught—even to an adult.

The observation that relative pitch can be taught is important since research has shown that individuals with higher musical aptitude and/or ability appear to have an enhanced competence in language perception (Besson, Schön, Moreno, Santos, and Magne, 2007; Magne, Schön, and Besson, 2006). Adults and children who have musical skills are better at pitch processing—even in language—than those individuals who do not have musical skills (e.g., Lee and Lee, 2010; Lee and Hung, 2008; Besson, Schön, Moreno, Santos, and Magne, 2007; Magne, Schön, and Besson, 2006; Slevc and Miyake, 2006; Tanaka and Nakamura, 2004). Musicians and individuals with higher musical aptitude outperform non-musicians and individuals with lower musical aptitude when it comes to L2 acquisition overall (e.g., Milovanov, Tervaniemi, and Gustafsson, 2004; Burnicka, 1999), especially with the perception and production of lexical tone (e.g., Lee and Lee, 2010; Lee and Hung, 2008; Slevc and Miyake, 2006; Tanaka and Nakamura, 2004). Furthermore, musical training or proficiency may improve pitch perception in speech contexts (Delogue et al, 2010; Lee and Hung 2008; Wong et al 2007; Magne et al., 2006; Alexander et al., 2005; Schwanhäußer and Burnham, 2005; Schön, Magne, and Besson, 2004; Burnham and Brooker, 2002). Neuroscientific studies have shown that pitch sensitivity is more enhanced in musicians (instrumentalists and vocalists) than in non-musicians (Tervaniemi et al., 2005; Koelsch et al., 1999; Crummler et al., 1994). Indeed, musical skill, ability, and/or training may thus play a role in the non-tone language speaking adult’s L2 acquisition of particularly a tone language.
(e.g., Lee and Hung, 2008; Alexander, Wong, and Bradlow, 2005).

This present study thus endeavors to investigate the possible relationship between relative pitch proficiency and the perception of lexical tone by adult non-tone language speakers. This undertaking is based on the following conclusions: (1) the particular music feature pitch variation is an integral feature of language in general; (2) lexical tone appears to present a greater challenge to adults whose L1 is a non-tone language and who are learning a tone language; (3) musicians appear to outperform non-musicians in lexical tone perception and production tasks; (4) there appears to be a relationship between pitch proficiency and L2 perception and production ability; (5) non-tone language speakers process lexical tone based on pitch height—a feature that requires context; and (6) the specific music pitch competence that allows an individual to distinguish pitches in context (and the relationships thereof) is relative pitch perception—a competence that can be taught.
CHAPTER 2
LITERATURE REVIEW

In the spirit of providing a clearer foundation for the purpose of this study, Chapter 2 offers a review of current literature regarding four areas relevant to this research. This chapter opens with a general look at literature pertaining to language and music and the relationship between the two competencies. This chapter continues with defining pitch, fundamental frequency, and lexical tone and showing how these subjects relate. Further, Chapter 2 continues with a look at lexical tone and adult L2 acquisition as well as at a few studies that examine musicians’ versus non-musicians’ language acquisition skills. Chapter 2 concludes with the motivation for this paper’s current study.

2.1 Language and Music

As two of the multiple intelligences proposed by Developmental Psychologist Howard Gardner (1983), language and music (capabilities through which individuals may demonstrate intellectual abilities) are two competencies with a few general similarities. To begin, as briefly discussed earlier, pitch is an integral facet of language and music. There are two avenues that pitch is an important feature of both spoken language and music. The first avenue through which pitch is relevant to both language and music is this: In the most basic of terms, humans perceive both the phonemes that make up spoken language and the notes that make up music based on frequency or “a finite set of sounds (notes or phonemes) carved out of a larger possible set of
sounds [...that] are organized out of a larger possible set of sounds (McMullen and Saffran, 2004, p. 291). Thus, just as humans have the ability to utilize a finite number of phonemes to generate an infinite number of words and sentences (Chomsky, 1957), humans have the ability to utilize a finite number of notes to generate an infinite number of melodic sequences (Lerdahl and Jackendoff, 1996). Furthermore, there is a set of “rules” that govern language (e.g., word order necessary to comprise a sentence) just as there is a set of rules that govern music (e.g., the different rules that create a major third versus a minor third chord). The second avenue through which pitch is important in both language and music is the observation that pitch is utilized to communicate important information in both competencies. As Pfordresher and Brown (2009) point out, “pitch conveys information about tonality (e.g., Krumhansl, 1990), harmonic changes (e.g., Holleran, Jones, and Butler, 1995), phrase boundaries (e.g., Deliège, 1987), rhythm and meter (e.g., Jones, 1987), and the likelihood of future events (e.g., Narmour, 1990). For speech, it conveys information about word stress, utterance accent, pragmatic stance, and speaker emotion (Cruttenden, 1997). For the large class of languages known as tone languages, pitch conveys information about lexical meaning as well (Yip, 2002)” (p. 1385). These two general observations provide a foundation for examining language and music concurrently; however, there is further justification to consider these two competencies together: the parallel in the development of as well as the biological mechanisms responsible for these two competencies.

2.1.1 The Development of Language and Music

Initial acquisition as well as competent comprehension and production of both language and music are gained “implicitly” from environmental input. For children, language
development is the result of interaction between biology and environment (e.g., Kuhl, 2004); the same interaction appears to be true for the acquisition of music ability (Gordon, 1967). Just as children begin to develop an ear for the phonemes in their first language, they also begin to be able to detect changes in music based on the musical structure (e.g., Western music) to which they were exposed. Babies begin their language and music acquisition process with sound—crying. Before a baby can pronounce a single word, he or she quickly is able to imitate musical features such as rhythm and melodic contours which subsequently are replaced by the phonemes of the language (Mora, 2000). McMullen and Saffran (2004) point out that “infants are able to learn which sounds mark meaningful differences in their language before they have extensive access to word meanings” (p. 293). Interestingly, infants can distinguish lexical tone before they have learned speech (Ioup and Tansomboon, 1987). According to studies by Gao, Shi, and Aijun (2010), infants in an environment where Mandarin Chinese is spoken can distinguish two of the four Mandarin lexical tones before learning speech. This ability is because the child’s first exposure to language is the pitched vocal sounds of intonation and “musical vowel sounds” are what the fetus hears through the mother’s body, womb and amniotic fluid, and only after environmental linguistic exposure does the child pick up “not only the musicality of language, but also the necessary communication skills” (Mora, 2000, p. 149). As McMullen and Saffran (2004) write, “Moving beyond the segmental level (phonemes and tones), it is clear that the suprasegmental cues [e.g., patterns of rhythm, stress, intonation, phrasing and contour] in both language and music are highly salient to infants” (p. 294). Further, just as the phonemes in the language of the child’s environment become more salient and fixed as the child ages, the same is true for the music structure (e.g., Western) to which the child was exposed (cf. Kuhl, 2004 and McMullen and Saffran, 2004). Thus, for example, research has shown that “[i]n many instances,
young infants show little to no effect of Western tonal conventions on their perception of and memory for musical stimuli […]. By the time they start school, however, Western listeners’ responses to stimuli begin to show influences of typical Western musical structure” (McMullen and Saffran, 2004, p. 293). As Figure 1 illustrates, for both language and music, it seems that perception occurs prior to full production ability (Patel 2003).

There appears to be a parallel in the development of language and music. More specifically, there seems to be a parallel in the time in which language acquisition occurs most robustly (Kuhl et al., 1992) and absolute pitch (the ability to identify a pitch without a reference note) development (Lee, Lee, and Shi, 2011) since the acquisition of lexical tones and absolute pitch appear to occur in the same time frame (Lee, Lee, and Shi, 2011). There is also a possibility that pitch ability may play a role in the actual acquisition of lexical tone (Lee, Lee, and Shi, 2011; Tillmann et al., 2011; Deutsch et al., 2009 and 2006). Deutsch (2010) posits that “[t]he language to which we are exposed can greatly influence the chances of developing perfect pitch” (p. 41). Three particular studies that discuss the evident parallel in language acquisition and the development of absolute pitch development are by Deutsch et al. (2006), Lee and Lee (2010), and Lee, Lee, and Shi (2011). In the first study, Deutsch, Henthorn, Marvin, and Xu (2006) tested English-speaking and Mandarin Chinese-speaking first-year music conservatory students to explore the possibility of what the researched categorized as “a speech-related critical period for acquiring absolute pitch.” The two groups of participants (with an age range of 17 to 34) were asked to identify the notes of a set of 36 piano tones. This study found that the earlier a participant began musical training, the more likely he or she had evidence of absolute pitch. Specifically, participants who began musical training between the ages of 4 and 5 were more likely to have absolute pitch ability than participants who began
music training between the ages of 9 and 13. In the second study (Lee and Lee, 2010), 72 Taiwanese music students were asked to listen to pairs where one token was a specific syllable with lexical tone and the other token was a musical representation (played on piano, viola, or pure tone). Participants were required to determine if the two tokens in the pair were the same or different. These researchers also found that the earlier the onset of music training, the more likely the presence of absolute pitch. Lee, Lee, and Shi (2011) conducted two experiments to explore the relationship between absolute pitch and lexical tone perception. In the first experiment, a pool of Taiwanese-speaking musicians were asked to identify a musical note without the aid of a reference note. This experiment helped to determine absolute pitch ability. In the second experiment, which was used to determine lexical tone perception ability, the same participants were instructed to listen to four Taiwanese words spoken by 30 different speakers and to identify the corresponding logographic Chinese character as quickly as possible. Lee, Lee, and Shi (2011) concluded that there was a negative correlation “between occurrence of absolute pitch and age of onset of musical training, suggesting that the acquisition of absolute pitch resembles the acquisition of speech” (p. 526). This observation by Lee, Lee, and Shi (2011) may be because, as Deutsch (2002) and Deutsch et al. (2006) have suggested, absolute pitch (a skill which is present in only a small minority of individuals unlike the ability to speak a tone language) was originally packaged with other features of speech. In all three instances, researchers concluded that the time periods for the development of both speech and pitch processing must be related.

It has also been argued that there may be a link between genes and tone language speaking ability (Deidu and Ladd, 2007), leading to the conclusion that certain populations have a heritable advantage to acquiring a tone language. The results of a statistical study
conducted by Deidu and Ladd (2007) suggest that there may be a link between genes and tone language speaking ability since, according to Deidu and Ladd (2007), there is a relationship between the geographical distribution of two genes (ASPM and Microcephalin) and the geographical distribution of tone languages.
Music

Language

Figure 1: Music and Language Development Timeline

(Adapted from http://www.frontiersin.org/files/Articles/31563/fpsyg-03-00327-r2/image_m/fpsyg-03-00327-g002.jpg; Retrieved December 5, 2014)
2.1.2 The Biological Mechanisms Responsible for Language and Music

The given observations regarding parallels between language and music development indicate not only a possible relationship between language and music in general, but also a possible biological connection in the brain. To understand the possible connection, we begin with the theory of brain lateralization, developed by scientists including Robert Sperry (1974). This theory argues that the brain’s cortex is divided into two halves (left and right) and that each half (or cerebral hemisphere) is responsible for two separate general functions—the left hemisphere is for logical thinking while the right hemisphere is for intuitive thinking. The brain is further divided into lobes (frontal, parietal, temporal, occipital, and limbic) and specific areas (e.g., Broca’s area, Wernicke’s area). Each lobe and area is responsible for more particular functions. Thus, the subsequent postulation is that language is generally ruled by the left cerebral hemisphere while music is generally ruled by the right cerebral hemisphere since the cortical areas Broca’s area (in the left inferior frontal gyrus) and Wernicke’s area (in the left posterior superior temporal gyrus) are both located in the left cerebral hemisphere. Research has shown that these two areas (Broca’s area and Wernicke’s area) are responsible for language skills (language production and language perception, respectively) (e.g., Boeree, 2004). Much research has shown that various types of language impairment, or aphasia, are seen most often resulting damage to the left temporal lobe rather than to the right temporal lobe (e.g., McMullen and Saffran, 2004).

Many researchers maintain that music and language are cognitive functions of two separate temporal lobes in the brain and that musical intelligence is thus independent from the language faculty (e.g., Steinke, Cuddy, and Holden, 1997; Jourdain, 1997). Jourdain (1997) further crystalizes this argument for the lateralization of music and language in the brain’s two
temporal lobes, claiming that the left lobe is 90% better at recognizing words whereas the right lobe is about 20% better at recognizing melodic patterns. Research has documented various cases of musicians who suffered language or music impairment. For example, a composer who suffered a stroke rendering him with no speech perception or production abilities was still able to communicate via music by teaching music as well as composing an array of complex musical selections (Peretz and Coltheart, 2003). It has also been argued that imaging evidence supports separation of language and music—evidence which supports, as McMullen and Saffran (2004) observe, brain modularity as pursuant to Fodor (1983). Brain studies utilizing imaging techniques also show lateralization of the auditory discrimination of chords and syllables (Tervaniemi et al., 1997).

Although there have been arguments for a distinct dissociation of language and music in the brain’s cerebral hemispheres, mounting neuro-scientific evidence indicates that these two competencies are actually related cognitive abilities that are processed, not by two distinct neural domains, but rather by an overlapping of the two neural substrates (Sammler et al, 2009; Patel, 2003; Patel, Gibson, Ratner-Besson, and Holcomb, 1998; Wong, Skoe, Russo, Dees, and Kraus, 2007; Deutsch, Henthron, and Holson 2004). Patel (2003) maintains that neuroimaging evidence supports an overlap in linguistic and musical syntax processing, for example, since “musical syntax processing activates language areas of the brain” (p. 675). Additional neurophysiological evidence such as that gathered by Sammler et al. (2009) shows that music and language activate the same areas of the brain, leading to the observation that the brain may treat language and music in a song, for instance, as one signal.

Other researchers have argued that neuropsychological evidence reveals that some regions of the brain such as Broca’s area which are usually considered language-specific are also
linked to musical processing (Levitin and Menon, 2003; Maess, Koelsch, Gunter, and Friederici, 2001; Tillmann, Janata, and Bharucha, 2003; Patel, Gibson, Ratner, Besson, and Holcomb, 1998). Neuroscientist Stefan Koelsch (2002) found, through the use of Magnetic Resonance Imaging, that when subjects were presented with sequences of chords, activity was noted in both sides of the brain, most especially in the Broca and Wernicke areas—the two regions in the left hemisphere vital for language processing (cited in Deutsch, 2010, p. 36). Research shows that there is an overlap in the neural networks utilized for speaking and singing—a logical overlap since both activities are governed by a grammar, and both utilize smaller units to form longer elements (sentences or melodies) (Deutsch, 2010, p. 36). Magne, Schön, and Besson (2006) contend that their research has shown that there is behavioral evidence that supports the idea of a shared pitch processing mechanism in the perception of both music and language. As Jourdain (1997) states, “The two temporal lobes compete fiercely, and failure on one side can make the other stumble” (p. 291). Language and music are hence closely-related faculties that ultimately aid one another in human expression.

2.2 Pitch, Fundamental Frequency, and Lexical Tone

Both music and lexical tone have essentially similar properties—pitch and fundamental frequency which shall be defined and discussed in this section.

2.2.1 Pitch

As a property of sound overall, pitch is “a perceptual attribute which allows the ordering of sounds on a frequency-related scale extending from low to high. More exactly, pitch is
defined as the frequency of a sine wave that is matched to the target sound by human listeners” (Klapuri and Davy, 2006, p. 8). As mentioned earlier, all languages utilize pitch as a feature in general “but differ in the ways in which modifications of pitch are used and how many different types of functions are served by pitch variations” (Maddieson, 2011). In particular, pitch variation in language can be used contrastively to distinguish either utterance types (intonation), individual word meanings (lexical tone), or even grammatical function such as tense, mood, aspect, or case (grammatical tone). In music, pitch can be used not only to distinguish the key scale upon which the music is based, but also attributes such as affect (e.g., sadness induced by a composition performed in a minor key [Juslin and Västfjäll, 2008]).

2.2.2 *Fundamental Frequency*

In the most basic terms, musical pitch, a perceptual feature based on human perception and interpretation, is related to changes in features including fundamental frequency, intensity, and duration (Gerhard, 2003). Fundamental frequency is a sound wave that is introduced by a source or vibrating object. For human speech, the source is the set of vocal chords. Hertz (commonly abbreviated as Hz) is the usual unit utilized to measure and illustrate frequency and represents the number of vibrations per second. Therefore, for example, one Hz equates one vibration per second (Henderson, 2017). Humans perceive pitch based on perceptual interpretation of frequency. Haas (2003) states that the best range for human hearing is 20 to 20,000 Hz with the most optimal range between 200 and 2,000 Hz. As the sound waves travel through the air, there is a disturbance that is created which the human ear detects. As illustrated in Figure 2, the correlating frequency is then associated with a pitch (Henderson, 2017). However, human pitch perception is impacted not only by frequency, but also by duration and
intensity of a sound; thus, according to Haas (2003), a possible problem of intonation difficulties, for example, may be due to dynamics since “[s]ustained sounds above 2 kHz may be perceived as rising in pitch as their intensity increases, whereas sounds below 2 kHz may be perceived as dropping in pitch” (Haas, 2003). Musically trained individuals can detect as little as a 2 Hz difference in frequency between two separate sounds (Henderson, 2017).

![Musical notes and their general corresponding frequency (or Hz)](Based on Suits, 2013).

There is a growing body of literature that substantiates the position that fundamental frequency can be used differently by speakers of different dialects of the same language (Deutsch et al., 2009). In one study, Deutsch et al. (2009) compared the dialects of two groups of Mandarin Chinese speakers from two different villages and found that “overall pitch levels” differed by about 30 Hz. In addition, fundamental frequency can be used differently by speakers of various languages (e.g., Podesva, 2007; Xue et al., 2002; Ohara, 1992; Yamazawa and Hollien, 1992; Loveday, 1981; Crystal, 1969). Several studies found that the overall F0 range of
the speech of native Mandarin speakers appears to be greater than that of native English speakers (e.g., Keating and Kuo, 2012; Mang, 2001; Eady, 1982)1 According to Keating and Kuo (2012), there are several possible explanations for this difference:

“A possible source of language differences is phonological inventory differences. For example, if one language has more (or more frequent) voiceless obstruents or high vowels than another, and if voiceless obstruents and high vowels have a raising effect on F0 (Lehiste, 1970), then those two languages might well have characteristic F0 differences, though they would likely be small. Yet another proposal (e.g., Chen, 2005) is that tone languages will have larger F0 ranges and/or higher average F0s than non-tone languages, on the assumption that lexical tones require a greater extent of F0 than intonation alone does [though see Xu and Xu (2005) for a contrary assumption]. Chen’s discussion suggests that the locus of such a tone-language effect could lie specifically with the occurrence of a high level tone, as in Mandarin and Min. The typical F0 of a high level lexical tone might be systematically higher than that of high intonational tones in a language like English, or there might simply be more lexical high tones in running speech in a tone language than there are intonational high tones in running speech in a non-tone language. Liu (2002) looked at the F0 range in Mandarin by tone, and found that from Tone 1 through Tone 4, the tones have progressively larger F0 ranges. Thus an alternative scenario is that the more Tone 4s in Mandarin speech, the more likely a larger F0 range” (Keating and Kuo, 2012, p. 1051).

