The role of musical training and linguistic experience in non-native lexical tone processing

Tan, Shen Hui

2018

Tan, S. H. (2018). The role of musical training and linguistic experience in non-native lexical tone processing. Master's thesis, Nanyang Technological University, Singapore.

<https://hdl.handle.net/10356/103512>

<https://doi.org/10.32657/10220/47388>

**NANYANG
TECHNOLOGICAL
UNIVERSITY**

**THE ROLE OF MUSICAL TRAINING AND
LINGUISTIC EXPERIENCE IN NON-NATIVE
LEXICAL TONE PROCESSING**

**TAN SHEN HUI
SCHOOL OF HUMANITIES**

2018

The Role of Musical Training and Linguistic Experience in Non-Native Lexical Tone Processing

Tan Shen Hui

School of Humanities

A thesis submitted to the Nanyang Technological
University in partial fulfilment of the requirement for the
degree of Master of Arts

2018

Acknowledgements

I would like to express my gratitude and appreciation for the following people who have been an indelible part of my postgraduate life journey, without whose support and encouragement this thesis would not have been possible.

I would first like to thank my supervisor, Associate Professor Alice Chan Hiu Dan, for being the reason this research came to fruition. Had it not been for her encouragement and support, I would not have had the courage to step on this journey. I would like to express my deepest gratitude to her for her kind guidance and insightful discussions throughout this process, and especially for her patience, care, and faith in me as I encountered seemingly insurmountable setbacks. Thank you for seeing something in me, and for giving me this opportunity to grow and learn more about research.

I would also like to extend my gratitude and appreciation to Professor Denis Burnham for being so kind as to share his experimental stimuli and data with me, not to mention his endless patience, constructive feedback, and helpful discussions especially in the process of data analysis and interpretation. I have greatly benefited from his passion for research and nurturing spirit towards those learning to do research like me.

A special mention goes to Associate Professor Tan Ying Ying for her kind advice, empathy, and moral support during my literature review writing process, which helped me overcome a particularly trying writer's block.

I am also thankful to the administrative staff of the Research and Graduate Education Office for their kind help with disseminating my flyers when I required Thai participants. In particular, I am grateful to Ms Ann-Marie Chua Mei Fen for her

additional help with the proofreading of my Thai flyers, and for so graciously lending her voice for the recording of my auditory Thai stimuli.

To my fellow lab members of the Neurolinguistics and Cognitive Science Lab, with whom I have made many treasured memories, Lau Fun, Kastoori Kalaivanan, Samantha Lim, Galston Wong, Joey Ong, Yvonne Lai, Wendy Toh, and Cecilia Zhao: A big thank you for all your kindness, support, and friendship. I am particularly grateful to Lau Fun for being my sounding board, and I am indebted to her for her insightful comments and suggestions throughout my research process, especially in the difficult times. I would also like to specially thank Kastoori and Samantha for the generous support I received from them when recruiting participants and running experiments, Joey and Cecilia for their help with statistics, and Galston, our lab manager, for his assistance with the lab equipment.

Special thanks also to my teaching partner, Christina Low, who encouraged me through my writer's block by celebrating my every little success, and to my teacher friend, Ng Hui Qi, who generously volunteered to proofread my literature review.

Last but not least, I would like to thank my parents and my god-aunt, Fong Cheng Yee, for their love and moral support throughout my postgraduate journey.

Table of Contents

Acknowledgements.....	i
Table of Contents	iii
List of Tables	vi
List of Figures	vii
Summary	x
CHAPTER 1: INTRODUCTION.....	1
CHAPTER 2: BACKGROUND	8
2.1 Pitch and lexical tone	8
2.2 Pitch processing mechanisms	9
2.3 Musical training and pitch processing	11
2.4 Tone language experience and pitch processing.....	13
2.5 Disentangling the effects of musical training and tone language experience	17
2.6 Effects of musical training and tone language experience on pitch production	22
2.7 Summary of research gaps and rationale	23
2.8 Research aims and objectives	25
CHAPTER 3: THE PRESENT RESEARCH.....	28
3.1 Overview and research questions.....	28
3.2 Research approach	29
3.3 Hypotheses and predictions	30
CHAPTER 4: METHOD	35
4.1 General procedure	35
4.2 Participants.....	35
4.2.1 Recruitment criteria	35
4.2.2 Participant distribution.....	36
4.2.3 Language background	37
4.2.4 Musical background.....	38
4.2.4.1 Classification Criteria	38
4.2.4.2 Sample Characteristics.....	39
4.3 Supplementary tasks	41
4.3.1 Musical ability test.....	41
4.3.1.1 Stimuli.....	41

4.3.1.2 Apparatus and procedure	42
4.3.1.3 Results and conclusion.....	43
4.3.2 Absolute pitch test.....	46
4.3.2.1 Stimuli.....	47
4.3.2.2 Apparatus and procedure	48
4.3.2.3 Data analysis	50
4.3.2.4 Results and conclusion.....	51
CHAPTER 5: EXPERIMENT 1.....	55
5.1 Overview.....	55
5.2 Stimuli.....	57
5.3 Apparatus and procedure	59
5.4 Data analysis	61
5.5 Results.....	63
5.5.1 d-prime data	63
5.5.1.1 Planned contrasts: Linguistic tones vs. non-linguistic tones	66
5.5.1.2 Planned contrasts: Clear speech vs. filtered speech.....	68
5.5.2 Reaction time data.....	69
5.6 Discussion	70
CHAPTER 6: SECONDARY DATA ANALYSIS.....	77
6.1 Overview.....	77
6.2 Data analysis	78
6.3 Results.....	78
6.3.1 Follow-up analyses	81
6.4 Discussion	83
CHAPTER 7: EXPERIMENT 2.....	87
7.1 Overview.....	87
7.2 Tone production task.....	88
7.2.1 Stimuli.....	88
7.2.2 Apparatus and procedure	88
7.3 Tone evaluation tasks.....	89
7.3.1 Informants	89
7.3.2 Stimuli and apparatus.....	89
7.3.3 Procedure	89
7.4 Data analysis and results.....	92
7.4.1 Tone identification task.....	92

7.4.1.1 Accuracy data.....	92
7.4.1.2 Confusion matrices	94
7.4.2 Tone rating task.....	101
7.5 Discussion	102
CHAPTER 8: GENERAL DISCUSSION	107
8.1 Summary of main results and findings	107
8.2 Influence of task difficulty and sensitivity of analysis used	108
8.3 The double-edged sword of tone language experience.....	111
8.4 The cross-domain interaction effect of musical training with linguistic experience	114
8.5 Limitations and future areas of research	115
8.5.1 Related to the scope	115
8.5.2 Related to the tasks	116
8.5.3 Related to the sample	117
8.5.4 Related to the analyses	119
8.5.5 Related to the findings	120
CHAPTER 9: CONCLUSION.....	122
REFERENCES.....	124
APPENDIX A: Laboratory questionnaire	133
APPENDIX B: Screening questionnaire	135
APPENDIX C: Individual musician backgrounds.....	139
APPENDIX D: Absolute pitch test stimuli.....	141
APPENDIX E: Thai short story for speech and filtered speech familiarisation contexts	142

List of Tables

Table 1	General Experimental Design	28
Table 2	Distribution of Participants	37
Table 3	Characteristics of Tone (Tm) and Non-Tone Musicians (NTm)	39
Table 4	Comparison of the Musical Ear Test (MET) scores of musicians and non-musicians in the current study and in Experiment 3 of the study by Wallentin et al. (2010)	45
Table 5	Thai Speech Stimuli with IPA Transcription and Gloss	58
Table 6	Mean percentage discrimination accuracies of tone musicians (Tm), tone non-musicians (Tnm), non-tone musicians (NTm), and non-tone non-musicians (NTnm) for each stimulus type, with the respective standard deviations in parentheses	64
Table 7	Characteristics of non-absolute-pitch musicians in the current study and in the study by Burnham et al. (2015)	72
Table 8	Confusion matrices (in percentages) of Thai tones (mid 32 (M), low 21 (L), falling 51 (F), high 45 (H), rising 214 (R)) produced by (a) 18 tone musicians (Tm), (b) 19 tone non-musicians (Tnm), (c) 14 non-tone musicians (NTm), and (d) 19 non-tone non-musicians (NTnm), as judged by 17 native Thai listeners	97
Table 9	Summary of tone confusion patterns across tone musicians (Tm), tone non-musicians (Tnm), non-tone musicians (NTm), and non-tone non-musicians (NTnm) for each of the Thai lexical tone categories (mid (32), low (21), falling (51), high (45), rising (214))	100