2.2.3  Lexical Tone

Of the more than 6,000 languages in the world today, it is estimated that approximately 70% of them have lexical tone as a prominent feature (Yip, 2002) and over 50% of the world’s populations speak a tonal language (Mattock and Burnham, 2006).2 Languages such as Mandarin Chinese, Vietnamese, and Thai are major tone languages with millions of speakers. Yip (2002) points out, a tone language is one that uses the suprasegmental feature, tone (or pitch variation), to provide information other than segmental features (consonants or vowels). For tone-language speakers, tone is used to change a

---

1It should be noted that Keating and Kuo (2012) point out that the given studies are somewhat limited.
2As of the time of this study, WALS and Ethnologue cite fewer languages as tonal than non-tonal.
word’s lexical meaning and/or grammatical form. Tone has two perceptual aspects—the physical properties (frequency, duration, and intensity) and the psychological response to the acoustic stimuli (e.g., Chan, Chuan, and Wang, 1975). Hence, the physical properties of frequency as earlier discussed as well as duration and intensity are important contributors to tone perception. Mandarin Chinese, for example (a widely-spoken tone language in the world today with over 900 million speakers around the world), has four lexical tones, as illustrated in Table 1 (Ladefoged and Maddieson, 1996). These tones are used to distinguish four different lexical meanings of what would appear to be the same word (the one-syllable word ma) if it were not for the suprasegmental distinction of lexical tone (see Table 1).

Lexical tones are not only assigned a Tone Description (e.g., High-Level) but also a Pitch Value (a number or series of numbers that indicates the rating of lexical tones, with 1 as the lowest pitch and 5 as the highest pitch). Thus, for example, the pitch value 55 would be 5 + 5 or a High-Level lexical tone. Pitch Value helps to clarify the distinct quality of lexical tone.

According to the *World Atlas of Language Structures*, tone languages can be divided into two basic categories: “those with a simple tone system—essentially those with only a two-way basic contrast, usually between high and low levels—and those with a more complex set of contrasts” (Maddieson, 2011). These categories can be further defined as level or contour patterns. The tonal language Yoruba (Niger-Congo), for example, has three level lexical tones: High, Mid, and Low (Orie, 2006). However, many of the East and Southeast Asian languages such as Thai, Vietnamese, and Chinese are contour tone languages (Maddieson, 2011) with more complex contrasts (e.g., Mandarin Chinese’s Tone 3 which is Falling-Rising). Generally, there is a significant difference from speaker to speaker in the lexical tone production pitch range since lexical tones are expressed physically by different F0 values with F0 height and F0 contour as the
primary acoustic parameters (Howie, 1976; Wu, 1986). Figure 3 illustrates the F0 contours of the four Mandarin Chinese lexical tones. Thus, for example, not all Mandarin speakers use exactly the same pitch to produce Tone 1. However, there is a consistency in the pitch range of the production of lexical tone, from speaker to speaker. Studies by Deutsch, Henthorn, and Dolson (2004) have concluded that although tone language speakers may not produce exactly the same pitch when producing a lexical tone, they have a consistent pitch template that appears to determine the pitches of the lexical tones and their relationships.

Researchers point out that two issues that may render discerning the four given Mandarin lexical tones somewhat challenging. First, in some contexts of continuous speech, the realization of Tone 2 and Tone 3 is often indistinguishable by some speakers. Tone 2 is High-Rising, whereas Tone 3 is Falling-Rising. Since both contours end with a rising element, the distinction may be somewhat hazy to some listeners—even native Mandarin speakers—who ultimately may depend on other elements of the discourse such as phonological environment and semantic context to determine the lexical entry. Second, as a natural characteristic of rapid speech, native speakers do not always produce exact and what could be defined as “typical” tokens of the various tones on a regular basis. The length of the segment, the actual F0 slope, and exact pitch may differ in various contexts. Hence, many speakers—even native Mandarin speakers—can misperceive Tone 2 and Tone 3 (Gao, Shi, and Aijun, 2010). With this difficulty present—even for native speakers of Mandarin—the challenge met by adult L2 learners of Mandarin can be even more difficult.
### TABLE 1: Tone to Indicate Lexical meaning in Mandarin Chinese

<table>
<thead>
<tr>
<th>Tone Number</th>
<th>Classical Tone Category Name</th>
<th>Pitch Value</th>
<th>Tone Description</th>
<th>Example in IPA</th>
<th>Lexical Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tone 1</td>
<td>Yinping</td>
<td>55</td>
<td>High-Level</td>
<td>[mā]</td>
<td>‘mother’</td>
</tr>
<tr>
<td>Tone 2</td>
<td>Yangping</td>
<td>35</td>
<td>High-Rising</td>
<td>[mā]</td>
<td>‘hemp’</td>
</tr>
<tr>
<td>Tone 3</td>
<td>Shangsheng</td>
<td>214</td>
<td>Falling-Rising</td>
<td>[mā]</td>
<td>‘horse’</td>
</tr>
<tr>
<td>Tone 4</td>
<td>Qusheng</td>
<td>51</td>
<td>Falling</td>
<td>[mā]</td>
<td>‘scold’</td>
</tr>
</tbody>
</table>

---

**Diacritics Key for Mandarin Examples**

- Macron [ā]: High Level Tone
- Acute Accent [á]: High-rising Tone
- Caron [ã]: Falling-rising Tone
- Grave [à]: Falling Tone
As mentioned in the introduction, tone can also distinguish grammatical features. As
illustrated in Table 2 and Table 3, Ngiti (Central Sudanic of Congo) utilizes tone to distinguish
verbs and to distinguish singular/plural nouns, respectively.

Figure 3: Mandarin Lexical Tones F0 contours
Mean F0 contours (averaged over speakers and tokens; n 5 and 48) of the
four Mandarin tones in the monosyllable /m a/ produced in isolation.
The time is normalized, with all tones plotted with their average duration
proportional to the average duration of Tone 3 (Adapted from Xu, 1997).
Table 2: Tone to Indicate Grammatical Function in Ngiti Verbs

Verbs in Ngiti, Central Sudanic, Zaire
(Taken from Summer Institute of Linguistics [SIL International], 2012)

These examples are of grammatical tone in the verb system

<table>
<thead>
<tr>
<th>Ngiti</th>
<th>Gloss</th>
<th>Verb Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>ma màkpěnà</td>
<td>‘I whistled’</td>
<td>recent past</td>
</tr>
<tr>
<td>ma màkpěnà</td>
<td>‘I whistled’</td>
<td>intermediate past</td>
</tr>
<tr>
<td>makpěnà</td>
<td>‘I will whistle’</td>
<td>near future</td>
</tr>
<tr>
<td>ma makpěnà</td>
<td>‘I used to whistle’</td>
<td>past habitual</td>
</tr>
<tr>
<td>ma múbhí</td>
<td>‘I walk’</td>
<td>present perfective</td>
</tr>
<tr>
<td>ma múbhi</td>
<td>‘I walked’</td>
<td>distant past</td>
</tr>
<tr>
<td>ma múbhi</td>
<td>‘I walked’</td>
<td>narrative past</td>
</tr>
</tbody>
</table>

Table 3: Tone to Indicate Grammatical Function in Ngiti Nouns

Nouns in
Ngiti
(Taken from The Grammar of Words: An Introduction to Linguistic Morphology page 12 By Geert Booij Oxford University Press 2012)

<table>
<thead>
<tr>
<th>Singular</th>
<th>Plural</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kamà</td>
<td>Kámá</td>
<td>‘chief(s)’</td>
</tr>
<tr>
<td>Målàyikà</td>
<td>Målàyiká</td>
<td>‘angel(s)’</td>
</tr>
</tbody>
</table>

Diacritics Key for Ngiti Examples

<table>
<thead>
<tr>
<th>Acute Accent</th>
<th>Grave Accent</th>
<th>Unmarked</th>
<th>Wedge</th>
</tr>
</thead>
<tbody>
<tr>
<td>[á]</td>
<td>[à]</td>
<td>[a]</td>
<td>[â]</td>
</tr>
<tr>
<td>High Tone</td>
<td>Low Tone</td>
<td>Mid Tone</td>
<td>Rising Tone</td>
</tr>
</tbody>
</table>
There are many languages that can utilize pitch to change lexical meaning but are still not tone languages. For example, languages such as Japanese, Korean, and certain Scandinavian and Slavic languages do utilize pitch variation, yet these languages are considered pitch accent languages and not tone languages. Although there is not clear general agreement on the actual delineations between some languages that utilize pitch variation to change lexical meaning, tone languages do differ somewhat from pitch accent languages overall. In pitch accent languages such as those aforementioned, the phonological environment in which the pitch variation occurs and the type of pitch variation are quite limited and sometimes play a very minor role, if any, in effective communication.

The fundamental characteristic that differentiates tonal and non-tonal languages is the fact that tonal languages use contrastive variations in pitch at the syllable or word level, whereas non-tonal languages may use pitch variations to signal intonation at the phrase or sentence level. The pitch content of lexical tones is perceived along at least three dimensions: height, or average F₀; the direction of the F₀ change, and the slope (pitch contour), or rate of F₀ change. As shall be discussed later, a tone language’s lexical tone inventory and the actual speakers’ sensitivity impact how much a speaker relies on the given dimensions. For example, non-tone language speakers tend to rely more on pitch height while tone language speakers tend to rely more on pitch movement and direction (Khouw and Ciocca, 2007; Gandour, 1983; Gandour and Harshman, 1978). Nonetheless, in the connected speech streams of both non-tone and tone language speakers, the given pitch variations can be somewhat reminiscent of melodies in music. This similarity is one reason to link music ability L2 proficiency.


2.3 *Lexical Tone and Adult L2 Acquisition*

Adults are frequently beleaguered with various difficulties when learning something new. It has been argued that these difficulties are due to deteriorating brain plasticity and a reduction in the number of synapses in the brain. However, although there is this widely-recognized generalization that adults do not have the capacity to learn as efficiently as children, research has revealed otherwise. Research has shown that adults actually can be more efficient at learning, especially when it comes to specific subject areas. As pointed out by the American Psychological Association, “[w]hile memorization skills and perceptual speed both start to decline in young adulthood, verbal abilities, spatial reasoning, simple math abilities and abstract reasoning skills all improve in middle age” (Phillips, 2011). One study in particular which supports this observation is the Seattle Longitudinal Study by Sherri Willis and colleagues. This particular study tracked the cognitive abilities of thousands of adults over a fifty-year span and found that older adults performed more efficiently than young adults on four out of six cognitive tests (Phillips, 2011). Neuroscience research has shown that adults’ highly-developed cognitive abilities often account for and support how adults outperform children when it comes to learning, in general. As stated by cognitive neuroscientist Patricia Reuter-Lorenz, “There is an enduring potential for plasticity, reorganization and preservation of capacities” (qtd. in Phillips, 2011, p. 38).

When it comes to language learning, generally the older an L2 learner, the more difficult it likely is for him or her to acquire native-like proficiency in an L2. There are two general postulations regarding maturational constraints or the observation that there is a negative correlation between age and L2 ultimate attainment (i.e., the older an L2 learner, the more difficult it likely is for him or her to acquire native-like proficiency). These two schools
of thought are founded upon the question of whether children have a particular advantage over adults at reaching ultimate attainment in a language before a certain age. In general, it is understood that adult L2 learners often do not attain the fluency of the monolingual native speaker. The first position is based on the hypothesis that language acquisition is innate and there is a specific period or age range (particularly prior to puberty) during which language ultimate attainment is possible (Lennenberg, 1967; Penfield and Roberts, 1959). According to this Critical Period Hypothesis, the neurocognitive changes due to maturation impact language development. Lennenberg (1967) posits that language learning is significantly reduced at puberty because there is cerebral lateralization and a reduction of brain plasticity. Supporters of the Critical Period Hypothesis maintain that it is impossible for adult learners to reach competency in an L2 since a mature human no longer has access to the innate language acquisition device which consists of the principles of universal grammar (UG) (e.g., Lennenberg, 1967). Although there are various adaptations of the Critical Period Hypothesis, one point upon which supporters of this position agree is that attempts at language development outside of that “critical period” would prove to be practically impossible, especially for L2 development (Johnson and Newport, 1989; Bley-Vroman, 1988).

There are many reasons that this conjecture regarding maturational constraints is somewhat questionable as there are several observations that strongly counter the theories of a critical period. First, there is no evidence to support such a strictly defined point where there is such an abrupt and extraordinary change in the brain that negatively impacts language learning. Instead, some studies have shown that there are age effects which equate a mere decline and not a specific point of discontinuity in reaching L2 ultimate attainment (e.g., Flege, Yeni-Komshian, and Liu, 1999). Second, evidence illustrates that individuals are able
to attain successfully an L2 outside of this “critical period” (Bongaerts, Planken, and Shils 1995; Birdsong, 1992). In fact, there is robust evidence that showed that adults are able to learn, for example, the grammar of a second language faster than children (Asher and Price, 1978, cited in McLauglin, 1992).

Furthermore, those researchers in opposition to the Critical Period Hypothesis argue that the problem of adult L2 acquisition is not the result of cerebral lateralization; instead, the problem is phonologically based since there is interference of the L1 (e.g., Best, 1995). Unlike the Critical Period Hypothesis position that there is a complete attenuation of speech learning ability at puberty, the L1 interference argument posits that there is a negative relationship between age and only some L2 sounds, likely because the perception and production of L2 sounds is contingent upon the L1 phonetic categories (Wang, Spence, Jongman, and Sereno (1999). In fact, some research shows that adults learning an L2 can perceive and/or produce L2 phonetic sounds outside of the L1 phonemic inventory if there is enough exposure and experience (e.g., Wang, Spence, Jongman, and Sereno, 1999; Flege, 1987).

There is indeed an added difficulty in reaching any type of fluency in an L2; however, the given obstacles are possibly amplified if the L1 is a non-tone language and the L2 is a tone language (e.g., Shen, 1989; Kiriloff, 1969). As discussed earlier, not only must the learner face general difficulties with acquiring segmental information, he or she must also learn lexical tone which is unfamiliar suprasegmental information. There are several reasons why the added task of recognizing lexical tone proves difficult for L2 learners. As mentioned in Section 2.1, native speakers of tone languages acquire lexical tone, for example, quite early in the language acquisition process (Iooup and Tansomboon, 1987) whereas the non-tone language speaker does
not. Hence, there is no level of representation for lexical tone in the native grammar to which the non-tone language speaker can relate the lexical tone (Miracle, 1989). Second, there are processing differences between tone language speakers and non-tone language speakers. Specifically, brain processing may be somewhat problematic for the learner since non-tone language speakers process the tones of the words in the right hemisphere of the brain shortly before processing the semantics in the left hemisphere of the brain, thus causing temporary confusion for the learner (University of California—Irvine, 2006). In addition, studies show that the non-tone language speaker may process tone in relation to his or her native intonational category; consequently, a mismatch ultimately arises due to L1 interference (Qin and Mok, 2011; Shen, 1989). For the non-tone language speaker learning a tone language, tone in the L1 is oftentimes a pragmatic cue, but in the L2, tone is a semantic cue; hence, (as mentioned earlier) the non-tone language speaker attempting to learn a tone language has to learn to use the tones for semantic or grammatical information and not just pragmatic information—training that takes additional time and effort that may need some focus on pitch (Wang et al, 2003).

Further, as discussed in Section 2.2, it appears non-tonal language speakers process F0 differently than tone-language speakers (Khouw and Ciocca, 2007; Gandour, 1983; Gandour and Harshman, 1978). Specifically, speakers of non-tone languages seem to be more sensitive to pitch height—or F0—and not direction or contour (Guion and Pederson, 2007), whereas tone language speakers place more emphasis on pitch direction and contour (Qin and Mok, 2011; Gandour and Harshman 1978).
2.4 Language Acquisition and Musicians/Non-Musicians

As discussed in Section 2.1, there appears to be a parallel in the general acquisition of language and music skills. As shall be discussed in this section, research has shown that there also seems to be a positive relationship between language acquisition and music training and/or proficiency.

2.4.1 L1 and Musicians/Non-Musicians

It has been argued that music skill, ability, and/or training may play an important role in human development in general (e.g., Granier-Deferre et al., 2010; Johnson-Green and Custodero, 2002). Studies show that music training may be significant in the development of overall intelligence (e.g., Rauscher, Shaw, and Ky, 1993). Sleve and Miyake (2006) point out that “musical ability can predict aspects of first-language (L1) verbal ability, such as reading ability in children (Anvari, Trainor, Woodside, and Levy, 2002; Atterbury, 1985)” (p. 675). Music proficiency may help in the processing of intonation contours and prosody (Deutsch, 2010), and small pitch changes in words and notes (e.g., Schön, Magne, and Besson, 2004). Research has shown that individuals with higher musical aptitude and/or ability appear to have an enhanced competence in language perception. In one study, Magne, Schön, and Besson (2006) investigated the impact of music training on other cognitive domains. In this study, 8 year old children were presented words or phrases that were categorized as congruous, weakly incongruous, or strongly incongruous in the final pitches. The final ERP data and behavior analysis revealed that that musician children outperformed non-musician children in the detection and discrimination of pitch in both music and language. In a similar study, Moreno et
al. (2009) divided a group of 8-year old participants into two groups; one group took music lessons for six months while the other group took painting lessons for six months. For pre-and post-training, participants listened to sentence pairs where one sentence’s final pitch was altered. Pre-testing showed that both groups performed about the same when it came to perceiving the differences; however, after the six-month instruction period, the group that received music lessons outperformed the group that received painting lessons. Furthermore, ERP studies show that musicians’ brainstems are more sensitive to pitch than the brain stems of non-musicians (e.g., Wong et al., 2007; Bidelman, Gandour, Krishnan, 2011).

2.4.2 Musicians/Non-Musicians and L2 Acquisition

Musical skill, ability, and/or training may also play a role in L2 learners’ acquisition process, no matter the age of the individual. One reason to expect a link between music proficiency and L2 development is the observation that both music and language are human universals that are comprised of perceptually discrete elements organized into hierarchically structured sequences (Patel, 2003; Sloboda, 1985). In a study of Polish elementary school children learning Russian, Burnicka (1999) determined that L2 learners with higher music proficiency are able to acquire an L2 more readily since they perceive and produce L2 segmental as well as suprasegmental more easily and have less L1 interference in L2 learning. Researchers Milovanov, Tervaniemi, and Gustafsson (2004) found that native Finnish-speaking high school students with higher a musical aptitude made fewer mistakes than the non-musical students when it came to the perception and production of English. In a review of various experiments, Besson, Schöen, Moreno, Santos, and Magne (2007) conclude that adults who have
musical skills are better at pitch processing—even in language—than those individuals who do not have musical skills.

Musical training or proficiency may improve specific aspects of L2 acquisition. For example, musical training or proficiency may improve pitch perception in speech contexts (Schön et al, 2004; Magne et al, 2006; Burnham and Brooker, 2002; Schwanhäußer and Burnham, 2005; Alexander et al, 2005; Wong et al, 2007). Music training may improve phonological production (e.g., Slevc and Miyake, 2006; Tanaka and Nakamura, 2004; Nakata, 2002). In one study, Slevc and Miyake (2006) examined four language domains (receptive phonology, productive phonology, syntax, and lexical knowledge) in adult L2 learners. In the receptive phonology portion of this study, 50 native Japanese speakers were given stimuli on the word, sentence, and passage level. The word task was comprised of minimal pairs that differed by one phoneme known to be difficult for the Japanese speaker (e.g., [l] and [r]) to discriminate; the sentence task required participants to listen to sentence pairs and identify which sentence (out of two that differed by one word) was presented in written form. For the passage task, the participants had to identify any mispronounced words in the passage. The productive phonology portion was similar to the receptive phonology portion, with there being a word, sentence, and passage level. In this portion, however, participants were recorded as they each read aloud the word minimal pairs, sentence minimal pairs, and a passage; these recordings were later judged by native English speakers. To test musical ability, these researchers had participants complete a tonal memory test and a chord analysis test (both subtests of the Wing’s Measure of Musical Intelligence). Slevc and Miyake (2006) ultimately conclude that the “popular conjecture that musical ability is associated with L2 proficiency is not a myth. Although the link may be restricted to L2 phonology, individuals who are good at analyzing, discriminating, and
remembering musical stimuli are better than other people at accurately perceiving and producing L2 sounds” (p. 679). In other research, Tanaka and Nakamura (2004) compared the auditory working memory and language proficiency of Japanese ESL speakers. In this study, participants were tested on their verbal memory, musical memory, and English pronunciation skills. Tanaka and Nakamura (2004) concluded that verbal memory and musical memory are related and these related skills impact second language pronunciation ability.

2.4.3 Musicians/Non-Musicians and Tone Language L2 Acquisition

The observation that tone language speakers are more musically proficient when it comes to musical pitch discrimination helps to explain why music proficiency may be useful in tone language L2 acquisition. As mentioned earlier, tone language speakers are more accurate at imitating musical pitch and discriminating intervals than speakers of non-tone languages. Research that illustrates this position was conducted by Pfordresher and Brown (2009) who conducted two studies to investigate whether pitch processing in music is impacted by tone language acquisition. The first study was a two-part task focusing on music perception and production. In the first task, tone language speakers and non-tone language speakers were required to imitate four-note sequences; in the second task, the same participants were asked to discriminate pairs of single pitches and then to discriminate intervals between two pitches. The second study was a replica of the first study; the only difference was the narrowing of criteria for the participant pool (fluent Mandarin speakers who “had lived in China for at least 10 years before coming to the United States, and had not learned English during their first 6 years” (p. 1391). Pfordresher and Brown (2009) concluded that tone-language speakers are more skilled at
imitating and discriminating musical pitch, indicating a possible symbiotic relationship or even a
reinforcement between tone-language and pitch discrimination. Some researchers argue that this
positive correlation between music skill and tone language may be due to the presence of
absolute pitch. As discussed earlier, some researchers have discovered that the earlier the onset
of music training, the more likely the presence of absolute pitch (e.g., Lee, Lee, and Shi, 2011;
Lee and Lee, 2010).

Music skill, proficiency and/or training may play a role in lexical tone development. Several studies support this observation. Many researchers argue that individuals with music
training can perceive and produce lexical tones more readily than individuals without music
training (Deutsch, 2010; Bidelman, Gandour, and Krishnan, 2010; Krishnan and Gandour, 2009;
Patel, 2008; Zatorre and Gandour, 2008; Musacchia, Sams, Skoe, and Kraus, 2007; Wong, Skoe,
Russo, Dees, and Kraus, 2007; Kraus and Banai, 2007; Magne, Schön, and Besson, 2006; Schön,
Magne, and Besson, 2004; Zatorre, Belin, and Penhune, 2002). Nguyen, Tillmann, Gosselin, and
Peretz (2009) concluded that musical pitch perception and lexical tone discrimination are
mediated by the same perceptual system. Deutsch (2010) observes that learning to sing or to
play an instrument may help a language learner be more attuned to the melody of speech. In
studies that measured the electrophysiological responses of the brain to various linguistic stimuli
(i.e., Event-Related Potentials or ERP studies), Deutsch (2010) found stronger responses in the
auditory brain stem of non-tone language speaking study participants who were exposed to a
tone language and had music backgrounds than in that same region of participants who had no
music background. Other neurophysiologic indicators support the idea that music training may
help pitch processing in language (Musacchia, Sams, Skoe, and Kraus, 2007; Wong, Skoe,
Russo, Dees, and Kraus, 2007; Magne, Schön, and Besson, 2006; Schön, Magne, and Besson,
Bidelman, Gandour, and Krishnan (2010) point out that the “neural representation of pitch may be influenced by one’s experience with music or language at subcortical as well as cortical levels of processing (Krishnan and Gandour, 2009; Patel, 2008; Zatorre and Gandour, 2008; Kraus and Banai, 2007; Zatorre, Belin, and Penhune, 2002). [...] Musical training sharpens subcortical encoding of linguistic pitch patterns” (p. 426). Lee and Hung (2008) contend that English-speaking musicians show better performance in the identification of lexical tones than non-musicians.

It has been furthermore hypothesized that certain musical abilities may impact L2 language perception and production abilities of particularly a tone language. Studies have shown that music ability helps in the perception of isolated lexical tones (Alexander, Wong, and Bradlow, 2005). The results of identification and discrimination tasks presented to 18 adult native English speakers (9 musicians and 9 non-musicians) by Alexander, Wong, and Bradlow (2005) reveal that musicians significantly outperformed non-musicians in such tasks. The musicians identified tones correctly 89% of the time and discriminated the difference between tones 87% of the time whereas the non-musicians identified tones correctly 69% of the time and discriminated the difference between tones 71% of the time, leading Alexander, Wong, and Bradlow (2005) to conclude that there is a possible link between music and language processing. As Alexander, Wong, and Bradlow (2005) aptly point out, the musicians are likely “able to transfer their music pitch-processing processing abilities to a speech-processing task.” Delogu, Lampis, and Belardinelli (2010) studied the musical ability and Mandarin tone discrimination competence of 48 Italian speakers with no tone-language experience. To allow the researchers to determine musical skill level, participants were administered a Melodic Test (a subtest of the Wing’s Standardized Tests of Musical Intelligence). In the linguistic portion, participants were
given a same/different task in which they had to listen to two lists (extracted from a bank of 504 monosyllabic Mandarin words) and determine if the list was the same or different (if the lists were different, they differed by only one item). When there was a difference, the participant was asked to identify which lexical item on the list was different. This study ultimately concluded that the speakers with greater music melodic skills were stronger at tonal discrimination. Furthermore, the combination of speech and music in songs seems to aid in the performance of tone (Schön, Boyer, Moreno, Besson, Peretz, and Kolinsky, 2008). Qin, Fukayama, Nishimoto, and Sagayama (2011) found that songs helped Chinese speakers from Japan learn the most difficult Mandarin tone for many native and non-native speakers, Tone 3 (falling-rising tone). Music only or speech only is not more effective when it comes to some tones (Qin, Fukayama, Nishimoto, and Sagayama, 2011; Schön, Boyer, Moreno, Besson, Peretz, and Kolinsky, 2008).

Although music ability seems to have a positive correlation with language acquisition in general and lexical tone perception and production in particular, the specific music ability appears to be relevant. On one hand, for example, Kirkham, Lu, Wayland, and Kaan (2011) found that the ability to play an instrument as opposed to the ability to sing does not appear to have any particular effect on language acquisition. On the other hand, it seems that pitch ability (having absolute pitch, for example) may possibly impact language acquisition in general (Burnham et al, 1996; Burnham and Brooker, 2002). Pitch ability can also impact tone language acquisition in particular (Lee, Lee and Shi, 2011) especially since the three basic dimensions (pitch height, direction, and contour) upon which a speaker/learner perceives lexical tone actually can be viewed as related to music.

Tillman et al. (2011) investigated the impact of an individual having a deficit in processing pitch in a musical context on lexical tone development. Since lexical tone is
contingent upon pitch variation to indicate meaning, a fundamental conjecture would be that an individual would need to be able to discriminate pitch to recognize lexical tone. Hence, it could be hypothesized that an individual who is not proficient at pitch perception would find lexical tone discrimination challenging. Tillmann et al. (2011) conducted two experiments based on a participant pool of native French-speaking individuals who had a deficit in pitch processing versus a control group. The first experiment was a same/different task in which participants listened to pairs of monosyllabic Mandarin words. The second experiment was a same/different task in which each pair was comprised of the syllable /ba/ produced with one of the five Thai lexical tones and a musical note played on the violin. The results of this study support the observation that the lexical tone perception ability of individuals who have a deficit in pitch processing is impaired, leading these researchers to posit that pitch ability may play a role in lexical tone development. Similar studies not only confirm this observation (Hyde and Peretz 2004; Peretz et al., 2002) but also support the point that (as earlier asserted) there may be an overlap in language and music processing, and thus a general domain for musical pitch and lexical tone discrimination.

2.5 In Summary

Contrary to absolute pitch ability, relative pitch proficiency can be taught to individuals, no matter their age. Evidence illustrating that L2 learners can improve pronunciation for some L2 target sounds suggests that there is brain plasticity for the adult learner. “According to prevailing views, brain organization is modulated by practice, e.g., during musical or linguistic training. […] [R]egular music practice may also have a modulatory effect on the brain’s
linguistic organization and alter hemispheric functioning in those who have regularly practiced
music for years” (Milovanov and Tervaniemi, 2011). In the practice of learning relative pitch, an
adult L2 learner may impact processing organization and yet also learn to attune to auditory cues
once not important in the linguistic discourse.

Considering an L2 learner’s pitch ability may be an important step in understanding L2
acquisition—especially of a tone language—since tone language and non-tone language speakers
appear to attach importance to different perceptual cues/auditory dimensions. Specifically,
speakers of non-tone languages seem to be more sensitive to pitch height—or F0—and not
direction or contour (Guion and Pederson, 2007), whereas tone language speakers place more
emphasis on pitch direction and contour (Qin and Mok, 2011; Gandour and Harshman 1978).
Guion and Pederson (2007) also maintain that many adult learners of a tone language may also
focus on both F0 and F0 slope of the tones, suggesting that L2 learners can learn to attend to the
cues that L1 speakers use for the tone distinction.

The discussion in this chapter has provided various reasons why considering whether
relative pitch ability may ultimately aid in the adult L2 acquisition of a tone language. Namely,
the following points have been argued. First, there is a relationship between language and music
in general as we see an overlap in the development of these two processes; furthermore, there is
apparently an overlap in the cognitive processing areas of music and language. Second, some
adult language learners are more successful at learning a new language than other adult learners.
Third, there appears to be a relationship between language acquisition and music ability. Fourth,
musicians outperform non-musicians when it comes to learning a second language, especially a
tone language. Fifth, there seems to be a correlation between tone language ability and pitch
ability, specifically absolute pitch; however, the depth of the association between tone language
ability and absolute pitch is still somewhat hazy. Sixth, unlike absolute pitch, relative pitch ability may contribute more specifically to tone language acquisition for the second language learner since the important processing cues for tone language ability are F₀ and F₀ slope—two cues that may be activated with the knowledge of relative pitch. Thus, once an adult learner is conceptually attuned to the relationship between notes in general, he or she may transfer this ability to tone language acquisition. Seventh, unlike absolute pitch ability, relative pitch ability can be acquired in adulthood.
CHAPTER 3

METHODS AND MATERIALS

This study endeavors to investigate the relationship between the specific music skill relative pitch (i.e., the ability to recognize the relationships and/or intervals between pitches) and the perception of Mandarin Chinese lexical tone (i.e., pitches used to differentiate lexical meaning in the language) in adult L2 tone and non-tone language speakers. This chapter presents the hypothesis and research questions and gives details about the participant pool and methodology. Chapter 3 concludes with the hypothesized results.

3.1 Hypothesis and Research Questions

The general two-part hypothesis for the present study is this:

a. The better a participant is at relative pitch perception, the better that participant will be at lexical tone perception, no matter the length of the token, the phonological environment in which the token is placed, or the voice quality, frequency range, and/or gender of speaker who produces the token.

b. Speakers whose L1 is a tone language will be more adept at relative pitch perception and at lexical tone perception in an L2 tone language than non-tone language speakers.
The research questions for this study are as follows:

1. Do tone-language speakers possess relative pitch ability more robustly than non-tone language speakers?

2. Are tone-language speakers stronger at identifying lexical tone in an L2 than non-tone language speakers?

3. Do non-tone language speakers who exhibit a higher level of relative pitch proficiency also exhibit greater skill in lexical tone perception, no matter the length of the token, the phonological environment in which the token is placed, or the voice quality, frequency range, and/or gender of speaker who produces the token?

4. Do non-tone language speakers who exhibit a higher level of relative pitch proficiency also exhibit greater skill in the perception of actual Mandarin Chinese words?

To determine the validity of the given hypothesis and correlated research questions, four short experiments were conducted with a pool of 60 participants (30 native English-only speakers with no tone-language experience and a group of 30 tone language speakers). The experiments were as follows:

1. The first experiment established a baseline by determining each participant’s general lexical tone perception proficiency;

2. The second experiment tested relative pitch proficiency by utilizing three tests (same/different, direction, and high/low).

3. The third experiment assessed each participant’s ability to perceive lexical tone with monosyllabic nonsense words; and
4. The fourth experiment assessed each participant’s ability to perceive lexical tone bound to actual Mandarin Chinese adjectives.

3.2 Participants

The participant pool was comprised of 30 native tone-language speakers and 30 English-only speakers. Participants were recruited randomly from the College of Staten Island, City University of New York. Of the 30 native tone-language speakers, there was a control group of 20 adult native Mandarin Chinese speakers and a test group of 10 native Thai speakers who had no experience with Mandarin Chinese. None of the 30 adult native English speakers had any training in or experience with a tone language. The inclusion criteria for all participants were self-identification as an adult (ages 18-40) and self-identification as either a native speaker of Mandarin Chinese, Thai, or English. The age limit for this study was set based on National Institute of Deafness and Other Communication Disorders (NIDCD) reports (2007) which argue that there is a significant increase in the numbers of individuals who report hearing loss by the age of 40. Exclusion criteria were based on the general parameters of this study. All participants were to have no self-reported hearing impairment, no formal music training, and no color blindness. Table 4 gives an overview of general demographics for the participants in this study (See Appendix 1 for the Participant In-take Questionnaire). Each participant was paid $40 for participating in this study.
Table 4: The Participant Pool

<table>
<thead>
<tr>
<th>Tone Language Speakers (n=30)</th>
<th>English Speakers (n=30)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>Age</td>
</tr>
<tr>
<td>18 - 25</td>
<td>18 - 25</td>
</tr>
<tr>
<td>26 - 33</td>
<td>26 - 33</td>
</tr>
<tr>
<td>34 - 40</td>
<td>34 - 40</td>
</tr>
<tr>
<td>Gender</td>
<td>Gender</td>
</tr>
<tr>
<td>Female</td>
<td>Female</td>
</tr>
<tr>
<td>Male</td>
<td>Male</td>
</tr>
</tbody>
</table>

Tone Language Speakers by Language

<table>
<thead>
<tr>
<th>Mandarin Chinese Speakers (n=20)</th>
<th>Thai Speakers (n=10)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>Age</td>
</tr>
<tr>
<td>18 - 25</td>
<td>18 - 25</td>
</tr>
<tr>
<td>26 - 33</td>
<td>26 - 33</td>
</tr>
<tr>
<td>34 - 40</td>
<td>34 - 40</td>
</tr>
<tr>
<td>Gender</td>
<td>Gender</td>
</tr>
<tr>
<td>Female</td>
<td>Female</td>
</tr>
<tr>
<td>Male</td>
<td>Male</td>
</tr>
</tbody>
</table>

3.3 Methodology

3.3.1 Pre-Experiment Training

To aid participants’ ability to identify the difference amongst the four Mandarin lexical tones, participants in the test groups (English-only speakers and Thai speakers) were given instruction on the four lexical tones used in Mandarin Chinese (see Appendix 2 for the Pre-Experiment Training Script). This instruction included defining the four basic tones while allowing the participants to observe orthographic representations and to hear recorded male and female vocalizations of each tone. To focus participants on lexical tone, tokens of each of the
four lexical tones were realized on the syllables la, li, lu, and lo (spoken by both a male and female speaker) to total thirty-two tokens (see Appendix 3 for the stimuli blocks). Participants were introduced to both phonological and orthographic realization of these tokens (See Appendix 4 for the Pre-Training Handout with orthographic representations). A short period of questions and answers was honored at the end of this fifteen- to twenty-minute instruction phase.

3.3.2 Experiment 1: Lexical Tone Identification Baseline

Experiment 1 established a general baseline by testing participants’ ability to recognize the four lexical tones in Mandarin Chinese. Based on Lee and Lee (2010) and Lee and Hung (2008), this experiment tested a control group of native Mandarin Chinese speakers as well as a test group of native Thai speakers with no experience with or training in Mandarin Chinese and a test group of English-only adult speakers, on their perception of lexical tone. The syllable sa, which is phonotactically acceptable in Mandarin, Thai, and English, was produced with all 4 lexical tones by 8 native Mandarin Chinese speakers (4 female and 4 male) resulting in a total of 32 stimuli of various frequencies due to the speaker and gender variation. Figure 4 presents Praat-generated waveform and spectrogram information that loosely illustrates frequency variation for the 16 stimuli of Part A. Figure 5 presents Praat-generated waveform and spectrogram information that loosely illustrates frequency variation for the 16 stimuli of Part B. The 32 stimuli were divided into two blocks of 16 stimuli, one token for each realization produced by each of the 8 speakers (See Appendix 5 for the stimuli). These two blocks contained a balance of male and female speakers and were randomized to prevent speaker normalization (Table 5 presents the stimuli blocks with the lexical tone and speaker gender).
There was a 3-second interval between each token and a 10-second interval between the two blocks. Participants were instructed to identify the lexical tone (1, 2, 3, or 4) as quickly as possible (See Appendix 6 for Experiment 1 Script). Participants were able to refer to their pre-training handout which contained orthographic representations of the Mandarin Chinese lexical tones. To ensure that participants understood the instructions, there was a Practice Test (See Appendix 6 for the Experiment 1 Script).
Figure 4: Waveform and Spectrogram representations of stimuli for Experiment 1 Part A

Praat-generated spectrogram evidence generally illustrates frequency variance. These variances are not only due to pitch contour of the lexical tone but also due to speaker variation. A more specific account of the token type and speaker gender is included in Table 5.
Praat-generated spectrogram evidence generally illustrates frequency variance. These variances are not only due to pitch contour of the lexical tone but also due to speaker variation. A more specific account of the token type and speaker gender is included in Table 5.
Table 5: Experiment One stimuli by lexical tone number and speaker gender.