List of Figures

- Figure 1.* A comparison of (a) Mandarin Chinese (M) tones and (b) Thai (T) tones produced by two native speakers of each gender, for each language (Wu, Munro, & Wang, 2014). T values were used to normalize for differences in pitch range between male and female speakers and ranged from 1 (lowest pitch of the speaker) to 5 (highest pitch of the speaker), as in Chao's (1948) notation scale. The abbreviations, HL (high level), R (rising), FR (falling rising), HF (high falling), ML (mid low), LF (low falling) describe the characteristics of the produced tones in each language. (Note: In the present study, the Thai tones are referred to in the traditional way, such that T-ML is labelled as the mid tone, T-LF as the low tone, T-HF as the falling tone, T-R as the high tone, and T-FR as the rising tone.).....32
- Figure 2.* Distribution of musician types. This figure presents the types of instruments played by tone musicians (Tm) ($n = 18$) and non-tone musicians (NTm) ($n = 14$), and the percentage of musicians ($N = 32$) who play them.40
- Figure 3.* Example trials from the Musical Ear Test (MET) (Wallentin et al., 2010). This figure illustrates “same” and “different” trials in each MET subtest (Melody, Rhythm), with differences between musical phrases in the “different” trials circled.42
- Figure 4.* Musicians ($n = 32$) and non-musicians' ($n = 38$) mean accuracy scores in the Musical Ear Test (MET), with standard error bars. Graph (a) shows participants' combined scores for both subtests. Graph (b) shows participants' scores in each subtest (Melody, Rhythm). (Note: $*p < .05$. $***p < .001$)44
- Figure 5.* Response screens (modified after those used by Bermudez and Zatorre (2009)) presented to the participants during the AP test. Panel (a) shows the note response screen and physical response key arrangement during the note name identification segment of each trial. Panel (b) shows the octave response screen during the octave identification segment of each trial, with each band of colour representing the note range of C to B within each octave.49
- Figure 6.* How mean absolute deviation (MAD) was computed. This figure presents two examples of how absolute deviation, as opposed to actual semitone deviation, is calculated. For visualization purposes, a piano keyboard with the notes F3, C4, F4, and C5 marked is included.50

<i>Figure 7.</i>	Results of AP test. Scatterplot (a) shows participants' mean percentage accuracy scores plotted against their mean absolute deviation (MAD) scores. Scatterplot (b) shows participants' mean reaction times plotted against their MAD scores. Scores of a known AP possessor on the AP test were included for the purposes of reference and comparison. Participants found to be outliers on either one of the accuracy measures based on boxplot analysis were labelled with their identification number.	53
<i>Figure 8.</i>	Boxplots of AP test scores. Boxplot (a) shows the spread of scores measured in percentage accuracy. Boxplot (b) shows the spread of scores measured in mean absolute deviation (MAD). Outliers with values between one and a half to three times the interquartile (IQ) range, i.e., beyond the whiskers are indicated with a circle. Extreme outliers, i.e., cases with values more than three times the IQ range are indicated with a star.	54
<i>Figure 9.</i>	Predicted patterns of non-native Thai lexical tone discrimination for tone musicians (Tm), tone non-musicians (Tnm), non-tone musicians (NTm), and non-tone non-musicians (NTnm) across the three stimulus types (speech, filtered speech, violin). The higher the score on the discrimination index, the better the perceptual discrimination.	56
<i>Figure 10.</i>	Mean d-prime scores of 18 tone musicians (Tm), 19 tone non-musicians (Tnm), 14 non-tone musicians (NTm), and 19 non-tone non-musicians (NTnm) for each of the three stimulus types (speech, filtered speech, violin), with standard error bars.	63
<i>Figure 11.</i>	Mean d-prime scores of musicians and non-musicians for the linguistic stimulus types (speech, filtered speech) and the non-linguistic (violin) stimulus type, with standard error bars.	67
<i>Figure 12.</i>	Mean reaction times (in milliseconds) of 18 tone musicians (Tm), 19 tone non-musicians (Tnm), 14 non-tone musicians (NTm), and 19 non-tone non-musicians (NTnm) for each of the three stimulus types (speech, filtered speech, violin), with standard error bars. Only reaction times of correct responses to "different" trials within ± 2 standard deviations of the overall mean are shown.	69
<i>Figure 13.</i>	Mean d-prime scores of 14 tone exposed non-tone musicians (NT ⁺ m), 12 tone unexposed non-tone musicians (NT ⁻ m), 19 tone exposed non-tone non-musicians (NT ⁺ nm), and 13 tone unexposed non-tone non-musicians (NT ⁻ nm), for each of the three stimulus types (speech, filtered speech, violin), with standard error bars.	79

- Figure 14.* Thai tone production accuracy of 18 tone musicians (Tm), 19 tone non-musicians (Tnm), 14 non-tone musicians (NTm), and 19 non-tone non-musicians (NTnm), based on the percentage of correct tone identification by 17 native Thai speakers. This figure illustrates the extent to which non-native Thai lexical tones (mid 32 (M), low 21 (L), falling 51 (F), high 45 (H), rising 214 (R)) produced by each participant group were correctly identified by native Thai listeners.93
- Figure 15.* Tone confusion rates of participants' Thai tone tokens (mid 32 (M), low 21 (L), falling 51 (F), high 45 (H), rising 214 (R)), as judged by 17 native Thai listeners. Graph (a) shows the tone confusion rates for tones produced by 18 tone musicians (Tm). Graph (b) shows the tone confusion rates for tones produced by 19 tone non-musicians (Tnm). Data labels are included for the percentages of tones that were correctly identified, as well as for error rates that were found to exceed the error threshold of 4.63%.98
- Figure 16.* Rating scores of Thai tone tokens (mid 32 (M), low 21 (L), falling 51 (F), high 45 (H), rising 214 (R)) produced by 18 tone musicians (Tm), 19 tone non-musicians (Tnm), 14 non-tone musicians (NTm), and 19 non-tone non-musicians (NTnm), as judged by 17 native Thai speakers. Rating scores are on a scale of 1 (very poor example of the intended tone category) to 7 (very good example of the intended tone category).101

Summary

Musical training has been found to enhance the perception of non-native lexical tones, however, not much is known about how it may interact with other contextual factors such as linguistic experience to affect non-native lexical tone processing, and whether it may also enhance the production of non-native lexical tones. To better qualify the extent of its effect on non-native lexical tone processing, the independent and joint effects of musical training and linguistic experience on the perception and production of non-native lexical tones are explored in the present study.

Four groups of participants were tested: Tone-language-speaking musicians (Tm), tone-language-speaking non-musicians (Tnm), non-tone-language-speaking musicians (NTm), and non-tone-language-speaking non-musicians (NTnm). In Experiment 1, participants discriminated non-native Thai lexical tones in three forms: normal Thai speech, low-pass filtered Thai speech, and violin analogues of the Thai speech. In Experiment 2, participants imitated Thai lexical tones, and their tone productions were evaluated via a tone identification task and a tone rating task by a separate group of native Thai informants.

Results revealed no statistically significant independent or combined effects of musical training or language experience in both experiments. Across groups, participants generally had comparably good perception and production of the non-native Thai tones. As this was contrary to past findings, further investigations were conducted. Secondary data analysis using data from the study by Burnham et al. (2015) indicated that the lack of group differences in the perception task could not be attributed to enhancements in non-native lexical tone perception among non-tone participants as a result of long-term ambient exposure to tone languages. Meanwhile, further analysis

conducted on the production task results using confusion matrices revealed that tone language speakers produced more confusable tones, especially if they were also non-musicians. This suggests an interaction effect of musical training with tone language experience in non-native lexical tone production, whereby the former mitigates disadvantages caused by the latter. Given this finding, the statistically non-significant results of both experiments were most probably due to insufficient task difficulty, as well as a lack of sensitivity of accuracy rates as a measure of participants' performance.

Overall, although the current study did not replicate past findings of a facilitatory effect of musical training on non-native lexical tone perception, it shed some light on the conditions for this effect to be manifested instead. Notably, it presents novel findings of an interaction between musical training and tone language experience during non-native lexical tone production which demonstrates that the benefits of musical training to non-native lexical tone processing not only extend beyond perception to production but can also be experienced by tone language speakers. The current study thus contributes to our understanding of when and how musical training may interact with linguistic experience to affect non-native lexical tone processing, which has practical implications on the second language (L2) acquisition of tone languages.

Keywords: musical training, tone language experience, non-native lexical tone processing, second language acquisition, pitch processing

CHAPTER 1: INTRODUCTION

One fascinating question in the study of human cognition is whether skills in one cognitive domain have transferable effects in another, be it in a facilitatory or inhibitory sense. On one hand, theories of domain-specific learning posit that the process of knowledge acquisition involves specialized mechanisms, such that training in one domain does not influence performance in another. On the other hand, theories of domain-general learning hold the view that there are overlaps in processing mechanisms across cognitive domains, which enable training in one domain to enhance performance in another. A clear understanding of whether and when certain skills are domain-general or domain-specific thus has significant theoretical implications on our understanding of how we acquire knowledge, not to mention practical implications on how we devise teaching and intervention programmes.

The focus of this thesis is on one general cognitive domain in particular – pitch processing, and the investigation of transfer effects that may occur within it. The ability to process pitch plays an important role in our lives: In all languages, pitch is used to convey a variety of pragmatic functions, for instance, to differentiate statements from questions, express feelings and attitudes, and for emphasis. In a subset of languages known as tone languages, pitch has an even more integral role as it takes on the function of a lexical tone, where pitch variations occur at a lexical level and fundamentally affect word meaning. Last but not least, pitch is the building block of melody and harmony, and is one of the main contributing factors to our enjoyment of music.