<table>
<thead>
<tr>
<th>PART A</th>
<th>Num</th>
<th>Tone</th>
<th>Gen</th>
<th>Num</th>
<th>Tone</th>
<th>Gen</th>
<th>Num</th>
<th>Tone</th>
<th>Gen</th>
<th>Num</th>
<th>Tone</th>
<th>Gen</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>1</td>
<td>M</td>
<td></td>
<td>5.</td>
<td>4</td>
<td>F</td>
<td>9.</td>
<td>1</td>
<td>M</td>
<td>13.</td>
<td>4</td>
<td>F</td>
</tr>
<tr>
<td>2.</td>
<td>3</td>
<td>M</td>
<td></td>
<td>6.</td>
<td>4</td>
<td>F</td>
<td>10.</td>
<td>3</td>
<td>M</td>
<td>14.</td>
<td>3</td>
<td>F</td>
</tr>
<tr>
<td>3.</td>
<td>2</td>
<td>F</td>
<td></td>
<td>7.</td>
<td>1</td>
<td>M</td>
<td>11.</td>
<td>1</td>
<td>F</td>
<td>15.</td>
<td>4</td>
<td>M</td>
</tr>
<tr>
<td>4.</td>
<td>3</td>
<td>M</td>
<td></td>
<td>8.</td>
<td>4</td>
<td>F</td>
<td>12.</td>
<td>4</td>
<td>M</td>
<td>16.</td>
<td>1</td>
<td>F</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>PART B</th>
<th>Num</th>
<th>Tone</th>
<th>Gen</th>
<th>Num</th>
<th>Tone</th>
<th>Gen</th>
<th>Num</th>
<th>Tone</th>
<th>Gen</th>
<th>Num</th>
<th>Tone</th>
<th>Gen</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>3</td>
<td>F</td>
<td></td>
<td>5.</td>
<td>1</td>
<td>M</td>
<td>9.</td>
<td>4</td>
<td>M</td>
<td>13.</td>
<td>2</td>
<td>M</td>
</tr>
<tr>
<td>2.</td>
<td>4</td>
<td>F</td>
<td></td>
<td>6.</td>
<td>2</td>
<td>F</td>
<td>10.</td>
<td>2</td>
<td>F</td>
<td>14.</td>
<td>2</td>
<td>F</td>
</tr>
<tr>
<td>3.</td>
<td>2</td>
<td>M</td>
<td></td>
<td>7.</td>
<td>4</td>
<td>M</td>
<td>11.</td>
<td>3</td>
<td>M</td>
<td>15.</td>
<td>3</td>
<td>M</td>
</tr>
<tr>
<td>4.</td>
<td>1</td>
<td>F</td>
<td></td>
<td>8.</td>
<td>1</td>
<td>F</td>
<td>12.</td>
<td>2</td>
<td>M</td>
<td>16.</td>
<td>3</td>
<td>F</td>
</tr>
</tbody>
</table>

This table lists the specific stimuli blocks by Mandarin lexical tone and the gender of the speaker who recorded the lexical tone.
3.3.3 Experiment 2: Relative Pitch Perception Proficiency Test

To determine the level of relative pitch ability, three music tests were constructed based loosely on the MBEA (the Montreal Battery of Evaluation of Amusia) which was a series of six tests that could be used to assess various music abilities based on six components of music processing: Scale, Contour, Interval, Rhythm, Meter, and Musical Memory. Considering one of the main objectives of this study (i.e., measuring relative pitch proficiency), we determined to structure this portion of the experiment somewhat similar to the MBEA melodic tests—the first three tests which considered Contour, Scale, and Interval.

For this study, the three tasks were Same/Different, Direction (Up/Down), and Higher/Lower. For the Same/Different task, 12 pairs of notes were recorded using the Kurzweil K2500. In this A/X task, participants listened to two notes that were produced by the same instrument voicing (piano, bass, flute). Of the 12 stimuli, there was a randomized mixture of piano, bass, flute voicings. For each pair of notes, participants had to determine if the two notes were the same or different. For the direction task, each token consisted of a short series of five piano notes that either ascended or descended the heptatonic scale. For this task, participants had to determine if the notes were ascending or descending the scale. For the Higher/Lower task, each token consisted of a pair of notes produced by one of the following voicings: piano, bass, flute. Participants were instructed to determine if the second note was higher or lower than the first note. To ensure that participants understood the directions, there was a short practice test prior to each of the three tasks (See Appendix 7 for Experiment 2 Script with Practice Tests).

---

3This researcher concedes that this test may be considered profoundly culturally biased since “[m]ajor/minor scales and the keys they may be taken to represent are [...] far from cross-cultural phenomena. Every musical culture that uses pitch in some way has a concept of contour [...] and] the familiar metaphor of higher vs. lower is less widespread than [one] might think” (J. Straus, personal communication, November 13, 2014). However, since the test group of speakers in this experiment is comprised of native English speakers likely grounded in the Western music tradition, this researcher has elected to pattern these tests after the MBEA.
Table 6: Stimuli Blocks for Experiment Two

<table>
<thead>
<tr>
<th>Num</th>
<th>Voice</th>
<th>Ans</th>
<th>Num</th>
<th>Voice</th>
<th>Ans</th>
<th>Num</th>
<th>Voice</th>
<th>Ans</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.</td>
<td>Piano</td>
<td>D</td>
<td>2.</td>
<td>Piano</td>
<td>D</td>
<td>2.</td>
<td>Piano</td>
<td>L</td>
</tr>
<tr>
<td>5.</td>
<td>Piano</td>
<td>F</td>
<td>5.</td>
<td>Piano</td>
<td>U</td>
<td>5.</td>
<td>Piano</td>
<td>H</td>
</tr>
</tbody>
</table>

3.3.4 Experiment 3: Monosyllabic Nonsense Words and Lexical Tone

Experiment 3 was comprised of two tests—Test A and Test B. In Experiment 3, participants were taught a mini-lexicon of four nonsense words that do not violate the phonotactics of English but do carry the four Mandarin Chinese lexical tones (rā, rá, vǔ, and vù; see Table 7 for the representations and glosses and Figure 6 for sample stimuli). Participants were introduced to these nonsense (or nonce) words and then subsequently tested on these words.
presented in pairs. Prior to the administration of Test A and Test B, there was a short training session designed to ease the memory load requirements of the experiment by allowing participants the opportunity to learn the meanings and make the connections between the words and their picture representations (see Appendix 8 for Experiment 3 Pre-Experiment Training Script and materials).

For Test A, participants were given a reference card with the orthographic representations of \( r\ddot{a} \) (scarf) and \( r\acute{a} \) (sunglasses). Participants were then be asked to view a series of 16 pictures (in random order) based on these two nonce words. Each picture was accompanied by a set of two tokens (tokens which differ only in lexical tone) realized in the English carrier declarative sentence, “The Pooh Bear is wearing a _____” (e.g., A Pooh Bear wearing sunglasses is shown. To illustrate, two sentences that the participant heard were: AThe Pooh Bear is wearing \( r\ddot{a} \). BThe Pooh Bear is wearing \( r\acute{a} \). (These two sentences present Tone 1 \([r\ddot{a}]\) and Tone 3 \([r\acute{a}]\), respectively.) Each participant was asked to determine which sentence out of each set correctly identified the picture. Hence, in reference to the previous example, if the picture showed a Pooh Bear waring a pair of sunglasses, then the participant had to recognize that “BThe Pooh Bear is wearing \( r\acute{a} \)” (Tone 3) is the correct answer. The pictures were presented in full color via overhead projector (see Appendix 9 for Experiment 3 Script and Appendix 10 for Experiment 3 stimuli).

<table>
<thead>
<tr>
<th>Nonsense Word</th>
<th>Tone</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>( r\ddot{a} )</td>
<td>1</td>
<td>scarf</td>
</tr>
<tr>
<td>( r\acute{a} )</td>
<td>3</td>
<td>sunglasses</td>
</tr>
<tr>
<td>( v\ddot{u} )</td>
<td>2</td>
<td>bracelet</td>
</tr>
<tr>
<td>( v\acute{u} )</td>
<td>4</td>
<td>hat</td>
</tr>
</tbody>
</table>

Table 7: Nonsense Words Lexicon with Mandarin Chinese Lexical Tone
Figure 6: Example Tokens for Experiment 3 Part A

rā Scarf  rā Sunglasses  vù Bracelet  vù Hat
Test B was structured exactly as Test A, differing only in tokens; participants were instructed regarding Tone 2 with *vú* (nonsense word meaning *bracelet*) and Tone 4 *vù* (nonsense word meaning *hat*) and given a card with representations. For Part A and Part B, a Practice Test was given prior to each test in preparation for the experiment (See Appendices 9 and 10).

3.3.5 Experiment 4: Mandarin Chinese Adjectives

Experiment 4 was quite similar to Experiment 3 and was structured fundamentally in the same manner. One difference was the stimuli, which were tokens of actual Mandarin Chinese adjectives that do not violate the phonotactics of English. However, another difference was that the incorrect stimulus could be realized with either of the other three tones. In this experiment, participants were taught a mini-lexicon of four Mandarin Chinese adjectives (see Table 8). The Experiment 4 Script (along with the training materials) is in Appendix 11.

For Test A, Tone 1 and Tone 2 were tested with the words *hēi sè* (black) and *hóng sè* (red) which were taught. For Test B, Tone 3 and Tone 4 were tested with the words *zǐ sè* (purple) and *lǜ sè* (green) which were taught. Participants were given a card with orthographic representations of these two words (see Appendix 11). The stimuli for Experiment 4 may be found in Appendix 12.

<table>
<thead>
<tr>
<th>Pinyin</th>
<th>Tone</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>hēi sè</td>
<td>1</td>
<td>black</td>
</tr>
<tr>
<td>hóng sè</td>
<td>2</td>
<td>red</td>
</tr>
<tr>
<td>zǐ sè</td>
<td>3</td>
<td>purple</td>
</tr>
<tr>
<td>lǜ sè</td>
<td>4</td>
<td>green</td>
</tr>
</tbody>
</table>
For each test, participants were instructed to view a series of 16 pictures (in random order) that differed based on the Mandarin color adjectives that were learned. Each picture was accompanied by a set of two tokens (tokens which differed only in lexical tone) realized in the English carrier declarative sentence, “The ____ is ______.” Thus, for example, a picture of a red circle was shown and the two sentences that the participant heard were: AThe circle is hŏng sè. BThe circle is hóng sè.) with Tone 3 and Tone 2, respectively. Each participant was asked to determine which sentence out of each set correctly identified the color of the picture. Hence, in reference to the previous example, the participant had to recognize that “BThe circle is a hóng sè” with Tone 2, is the correct answer. As in Experiment 3, there was a Practice Test administered prior to Part A and Part B for this experiment (See Appendix 11).

3.4 Hypothesized Results

This researcher speculated that this study would support the following general observations:

1. Speakers whose L1 is a tone language are more likely to have stronger relative pitch skills than English-only speakers.

2. Speakers whose L1 is a tone language (regardless the actual tone language) will be stronger at identifying Mandarin lexical tone than English-only speakers; however, Mandarin speakers will be stronger at identifying Mandarin lexical tone than other speakers whose L1 is a different tone language.

3. Participants who have stronger relative pitch perception skills will be more successful at lexical tone perception, no matter the length of the token, the
phonological environment in which the token is placed, or the voice quality,
frequency range, and/or gender of speaker who produces the token.

4. Non-tone language speakers who exhibit a higher level of relative pitch proficiency
will exhibit greater skill in the perception of actual Mandarin Chinese words.
CHAPTER 4
RESULTS AND DISCUSSION

Chapter 4 presents the detailed objective findings of this study and a discussion regarding these findings. This chapter concludes with limitations and suggested further research.

4.1 Results

The Results section begins with a report of descriptive statistics as well as additional preliminary univariate analyses of data and then continues with the results of inferential statistics based on the findings of various bivariate analyses that were utilized to answer the research questions employed to guide this study. In addition, reliability estimates (Kuder and Richardson Formula 20) are provided for the dependent measures.

4.1.1 Descriptive Statistics

To begin, Table 9 provides descriptive performance statistics, expressed in percentage scores, for all four experiments. As Table 9 illustrates, the mean performance appears to differ significantly between tone language speakers and English speakers, with the same difference noted in the standard deviation from the mean.

Table 10 provides descriptive performance statistics, expressed in percentage scores, for tone identification tasks. In Table 10, it is clear that, based on the three tone identification tasks,
the Mandarin Chinese speakers performed the best, whereas Thai speakers outperformed the English speakers when it came to Mandarin Chinese lexical tone identification as well as identification of Mandarin Chinese adjectives. There is a significant difference between the Thai speakers and the English speakers, with the same noted difference reflected in the standard deviation from the mean.

Table 11 provides descriptive performance statistics, expressed in percentage scores, for the specific music proficiency tasks that comprised Experiment Two. In Table 11, we see that, based the three sub-tests and the overall totals for the music proficiency experiment, once again the Mandarin Chinese speakers performed the best, whereas Thai speakers outperformed the English speakers. There is a significant difference between the Thai speakers and the English speakers, with the same noted difference reflected in the standard deviation from the mean.
Table 9: Descriptive Statistics (%) on all four Experiments

<table>
<thead>
<tr>
<th>Experiment</th>
<th>Group (n = 60)</th>
<th>M (SD)</th>
<th>Var</th>
<th>Min</th>
<th>Max</th>
<th>Conf.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Experiment One:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Tone Identification</em></td>
<td>Tone Language (n = 30)</td>
<td>0.74 (0.23)</td>
<td>0.05</td>
<td>0.41</td>
<td>1.00</td>
<td>0.09</td>
</tr>
<tr>
<td></td>
<td>English (n = 30)</td>
<td>0.30 (0.11)</td>
<td>0.01</td>
<td>0.13</td>
<td>0.56</td>
<td>0.04</td>
</tr>
<tr>
<td><strong>Experiment Two:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Tone Language (n = 30)</td>
<td>0.86 (0.14)</td>
<td>0.02</td>
<td>0.64</td>
<td>1.00</td>
<td>0.05</td>
</tr>
<tr>
<td></td>
<td>Music Proficiency English (n = 30)</td>
<td>0.49 (0.20)</td>
<td>0.04</td>
<td>0.33</td>
<td>1.00</td>
<td>0.07</td>
</tr>
<tr>
<td><strong>Experiment Three:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Nonsense Words</em></td>
<td>Tone Language (n = 30)</td>
<td>0.89 (0.09)</td>
<td>0.01</td>
<td>0.75</td>
<td>1.00</td>
<td>0.03</td>
</tr>
<tr>
<td></td>
<td>English (n = 30)</td>
<td>0.63 (0.20)</td>
<td>0.04</td>
<td>0.38</td>
<td>0.97</td>
<td>0.08</td>
</tr>
<tr>
<td><strong>Experiment Four:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Mandarin Color Adj.</em></td>
<td>Tone Language (n = 30)</td>
<td>0.93 (0.09)</td>
<td>0.01</td>
<td>0.78</td>
<td>1.00</td>
<td>0.03</td>
</tr>
<tr>
<td></td>
<td>English (n = 30)</td>
<td>0.48 (0.10)</td>
<td>0.01</td>
<td>0.31</td>
<td>0.66</td>
<td>0.04</td>
</tr>
</tbody>
</table>
Table 10: Descriptive Statistics (%) on Tone Identification Tasks (*Exps One, Three, and Four*)

<table>
<thead>
<tr>
<th>Experiment</th>
<th>Group (n = 60)</th>
<th>M (SD)</th>
<th>Var</th>
<th>Min</th>
<th>Max</th>
<th>Conf.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Experiment One:</strong></td>
<td>Mandarin (n = 20)</td>
<td>0.87 (0.16)</td>
<td>0.03</td>
<td>0.56</td>
<td>1.00</td>
<td>0.07</td>
</tr>
<tr>
<td></td>
<td>Thai (n = 10)</td>
<td>0.48 (0.07)</td>
<td>0.01</td>
<td>0.41</td>
<td>0.63</td>
<td>0.05</td>
</tr>
<tr>
<td></td>
<td>English (n = 30)</td>
<td>0.30 (0.11)</td>
<td>0.01</td>
<td>0.12</td>
<td>0.56</td>
<td>0.04</td>
</tr>
<tr>
<td><strong>Experiment Three:</strong></td>
<td>Mandarin (n = 20)</td>
<td>0.94 (0.07)</td>
<td>0.01</td>
<td>0.75</td>
<td>1.00</td>
<td>0.03</td>
</tr>
<tr>
<td></td>
<td>Thai (n = 10)</td>
<td>0.80 (0.04)</td>
<td>0.00</td>
<td>0.75</td>
<td>0.84</td>
<td>0.03</td>
</tr>
<tr>
<td></td>
<td>English (n = 30)</td>
<td>0.63 (0.20)</td>
<td>0.04</td>
<td>0.38</td>
<td>0.97</td>
<td>0.08</td>
</tr>
<tr>
<td><strong>Experiment Four:</strong></td>
<td>Mandarin (n = 20)</td>
<td>0.99 (0.02)</td>
<td>0.00</td>
<td>0.94</td>
<td>1.00</td>
<td>0.01</td>
</tr>
<tr>
<td></td>
<td>Thai (n = 10)</td>
<td>0.81 (0.04)</td>
<td>0.00</td>
<td>0.78</td>
<td>0.91</td>
<td>0.03</td>
</tr>
<tr>
<td></td>
<td>English (n = 30)</td>
<td>0.48 (0.09)</td>
<td>0.01</td>
<td>0.31</td>
<td>0.66</td>
<td>0.04</td>
</tr>
</tbody>
</table>
Table 11: Descriptive Statistics (%) on Music Proficiency

<table>
<thead>
<tr>
<th>Experiment</th>
<th>Group</th>
<th>M (SD)</th>
<th>Var</th>
<th>Min</th>
<th>Max</th>
<th>Conf.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Experiment Two</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Same/Different</td>
<td>Mandarin (n = 20)</td>
<td>0.87 (0.23)</td>
<td>0.05</td>
<td>0.33</td>
<td>1.00</td>
<td>0.11</td>
</tr>
<tr>
<td></td>
<td>Thai (n = 10)</td>
<td>0.57 (0.14)</td>
<td>0.02</td>
<td>0.50</td>
<td>0.83</td>
<td>0.10</td>
</tr>
<tr>
<td></td>
<td>English (n = 30)</td>
<td>0.45 (0.22)</td>
<td>0.05</td>
<td>0.25</td>
<td>1.00</td>
<td>0.08</td>
</tr>
<tr>
<td>Up/Down</td>
<td>Mandarin (n = 20)</td>
<td>1.00 (0.00)</td>
<td>0.00</td>
<td>1.00</td>
<td>1.00</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>Thai (n = 10)</td>
<td>0.77 (0.16)</td>
<td>0.03</td>
<td>0.67</td>
<td>1.00</td>
<td>0.12</td>
</tr>
<tr>
<td></td>
<td>English (n = 30)</td>
<td>0.47 (0.19)</td>
<td>0.04</td>
<td>0.33</td>
<td>1.00</td>
<td>0.07</td>
</tr>
<tr>
<td>High/Low</td>
<td>Mandarin (n = 20)</td>
<td>0.91 (0.11)</td>
<td>0.01</td>
<td>0.75</td>
<td>1.00</td>
<td>0.05</td>
</tr>
<tr>
<td></td>
<td>Thai (n = 10)</td>
<td>0.82 (0.08)</td>
<td>0.01</td>
<td>0.75</td>
<td>0.92</td>
<td>0.05</td>
</tr>
<tr>
<td></td>
<td>English (n = 30)</td>
<td>0.55 (0.23)</td>
<td>0.05</td>
<td>0.25</td>
<td>1.00</td>
<td>0.09</td>
</tr>
<tr>
<td>All Three</td>
<td>Mandarin (n = 20)</td>
<td>0.92 (0.10)</td>
<td>0.01</td>
<td>0.69</td>
<td>1.00</td>
<td>0.05</td>
</tr>
<tr>
<td></td>
<td>Thai (n = 10)</td>
<td>0.73 (0.10)</td>
<td>0.01</td>
<td>0.64</td>
<td>0.91</td>
<td>0.07</td>
</tr>
<tr>
<td></td>
<td>English (n = 30)</td>
<td>0.49 (0.20)</td>
<td>0.04</td>
<td>0.33</td>
<td>1.00</td>
<td>0.07</td>
</tr>
</tbody>
</table>
Levene’s Test was run to assess the homogeneity of variances. As Table 12 illustrates, the Levene’s Test was clearly violated with respect to Experiments Two, Three, and Four (the crux of this study).