With respect to these different aspects of pitches, a question that has naturally arisen is whether they are processed differently. In response to this, past research has shown that linguistic pitch is processed differently from non-linguistic pitch (Van

Lancker & Fromkin, 1973, 1978; Wang, Jongman, & Sereno, 2001), and the same for native as opposed to non-native pitch (Gandour et al, 2002). This points to a certain extent of domain-specificity in pitch processing. Interestingly, evidence pointing to domain-generalty in pitch processing has also been found. Specifically, research suggests an overlap between the processing of musical pitch and that of linguistic pitch. Even though music and language are conceived of as distinct cognitive domains, musicians were found to have enhanced neural representation and perceptual discrimination of non-native lexical tones (Alexander, Wong, & Bradlow, 2005; Wong, Skoe, Russo, Dees, & Kraus, 2007), which was attributable to enhanced musical pitch ability from their years of musical training (Delogu, Lampis, & Belardinelli, 2010).

In this thesis, we narrow our focus to this particular aspect of pitch processing: its overlap in the processing of musical pitch and lexical tone, in the form of the aforementioned cross-domain transfer benefit of musician training on non-native lexical tone processing. The study of this cross-domain transfer effect has significant implications in the second language (L2) acquisition of tone languages; with such evidence of commonalities in the processing of musical pitch and lexical tones, the integration of musical, pitch-related interventions may be a viable option to enhance non-native lexical tone learning. In L2 acquisition, however, many other factors come into play for the mastery of an L2, e.g., one's linguistic background, levels of self-motivation, acquisition settings, and so on. The central question is thus: To what extent does musical training influence the processing of non-native lexical tones?

In relation to this question, one research gap with regards to this cross-domain transfer benefit of musical training on non-native lexical tone processing is that these findings have been largely based upon investigations of musicians and non-musicians with non-tone language backgrounds. Therefore, even though tone languages, in fact,

form the majority of the world's languages (Yip, 2002), whether musicians who are native speakers of a tone language also experience enhanced perception of non-native lexical tones due to transfer effects of musicianship is not well studied. This is problematic, firstly, in terms of sample representation. Without a more inclusive sample, which is to be achieved by also comparing musicians and non-musicians with tone language backgrounds, the argument for a robust effect of musical training on lexical tone perception, and of there being a clear overlap between musical pitch and lexical tone processing, is weakened.

To make things more complicated, studies in the field of L2 acquisition have established that the perception of L2 sounds is hugely influenced by prior linguistic experience (Best, 1995; Flege, 1995). Thus, depending on the similarities between their L1 and L2 tonal inventories, native tone language speakers can either benefit or be hindered by their tone language experience when attempting to discriminate non-native lexical tones. In effect, this means that the manifestation of the cross-domain transfer effect of musical training on non-native lexical tone processing may not be straightforward because it may be affected by prior experience with tone languages as another factor. For instance, if having native experience in a particular tone language enhances the perception of lexical tones in a particular non-native tone language, musical training may add on to this effect, resulting in an even greater perceptual enhancement. On the other hand, if that native tone language experience is hindering the perception of the non-native lexical tones instead, musical training may compensate for it, reducing the hinderance to perceptual discrimination.

Despite this, most studies have investigated the effect of musical training on non-native lexical tone processing independently of tone language experience as a factor. As a result, the extent of influence of musical training on non-native lexical tone processing

relative to that of tone language experience is unclear. For instance, we do not know whether the two factors interact, such that one factor would have more influence on non-native lexical tone processing than the other. In other words, would it be more advantageous to be a tone-language-speaking non-musician or a non-tone-language-speaking musician when processing non-native lexical tones? Another question of interest is whether the effects of both factors on non-native lexical tone processing are additive: Would tone-language-speaking musicians, i.e., those with both musical training and tone language experience, have an even greater advantage compared to tone-language-speaking non-musicians and non-tone-language-speaking musicians, i.e., those with only one of the two experiences?

Answering these questions is important in order to refine our understanding of the relationship between musical pitch and non-native lexical tone processing by contextualizing it against the ubiquitous backdrop of prior linguistic experiences. This is especially so considering the importance of tone language experience as a factor which can either enhance or hinder non-native lexical tone processing. Through an investigation which takes into account the possible interactions between musical training and tone language experience on non-native lexical tone processing, there can be practical implications on the design of strategies to aid non-native lexical tone learning, depending on the kind of interactions that we find.

In addition to the lack of integration of tone-language-speaking populations as well as tone language experience as a factor in studies investigating the effect of musical training on non-native lexical tone processing, there is another research gap. Most studies which have arrived at the conclusion that musical training can enhance non-native lexical tone processing have mainly demonstrated this at the perceptual level. This is also the case among studies indicating an effect of tone language experience on

non-native lexical tone processing. Our understanding of the cross-domain transfer of musical pitch to non-native lexical tone processing is thus limited to perception; little is known about whether musical training also enhances the production of non-native lexical tones, much less whether it interacts with tone language experience in this context. However, there is a general consensus that perception generally precedes production in speech development (Werker & Tees, 1984a), and there is also some evidence that enhanced perception abilities can lead to enhanced production abilities (Wang, Jongman, & Sereno, 2003). In order to qualify the extent of influence of musical training on non-native lexical tone processing, and therefore broaden our understanding of the overlaps between musical pitch and non-native lexical tone processing, further investigations of whether the beneficial transfer effects of musical training (and tone language experience) on non-native lexical tone perception also benefits non-native lexical tone production would be meaningful. In turn, this would have further practical implications in the L2 acquisition of tone languages, given that effective communication relies not only on good perception abilities but on good production abilities as well.

A more complete understanding of the cross-domain transfer effects of musical training on non-native lexical tone processing in terms of its extent of influence would thus warrant the following: (1) the inclusion of a more representative sample of musicians and non-musicians, (2) a simultaneous investigation of its independent and combined effects with tone language experience as an additional factor, and (3) an investigation of these effects from the aspects of both perception and production.

This thesis thus attempts to address these research gaps through a comparison of musicians and non-musicians of both tone-language-speaking and non-tone-language-speaking backgrounds in their perception and production of non-native lexical tones. By including tone-language-speaking musicians and non-musicians, a more representative

sample is achieved. Subsequently, a stronger argument can be made for the effect of musical training on non-native lexical tone processing, should tone-language-speaking musicians also prove to have enhanced non-native lexical tone processing compared to tone-language-speaking non-musicians. By comparing the performance of tone-language-speaking musicians, tone-language-speaking non-musicians and non-tone-language-speaking musicians, we would be able to determine which experience, musical training or native tone language experience, is more crucial in affecting non-native lexical tone processing, as well as whether the two experiences are additive. Finally, by including a production task, we can observe whether the two experiences have further transfer effects on the processing of non-native lexical tones beyond the perceptual level.

In short, the central aim of this thesis is to investigate the independent and combined effects of musical training and tone language experience on the perception and production of non-native lexical tones, so the benefits of musical training to non-native lexical tone processing can be better contextualized. With this in mind, the following research questions are asked:

- (1) Do the perceptual benefits of musical training apply to tone language speakers?
- (2) Between musical training and tone language experience, does one bear more weight in non-native lexical tone processing?
- (3) Are the effects of musical training and tone language experience additive?
- (4) In addition to the perceptual benefits, does musical training and tone language experience also benefit the production of non-native lexical tones?

In the next chapter, we review past literature on pitch processing, ranging from evidence of its domain-generalness and its domain-specificity, to seminal work as well as more recent studies on how it can be affected by musical training and tone language experience. Subsequently, the research aim and objectives derived based on the literature review are detailed. Following that, Chapter 3 lays out the research questions, approach, hypotheses, and predictions of the study. In Chapter 4, the overall methodology is presented, along with the methodology and results of two supplementary tasks carried out to control for the confounds of musical aptitude and absolute pitch ability. The perception task conducted in Experiment 1 is reported in Chapter 5, and this is followed by an additional, secondary data analysis which is presented in Chapter 6. In Chapter 7, the production task conducted in Experiment 2 is reported. A general discussion of the findings from Experiment 1, the secondary data analysis, and Experiment 2 is then presented in Chapter 8. The limitations of the current work as well as some future areas of research are also raised in this chapter. Finally, Chapter 9 presents the conclusion of this thesis, with a synthesis of key points and significance of this work.

CHAPTER 2: BACKGROUND

2.1 Pitch and lexical tone

Pitch, the psychological correlate of the fundamental frequency (F0), is a critical perceptual dimension in the domains of music and language. In music, an awareness of and ability to generate abstract representations of pitch changes enable us to recognize, enjoy, and reproduce melodies and harmonies. In language, pitch is used to convey a variety of cues, which can be social (e.g., biological sex of a speaker), emotional (e.g., state of mind of a speaker), intonational (e.g., whether an utterance is a question, a statement, or an order), or lexical (e.g., to distinguish word meaning) in nature.

In a subset of languages known as tone languages, pitch is especially crucial as a perceptual cue to distinguish lexical meaning (Yip, 2002). In tone languages, pitch differences in aspects such as height, direction, and contour are the basis for the forming of stable phonological categories known as lexical tones, and it is these pitch-based phonological categories that subsequently play an important role in distinguishing word meaning. For example, in Mandarin Chinese, the same syllable /ma/ carrying the four lexical tones, high level (55), high rising (35), low falling rising (214), and high falling (51) denote four different meanings respectively: /ma⁵⁵/ ‘mother’, /ma³⁵/ ‘hemp’, /ma²¹⁴/ ‘horse’, and /ma⁵¹/ ‘scold’ (Ladefoged & Johnson, 2015).¹ The role of pitch in tone

¹ The numbers in superscript are a description of the tones based on the Chao (1948) notation system. They represent the relative pitch onset and offset each tone on a scale of 1 (lowest pitch of the speaker) to 5 (highest pitch of the speaker).

languages is thus distinct from that in non-tone languages, as pitch differences do not merely occur at the phrasal level but are also encoded lexically to convey meaning.

2.2 Pitch processing mechanisms

As with the study of other aspects of cognition, researchers in the field of pitch processing have been interested in finding out whether pitch processing involves more domain-specific or domain-general mechanisms. In recent years, research using neuroimaging techniques has suggested the existence of brain regions specialized for pitch processing, namely the lateral Heschl's gyrus and the anterolateral planum temporale (Hall & Plack, 2008; Plack, Barker & Hall, 2014). Studies on dichotic listening further point to a hemispheric specialization for the processing of different aspects of pitch: When listening to lexical tones, a right-ear advantage was found for tone language listeners, but not for non-tone language listeners (Van Lancker & Fromkin, 1973, 1978; Wang et al., 2001), suggesting that aspects of pitch which are linguistically relevant are processed in the left hemisphere. This was further confirmed by results showing that when the same lexical tones were presented in hummed versions, the right-ear advantage observed in tone language listeners disappeared. In addition, processing mechanisms for linguistic pitch may also differ depending on whether the pitch is native or non-native: Despite both groups being tone language speakers, Thai listeners but not Mandarin Chinese listeners showed a left hemisphere dominance when processing Thai lexical tones, suggesting that native pitch engages the left hemisphere to a greater degree than non-native pitch (Gandour et al, 2002). Thus, some level of domain-specificity appears to be at play in pitch processing, whereby the context and interpretation of a pitch signal (i.e., linguistically relevant vs. irrelevant, native vs. non-native) affect how it is processed.

On the other hand, evidence for the shared processing of linguistic and non-linguistic pitch has been offered through studies on congenital amusia, also commonly referred to as “tone deafness”. This is mainly based on the finding that amusics not only face difficulties in the perception of musical pitch but also have trouble discriminating linguistic pitch, even if they are native speakers of tone languages, where discrimination of pitch is crucial for phonological categorization. For example, relative to non-amusic controls, native Mandarin Chinese amusics have been found to be impaired not only in the discrimination and identification of melodic contours but also in the identification of Mandarin Chinese speech intonation (Jiang et al., 2010). This is similar to amusics of non-tone languages such as British English and Canadian French, who also faced difficulty with intonation processing (Liu, Patel, Fourcin, and Steward, 2010; Patel, Wong, Foxton, Lochy, and Peretz, 2008). Furthermore, in addition to impairments in the perception and production of musical pitch, another study found that native Cantonese amusics also showed impairment in the perception of Cantonese tones when compared to their non-amusic counterparts (Liu et al., 2016).

Besides studies on congenital amusia, the idea of domain-generalty in pitch processing has also been supported by studies showing potential bidirectionality in the effects of musical pitch expertise and linguistic pitch expertise. For example, the ability to distinguish musical pitch has been observed to correlate consistently with the ability to distinguish linguistic pitch at a sentence level even after taking into account basic pitch and auditory discrimination skills, as well as individual differences in attention and motivation (Perrachione, Fedorenko, Vinke, Gibson, & Dilley, 2013). Similarly, long-term expertise with linguistic pitch through native experience with a tone language appears to have transfer benefits to the processing of musical pitch: Native young adult tone language speakers of Mandarin Chinese, Cantonese, and Vietnamese were found,

as a group, to be more accurate at imitating and discriminating musical pitch intervals compared to their American English counterparts (Pfordresher & Brown, 2009). Moreover, this cross-domain language-to-music transfer benefit is also observed among children as young as under four years old: By comparing Mandarin Chinese children with American English children, Creel, Weng, Fu, Heyman, and Lee (2018) found the former to show enhanced musical pitch discrimination skills than the latter, even at an age prior to literacy instruction (which can potentially facilitate awareness of lexical tone, based on the findings of Burnham et al. (2011) on native Thai children and young adults)) and after taking into account their prior musical experiences.

2.3 Musical training and pitch processing

Overall, the literature on pitch processing shows that pitch processing can be domain-specific in some ways, and domain-general in others. In the quest to further understand how and when pitch processing mechanisms may be domain-specific or domain-general, recent decades have seen a growing interest in the study of music processing, particularly on how training in the musical domain may benefit pitch processing in the linguistic domain. Given the salience of pitch in music, if pitch was processed in a domain-general manner, then expertise in musical pitch from musical training would have cross-domain transfer effects on linguistic pitch processing. This would have practical implications on interventions designed to aid linguistic pitch processing in contexts such as language acquisition or even amusia.

Accordingly, one of the seminal findings from this line of inquiry has been a musician advantage for the processing of non-native lexical tones. Behaviourally, musicians showed higher accuracies and faster reaction times in discriminating and identifying Mandarin Chinese lexical tones (Alexander et al., 2005; Lee & Hung, 2008),

and also tended to be more successful in learning to use non-native Mandarin Chinese-like lexical pitch patterns for word-meaning mappings (Wong & Perrachione, 2007). Furthermore, when presented with non-native Thai lexical tones in different contexts on a speech to non-speech continuum, musicians showed more comparable discrimination abilities across the contexts compared to non-musicians, indicating greater commonality in their linguistic and non-linguistic pitch processing (Burnham, Brooker, & Reid, 2015). From a neural perspective, musicians have also been found to show greater mismatch negativity responses than non-musicians in response to non-native Mandarin Chinese tone homologues in a passive oddball paradigm, which indicates greater sensitivity to changes in tone at the cortical level (Chandrasekaran, Krishnan, & Jackson, 2009). At the subcortical level, musicians also have more robust and faithful brainstem encoding of non-native Mandarin Chinese lexical tones compared to non-musicians (Bidelman, Gandour, & Krishnan, 2011a; Wong et al., 2007). These advantages of musicianship can most likely be attributed to enhanced musical abilities as a result of musical training, which has, in turn, led to enhanced pitch discrimination skills transferable to the domain of language. Evidence for this comes from a study by Delogu et al. (2010), who found Italian adults and children with better musical pitch discrimination abilities to be better at discriminating tonal variations in Mandarin Chinese monosyllabic sequences, regardless of whether they were musicians or not.

However, in most studies on the effect of musical training on non-native lexical tone processing, testing has only been conducted on non-tone language speakers, which limits the generalizability of the effect of musicianship. For instance, all of the abovementioned studies only compared non-tone-language-speaking musicians and non-musicians, and not musicians and non-musicians who were native tone language speakers. Taking the latter comparison into consideration would yield a more

comprehensive and representative picture of the effect of musical training on non-native lexical tone processing, not least because tone languages form 70% of the world's languages (Yip, 2002). In other words, it would not be unreasonable to expect native tone language speakers to be acquiring and learning other non-native tone languages in their lifetime. There are thus theoretical and practical implications for studying the effect of musical training on tone language speakers as well when it comes to non-native lexical tone processing. Specifically, it would help shed light on whether musical training also benefits the non-native lexical tone processing of native tone language speakers (i.e., by comparing musician tone language speakers and non-musician tone language speakers), which would better qualify the effects of musical training.

2.4 Tone language experience and pitch processing

Including tone language speakers as sample populations in the study of how musical training may affect non-native lexical tone processing is also crucial for another reason: to study whether this transfer effect of musical training on linguistic pitch processing is further influenced by other contextual factors such as linguistic experience itself, specifically tone language experience. This is important to refine our theoretical understanding of the effect of musical training in this context of pitch processing (i.e., so that we can find out whether this effect is not only independent but also interactive with linguistic experience), and subsequently inform interventions designed to aid the acquisition of non-native tone languages.