Table 12: Levene’s Test for Experiments One through Four

<table>
<thead>
<tr>
<th>Measure</th>
<th>Total %</th>
<th>F</th>
<th>P-Value</th>
<th>F-crit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experiment One:</td>
<td></td>
<td>1.102371</td>
<td>0.298256</td>
<td>4.012973</td>
</tr>
<tr>
<td>Tone Identification</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Experiment Two:</td>
<td>16.05306</td>
<td>0.000184</td>
<td>4.012973</td>
<td></td>
</tr>
<tr>
<td>Music Proficiency</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Experiment Three:</td>
<td>26.35231</td>
<td>3.71E-06</td>
<td>4.012973</td>
<td></td>
</tr>
<tr>
<td>Nonsense Words</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Experiment Four:</td>
<td>44.83373</td>
<td>1.1E-08</td>
<td>4.012973</td>
<td></td>
</tr>
<tr>
<td>Mandarin Color Adj.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*p value is significant at 0.05

Next, a series of t-tests was run to evaluate the significance of the average difference between tone language speakers and English speakers. The noted difference between the sample means (Exp 1: 0.7375 - 0.304167; Exp 2: 0.856481 - 0.488889; Exp 3: 0.891667 - 0.629167; Exp 4: 0.929167 - 0.484375) is substantial enough to say that the scores differ significantly between tone language and English speakers. Furthermore, there is rejection of the null (H0: μ1 - μ2 = 0) based on t stat > t crit 2 tail (Experiment One, 9.30 > 2.02; Experiment
Two, 8.44 > 2.01; Experiment Three 6.43 > 2.02; Experiment Four, 18.18 > 2.01), leading thus to the alternative hypothesis (H1: μ1 - μ2 ≠ 0).

A series of one-way ANOVAs was run with Group (Mandarin, Thai, English) as the Independent variable (see Table 13). The null (H0: μ1 - μ2 = 0) is rejected since f > f crit (3.16) in each instance (Experiment One, 120.01; Experiment Two, 46.88; Experiment Three, 24.96; Experiment Four, 309.39).

For the two test groups (Thai speakers and English speakers), a series of t-tests were run to determine the significance where, for each test, t stat > t crit two tail, leading to rejection of the null (Experiment 1, 5.74 > 2.06; Experiment 2, 14.42 > 2.03; Experiment 3, 4.30 > 2.02; Experiment 4, 14.42 > 2.03). The given results up to this point could be interpreted as showing that not only were the English speakers less efficient in the lexical tone identification tasks, English speakers were also less proficient in the music tasks than the tone language speakers.
Table 13: One-Way ANOVA Results with Group as Independent Variable

<table>
<thead>
<tr>
<th>Experiment</th>
<th>F</th>
<th>P-Value</th>
<th>F-crit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experiment One</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Parts A and B)</td>
<td>120.0116</td>
<td>3.7E-21</td>
<td>3.158843</td>
</tr>
<tr>
<td>Experiment Two</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Same/Different)</td>
<td>23.57231</td>
<td>3.47E-08</td>
<td>3.158843</td>
</tr>
<tr>
<td>(Up/Down)</td>
<td>75.5572</td>
<td>9.35E-17</td>
<td>3.158843</td>
</tr>
<tr>
<td>(High/Low)</td>
<td>25.65941</td>
<td>1.13E-08</td>
<td>3.158843</td>
</tr>
<tr>
<td>(All Three)</td>
<td>46.87846</td>
<td>9.15E-13</td>
<td>3.158843</td>
</tr>
<tr>
<td>Experiment Three</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Parts A and B)</td>
<td>24.96027</td>
<td>1.64E-08</td>
<td>3.158843</td>
</tr>
<tr>
<td>Experiment Four</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Parts A and B)</td>
<td>309.3932</td>
<td>2.47E-31</td>
<td>3.158843</td>
</tr>
</tbody>
</table>

p value is significant at 0.05
4.1.2 Test Reliability: Assessment Tasks

Estimates of test reliability for the dependent measures were calculated using KuderRichardson Formula 20 (ρKR20). For the Nonsense Words (ρKR20 = 0.90856), whereas for the Mandarin Color Adjectives (ρKR20 = 0.925318); thus, it can be determined that these measures have high internal consistency and hence a high reliability.

4.1.3 Inferential Statistics

To begin the assessment of possible relationships amongst variables, this researcher ran a series of Pearson’s Correlations (see Table 14) to determine the strength of association between performance on the various assessments for English speakers. Table 14 reveals that for English speakers, there appears to be a few moderate relationships (baseline and music proficiency, .51; baseline and tone identification in nonsense words, .61; music and tone identification in nonsense words, .55).

<table>
<thead>
<tr>
<th>ENGLISH SPEAKERS</th>
<th>MUSIC</th>
<th>BASELINE</th>
<th>NONSENSE</th>
<th>COLORS</th>
</tr>
</thead>
<tbody>
<tr>
<td>MUSIC</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BASELINE</td>
<td><strong>0.509896915</strong></td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NONSENSE</td>
<td><strong>0.551097662</strong></td>
<td><strong>0.614673158</strong></td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>COLORS</td>
<td>0.324479695</td>
<td>0.294733395</td>
<td>0.110367938</td>
<td>1</td>
</tr>
</tbody>
</table>
4.1.4 Results: Research Question One

The first research question is as follows:

Do tone-language speakers possess relative pitch proficiency more robustly than non-tone language speakers?

As we saw in Table 11, the tone language speakers performed more proficiently in the music proficiency tasks than English speakers did (see Table 15 for a recap of the means based on score percentages); specifically, Mandarin speakers performed better than the Thai speakers who performed better than the English speakers (see Figure 7). To begin to assess this particular research question more closely, this researcher ran a series of t-tests for unequal variances regarding the three specific relative pitch tasks in Experiment Two as well as the totals to evaluate the statistical difference between the means of the tone language speakers versus the English speakers. We find that the null is rejected as, in each instance, t stat > t crit two tail (same/different, 5.248039 > 2.002465; up/down, 10.38517 > 2.002465; high/low, 6.995103 > 2.018082; all three tasks, 8.442856 > 2.006647).

<table>
<thead>
<tr>
<th></th>
<th>Same/different</th>
<th>up/down</th>
<th>high/low</th>
<th>All three Tasks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mandarin</td>
<td>0.87</td>
<td>1</td>
<td>0.91</td>
<td>0.92</td>
</tr>
<tr>
<td>Thai</td>
<td>0.57</td>
<td>0.77</td>
<td>0.82</td>
<td>0.73</td>
</tr>
<tr>
<td>English</td>
<td>0.45</td>
<td>0.47</td>
<td>0.55</td>
<td>0.49</td>
</tr>
</tbody>
</table>
### Table 16: Test Groups’ Means (%) in Tone Identification and Mandarin Adjectives Tasks

<table>
<thead>
<tr>
<th></th>
<th>Baseline</th>
<th>Nonsense Words</th>
<th>Mandarin Adjectives</th>
</tr>
</thead>
<tbody>
<tr>
<td>Thai</td>
<td>0.48</td>
<td>0.8</td>
<td>0.81</td>
</tr>
<tr>
<td>English</td>
<td>0.3</td>
<td>0.63</td>
<td>0.48</td>
</tr>
</tbody>
</table>

### Figure 7: Groups’ Music Proficiency Means

![Figure 7: Groups’ Music Proficiency Means](image-url)
4.1.5 Results: Research Question Two

The second research question is as follows:

*Are tone-language speakers stronger at identifying the lexical tone in an L2 than non-tone language speakers?*

As we saw in Table 10, Thai speakers’ means were higher than English speakers when it came to the three tone identification tasks (see Table 16 and Figure 8 for a recap).

**Figure 8: Test Groups’ Means (%) in Tone Identification Tasks**

![Bar chart showing test groups' means in tone identification tasks.](image)

To further address this question, a series of t-tests was run regarding the three tone identification tasks and comparing the means between Thai versus English speakers. We find that the null is rejected as, in each instance, t stat > t crit two tail (Baseline, 5.740605 > 2.059539; Nonsense words, 4.298171 > 2.028094; Mandarin Adjectives, 14.42397 > 2.032245).
4.1.6 Results: Research Question Three

The third research question is as follows:

*Do non-tone language speakers who exhibit a higher level of relative pitch proficiency also exhibit greater skill in lexical tone perception, no matter the length of the token, the phonological environment in which the token is placed, or the voice quality, frequency range, and/or gender of speaker who produces the token?*

To address this question, a set of Pearson’s Correlations was run for the strength of association between specific aspects of English speakers’ music proficiency. As Table 17 illustrates, there is a moderate positive relationship between relative pitch proficiency overall and Nonsense Words (0.55) but a weak negative relationship between music proficiency and Mandarin Color Adjectives (-0.32). Regarding the three particular relative pitch assessment tasks that comprised the music proficiency experiment (ability to note same/different, up/down, and high/low with pairs of musical notes) and lexical tone identification proficiency as assessed in Experiments Three and Four, there are various moderate positive associations amongst assessments. What should be noted, however, is that the Pearson’s coefficients illustrate a stronger relationship between the three specific tests for music proficiency (same/different, up/down, and high/low) and nonsense words, with there being a moderate positive relationship (.56, .49, and .45, respectively) than between music proficiency and Mandarin color adjectives. Specifically, there is a weak negative relationship between the same/different and high/low tasks and Mandarin Color Adjectives (-0.29, -0.21), but there is a moderate negative relationship between up/down and Mandarin color adjectives (-0.43).
Table 17: Correlation Matrix: English Speakers’ Music Proficiency and Lexical Tone Identification

<table>
<thead>
<tr>
<th></th>
<th>All Three Music Tests</th>
<th>Same/Different</th>
<th>Up/Down</th>
<th>High/Low</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nonsense Words</td>
<td>0.55</td>
<td>0.56</td>
<td>0.49</td>
<td>0.45</td>
</tr>
<tr>
<td>Mandarin Color Adjectives</td>
<td>-0.32</td>
<td>-0.29</td>
<td>-0.43</td>
<td>-0.21</td>
</tr>
</tbody>
</table>
To further explore whether music proficiency can predict tone identification ability, a series of Regression Analyses was run to evaluate each relative pitch assessment task in the music proficiency test (Experiment Two) and its possible predictability. For English speakers, it appears that there is statistical significance (see Table 18) in the ability to ascertain lexical tone in nonsense words and the ability to distinguish between musical pitches (same/different task), the direction of pitches (up/down task), and the height of pitches (high/low task); specifically, low $p$ values ($p$ is significant at 0.05) indicate significance (same/different, $p = 0.001395689$; up/down, $p = 0.005835513$; high/low, $p = 0.013475356$) which allows for rejection of the null hypothesis. These $p$ values indicate that changes in the various predictors’ values (same/different, up/down, and high/low) appear to be meaningfully related to English speakers’ ability to recognize lexical tone.

4.1.7 Results: Research Question Four

The final research question is as follows:

*Do non-tone language speakers who exhibit a higher level of relative pitch proficiency also exhibit greater skill in the perception of actual Mandarin Chinese words?*

To address this question, a set of Pearson’s Correlations was run for the strength of association between specific aspects of English speakers’ music proficiency and the perception of actual Mandarin Chinese words (as in Experiment Four, Mandarin Chinese Adjectives). As Table 17: Correlation Matrix illustrates, there is a weak negative relationship between music proficiency and Mandarin Color Adjectives (-0.32). Regarding the three particular relative pitch
assessment tasks that comprised the music proficiency experiment, (ability to note same/different, up/down, and high/low with pairs of musical notes) and Mandarin adjectives, there are various moderate positive associations amongst assessments. Pearson’s coefficients illustrate a moderate positive relationship (.56, .49, and .45, respectively) between music proficiency and Mandarin color adjectives. Specifically, there is a weak negative relationship between the same/different and high/low tasks and Mandarin Color Adjectives (-0.29, -0.21), but there is a moderate negative relationship between up/down and Mandarin color adjectives (-0.43).

Another look at the Regression Analyses (Table 18) indicates that the relationship (if any) between the various music proficiency assessment tasks and English speakers’ ability to recognize Mandarin Chinese adjectives is weak to insignificant as the large p values indicate. English speakers’ ability to ascertain the difference between two notes and whether one note is higher or lower than the other one does not seem to impact English speakers’ ability to recognize Mandarin Chinese adjectives (same/different, 0.125181515; high/low, 0.254942571). However, English speakers’ ability to determine direction that a note is going appears to be significantly related to their ability to recognize Mandarin Chinese adjectives (up/down, 0.018566393).
Table 18: Regression Analysis: Music Proficiency and Tone Identification

<table>
<thead>
<tr>
<th></th>
<th>Adj. R Square</th>
<th>Significance F</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Same/Different and Nonsense Words</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thai Speakers (N=10)</td>
<td>0.266304348</td>
<td>0.072725577</td>
<td>0.072725577</td>
</tr>
<tr>
<td>English Speakers (N=30)</td>
<td>0.285352558</td>
<td><strong>0.001395689</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Same/Different and Mandarin Adj.</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thai Speakers (N=10)</td>
<td>-0.022727273</td>
<td>0.397203841</td>
<td>0.397203841</td>
</tr>
<tr>
<td>English Speakers (N=30)</td>
<td>0.049135768</td>
<td>0.125181515</td>
<td>0.125181515</td>
</tr>
<tr>
<td><strong>Up/Down and Nonsense Words</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thai Speakers (N=10)</td>
<td>-0.030667702</td>
<td>0.417060328</td>
<td>0.417060328</td>
</tr>
<tr>
<td>English Speakers (N=30)</td>
<td>0.214257462</td>
<td><strong>0.005835513</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Up/Down and Mandarin Adj.</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thai Speakers (N=10)</td>
<td>0.050324675</td>
<td>0.258902175</td>
<td>0.258902175</td>
</tr>
<tr>
<td>English Speakers (N=30)</td>
<td>0.15323979</td>
<td><strong>0.018566393</strong></td>
<td></td>
</tr>
<tr>
<td><strong>High/Low and Nonsense Words</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thai Speakers (N=10)</td>
<td>-0.098934497</td>
<td>0.674638224</td>
<td>0.674638224</td>
</tr>
<tr>
<td>English Speakers (N=30)</td>
<td>0.170406655</td>
<td><strong>0.013475356</strong></td>
<td></td>
</tr>
<tr>
<td><strong>High/Low and Mandarin Adj.</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thai Speakers (N=10)</td>
<td>-0.076555024</td>
<td>0.565110058</td>
<td>0.565110058</td>
</tr>
<tr>
<td>English Speakers (N=30)</td>
<td>0.011953309</td>
<td>0.254942571</td>
<td>0.254942571</td>
</tr>
</tbody>
</table>

*p value is significant at 0.05
4.2 Discussion

The aim of this current study has been to assess whether there is a relationship between music proficiency, particularly relative pitch proficiency, and lexical tone perception. In more specific terms, this study has endeavored to investigate the relationship between relative pitch proficiency and the perception of Mandarin Chinese lexical tones by adult second language learners endeavoring to learn a tone language. The rationale for such an endeavor is postulated based on the following points: First, there is a relationship between language and music in general as we see an overlap in the development of these two processes; furthermore, there is apparently an overlap in the cognitive processing areas of music and language. Second, some adult language learners are more successful at learning a new language than other adult learners. Third, there appears to be a relationship between language acquisition and music ability. Fourth, musicians outperform non-musicians when it comes to learning a second language, especially a tone language. Fifth, there seems to be a correlation between tone language ability and pitch ability, specifically absolute pitch; however, the depth of the association between tone language ability and absolute pitch is still somewhat hazy. Sixth, unlike absolute pitch, relative pitch ability may contribute more specifically to tone language acquisition for the second language learner since the important processing cues for tone language ability are $F_0$ and $F_0$ slope—two cues that may be activated with the knowledge of relative pitch. Thus, once an adult learner is conceptually attuned to the relationship between notes in general, he or she may transfer this proficiency to tone language acquisition. Seventh, unlike absolute pitch ability, relative pitch proficiency can be acquired in adulthood.

To evaluate the possibility of a relationship between relative pitch ability and lexical tone perception, this research project assessed the relative pitch proficiency and lexical tone
perception ability/tone language comprehension of one control group (Mandarin Chinese speakers) and two test groups (Thai speakers and English speakers). Currently, there are many studies that evaluate the relationship between absolute pitch and lexical tone; however, this study presents a newly-developing approach since it seeks to consider relative pitch instead of absolute pitch.

The general two-part hypothesis for this present study is this:

1. The better a participant is at relative pitch perception, the better that participant will be at lexical tone perception, no matter the length of the token, the phonological environment in which the token is placed, or the voice quality, frequency range, and/or gender of speaker who produces the token.

2. Speakers whose L1 is a tone language will be more adept at lexical tone perception than non-tone language speakers.

Correlational and regression analyses revealed the validity of this hypothesis as several expected outcomes were observed. However, as shall be discussed, there were also several unexpected outcomes.

4.2.1 Relative Pitch Proficiency in Tone Language and Non-Tone Language Speakers

Results from this study show that tone language speakers are more proficient at the relative pitch tasks than non-tone language speakers. This observation is supported by the significant difference in performance by Mandarin Chinese and Thai speakers versus the English speakers’ performance on the three relative pitch tasks. This finding is as expected. As has been discussed in Chapter 1, previous research has suggested that tone language speakers have a
propensity for music proficiency. One explanation is that, for example, there may be a possible relationship or reinforcement between speaking a tone language and pitch discrimination (Lee, Lee, and Shi, 2011; Tillmann, Burnham, Nguyen, Grimault, Gosselin, and Peretz, 2011; Deutsch et al, 2006 and 2009). Furthermore, evidence from prior studies has shown that tone language speakers are more accurate at imitating musical pitch and discriminating intervals than non-tone language speakers (Pfordresher and Brown, 2009).