Based on past literature, linguistic experience is a defining factor in the acquisition of non-native languages. Once perceptual reorganization has taken place in the first year of life, discrimination of non-native contrasts, be it consonants, vowels, or even tones begins to deteriorate (though not to the point of it being lost) (Kuhl, Williams,

Lacerda, Stevens, & Lindblom, 1992; Mattock & Burnham, 2006; Polka & Werker, 1994; Werker, Gilbert, Humphrey, & Tees, 1981; Werker & Tees, 1983, 1984a), and the categories that have been formed in our native language(s), also known as our first language(s) (L1) become the lens through which non-native sounds are perceived. Models of speech perception in second language (L2) acquisition expound on this phenomenon, positing that phonetic similarity and dissimilarities between sounds in L1 and L2 play an important role in determining whether two L2 sounds are assimilated into one or two L1 categories (Best, 1995; Flege, 1995). How L2 sounds are assimilated into L1 categories then determine whether or not an L2 learner is able to perceptually discriminate the two L2 sounds: If two L2 sounds are assimilated into two L1 categories, discrimination is easier than if the two L2 sounds are assimilated into a single L1 category.

Crucially, studies conducted on native tone language speakers show tone language experience to be an important factor affecting non-native lexical tone processing. Behaviourally and neurally, speakers of tone languages appear to process lexical pitch of non-native languages differently from speakers of non-tone languages. In some studies, it was found that tone language speakers had enhanced discrimination of non-native lexical tones compared to non-tone language speakers. For example, Mandarin Chinese and Taiwanese speakers outperformed English speakers in the discrimination of non-native Thai tones (mid vs. low tone) both pre- and post-training, and in the identification of the tones post-training (Wayland & Guion, 2004). Similarly, Burnham et al. (1996) found Cantonese speakers to outperform English speakers in the discrimination of non-native Thai tones across different contexts (speech, filtered speech, and music): Cantonese speakers had equivalent discrimination abilities across the three contexts, but English speakers had attenuated discrimination abilities for the

non-native Thai tones in the speech context. Aside from behavioural findings, this advantage of tone language experience is also apparent at the subcortical level of brainstem encoding: Mandarin Chinese and Thai speakers had superior pitch-tracking accuracies in response to non-native Thai and Mandarin Chinese tones (a set of mid, rising, and falling tones per language) respectively compared to English speakers (Krishnan, Gandour, & Bidelman, 2010). These suggest that long-term experience with lexical pitch results in attunement to F0 differences, and that this familiarity with the use of pitch contrastively facilitates the same process in a non-native context.

On the other hand, some studies did not observe such facilitatory benefits of tone language experience. For example, Mandarin Chinese speakers in Francis, Ciocca, Ma, and Fenn's (2008) study did not significantly differ from English speakers in the identification of non-native Cantonese tones both pre- and post-training. Similarly, Cantonese learners did not outperform English learners in the identification and mimicry of Mandarin Chinese tones in Hao's (2012) study. In the face of these mixed results, research suggests that the effect of tone language experience is also contingent on other factors, such as the perceptual similarity between L1 and L2 tones, the relative complexity of L1 and L2 tonal inventories, and cross-linguistic differences in attention to pitch height vs. pitch contour.

As a case in point, Mandarin Chinese speakers in a study by Francis et al. (2008) were found to be best at identifying the following Cantonese tones: high level (55), high rising (25), and mid level (33). This can be explained by the perceptual similarity between these tones and their Mandarin Chinese L1 tone counterparts; the Cantonese high level (55) tone and high rising (25) tone could have been easily mapped to the Mandarin Chinese high level (55) tone and high rising (35) tone respectively, while the Cantonese mid level (33) tone could have been mapped either to the Mandarin Chinese

low falling rising (214) tone or the neutral tone category. Similarly, perceptual similarity between L1 and L2 tones helped to explain why in Experiment 1 of Hao's (2012) study, Cantonese learners did not outperform English learners in Mandarin Chinese tone identification. In Experiment 2 of Hao's (2012) study, it was found that Cantonese speakers frequently assimilated the Mandarin Chinese high level (55) tone and high falling (51) tone to the same Cantonese high level (55) tone category. Based on the Perceptual Assimilation Model (Best, 1995), two L2 sounds which are assimilated into the same L1 category tend to hinder perceptual discrimination, which therefore explains the Mandarin Chinese tone identification performance of the Cantonese learners in Experiment 1.

In other instances, the relative complexity of the L1 and L2 tonal inventories comes into play. For example, while Mandarin Chinese speakers in Lee, Vakoch, and Wurm's (1996) study also did not perform better than English speakers in the discrimination of Cantonese tones, Cantonese speakers in the study performed better than English speakers in the discrimination of Mandarin Chinese tones. The implication is that having tone language experience may benefit the processing of non-native tones when the L1 tonal inventory is more complex than that of the L2; Cantonese has six tones (three contour tones and three level tones) while Mandarin Chinese only has four tones (three contour tones and only one level tone). From another perspective, Gandour (1983) adds that it is not that non-tone language speakers are unable to process the fundamental acoustic dimensions of F0 height and direction of change, but rather that they give different weights to the dimensions compared to tone language speakers. For example, among Mandarin Chinese, Taiwanese, Thai and English speakers, it was found that English speakers paid more attention to the dimension of pitch height than pitch contour compared to most of the other tone language speakers. Thus, depending on the

type of non-native pitch contrasts that are tested, non-tone language speakers like English speakers may still perform well.

Based on existing findings, it is clear that language background has significant implications on the acquisition of a non-native tone language. Given this, it is possible that even with the positive influences of musical training on non-native tone processing, this positive influence may be further facilitated or even inhibited depending on the musician's tone language background. A comparison of musician tone language speakers and musician non-tone language speakers would shed further light on this. However, as previous research has mainly focused on the respective influences of musical training and tone language experience on non-native tone processing separately, it is not clear whether musicians with tone language experience would show an additive effect of having both experiences, and thus outperform musicians without tone language experience and non-musicians with tone language experience in the processing of non-native tones. It is also not clear whether the effects of each experience are comparable, such that musicians without tone language experience would have equivalent performance as non-musicians with tone language experience. There is thus a need to examine whether these two factors of non-native lexical tone processing, namely musical training and tone language experience, have any interactive or additive effects so that we can deepen our understanding of the relative weights of musical training and tone language experience in the context of non-native pitch processing.

2.5 Disentangling the effects of musical training and tone language experience

While there have been attempts in the recent years to disentangle the effects of musical training and tone language experience on pitch processing, so far, the attempts have been few, and the results, generally mixed. In effect, there are three main questions

to be addressed: Firstly, do the transfer benefits of musical training to pitch processing also apply to tone-language-speaking populations? Secondly, are the two factors, musical training and tone language experience, additive? Thirdly, do the two factors have equivalent effects on pitch processing?

With reference to these questions, Dawson, Tervaniemi, and Aalto's (2018) study is one of the few to study the effects of musical training on pitch processing among a tone-speaking population. Through brainstem recordings and listening tasks conducted on adult Mandarin Chinese musicians and non-musicians, the study provided evidence of perceptual benefits of musical expertise among tone language speakers. Behaviourally, Mandarin Chinese musicians had lower discrimination thresholds for the frequency of non-speech stimuli compared to Mandarin Chinese non-musicians, reflecting enhanced pitch discrimination abilities. Similarly, in a longitudinal piano training study on Mandarin Chinese children for six months, Nan et al. (2018) found a musician advantage among tone language speakers even at the young age of four to five. Relative to two control groups (a reading training group and a no training control group), the piano training group had significantly larger positive mismatch responses (pMMRs) post-test in response to musical pitch and Mandarin Chinese tones. Furthermore, the increase in pMMRs in response to musical pitch was significantly correlated to improvements in word discrimination based on consonantal differences only in the piano training group. These results indicate that even within a tone-language-speaking population, musical training can further enhance pitch processing behaviourally and neurally, as well as in speech and non-speech contexts. In sum, it is postulated that there can be an additive effect of the two factors, and that both experiences are not necessarily equivalent.

Nonetheless, evidence to the contrary has also been found in past studies which have included musicians and non-musicians who are tone language speakers. For instance, despite observing a behavioural advantage of musicianship among tone language speakers, Dawson et al. (2018) found no differences between tone-language-speaking musicians and non-musicians at the subcortical level. Both Mandarin Chinese musicians and non-musicians demonstrated comparable pitch tracking of a non-native Finnish pitch contour (which is not lexically contrastive as in Mandarin Chinese, but rather, co-signals long vowels). This suggests that having both experiences does not consistently result in an additive effect. It is possible that equivalent linguistic pitch expertise among tone language speakers led to ceiling effects in this case. Alternatively, it is also possible that the use of non-native speech stimuli was perceived non-linguistically, thereby facilitating the pitch tracking of the non-musicians and resulting in reduced differences in pitch processing abilities between the tone-language-speaking musicians and non-musicians. Thus, it can be inferred that at least under certain conditions, it can be possible for no additive benefits of musical training to be manifested, such that having musical training in addition to tone language experience does not result in greater perceptual enhancement in pitch processing compared to having just tone language experience.