When comparing Eastern music to Western music, there are fewer melodic major thirds in the music of tone language speakers than non-tone language speakers (Han et al., 2011, p. e20160). The music of tone language speakers most often utilizes the pentatonic (five-note) scale (Ho and Han, 1982) whereas the music of non-tone language speakers relies on the heptatonic (seven-note) scale (Barry, 1909). There are thus larger intervals in the music of, for example, Eastern music than in Western music. As Han et al. (2011) point out, “Thus in comparison with the major heptatonic scale, the major pentatonic scale offers approximately 6% fewer opportunities for major thirds, despite an 11% increase in opportunities for larger melodic intervals” (p. e20160). In one study, Han et al. (2011) compared the pitch intervals in the tradition music of Chinese, Thai, and Vietnamese (three tone language speaking cultures) to American, French, and German (three non-tone language speaking cultures) traditional music and found that “[c]hanges in pitch direction occur more frequently and pitch intervals are larger in the music of tone compared to non-tone language cultures. More frequent changes in pitch direction and larger pitch intervals are also apparent in the speech of tone compared to non-tone language cultures” (p. e20160). For the tone language speaker, the frequency of pitch changes in music may also transfer to the weight of importance in language. As Han et al. (2011) point out, “[A]djacent syllables often have different pitch contours and levels, resulting in more
frequent changes in pitch direction and larger pitch change between syllables. In contrast, given that very few syllables in non-tone languages are distinguished in this way, changes in pitch direction should be less frequent and pitch changes between syllables should be smaller” (e20160). As aforementioned, tone language speakers are more sensitive to pitch direction in lexical tone perception (Khouw and Ciocca, 2007; Gandour, 1983; Gandour and Harshman, 1978). In the prosody of tone language speakers, there are not only larger pitch intervals, there are also more frequent changes in pitch direction than in the prosody of non-tone language speakers (Han et al., 2011). Thus, for the tone language speaker, it is important for there to be a level of pitch discrimination as well as production competence for successful language perception and production to occur. The Mandarin Chinese or Thai speaker, for example, must able to discern the pitch variations of the lexical tones since the crucial and only distinction for the same syllable in a tone language is the pitch variation. If a tone language speaker is not able to discriminate the pitch variations of the linguistic tone, then serious utterance confusion is the likely outcome (e.g., the same two phonemes in the syllable /ma/ could mean horse or mother and are distinguished only by lexical tone4). For the English speaker, however, pitch discrimination is not necessarily as important. Although there may be some confusability regarding the pragmatic meaning, an English speaker can understand the general meaning of an utterance. There is no reason for an English speaker to confuse the words for horse and mother since they are made up of a different set of phonemes.

A closer look at the relative pitch proficiency of the two test groups in particular (native Thai speakers and native English speakers) reveals that for the English speakers, although they did not perform as well as the Thai language speakers on the relative pitch tasks overall, more

---

4This researcher acknowledges that contextual clues, especially in connected speech play an important role; however, this paper will not venture into this discussion.
English speakers performed their best specifically on the High/Low task. Approximately half of the English speakers did well on the High/Low task. The High/Low task required participants to listen to a pair of notes and determine if the second note was higher or lower than the first note. To explain why English speakers did their best on the High/Low task, it could be argued that, as research has shown regarding lexical tone perception, these speakers are attuned to F₀ onset and offset values or pitch height (Khouw and Ciocca, 2007; Gandour, 1983; Gandour and Harshman, 1978). With the High/Low task, there were two distinct notes which facilitated in the perception of onset and offset of the notes, which may have facilitated perception and identification.

Substantiated by the previous research aforementioned, the observation that the two groups of tone language speakers outperformed non-tone language speakers on the relative pitch tasks is indeed expected. However, a relevant unanticipated finding of this present study is that the two groups of tone language speakers performed differently from one another on the relative pitch tests. More specifically, Mandarin Chinese speakers were more proficient than the Thai speakers. Thai speakers performed more like English speakers when it came to discriminating whether two notes were the same or different (the Same/Different task). However, when it came to discriminating whether one note is higher than another note (the High/Low task), the Thai speakers did perform more like (although not as well as) the Mandarin Chinese speakers. It is quite remarkable that the greatest performance difference between Mandarin Chinese speakers and Thai speakers was on the Same/Different task, but the least difference between these two groups of tone language speakers was on the High/Low task. Furthermore, it is interesting to note that all of the Mandarin Chinese speakers scored a perfect score on the Up/Down task (a task which required participants to determine if a short series of notes was ascending or descending the scale), but the Thai speakers’ performance means are almost equidistant between
the Mandarin Chinese speakers and English speakers for the direction task. Since research has indicated that native tone language speakers tend to focus on pitch contour, or F0 slope (pitch movement and direction), when it comes to processing lexical tone (Qin and Mok, 2011; Khouw and Ciocca, 2007; Gandour, 1983; Gandour and Harshman, 1978), it is quite unusual that the Mandarin Chinese and Thai speakers did not perform more similarly on the Up/Down task. A falling or rising lexical tone, for example, is not two separate “notes” but rather a “slide” of the pitch on the scale—an effect that would seem to emulate pitch contour and hence the fundamental proficiency that was being assessed in the Up/Down task. However, one explanation may be that there is evidence that supports a difference between native Mandarin Chinese speakers and native Thai speakers in F0 processing. For example, in a comparison of lexical tone processing by native English, Thai, and Chinese speakers, Kaan et al. (2008) found that “Thai speakers differed from the Chinese and English speakers with respect to the processing of late F0 contour differences (high-rising versus mid-even tokens).” This difference may thus account for the difference in the musical pitch processing between the two given tone-language speaking participant groups.

4.2.2 Relative Pitch Proficiency and Lexical Tone Perception

The results of this study indicate that more efficient relative pitch proficiency overall moderately facilitated more efficient lexical tone perception for non-tone language speakers. As this study’s results indicate, English speakers who were better at the relative pitch tasks overall were better at determining lexical tone in the nonsense words task but were not necessarily better at identifying Mandarin words (the color adjectives). Furthermore, there was a weak negative
relationship between performance on the relative pitch tasks overall and the perception of Mandarin Color adjectives by English speakers. A closer look at the participant pool clearly reflects these findings. Of the 30 English-speaking participants, 5 performed exceedingly well on Experiment 3 (the nonsense words task), earning 90% or greater; however, these participants performed poorly on Experiment 4 (the Mandarin adjectives task), earning within the range of 31%-48% correct. Nonetheless, of these 5 participants, 4 of them did very well on the relative pitch tasks, scoring 80% or more correct on one or more of the relative pitch tasks. Furthermore, 4 additional English-speaking participants who identified 72% - 88% of the nonsense task tokens correctly performed somewhat poorly on the relative pitch tasks, identifying from 33% - 67% of the tokens correctly in one or more of the three tasks. These findings may possibly be the result of the phonological environment of the lexical tone. For the nonsense words task (Experiment 3), the lexical tone was placed on a sentence-final one-syllable nonsense word. For English speakers, sentence-final tone distinctions are quite salient since, for example, English speakers use sentence-final pitch variations to signal important pragmatic cues. Hence, the English speaker may already have been attuned to the idea of focusing on the sentence-final pitch variation and mapped the results of this focus on what the speaker knew about lexical tone in Mandarin Chinese.

The phonological environment of the Mandarin Chinese words, on the other hand, was completely different from the phonological environment of the nonsense words. To be specific, for the nonsense words task, the lexical tone was realized on sentence-final one-syllable words, but for the Mandarin Chinese words (adjectives which were sentence-final two-syllable words), the tested lexical tone was realized on the penultimate syllable of the sentence with the sentence-final syllable realized as a neutral -se. Hence, with the nonsense words, it may have been much
easier for the English speaker to perceive and identify the lexical tone since it was the last pitch that was followed by a few seconds of silence. However, if the English speakers had, for example, utilized a framework of lexical tone identification based on their understanding of sentence-final pitch variation (e.g., intonation) in the Mandarin Chinese words task, the English speaker may have been more focused on the neutral tone of the sentence-final –se, and this faulty approach may have subsequently lead to the mis-identification of the lexical tone that was to be assessed. Interestingly, since previous research maintains that non-tone language speakers are more sensitive to pitch height (e.g., Qin and Mok, 2011; Khouw and Ciocca, 2007; Gandour, 1983; Gandour and Harshman, 1978), the observation that Mandarin Chinese adjectives were more challenging appears to be somewhat problematic. The Mandarin Chinese adjectives task presented the participant with a tone melody that required the English speaker to recognize the tone of the first syllable in relationship to the neutral tone of the second syllable. It would seem that this relationship between the tones of the two syllables would have facilitated discriminating pitch height, but apparently not so.

The findings of this study further indicate that the Thai speakers’ perception of Mandarin Chinese lexical tone appears not to have been facilitated by relative pitch proficiency. This observation may be due to the strong possibility that for the tone language speakers, there are attributes other than relative pitch proficiency that contribute to lexical tone perception. As aforementioned, tone language speakers process lexical tone based on different cues than non-tone language speakers; hence, relative pitch proficiency may be irrelevant in the processing of lexical tone by the tone language speaker. As the tone language speaker negotiates perceiving lexical tone in a new tone language, he or she may be utilizing a perceptual system based on his or her understanding of lexical tone in the native language and not on relative pitch.
The given findings suggest that relative pitch proficiency likely helps the non-tone language speakers and not the tone language speakers when it comes to lexical tone perception in another language since relative pitch proficiency may possibly be used as avenue through which a non-tone language speaker’s pitch variation detection may be heightened. In addition, various studies have indicated that musical training (in both childhood and adulthood) may positively contribute to neuroplasticity, particularly relevant to language function (Kraus and Chandraskeran, 2010; Moreno and Bidelman, 2014). As mentioned, tone language speakers may already have a “heightened” music proficiency; thus, for the non-tone language speaker, music training is useful since it would create a more equal foundation upon which learning a tone language can be built.

4.2.3 Lexical Tone Perception of Tone Language Speakers and Non-Tone Language Speakers

The results of this study clearly indicate that Thai speakers were more adept at perceiving Mandarin Chinese lexical tone than the English speakers. As has been discussed in an earlier chapter, research has shown that second language learners’ approach to encoding a second language may be impacted by the native language (e.g., Gass and Selinker, 1992). An individual who already speaks a tone language has a framework for lexical tone which may possibly be transferred to another tone language; such L1 transfer may prove to quite useful for the native tone language speaker, whereas the non-tone language speaker who has never spoken a tone language lacks such “knowledge” afforded by the L1 lexical tone framework. Of course, for any L2 learner, there may be some L1 interference. However, although for the tone language speaker learning another tone language there may be L1 interference with lexical tone, the interference
may likely be on a different level than the L1 interference experienced by a non-tone language speaker. The tone language speaker already recognizes the use of pitch variation for lexical distinction, but the non-tone language speaker is most familiar with pitch variation for other cues at the phrase or sentence level (e.g., intonation). The result of these two distinct differences in the understanding and use of pitch in language would indeed lead to differing outcomes in relation to L1 transfer to the L2.

In Experiment One, lexical tone identification baseline, participants were presented with a variety of tokens that presented the four Mandarin Chinese lexical tones with differing frequencies, voice qualities, and by both male and female speakers. For this task, English speakers struggled to recognize all four Mandarin Chinese lexical tones while, quite interestingly, the Thai speakers struggled most with recognizing Tone 2 (high-rising tone) and Tone 4 (falling tone). A comparison of the lexical tone systems of Thai and Mandarin Chinese (see Table 19) shows some overlap between the two systems and substantiates the conjecture that L1 interference may have contributed to the Thai speakers’ difficulty with these two Mandarin Chinese lexical tones. As Table 19 shows, both Mandarin Chinese and Thai have a salient rising tone and a salient falling tone which should mean positive transfer from Thai to Mandarin Chinese for the Thai speakers; however, the features of these tones in the two separate systems appear to differ slightly. This subtle difference may have caused negative transfer. In the general representations of the Mandarin Chinese and Thai lexical tones in Figure 9 and Figure 10, respectively, we see that for Mandarin Chinese, Tone 2 (high-rising) ascends straight up the scale, but for Thai, the rising tone does not mimic the Mandarin Chinese rising tone as there is a slight “dip” in the pitch at the onset of the Thai rising tone. Regarding Tone 4 (falling) in Mandarin Chinese, the tone descends straight down the scale, but for Thai, there is a slight rise in the pitch
at the onset of the lexical tone (see Figures 9 and 10). The subtle differences may have caused some confusion for the Thai speakers on Experiment One (as well as in Experiment Three).

Of the Tone 2 tokens that were mis-identified, both English and Thai speakers mis-identified these tokens as Tone 1 more than any other tone. This observation is interesting since Tone 2 and Tone 1, though different when it comes to direction, are somewhat at the same pitch height (high). Since processing based on the speaker’s native intonation category, for example, may have lead to a mismatch due to L1 interference (cf., Qin and Mok, 2011; Bent, 2006; Shen, 1989), the Thai speakers may be confusing these two tones due to L1 interference. However, English speakers may be confusing these two tones because, as research has shown, pitch height is important in distinguishing lexical tone; if there is little difference in pitch height, then there is more difficulty in distinguishing the lexical tone. For the native English speaker, there likely needs to be a greater distinction in pitch height. Since non-tone language speakers are more attuned to onset and offset values of lexical tone (Gandour, 1983; Gandour and Harshman, 1978), the English speakers may have focused on the high offset value of Tone 2 without recognizing the pitch contour.

<table>
<thead>
<tr>
<th>Mandarin Chinese Lexical Tones</th>
<th>Thai Lexical Tones</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Tone</strong></td>
<td><strong>Example</strong></td>
</tr>
<tr>
<td>Tone 1</td>
<td>High Level</td>
</tr>
<tr>
<td>Tone 2</td>
<td>High-Rising</td>
</tr>
<tr>
<td>Tone 3</td>
<td>Falling-Rising</td>
</tr>
<tr>
<td>Tone 4</td>
<td>Falling</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Figure 9: Mandarin Chinese Lexical Tones

High Level   High-Rising   Falling-Rising   Falling

Figure 10: Thai Lexical Tones

High   Rising   Falling   Middle   Low

(Based on Handel, 2013)
Regarding the mis-identified Tone 4 tokens, more tone language speaking participants chose Tone 2 while more English speakers chose Tone 1. This observation is quite curious, for if pitch direction is an important cue for tone language speakers when it comes to lexical tone identification, Tone 4 (a falling tone) is a direct contradiction of Tone 2 (a rising tone). In addition, more English speaking participants chose Tone 1 (high-level tone) for Tone 4 (falling tone). It would seem that the pitch height should indeed reveal a difference in these two lexical tones for the English speakers and prevent such a mismatch, but pitch height difference has apparently not had any impact in this case for the English speakers. However, for the English speaker, it is possible that the falling tone did not provide enough pitch height distinction to help in pitch discrimination.

In Experiment One (the Baseline), the Thai speakers performed much like the English speakers; however, by Experiment 3 (the nonsense words task) and Experiment 4 (Mandarin Chinese words task), the Thai speakers performed more like Mandarin Chinese speakers. This finding may be because the baseline was somewhat a “training” period for the Thai speakers. By Experiment 3 and Experiment 4, the Thai speakers were able to negotiate the Mandarin Chinese lexical tones more efficiently after hearing these lexical tones repeated; this repetition likely eased the working memory load for the Thai speakers. On the other hand, the English speakers were likely negotiating a greater working memory load by Experiment 3 and Experiment 4 since the speakers were required to identify lexical tone in different phonological environments.

In the nonsense words task (Experiment 3) Part A, tokens of Tone 1 (high) and Tone 2 (high-rising) were presented and assessed; it is not surprising that many English speakers confused these two tones since they are close in sound due to pitch height. Experiment 3, Part B tested Tone 3 (falling-rising) and Tone 4 (falling). The English speakers missed almost ten times
more the number of tokens than the tone language speakers missed. For the tone language
speakers, who likely rely on direction as an important cue, it is interesting to point out that
these participants were more likely to note the quick rise at the end of Tone 3 due to pitch
direction. However, for English speakers, there is not a significant distinction in pitch height
between Tone 3 (falling-rising) and Tone 4 (falling) since both tones fall. Furthermore,
although each of the tested tokens in Experiment 3 were sentence-final, this phonological
environment did not seem to have aided the English speakers in discriminating the distinction
between Tone 3 and Tone 4. What should be noted, however, is since English speakers tend to
use onset or offset values more than contour, the onset and offset values of Tone 3 and Tone 4
are strikingly different.

For the Mandarin Chinese words (assessed in Experiment 4), only about half of the
participants in both groups of the tone language speakers (Mandarin Chinese speakers and Thai
speakers) earned a perfect score. For Part A, which tested Tone 1 (high-level) and Tone 2 (high
rising), the two groups of tone language speakers missed less than 2% of the tokens as opposed
to English speakers who missed almost half of the tokens presented. With Part B, which tested
Tone 3 (falling-rising) and Tone 4 (falling), tone language speakers missed approximately 10%
of the tokens whereas English speakers missed almost 60% of the tokens. One main question
arises regarding the given findings: Why did some of the Mandarin Chinese speakers miss some
of the words in their native language? This researcher posits two possible reasons why there was
this outcome. First, the sentences in which the Mandarin Chinese adjectives were placed were
English sentences; thus, codeswitching may have impacted the perception. Second, working
memory issues, test fatigue, or boredom may have possibly begun to affect the participants since
this task was the final task and completion of the four tasks took two hours for participants to
complete. In a study to determine if phonological information may be processed similarly to musical pitch information in the working memory, Williamson, Baddeley, and Hitch (2006) conducted a study endeavoring to assess English speakers found that “[m]elodies composed of notes that were close together were recalled less successfully compared to melodies where the notes were more distant. […] [A]coustic confusability is detrimental to the short-term recall of both verbal and musical materials in short term memory” (p. 1588). Thus, for the English speaker, if the pitches in the lexical tones were too close together, this feature may have negatively impacted these participants’ ability to perceive and identify the correct lexical tone since the working memory may have been over-extended. There may be also a difference in the speaking F₀ which may impact the perception difference between non-tone language speakers and tone-language speakers. As pointed out by Keating and Kuo (2012), “Speaking F₀ is to some extent an arbitrary aspect of speech, and a particular F₀ range may be part of the phonetic structure of a language, such that in the limit, a speaker would sound non-native (have a foreign accent) using a different F₀ range” (p. 1050). Hence, if F₀ can impact how listeners perceive accent, it may also impact how listeners perceive lexical tone. Furthermore, neuroscience research has shown that language background may affect both attentive and non-attentive processing of lexical tones (cf. Chandrasekaran, Krishnan, and Gandour, 2007; Kaan et al, 2007); it would seem that the Mandarin Chinese speakers (no matter how tired or bored) should have still been able to earn perfect scores on the Mandarin Chinese words task. It is important to note that the findings of this study reveal that having lexical tone as a characteristic of one’s native language does not necessarily mean that this speaker will be automatically and fluently competent at perceiving lexical tone in another language. This observation is substantiated by the Thai speakers’ performance on the tasks; although the Thai speakers performed better than
the English speakers, the Thai speakers did not perform as well as the Mandarin Chinese speakers.