Further evidence against additive effects of musical training on pitch processing comes from Cooper and Wang (2012), which is one of the few studies to compare musicians and non-musicians from both tone-language-speaking (Thai) and non-tone-language-speaking (English) groups. In this study, no behavioural evidence of an additive effect of musical training and tone language experience was observed in the context of non-native Cantonese tone word learning. After seven 30-minute training sessions on mapping 15 Cantonese tone words (3 syllables x 5 contrastive tones) to

meanings (represented by pictures), Thai musicians did not significantly outperform English non-musicians despite having both musical training and native experience with a tone language. Instead, English musicians and Thai non-musicians, i.e., those with only one of the two experiences, had the best tone word learning performance, and relative to them, Thai musicians, in fact, showed a trend towards lower word learning ability. Similarly, when tested on their Cantonese tone identification abilities, Thai musicians only performed significantly better than Thai non-musicians, and it was English musicians instead who performed the best among the four groups, be it pre- or post-training.

In like manner, attempts to investigate whether musical training and tone language experience have similar pitch processing effects have shown inconsistent results. For example, in a study comparing Mandarin Chinese non-musicians, English musicians, and English non-musicians, Bidelman et al. (2011a) found that in response to homologues of the Mandarin Chinese high rising (35) tone and a musical interval (major third), Mandarin Chinese non-musicians and English musicians both showed greater pitch-tracking accuracies compared to English non-musicians. This suggests that musical training and tone language experience have comparable effects in pitch processing. In contrast, other studies show some discrepancies in results, which point to overlapping but different underlying processing mechanisms of music and language. For instance, Bidelman, Gandour, and Krishnan (2011b) and Hutka, Bidelman, and Moreno (2015) found that while tone-language-speaking non-musicians and non-tone-language-speaking musicians both showed enhanced pitch abilities compared to non-tone-language-speaking non-musicians, the pitch abilities of each group were not always the same. In Bidelman et al. (2011b), Mandarin Chinese non-musicians and English musicians both had stronger brainstem representations of musical pitch

sequences compared to English non-musicians. They also had equally superior F0 encoding for the musical pitch sequences, regardless of whether the sequences contained a major/minor third that was in/out of tune. However, behaviourally, Mandarin Chinese non-musicians did not have the same level of perceptual sensitivity as the English musicians to minute changes in the musical pitch sequences, which suggests that a neural-behavioural disconnect is at play. This neural-behavioural disconnect was reversed in Hutka et al. (2015). Behaviourally, Cantonese non-musicians and English musicians had comparable pitch discrimination of the F0 of complex tones, with both groups performing better than English non-musicians. However, the English musicians had an additional advantage in timbral discrimination, and also had a neural advantage in pitch discrimination, which was reflected in enhanced mismatch negativity responses to musical stimuli. How comparable the effects of musical training and tone language experience are is thus not straightforward and may depend on other factors such as task type (behavioural vs. neural) and stimuli type (linguistic vs. non-linguistic).

Therefore, as we can see, not only have there been few studies investigating the combined and relative effects of musical training and tone language experience, but the results so far are mixed. In some cases, the effects of both factors on pitch processing seem to be equivalent, but in others, it is less clear. These suggest that the effects of the two factors may not be invariably additive, and that other factors such as stimulus types (speech vs. non-speech; native vs. non-native), methodologies used (neural vs. behavioural), as well as sample differences in terms of language background and participant groups compared may affect which processing mechanism comes into play. Thus, given the inconsistent findings across existing studies, whether and when musical training and tone language experience give rise to an additive effect on pitch processing,

not to mention the relative weights of each experience during pitch processing, is still unclear and requires further investigation.

2.6 Effects of musical training and tone language experience on pitch production

In addition, while most studies have focused on the perceptual effects of musical training and tone language experience on pitch processing, few have investigated whether these two factors also influence non-native pitch processing in broader contexts, such as pitch production. While enhanced perception does not necessarily guarantee enhanced production skills, perception generally precedes production in human language development, as can be seen from how infants acquire an L1. Even before they can speak, infants under 6-8 months old are able to discern both native and non-native phonetic contrasts (Werker & Tees, 1984a). However, when the linguistic input is flawed as in the case of hearing-impaired infants, speech production abilities can be degraded, especially in the absence of timely intervention (Ambrose et al., 2014). Conversely, perceptual training which improves perception abilities can subsequently improve production abilities, as shown in a study by Wang et al. (2003), where American English listeners improved in their production of non-native Mandarin Chinese lexical tones after a two-week training which also improved their perception of the said tones. Thus, given that musical training and tone language experience can facilitate the perception of non-native lexical tones, a subsequent question of interest would be whether this facilitation extends to the production of said tones.

In particular, research has also shown musicians to have better production skills in non-native non-tone languages compared to non-musicians. Given this, it would not be a far stretch to suppose that musical training may also benefit the production of non-native tone languages, specifically in the production of non-native lexical tones. This is

because pitch is a shared, central dimension in both music and tone languages, and if a general transfer effect from musical training to the production of non-native consonants is possible, a direct pitch-related transfer to the production of non-native tones would be equally, if not more possible. To elaborate, evidence of a musician advantage in the production of non-native non-tone languages has been observed in Milovanov, Pietilä, Tervaniemi, and Esquef's (2010) study on Finnish speakers. In this study, musically trained participants made fewer mistakes in a production task where they imitated English words containing target English phonemes which were challenging to produce because they did not exist in the Finnish sound inventory. A significant positive correlation was also observed between participants' general musical ability and English pronunciation skills. Besides this, in an older study, Slevc and Miyake (2006) tested Japanese speakers who had acquired English as an L2 and found that participants' general musical abilities predicted how well they were able to perceive and produce the English /r/-/l/ contrast. In light of this musician advantage in the production of non-native non-tone languages, as well as the bidirectionality of processing mechanisms between music and language, whereby tone language speakers have also been observed to have enhanced production of musical pitch (Pfordresher & Brown, 2009), it is entirely possible that musicians, in turn, would have enhanced production of non-native lexical tones.

2.7 Summary of research gaps and rationale

In summary, the effect of musical training on pitch processing, specifically in the aspect of non-native lexical tone processing, is lacking in two main ways. Firstly, it has been investigated independently of another major factor, i.e., linguistic experience. This refers to the need to consider the influence of prior experience with tone languages

on non-native lexical tone processing. Secondly, the effect of musical training on non-native lexical tone processing has also been mainly investigated from the aspect of perception, and not from the aspect of production.

Without a simultaneous investigation of the effect of musical training and linguistic experience on non-native lexical tone processing, specifically by testing both tone- and non-tone-language-speaking musician and non-musicians, we would not know whether tone language speakers also experience enhancement in pitch perception from musical training, the way past studies have shown non-tone language speakers to. The inclusion of tone-language-speaking musicians and non-musicians as part of the sample would thus improve the generalizability of past findings of a musicianship effect on non-native lexical tone processing, especially given that tone language speakers are not a rarity.

Furthermore, as prior experience with tone languages also affects pitch processing by either facilitating or inhibiting the perception of non-native lexical tones, including tone-language-speaking populations also enables an investigation of whether there are interactive effects between musical training and tone language experience in pitch processing. In particular, it allows us to find out whether musical training and tone language experience result in equivalent pitch processing enhancements, such that musicians without tone language experience would process non-native lexical tones as well as non-musicians with tone language experience. Besides this, we would also be able to find out whether musicians with tone language experience enjoy enhancements in pitch processing abilities that are greater than musicians without tone language experience and non-musicians with tone language experience, i.e., whether there are additive effects of musical training and tone language experience. Together, these would contribute to a finer understanding of how pitch processing of non-native tones is

affected by transfer effects from musical training as well as native tone language experience, which also has practical implications for the L2 acquisition of tone languages.

Given that perception abilities generally precede production abilities, and that musical training has been found to improve certain aspects of L2 speech production in past studies, it is a reasonable conjecture that the perceptual benefits of musical training and tone language experience to non-native lexical tone processing could have transfer effects from the domain of perception to that of production. Investigating whether musical training and tone language experience have such transfer effects on the production of non-native lexical tone processing will thus enrich our knowledge of their respective influences on pitch processing. Moreover, considering the fact that one major, if not ultimate goal in the context of L2 acquisition is for smooth communication with native speakers to take place, finding out whether perceptual benefits in non-native lexical tones extend to production benefits would have significantly broader implications for the L2 acquisition of tone languages.

2.8 Research aims and objectives

The aim of this thesis is thus to address these research gaps by investigating the effect of musical training on both the perception and production of non-native lexical tones in conjunction with the factor of tone language experience. This is to be achieved by the following research objectives: (1) to investigate the effects of musical training on non-native lexical tone processing with the inclusion of a more representative sample of musicians and non-musicians, (2) to investigate whether musical training interacts with tone language experience during non-native lexical tone processing, and (3) to

investigate whether musical training and tone language experience affect the production of non-native lexical tones in addition to its perception.

With respect to the first objective, both tone- and non-tone-language-speaking musicians will be included in the study, so that their pitch processing abilities may be contrasted with their tone- and non-tone-language-speaking non-musician counterparts. This will allow us to examine whether past findings of a beneficial transfer effect of musical training on pitch processing would also be observed among tone-language-speaking populations, which serves to improve generalizability of the present study's results.