4.2.4 A Final Word about Music Training and Lexical Tone Perception

There is evidence that the brain is positively modified by music training (e.g., Kraus and Chandresekaran, 20103). Furthermore, this neuroplasticity is not age-restricted (Skoe and Kraus, 2012; Bidelman and Alain, 2015). We have also presented evidence that supports previous research illustrating how musicians outperform non-musicians on language-related tasks. More specifically, we have shown that the non-tone language speaking adult who has a higher relative pitch proficiency is more adept at lexical tone perception. Since the brain is malleable and relative pitch is a teachable skill (no matter the age of the individual), it can be argued that parallel pitch instruction with tone language instruction may be essential for teaching linguistic tone to the non-tone language speaker.

4.3 Limitations and Future Research

4.3.1 Limitations

There were several limitations in the current study. One major limitation for this study was the length of time required for the four experiments. The study takes two hours from start to finish, and this researcher speculates that, although there was a break mid-way, the listening tasks were still likely tiring. Hence, there is the possibility that tone language speakers began to perform less than optimal by the end of this study due to fatigue. Another limitation of this study
was the failure to determine and include specific pitches categorized based on formant frequencies as a tested variable. Instead, this researcher chose, for example, to obtain a series of tokens from a variety of speakers without specifically categorizing them based on formant frequencies.

4.3.2 For Further Research

Since this study has proven to fit nicely into the current body of work assessing the relationship between music and language overall, and between music and tone language acquisition as a second language, further investigation of relative pitch ability and lexical tone perception is indeed needed. In the future, if such a study is attempted again, three major items should be revised. First, the relative pitch tests should utilize the human voice instead of (or in addition to) musical instruments; moreover, there should be a task that presents tokens that “slide” rather than represent distinct separate notes. Second, since there is the possibility that participants may have been confused by terminology used in relation to music, participants need to be trained more specifically on what is meant musically by “higher” and “lower” or “ascending” and “descending.” Third, since there was concern about the impact of the time commitment for this study (2 hours) and its possible impact of fatigue by the end, the order of the experiments should be randomized each time a new set of participants is tested.

Furthermore, although this study did present tokens of various frequencies, voice types, and by both genders, the assessment was general. Thus, one area that needs additional exploration is the impact of voice type. A higher pitched female voice may render a much different response than a lower pitched female voice; conversely, the same may prove to be true
not only for male speakers but also for female versus male speakers. A second area that may present interesting results is investigation by utilizing nonsense tones. Nonsense tones may allow for more uniform assessment of the native tone language speakers. Another area for further research is melody. The English-speaking participants in this study performed well on one-syllable sentence-final tokens but not on two-syllable words where the lexical tone is on the first syllable. Teaching a longer “melody” in a word (for example, the word for the calendar month December in Mandarin Chinese is andalone yuè – a three-syllable word with a melody based on Tone 2/Tone 4/Tone 4 (rising/falling/falling). Hence, a study investigating pitch ability and melody recognition may prove to be fruitful. Also, there should be further assessment of pitch ability based on a musical scale other than the Western scale (e.g., Eastern). Further, it would be interesting to include other tone language or even pitch-accent speakers from other regions (e.g., Yoruba or Cherokee). Moreover, although this study did not endeavor to study age and gender and the impact of these two demographics categories on tone identification, studies that consider these demographics, as well as other possible categories such as English dialect, may provide interesting results.

Three final areas warrant further exploration in relation to this study. First, the participant pool should be comprised of individuals who struggle to maintain pitch (e.g., while singing) or who cannot mimic pitch at all. The second area is based on neuroscience as running the pitch tasks and lexical tone tasks concurrent with fMRI studies on both musicians and non-musicians. Finally, it is quite curious that many participants struggled with determining if a series of notes was ascending or descending. Further investigation of the tritone paradox (the auditory illusion in which listeners confuse ascending and descending notes) (Deutsch et al., 1990 and Deutsch 1991) and its possible impact on lexical tone perception is needed.
Chapter 5

Conclusion

At the onset of this study, this research endeavored to investigate the possible relationship between music and language—specifically whether being proficient at relative pitch would enhance an English learner’s aptitude to learn a tone language. The results of this study show that non-tone language speaking participants who exhibit stronger relative pitch perception skills are more successful at lexical tone perception when the lexical tone is sentence-final; however, placing the same tokens in a different phonological environment may impact a non-tone language speaker’s perception ability. Furthermore, features such as voice quality and gender of the speaker who produced the token did not appear to impact outcomes. Finally, the results of this study indicate that tone language speakers are more adept at relative pitch as well as lexical tone perception in a second language than non-tone language speakers.

Intrigued by the general research that supports not only the relationship between music and other abilities but also between music and language in particular, this researcher hopes that this study makes a contribution (no matter how small) to the argument that music education is an integral part of the learning experience—especially for children—as it may contribute to overall learning. In addition, as a contributor to general neuroplasticity, music instruction may aid in the important endeavor of adult second language learning. In the global world in which we live today, learning a second, third, or even fourth language has become increasingly important. Native English-speaking adults who wish to learn a tone language may benefit from relative pitch instruction—a key point that this study sought to substantiate. Is it all relative? When it comes to learning a tone language, it just may be.
LIST OF APPENDICES

Appendix 1: Participant In-take Questionnaire
Appendix 2: Script for Pre-Experiment Training
Appendix 3: Stimuli Blocks for Pre-Experiment training
Appendix 4: Pre-Experiment Training Handout
Appendix 5: Experiment 1 Stimuli
Appendix 6: Experiment 1 Script
Appendix 7: Experiment 2 Script
Appendix 8: Experiment 3 Pre-Training materials
Appendix 9: Experiment 3 Script
Appendix 10: Experiment 3 Stimuli
Appendix 11: Experiment 4 Script
Appendix 12: Experiment 4 Stimuli
APPENDIX 1

Is it all Relative? Relative Pitch Ability and Lexical Tone Perception/Tone Language Comprehension by English-Speaking Adults

IN-TAKE QUESTIONNAIRE

Thank you for taking the time to respond to our request for participants in this study.

INFORMED CONSENT:

Introduction: My name is Sherri Celeste Wortes, and I am a doctoral candidate at The Graduate Center, City University of New York, where the actual research study will take place. The purpose of this research study is to determine if there is a relationship between specific music ability such as relative pitch and the ability to perceive and reproduce lexical tone. The results of this study may aid in future instruction of adult second language learners who have never had experience with a tone language as they try to learn a most important aspect of such languages—lexical tone. This study is likely to support the hypothesis that utilizing music activities as part of a second language curriculum will aid second language learners in their endeavor. This finding will hopefully increase the awareness of educators in the relationship between curricula and also enhance teaching methods in the second language classroom.

Procedures: Approximately 60 individuals are expected to participate in this study, which has four parts. All participants will initially be asked to complete a brief questionnaire. Next, non-tone language speaking individuals will be introduced to lexical tone and then asked to ascertain the difference in lexical tone carried on a simple syllable. Third, participants will take part of the Montreal Battery of Evaluation of Amusia. Fourth, participants will take a lexical tone discernment test utilizing nonsense words. Finally participants will take a listening test to discern Mandarin adjectives which have different lexical tones. The time commitment of each participant is expected to be one session of no more than two hours. At the conclusion of this session, each participant who has undergone the complete study will be compensated $40 for his or her time. This session will take place in Room 7393 at the CUNY Second Language Acquisition Lab, The Graduate Center, City University of New York, 365 Fifth Avenue, New York, NY 10016 OR Room 226A at The College of Staten Island, 2800 Victory Blvd., Staten Island, NY 10314.

Possible Discomforts and Risks: This study poses practically no risks you as a participant. However, to ensure there is no breach in confidentiality, personal identifiers such as your name will not be taken.

Benefits: This study poses no direct benefits to you. However, your participation in this study may increase general knowledge of second language acquisition didactics and abilities.

Voluntary Participation: Your participation in this study is voluntary, and you may decide not to participate without prejudice or penalty; however, to be compensated $40, you are must complete the study to its entirety.

Confidentiality: The data obtained from you will be collected via digital audio and written document. The collected data will be accessible to: me, the Principal Investigator, my Faculty Advisor, Dr. Gita Martohardjono;
and to the Institutional Review Board (IRB) Members and staff. I will protect your confidentiality by coding the data and securely storing the data.

**Statement of Consent:** Do you have any questions at this point?

Do you voluntarily agree to answer a few non-personal questions for this study? These questions should take no more than 3 minutes.    _____YES     _____NO

**PART ONE**
A. Gender:    ____Male  _____Female  ___Other

B. Age range:  
   _____17 years or less
   _____18-29
   _____30-39
   _____40+

C. What languages do you speak fluently.
   __________________________________________
   __________________________________________
   __________________________________________
   __________________________________________

D. Is English your first language?  _____Yes _____No     If yes, please skip to F. If no, please continue with E.

E. (If “no” to question C) What is your first language?   __________________________________

F. What countries have you traveled to?
   __________________________________________
   __________________________________________
   __________________________________________
   __________________________________________

G. Do you live and/or work in a neighborhood(s) where a tone language such as Mandarin Chinese or Cantonese is spoken? (If yes, please do NOT indicate which neighborhood.)
   ____Yes   _____No

**PART TWO**
H. Do you play a musical instrument or sing?  ____Yes  ____No     If no, skip to PART THREE. If yes, please continue with I.

I. Tell me the instrument(s) in which you are musically competent.
   __________________________________________
   __________________________________________
   __________________________________________
J. What is your principal instrument? ______________________________

K. If you are a vocalist, what is your vocal fach? ______________________________

L. How many years of continuous music instruction have you had?
   _____ less than 2 years
   _____ 2 – 4 years
   _____ 5 – 7 years
   _____ 8 or more years

M. If you are no longer receiving music instruction, what was the year in which your music instruction ended?
   ___________________________

N. Do you currently perform as a professional musician and/or vocalist?  ____Yes  _____No

O. If yes, when did you last perform?
   _____ Within the last 30 days
   _____ Three to six months ago
   _____ More than a year ago

P. Do you have perfect (absolute) pitch?  ____Yes  _____No  ____Do not know

Q. Do you have strong relative pitch ability?  ____Yes  _____No  ____Do not know

PART THREE
In this study, we are endeavoring to discern if there is a relationship between an individual’s pitch ability and his or her discrimination and reproduction of lexical tone. In the study, you will be asked to listen and respond; in addition, you will be given pictures which you must discern color.

R. Do you have any self-reported color blindness? ____Yes  _____No  ____Do not know

S. Do you have any self-reported hearing loss?  ____Yes  _____No  ____Do not know

PART FOUR
The tasks for this study should take no more than two hours of your time.

T. Do you have any questions at this point?  ____Yes  _____No  ____Do not know

We thank you for taking the time to answer our questions. Please hand this questionnaire to the proctor and await further instruction.
APPENDIX 2

Pre-Experiment Training Script

(Items in GREEN are instructions to the Proctor)

Thank you for agreeing to participate in this study! We truly appreciate your willingness to help us. Before we begin our experiments, I would like to introduce you to some important aspects of Mandarin Chinese, the language on which we will base the next few experiments. (Hand out the Pre-Experiment Training Handout.)

Mandarin Chinese is a tone language; that is, Mandarin uses what can be considered musical pitch as an important part of its language system. In Mandarin Chinese, one-syllable words carry different pitches, and these changes in pitch actually change the meaning. Consider, for example, the one-syllable word ma. Take a look at your training handout (Point out on your copy and ensure participant is following.) In PART ONE of the handout, we have two different versions of that syllable. Now, although both versions have the same two letters, they carry a different pitch. The first version, mā, means mother, but mă would mean horse. As you can see, changes in pitch on the same word seriously impact meaning. Thus, if you consider the two examples of ma, Mandarin speakers must definitely be very careful with pitch; otherwise, they could accidentally call their mothers a horse! (Chuckle.) These pitches that impact meaning in a language such as Mandarin are called lexical tones or tones. That is why I opened my introduction with the point: Mandarin Chinese is a tone language.

If you look at PART TWO on your handout, you will find that there are four lexical tones, or tones, in Mandarin. (Point to your copy and ensure that the participant is following along.)

Tone to Indicate Lexical meaning in Mandarin Chinese

<table>
<thead>
<tr>
<th>Tone Number</th>
<th>Tone Description</th>
<th>Example in IPA</th>
<th>Lexical Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tone 1</td>
<td>High-Level</td>
<td>[mā]</td>
<td>'mother'</td>
</tr>
<tr>
<td>Tone 2</td>
<td>High-Rising</td>
<td>[mă]</td>
<td>'hemp'</td>
</tr>
<tr>
<td>Tone 3</td>
<td>Falling-Rising</td>
<td>[mā]</td>
<td>'horse'</td>
</tr>
<tr>
<td>Tone 4</td>
<td>Falling</td>
<td>[mă]</td>
<td>'scold'</td>
</tr>
</tbody>
</table>

Please follow along in PART TWO on the Pre-Experiment Training Handout as I talk about the four lexical tones in Mandarin. Tone 1 is the High-Level Tone, and it sounds like this: mā. Now, you may remember that we used this word before with the same tone. Do you/Does anyone remember what mā means? (Give the participant a few moments to see if he or she recalls.) Mother. (If the participant remembers and is able to answer, please commend him or her and then continue. If the participant does not remember, just continue.) Tone 2 is the High-
Rising Tone, and it sounds like this: *má*. *Má* means *hemp* which is used for various products such as fiber for clothing. Tone 3 is the Falling-Rising tone, and it sounds like this: *mā*. Of course, you might remember that we used *mā* before. Do you/Does anyone remember what *mā* means? *Horse*. (Give the participant a few moments to see if he or she recalls.) *Mother*. (If the participant remembers and is able to answer, please commend him or her and then continue. If the participant does not remember, just continue.) Excellent. The final tone in Mandarin is Tone 4, the Falling Tone, and it sounds like this: *mà*. This word, *mà*, means *scold*.

I would like to point out that these lexical tones are represented by various diacritic markings. If you look closely at the representations on your handout, the marking above the *a* in *ma* differs with each tone. Tone 1 has a line. Tone 2 has what looks like an apostrophe. Tone three has a little “bowl.” Tone 4 has something that also looks like an apostrophe (but going in the opposite direction of the one for Tone 2). Fortunately, you will not have to worry about those markings. We want to make your experience very straightforward so that you can concentrate on what you *hear*. Let’s take a look at PART THREE on your handout. Instead of using the complex representations that you just saw, the tones are going to be represented by raised numbers. Thus, *ma* with Tone 1 will be *ma*₁; *ma* with Tone 2 will be *ma*₂; *ma* with Tone 3 will be *ma*₃; and *ma* with Tone 4 will be *ma*₄. I hope that this method is straightforward, yes? (Allow for the participant to respond. If the participant agrees, then continue. If the participant does not agree, pause and ask for any questions and/or concerns and respond accordingly.)

Of course, it may be quite straightforward to hear the differences in tone as I, a non-native Mandarin speaker, utter these syllables slowly and deliberately—a possibly with a weird accent! (Chuckle.) What happens when we listen to a native speaker? So, we shall listen to the four tones on syllables other than *ma*, and these syllables are spoken by native Mandarin speakers—one male and one female. As you listen to the recording, you will hear four different syllables: *la*, *li*, *lu*, and *lo*. Each syllable—with a particular tone—will be uttered two times. The first time, the syllable will be uttered by a male. The second time, the syllable will be uttered by a female. Thus, for example, you will hear *la*₁ uttered by a male and then *la*₁ uttered by a female speaker; there will be a five second break and then you will hear *li*₁ uttered by a male and then a female speaker, etc. You will, thus, hear a total of eight versions of Tone 1. Then, there will be a 10 second pause. After the 10 second pause, you will hear eight versions of Tone 2 (a male and a female speaker for the syllables *la*₂, *li*₂, *lo*₂, and *lu*₂). The same protocol will be followed regarding Tone 2, Tone 3, and Tone 4. If you refer to PART FOUR on your Pre-Experiment Training Handout, these syllables are spelled out for you so that you can follow along. (Point to your copy of the handout.) We are reading from left to right for each row in the chart. Does this chart make sense to you? (Pause to be sure that the participant understands; give him or her a chance to respond.) Do you have any questions before we listen to the tape? (Pause
and allow the participant to consider if he or she has any questions. If there are none, then continue. If there are questions, please answer accordingly.)

(Play the recording)

Now that you have been introduced to the four lexical tones in Mandarin Chinese, do you have any questions? (Pause and allow the participant to consider if he or she has any questions. If there are none, then collect the handout. If there are questions, please answer accordingly and then collect the handout.)
## APPENDIX 3

### Pre-Experiment Training Stimuli

<table>
<thead>
<tr>
<th></th>
<th>Male</th>
<th>Female</th>
<th>Male</th>
<th>Female</th>
<th>Male</th>
<th>Female</th>
<th>Male</th>
<th>Female</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tone 1</td>
<td>$la^1$</td>
<td>$la^1$</td>
<td>$li^1$</td>
<td>$li^1$</td>
<td>$lu^1$</td>
<td>$lu^1$</td>
<td>$lo^1$</td>
<td>$lo^1$</td>
</tr>
<tr>
<td>Tone 2</td>
<td>$la^2$</td>
<td>$la^2$</td>
<td>$li^2$</td>
<td>$li^2$</td>
<td>$lu^2$</td>
<td>$lu^2$</td>
<td>$lo^2$</td>
<td>$lo^2$</td>
</tr>
<tr>
<td>Tone 3</td>
<td>$la^3$</td>
<td>$la^3$</td>
<td>$li^3$</td>
<td>$li^3$</td>
<td>$lu^3$</td>
<td>$lu^3$</td>
<td>$lo^3$</td>
<td>$lo^3$</td>
</tr>
<tr>
<td>Tone 4</td>
<td>$la^4$</td>
<td>$la^4$</td>
<td>$li^4$</td>
<td>$li^4$</td>
<td>$lu^4$</td>
<td>$lu^4$</td>
<td>$lo^4$</td>
<td>$lo^4$</td>
</tr>
</tbody>
</table>
APPENDIX 4

Pre-Experiment Training Handout

PART ONE:   *mā* means *mother*   *mă* means *horse*.

PART TWO

tone to Indicate Lexical meaning in Mandarin Chinese

<table>
<thead>
<tr>
<th>Tone Number</th>
<th>Tone Description</th>
<th>Example in IPA</th>
<th>Lexical Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tone 1</td>
<td>High-Level</td>
<td>[mā]</td>
<td>‘mother’</td>
</tr>
<tr>
<td>Tone 2</td>
<td>High-Rising</td>
<td>[mā]</td>
<td>‘hemp’</td>
</tr>
<tr>
<td>Tone 3</td>
<td>Falling-Rising</td>
<td>[mā]</td>
<td>‘horse’</td>
</tr>
<tr>
<td>Tone 4</td>
<td>Falling</td>
<td>[mă]</td>
<td>‘scold’</td>
</tr>
</tbody>
</table>

PART THREE

<table>
<thead>
<tr>
<th>Tone</th>
<th>Representation</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tone 1</td>
<td><em>ma</em>¹</td>
<td>‘Mother’</td>
</tr>
<tr>
<td>Tone 2</td>
<td><em>ma</em>²</td>
<td>‘Hemp’</td>
</tr>
<tr>
<td>Tone 3</td>
<td><em>ma</em>³</td>
<td>‘Horse’</td>
</tr>
<tr>
<td>Tone 4</td>
<td><em>ma</em>⁴</td>
<td>‘Scold’</td>
</tr>
</tbody>
</table>

PART FOUR

<table>
<thead>
<tr>
<th></th>
<th>Male</th>
<th>Female</th>
<th>Male</th>
<th>Female</th>
<th>Male</th>
<th>Female</th>
<th>Male</th>
<th>Female</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tone 1</td>
<td><em>la</em>¹</td>
<td><em>la</em>¹</td>
<td><em>li</em>¹</td>
<td><em>li</em>¹</td>
<td><em>lu</em>¹</td>
<td><em>lu</em>¹</td>
<td><em>lo</em>¹</td>
<td><em>lo</em>¹</td>
</tr>
<tr>
<td>Tone 2</td>
<td><em>la</em>²</td>
<td><em>la</em>²</td>
<td><em>li</em>²</td>
<td><em>li</em>²</td>
<td><em>lu</em>²</td>
<td><em>lu</em>²</td>
<td><em>lo</em>²</td>
<td><em>lo</em>²</td>
</tr>
<tr>
<td>Tone 3</td>
<td><em>la</em>³</td>
<td><em>la</em>³</td>
<td><em>li</em>³</td>
<td><em>li</em>³</td>
<td><em>lu</em>³</td>
<td><em>lu</em>³</td>
<td><em>lo</em>³</td>
<td><em>lo</em>³</td>
</tr>
<tr>
<td>Tone 4</td>
<td><em>la</em>⁴</td>
<td><em>la</em>⁴</td>
<td><em>li</em>⁴</td>
<td><em>li</em>⁴</td>
<td><em>lu</em>⁴</td>
<td><em>lu</em>⁴</td>
<td><em>lo</em>⁴</td>
<td><em>lo</em>⁴</td>
</tr>
</tbody>
</table>
APPENDIX 5

Experiment 1 Stimuli Recorded

Key:
MSP1  \(sa^1\)  \(sa^2\)  \(sa^3\)  \(sa^4\)
MSP2  \(sa^1\)  \(sa^2\)  \(sa^3\)  \(sa^4\)
MSP3  \(sa^1\)  \(sa^2\)  \(sa^3\)  \(sa^4\)
MSP4  \(sa^1\)  \(sa^2\)  \(sa^3\)  \(sa^4\)
FSP1  \(sa^1\)  \(sa^2\)  \(sa^3\)  \(sa^4\)
FSP2  \(sa^1\)  \(sa^2\)  \(sa^3\)  \(sa^4\)
FSP3  \(sa^1\)  \(sa^2\)  \(sa^3\)  \(sa^4\)
FSP4  \(sa^1\)  \(sa^2\)  \(sa^3\)  \(sa^4\)

After recording completion, the above stimuli were randomized, allowing only for an equal number of male and female speakers in both blocks.

See Table 5 for the final result.
APPENDIX 6

Experiment 1 Script

Now that you have been introduced to the 4 lexical tones in Mandarin Chinese, we would like to test how well you can identify these tones.  *(Hand out the Experiment 1 Tokens card.)* In a moment, you will hear a recording with a series of the syllable sa.  On your card, you see four representations of sa:  sa\(^1\), sa\(^2\), sa\(^3\), sa\(^4\) *(First say, for example “sa-1” and then pronounce sā.*  Do this for each token.)  As you listen to the recording, we ask that you identify which tone each syllable carries by pressing the number that corresponds with the tone; in other words, press number 1 for Tone 1), number 2 for Tone 2, number 3 for Tone 3), or number 4 for Tone 4.  Thus, if you hear sā, for example, what number would you have to press?  *(Give the participants an opportunity to answer.  If the participants answer “one,” then commend them and then continue; if the participants pause a long time or answer incorrectly, then say “Since sā carries Tone 1, you would identify Tone 1 by pressing the number 1.”)*  Here is the fun part.  You will be awarded a penny for each syllable’s tone that you are able to identify correctly within five seconds or less!  If you take longer than five seconds to identify correctly a token, then that token will be counted correct, but you will not receive compensation for token.  Any token that is incorrect will be marked accordingly.  You will be told at the end of this particular experiment the number of items that you identified correctly within five seconds and compensated.

Do you have any questions?  *(Pause and allow the participants to consider if he or she has any questions.  If there are none, then continue with the practice question.  If there are questions, please answer accordingly and then continue with the practice question.)*

Let’s try this out to be sure that you feel comfortable with what you are about to do.  You are allowed to look at your reference card if you need a reminder.

You are about to hear three syllables.  Please determine which tone each syllable carries and press 1, 2, 3, or 4 to identify the corresponding tone that you hear.  *(Continue with the Practice Test.)*  That seemed pretty simple, right?  *(Pause and allow the participants to answer.  If he or she agrees, then continue with the test.  If anyone disagrees, please query to find out the issue/s and answer any questions and/or address any concerns.  After the test is complete, please collect the cards.)*
Experiment 1 Card Given to Participants

**EXPERIMENT 1 TOKENS**

| $sa^1$ | Tone 1  | High-Level Tone |
| $sa^2$ | Tone 2  | High-Rising Tone |
| $sa^3$ | Tone 3  | Falling-Rising Tone |
| $sa^4$ | Tone 4  | Falling Tone |
APPENDIX 7

Experiment 2 Script

For this experiment, which has three short parts of twelve questions each, you will listen musical notes. We will begin with the Same/Different Task. In this task you will hear a pair of notes and you must determine if the second note is the same as the first note. Let’s try:

Play the recording. Are these notes the same or different?

Let’s practice
Play the recording (be sure to pause after each note series).

STIMULI FOR THE PRACTICE

1. Piano  D⁴  C⁴
2. Flute  C⁴  E⁴

Determine whether to move forward based on participants’ responses.

Administer Experiment 2 Same/Different Task

For the Direction Task:

For this experiment, which has three short parts of twelve questions each, you will listen musical notes. We will begin with the Same/Different Task. In this task you will hear a pair of notes and you must determine if the second note is the same as the first note. Let’s try:

Play the recording (be sure to pause after each note series). Are these notes ascending up the scale (or going higher) or descending the scale (or going lower)?

Let’s practice:
Play the recording (be sure to pause after each note series).
STIMULI FOR THE PRACTICE

1. C⁴ D⁴ E⁴ F⁴ G⁴
2. G⁴ F⁴ E⁴ D⁴ C⁴
3. C⁴ B³ A³ G³ F³

Determine whether to move forward based on participants’ responses.

Administer Experiment 2 Direction Task

For the Higher/Lower Task:

Finally, for the Higher/Lower Task, you will

For this task you will hear a pair of notes and you must determine if the second note is higher or lower than the first note. Let’s try:

Play the recording (be sure to pause after each pair). Is the second note higher or lower than the first note?

Let’s practice
Play the recording (be sure to pause after each note series).

STIMULI FOR THE PRACTICE

1. Bass C⁴ C⁴
2. Piano D⁴ C⁴
3. Flute C⁴ E⁴

Administer Experiment 2 Higher/Lower Task
Now let’s see if we can take the idea of tones and extend them to words. For this experiment, which has two parts, you will learn four make believe words that you will later practice identifying.

We are going to begin, however, by teaching you our first two words: rā and rā. (Give the participants the EXP 3 TEST A card with the corresponding words and pictures). Rā, which is Tone 1 (as indicated by ra-1) is a make-believe word for a scarf. Rā, which is Tone 3 (as indicated by ra-3) is a make-believe word for sunglasses. So, rā, Tone 1, is a scarf, and rā, Tone 3, is sunglasses. Do you have any questions about these two words? (Pause and allow the participants to ask any questions, answer accordingly).

Let’s try this out. (Show the training pictures for scarf--and allow the participants to answer. Go through the pictures several times until it is clear that the participants understands that rā, Tone 1, is a scarf. Next, show the participants the training pictures of sunglasses—and allow the participants to answer. Once again, go through these pictures several times until it is clear the participants understands that rā, Tone 3, is sunglasses.

We are going to try this practice again, but this time the pictures will be mixed up (for example, you may see sunglasses and then a scarf). I want to be sure that you feel comfortable recognizing that rā, Tone 1, is a scarf and rā, Tone 3, is sunglasses. Okay? Here we go. (The pictures are now randomized—each picture will be shown twice. Go through this pre-experiment training with the participants and commend them along the way as positive reinforcement.)

So let’s take a break and then we will continue. In about 3 minutes.

After the break, proceed to Test A.
PRE-EXPERIMENT TRAINING FOR EXPERIMENT 3 TEST B

Now for the next two words: vú and vù. (Give the participant the card) Vú, which is Tone 2 (as indicated by vu-two) is a make-believe word for bracelet. Vù, which is Tone 4, as indicated by vu-4) is a make-believe word for a hat. So, vú, Tone 2, is a bracelet, and vù, Tone 4, is a hat. Do you have any questions about these two words? (Pause and allow the participant to ask any questions, answer accordingly, then proceed.)

Just like the first time, let’s try this out. (Show the pictures—bracelet—and allow the participants to answer. Go through these pictures several times until it is clear that the participants understand that vú, Tone 2, is bracelet. Next, show the participants the second three pictures—hat—and allow the participant to answer. Once again, go through these 3 pictures several times until it is clear the participants understand that vù, Tone 4, is a hat.)

We are going to try this practice again, but this time the pictures will be mixed up (for example, you may see a hat and then a bracelet). I want to be sure that you feel comfortable recognizing that vù, Tone 4, is a hat and vú, Tone 2, is a bracelet.

Okay? Here we go. (The pictures are randomized. Go through this pre-experiment training with the participants and commend them along the way as positive reinforcement.)

Once again, we will take a quick break and then continue. (Break for about 3 minutes After the break, proceed to Test B.)
### EXPERIMENT 3 TEST A CARD

<table>
<thead>
<tr>
<th>ID</th>
<th>Item</th>
</tr>
</thead>
<tbody>
<tr>
<td>ra-1</td>
<td>scarf</td>
</tr>
<tr>
<td>ra-3</td>
<td>sunglasses</td>
</tr>
</tbody>
</table>

### EXPERIMENT 3 TEST B CARD

<table>
<thead>
<tr>
<th>ID</th>
<th>Item</th>
</tr>
</thead>
<tbody>
<tr>
<td>vu-2</td>
<td>bracelet</td>
</tr>
<tr>
<td>vu-4</td>
<td>hat</td>
</tr>
</tbody>
</table>
PRACTICE STIMULI FOR EXPERIMENT 3 TEST A

What is the Pooh bear wearing?
PRACTICE STIMULI FOR EXPERIMENT 3 TEST B

What is the Pooh bear wearing?
**APPENDIX 9**

**Experiment 3 Script**

For this experiment, which has two parts, we will begin by teaching you four make-believe words that you will later practice identifying. (Hand out the *Experiment 3 Vocabulary* card with the four nonce words.) We will begin with two words, then give you Test Part A, and then teach you two more words, and then give you Test Part B. Okay?

So, the two words that we will begin with are rā and rǎ. Rā, which is Tone 1 (as indicated by ra-one) is a make-believe word for a scarf. Rǎ, which is Tone 3 (as indicated by ra-3) is a make-believe word for sunglasses. So, rā, Tone 1, is a scarf, and rǎ, Tone 3, is sunglasses. Do you have any questions about these two words? *(Pause and allow the participant to ask any questions, answer accordingly, then proceed with the Practice Test Part A.)*

In this experiment, you will see a series of pictures. With each picture, you will hear a pair of sentences. You are asked to determine which sentence correctly identifies what you see in the picture by pressing A or B. For example, look at this picture. Now listen to these two sentences: A The Pooh Bear is wearing rā. B The Pooh Bear is wearing rǎ. Which is the correct sentence, Sentence A or Sentence B? *(Pause and allow the participant an opportunity to answer. If the participant answers B, commend him or her and continue with the test. If the participant answers A, pause and state the following: “Let’s listen to the two sentences again: A The Pooh Bear is wearing rā. B The Pooh Bear is wearing rǎ. If you notice, ra in sentence A is actually Tone 4 and not Tone 1; we need the word for frizzy hairdo which is rā—Tone 1.” Reaffirm the participant and continue with the test.)*

*(Administer Experiment 3 Part A. See Appendix L.)*

**AFTER EXPERIMENT 3 PART A, PLEASE COLLECT THE CARD AND PROCEED TO THE PART B TRAINING (SEE APPENDIX J-3)**

We are going to continue with Experiment 3, in which you are learning a mini-vocabulary and then testing to see if you can recognize the tone. Our final two words in this experiment are vú and vù. Vú, which is Tone 2 (as indicated by vu-two) is a make-believe word for bracelet. Vù, which is Tone 4, as indicated by vu-four) is a make-believe word for hat. So, vú, Tone 2, is bracelet, and vù, Tone 4, is a hat. Do you have any questions about these two words? *(Pause and allow the participant to ask any questions, answer accordingly, then proceed with the Practice Test Part B.)*
In this experiment, you will see a series of pictures. With each picture, you will hear a pair of sentences. You are asked to determine which sentence correctly identifies what you see in the picture by pressing A or B. For example, look at this picture (hand out card for Experiment 3 Practice Test A). Now listen to these two sentences: 

A. The Pooh bear is wearing vú.
B. The Pooh Bear is wearing vù.

Which is the correct sentence, Sentence A or Sentence B? (Pause and allow the participants an opportunity to answer. If the participants answer A, commend them and continue with the test. If the participants answers B, pause and state the following: “Let’s listen to the two sentences again: A. The Pooh Bear is wearing vú. B. The Pooh Bear is wearing vù. If you notice, vu in sentence B is actually Tone 4 and not Tone 2. Vú, which carries Tone 2, is the word for bracelet.” Be sure to reaffirm the participant and continue with the test.)

Administer Experiment 3 Part B.

Once the Experiment Part B is complete, please collect the cards.
EXPERIMENT 3 PRACTICE TEST PART A

EXPERIMENT 3 PRACTICE TEST PART B
APPENDIX 10

EXPERIMENT THREE STIMULI

Experiment 3 TEST A Stimuli

rā Tone 1 scarf rā Tone 3 sunglasses

© 2015 Sherri Celeste Wortes aka Sloane Celeste von Wertz
Experiment 3 TEST A Stimuli

vú  Tone 2  bracelet  vú  Tone 4  hat

© 2015 Sherri Celeste Wortes aka Sloane Celeste von Wertz
APPENDIX 11

Experiment 4 Script

For this experiment, which has two parts, you will learn four Mandarin Chinese color words. We will begin with two words (red and black), then give you Test Part A, and then teach you two more words (green and purple), and then give you Test Part B. Okay?

(Hand out the stimuli card).

The word HONE-se₁ is black in Mandarin and the word HONG-se² is red in Mandarin. Let’s listen to two of our speakers say the Mandarin word for black for us (play the recording). Now, let’s listen to our speakers say the Mandarin word for red for us (play the recording). So now we know that Hĕi sè is black in Mandarin while Hòng sè is red in Mandarin. Let’s listen to these two words in sentences. (Hold up the card with the corresponding pictures as you play the recording.) Do you have any questions about these two words? (Pause and allow the participants to ask any questions, answer accordingly, then proceed with the Practice Test Part A.)

In this experiment, you will see a series of pictures. With each picture, you will hear a pair of sentences. You are asked to determine which sentence correctly identifies what you see in the picture by pressing A or B. For example, look at this picture. Now listen to these two sentences: ¹The circle is Hĕi sè. ²The the circle is Hĕi sè. Which is the correct sentence, Sentence A or Sentence B? (Pause and allow the participants an opportunity to answer. If the participants answer B, commend him or her and continue with the test. If the participants answer B, pause and state the following: “Let’s listen to the two sentences again: (play the two sentences). If you notice, Hĕi sè in sentence B is actually Tone 3 and not Tone 1; we need the word for black which is Hĕi sè—Tone 1.” Reaffirm the participants and continue with the test.)

(Administer Experiment 4 Part A.)

AFTER EXPERIMENT 4 A, continue with Experiment 4 B.
EXPERIMENT 4 PART B

We are going to continue with Experiment 4, in which you are learning a mini-vocabulary and then testing to see if you can recognize the tone. Our final two words in this experiment are zǐ sè and lǜ sè.

The word zǐ sè³ is purple in Mandarin and the word lǜ sè⁴ is green in Mandarin. Let’s listen to a Mandarin speaker say these words. (play the recording). So now we know that The word zǐ sè³ is purple in Mandarin and the word lǜ sè⁴ is green in Mandarin. Let’s listen to these two words in sentences.

Do you have any questions about these two words? (Pause and allow the participants to ask any questions, answer accordingly, then proceed with the Practice Test Part B.)

In this experiment, you will see a series of pictures. With each picture, you will hear a pair of sentences. You are asked to determine which sentence correctly identifies what you see in the picture by pressing A or B. For example, look at this picture. Now listen to these two sentences: A The circle is zǐ sè. B The the circle is zǐ sè. Which is the correct sentence, Sentence A or Sentence B? (Pause and allow the participants an opportunity to answer. If the participants answer A, commend them and continue with the test. If the participants answer B, pause and state the following: “Let’s listen to the two sentences again: (play the two sentences). If you notice, zǐ sè in sentence B is actually Tone 1 and not Tone 3; we need the word for purple which is zǐ sè–Tone 3.” Reaffirm the participants and continue with the test.)

Administer Experiment 4 Part B.

Once Experiment 4 Part B is complete, please collect the cards.
Experiment 4 Card Given to Participants for Test A

**TWO Mandarin Chinese Color Adjectives**

<table>
<thead>
<tr>
<th>Word</th>
<th>Color</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hēi-se¹</td>
<td>black</td>
</tr>
<tr>
<td>Hóng-se²</td>
<td>red</td>
</tr>
</tbody>
</table>

Experiment 4 Card Given to Participants for Test B

**TWO Mandarin Chinese Color Adjectives**

<table>
<thead>
<tr>
<th>Word</th>
<th>Color</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zǐ -se³</td>
<td>purple</td>
</tr>
<tr>
<td>Lǜ-se⁴</td>
<td>green</td>
</tr>
</tbody>
</table>
APPENDIX 12

Stimuli for Experiment 4 TEST A

The star is _____  The square is _____

The circle is _____  The triangle is _____
The circle is _____

The triangle is _____

The square is _____

The arrow is _____
Stimuli for Experiment 4 TEST B

The square is _____  The star is _____

The circle is _____  The arrow is _____
The square is _____

The circle is _____

The triangle is _____

The arrow is _____
REFERENCES


Language Awareness in Speech Perception and Production: A Festshrift in honour of James E. Flege (pp. 57-77) Amsterdam: John Benjamin.


Graduate School of Information Science and Technology, the University of Tokyo, Japan. Accessed October 29, 2011. Web.