To achieve the second objective of investigating possible interactions between musical training and tone language experience on non-native lexical tone processing, comparisons between the performances of tone-language-speaking musicians, non-tone-language-speaking musicians, and tone-language-speaking non-musicians during non-native lexical tone processing will be made. A comparison of non-tone-language-speaking musicians with tone-language-speaking non-musicians will allow us to examine whether musical training confers more advantages than tone language experience in non-native lexical tone processing. Meanwhile, a comparison of tone-language-speaking musicians with the other two groups will allow us to examine whether musical training and tone language experience have additive effects.

Finally, to achieve the third objective, non-native lexical tone processing abilities will be measured via both a perception and production task. This inclusion of a production task will allow us to study whether musical training and tone language experience also confer advantages to the production of non-native lexical tones in addition to its perception.

Through such an investigation of the independent and interactive effects of musical training and tone language experience on the perception and production of non-native lexical tones, this study examines the central question of the extent to which musical training influences the processing of non-native lexical tones. Namely, this is achieved by asking whether this influence of musical training also holds true among tone language speakers, whether its magnitude changes depending on one's prior tone language experience, and whether it also extends to the realm of production in addition to that of perception. In this way, a further contribution to our understanding and application of the transfer effects between the processing of musical pitch and non-native lexical tones is made.

CHAPTER 3: THE PRESENT RESEARCH

3.1 Overview and research questions

The purpose of this study is to investigate how musical training and linguistic experience may independently and interactionally affect the perception and production of non-native lexical tones. To this end, a 2 x 2 factorial design was employed, with musical training and tone language experience as the two independent variables, so that a comparison of musicians and non-musicians from both tone-language-speaking and non-tone-language-speaking backgrounds could be conducted (see Table 1). In effect, four groups of participants were studied: Tone-language-speaking musicians, hereafter tone musicians (Tm), tone-language-speaking non-musicians, hereafter tone non-musicians (Tnm), non-tone-language-speaking musicians, hereafter non-tone musicians (NTm), and non-tone-language-speaking non-musicians, hereafter non-tone non-musicians (NTnm).

Table 1

General Experimental Design.

		Independent Variable A: Musical Training	
		(A1) Musician	(A2) Non-musician
Independent Variable B: Tone Language Experience	(B1) Tone language speaker	Tone musician (Tm)	Tone non-musician (Tnm)
	(B2) Non-tone language speaker	Non-tone musician (NTm)	Non-tone non-musician (NTnm)

In this study, the following research questions are addressed:

- (1) Do the perceptual benefits of musical training apply to tone language speakers?
- (2) Between musical training and tone language experience, does one bear more weight in non-native lexical tone processing?
- (3) Are the effects of musical training and tone language experience additive?
- (4) In addition to the perceptual benefits, does musical training and tone language experience also benefit the production of non-native lexical tones?

To answer the first research question, we compared Tm and Tnm in a non-native lexical tone perception task. If the perceptual benefits of musical training also apply to tone language speakers, Tm should perform better than Tnm. To answer the second question, we compared Tm and NTm, i.e., those with only one of the two experiences, to see if they have comparable performances. This leads to the answer to the third question: If Tm outperforms Tnm and NTm in both cases, then we can conclude that musical training and tone language experience have additive effects. To answer the fourth question, a non-native lexical tone production task was conducted.

3.2 Research approach

The effect of musical training and tone language experience on the perception of non-native lexical tones was investigated in Experiment 1, using an AX discrimination task modelled after that used in Burnham et al. (2015). In this experiment, participants listened to and attempted to discriminate non-native Thai lexical tones in three forms: speech, low-pass filtered speech, and violin analogues of speech. This paradigm thus enabled a fine-grained examination of how well non-native pitch can be discriminated, by having it presented on a speech to non-speech continuum.

Based on the results in Experiment 1, a follow-up investigation was conducted to examine whether non-tone language speakers with long-term ambient exposure to tone languages would nonetheless have enhanced perception of non-native lexical tones. Using secondary data analysis, the perceptual discrimination results of non-tone participants in Experiment 1 were compared with those of non-tone participants from Burnham et al. (2015).

In Experiment 2, the effect of musical training and tone language experience on the production of non-native lexical tones was investigated. Here, participants listened to the same Thai stimuli used in the perception task and attempted to reproduce them as accurately as possible. To evaluate the quality of their imitation abilities, a separate group of native Thai speakers was recruited as informants to identify the tone category of each produced token, and rate how well each token has been produced.

In addition to the non-native lexical tone perception and production tasks, two supplementary tasks were conducted to control for possible confounds of innate musical ability as well as absolute pitch (AP) ability (see Chapter 4 Section 4.3). The first was a musical ability test designed by Wallentin et al. (2010), i.e., the Musical Ear Test (MET), and the second was an AP test modelled after that of Bermudez and Zatorre (2009). Participants' scores on the MET were used to validate their classification into musician and non-musician groups, whereas the AP test was used to screen for and exclude AP possessors.

3.3 Hypotheses and predictions

In general, both musical training and tone language experience were hypothesized to enhance non-native lexical tone perception; the former by enhancing

general pitch discrimination skills, and the latter because of familiarity with the phonological use of pitch. This resulting perceptual attunement to pitch was further hypothesized to enhance the auditory feedback system, thereby enhancing the ability of musicians and tone language speakers to produce non-native lexical tones. Thus, musicians were predicted to perform better than non-musicians, and likewise, tone language speakers were predicted to perform better than non-tone language speakers. This would be reflected in higher perceptual discrimination scores for musicians and tone language speakers in the speech condition when compared to non-musicians and non-tone language speakers respectively, and in higher identification rates and rating scores for their produced tone tokens.

While past studies have shown that it is possible for tone language experience to hinder subsequent perception of non-native tones (Francis et al., 2008; Hao, 2012; Lee et al., 1996), we expected a facilitatory effect here based on the results of Wayland and Guion's (2004) study. In their study, a facilitatory effect of tone language experience on non-native lexical tone discrimination and identification was found among participants with language backgrounds similar to those tested in the current study (Mandarin Chinese tone language speakers vs. English non-tone language speakers), while also using Thai tones as the non-native stimuli, as in the current study. It is likely that this facilitatory effect was due to the similar tonal inventories between the two languages: Mandarin Chinese has four tones, Thai has five; both languages have level and contour tones; the tones in both languages have overlapping acoustic characteristics (see Figure 1). Based on Best's Perceptual Assimilation Model (Best, 1995), having overlapping tonal inventories between two tone languages would aid perceptual discrimination, which would mean that it should be easier for tone participants to assimilate the Thai tones into Mandarin Chinese tone categories compared to non-tone participants.

Figure 1. A comparison of (a) Mandarin Chinese (M) tones and (b) Thai (T) tones produced by two native speakers of each gender, for each language (Wu, Munro, & Wang, 2014). T values were used to normalize for differences in pitch range between male and female speakers and ranged from 1 (lowest pitch of the speaker) to 5 (highest pitch of the speaker), as in Chao's (1948) notation scale. The abbreviations, HL (high level), R (rising), FR (falling rising), HF (high falling), ML (mid low), LF (low falling) describe the characteristics of the produced tones in each language. (Note: In the present study, the Thai tones are referred to in the traditional way, such that T-ML is labelled as the mid tone, T-LF as the low tone, T-HF as the falling tone, T-R as the high tone, and T-FR as the rising tone.)

Perceptually, it was hypothesized that the perceptual benefits of musical training would also be observed among tone language speakers. Prior research comparing Tm and Tnm has shown that behaviourally, Tm has an advantage over Tnm in discriminating the pitch of non-speech stimuli (Dawson et al., 2018) and identifying non-native Cantonese tones (Cooper and Wang, 2012), even though neural evidence was somewhat mixed (Dawson et al., 2018; Nan et al., 2018). Therefore, it was possible that these perceptual benefits that have been found for Tm over Tnm under different testing conditions would also extend to the discrimination and production of non-native speech stimuli, i.e., Thai tones in the present study.

With respect to the relative weights of each factor, tone language experience was hypothesized to have a greater facilitative effect than musical training specifically in the context of non-native lexical tone perception. In other words, it was predicted that Tnm would have better perceptual discrimination scores in the perception task compared to NTm, at least in the speech condition. This was because firstly, both native and non-native lexical tone perception would occur within the shared domain of language, which should give rise to a nearer (and therefore likelier to be stronger) transfer effect as compared to musical training. Moreover, the similarities in tonal inventories between Mandarin Chinese and Thai would likely aid perceptual discrimination for the Mandarin Chinese Tnm in the study more compared to NTm. Although musicians were predicted to have enhanced perception of non-native tones due to enhanced pitch discrimination abilities from long-term musical training, the experimental design of the perception task in this study was such that different tokens were tested within each tone category, which more closely approximated the real-life context of listening and learning non-native lexical tones, where learners would seldom hear the same token for each tone category. Accordingly, accurate discrimination of the non-native Thai tones would require the

ability to abstract each tone category's main pitch features, which tone language speakers would have more experience in compared to musicians. Thus, this led to the prediction that Tnm would outperform NTm in the non-native tone perception task.

In the context of non-native lexical tone production, however, it was less clear whether tone language experience would have a greater facilitative effect than musical training. On the one hand, Tnm may benefit from the near transfer effect and similar tonal inventories between Mandarin Chinese and Thai when attempting to produce non-native Thai tones. On the other hand, it was possible that they would produce more Mandarin Chinese-like tones instead, as a result of perceptual assimilation of Thai tones to Mandarin Chinese tone categories. In contrast, NTm may experience less of this interference due to their non-tone language background. It seemed possible that either Tnm would also outperform NTm, or NTm would outperform Tnm, or that both groups would receive equally favourable evaluations from native Thai listeners.

Lastly, it was hypothesized that the combined effects of musical training and tone language experience on non-native lexical tone processing would not be additive. This was tentatively hypothesized based on Cooper and Wang's (2012) findings that musical training and tone language experience had no additive effects on non-native lexical tone word learning and tone identification. It should be noted, however, that these results were based on native English and native Thai participants learning to identify and map non-native Cantonese lexical tones to meaning, which is different from the current study in terms of the participants' language background, non-native stimuli used, as well as experimental paradigm. Thus, it was possible that different results would be observed in the current study. Tentatively, it was predicted that Tm would perform similarly as Tnm and also as NTm in both perception and production tasks.

CHAPTER 4: METHOD

4.1 General procedure

After an audiometric screening assessment to check for normal hearing, participants completed experimental tasks in the following order: (1) The AX non-native lexical tone discrimination task, (2) the musical ability test, (3) the non-native lexical tone imitation task, and (4) the absolute pitch (AP) test, which was only administered on participants in the musician group. Based on the participants' produced Thai tone tokens, an independent group of native Thai-speaking informants completed a tone identification task, followed by a tone rating task.

For both participants and informants, the entire experiment took place in a sound-attenuated room, after informed consent was obtained. At the end of the experiment, participants and informants completed a standard laboratory questionnaire for contact and profiling purposes (see Appendix A) and were monetarily compensated at SGD10 per hour of their participation.

4.2 Participants

4.2.1 Recruitment criteria

A total of 87 participants took part in the study. Aside from one participant from National University of Singapore, the rest were recruited from Nanyang Technological University of Singapore. Participants were recruited on the basis that firstly, none of them spoke Thai as a native language or learned it as a second language (L2). This was important as the main experimental stimuli were Thai tones. Secondly, participants were recruited only if they could be classified into one of the four experimental groups: tone

musicians (Tm), tone non-musicians (Tnm), non-tone musicians (NTm), and non-tone non-musicians (NTnm). Lastly, participants had to have normal hearing.

To assess participants' eligibility for the study, a short online screening questionnaire was administered (see Appendix B). Participants were asked questions on their language and musical backgrounds to determine whether they had any tone language experience, and whether they played any musical instruments. Those who reported knowing how to play a musical instrument were asked further questions on their musical background, e.g., how they learned to play their musical instrument(s), and the age onset and number of years of learning. Responses from the screening questionnaire were used to determine which group each participant belonged to. In addition to the screening questionnaire, participants also completed an audiometric screening conducted at 25dB for four frequencies (0.5kHz, 1kHz, 2kHz, and 4kHz), using a Grason-Stadler GSI-18 screening audiometer. All participants were assessed to have normal hearing.

4.2.2 Participant distribution

Seventeen participants were excluded from the final sample because of classification issues which posed potential confounds: One of them was tested to have AP (see Section 4.3.2.4), while the other 16 were non-tone participants who belatedly reported some formal training or receptive proficiency in a non-native tone language, Mandarin Chinese. The final sample thus comprised 70 participants aged between 18 to 28 years old (44 females, 26 males, $M_{age} = 22.21$ years, $SD = 2.64$), with a distribution as presented in Table 2.

Table 2

Distribution of Participants.

Group	<i>n</i>	Age ^a	Gender	
		<i>M (SD)</i>	<u>Female</u>	<u>Male</u>
Tone musicians (Tm)	18	21.56 (2.55)	14	4
Tone non-musicians (Tnm)	19	22.84 (2.03)	12	7
Non-tone musicians (NTm)	14	20.64 (2.41)	9	5
Non-tone non-musicians (NTnm)	19	23.37 (2.85)	9	10

Note. *N* = number of participants. *M* = Mean. *SD* = Standard Deviation.

^aA statistically significant difference in age was observed between groups, as determined by a one-way between subjects ANOVA, $F(3, 66) = 4.07, p = .010$. A Tukey post hoc test revealed that NTm were younger than NTnm ($p = .067$) and Tnm ($p = .014$).

4.2.3 Language background

Tone participants ($n = 37$) were all native English-Mandarin bilinguals, with most of them ($n = 28$) reporting some receptive proficiency in tone languages other than Mandarin Chinese: Hokkien ($n = 20$), Cantonese ($n = 8$), and Teochew ($n = 7$). About one third of the tone participants ($n = 11$) also reported having exposure to the following non-tone languages as an L2: Malay ($n = 8$), Korean ($n = 6$), Japanese ($n = 5$), Italian ($n = 3$), French ($n = 2$), German ($n = 1$).

Non-tone participants ($n = 33$) consisted of five kinds of bilinguals: English-Malay ($n = 18$), English-Tamil ($n = 7$), English-Indonesian ($n = 4$), English-Hindi ($n = 2$), and English-Bengali ($n = 2$). About half of the non-tone participants ($n = 14$) also reported having additional exposure to the following non-tone languages: Hindi ($n = 3$), Telugu ($n = 3$), Malay ($n = 2$), Arabic ($n = 2$), Indonesian ($n = 2$), Malayalam ($n = 2$),

Javanese ($n = 1$), Punjabi ($n = 1$), Gujarati ($n = 1$), Rajasthani ($n = 1$), French ($n = 2$), German ($n = 2$), Japanese ($n = 2$), Spanish ($n = 1$), and Korean ($n = 1$). Importantly, non-tone participants reported no native acquisition, formal training, or receptive proficiency in any tone languages.

4.2.4 Musical background

4.2.4.1 Classification Criteria

Musicians were defined as individuals who (1) had received musical training from taking private music lessons or engaging in extracurricular musical activities (e.g., band, choir) (2) for at least five years, (3) with an onset between childhood and early adolescence, and (4) were still playing their musical instruments regularly in the past 25 months. Accordingly, participants who had not received any such musical training nor played any musical instruments on their own were classified as non-musicians ($n = 38$). Aside for one who reported two to three childhood violin lessons, non-musicians reported no musical training apart from the compulsory general music lessons in the local school curriculum.

Due to the difficulty in finding NTm, much fewer NTm who had had at least five years of musical training, the second criterion was relaxed to a minimum of three years provided that the other criteria were met. There were four such exceptions in the non-tone musician group. Of the four, two had self-taught themselves another musical instrument for more than 10 years despite having only three and four years of formal musical training respectively.

4.2.4.2 Sample Characteristics

In general, musicians ($n = 32$) began musical training at 4-14 years old and were playing their musical instruments for an average of 0.5-23.0 hours per week in the past 25 months. In terms of frequency of playing, most musicians reported playing their instruments about once a week ($n = 14$) or more than once a week ($n = 12$). A few reported playing their instruments almost every day ($n = 4$) or about once a month ($n = 2$).

Table 3 summarizes the characteristics of the two musician groups. Overall, Tm ($n = 18$) and NTm ($n = 14$) were generally similar and only differed on two measures: Tm had received more years of musical training (range = 6-20 years) than NTm (range = 3-16 years) and had also stopped musical training more recently (range = 0-7 years and 0-11 years respectively).

Table 3

Characteristics of Tone (Tm) and Non-Tone Musicians (NTm).

	Tm ^a	NTm ^b	<i>t</i>	<i>df</i>
	<i>M (SD)</i>	<i>M (SD)</i>		
Age started musical training	7.06 (2.29)	8.57 (2.71)	-1.72	30
Years of musical training	11.89 (3.83)	7.43 (3.69)	3.32**	30
Years since musical training	2.40 (2.25)	4.86 (3.23)	-2.53*	30
Average playing hours per week	4.82 (3.35)	5.36 (5.51)	-0.34	30
Number of instruments played	2.44 (0.92)	2.29 (1.27)	0.41	30

Note. *M* = Mean. *SD* = Standard Deviation. *N* = number of participants. Age started musical training is in years. Average playing hours per week are for the past 25 months.

^a $n = 18$, ^b $n = 14$. * $p < .05$. ** $p < .01$.

In terms of musician type, most musicians were entirely non-percussionists ($n = 27$), with keyboard (mainly piano) being the most commonly played instrument across both musician groups ($n = 25$) (see Figure 2). Of the five musicians who played percussion instruments, only one played it as their primary (and only) instrument type. The main difference between the two musician groups was that NTm had a relatively larger proportion of vocalists and percussionists compared to Tm. Overall, however, the two musician groups were similar in their distribution of instruments played.

Figure 2. Distribution of musician types. This figure presents the types of instruments played by tone musicians (Tm) ($n = 18$) and non-tone musicians (NTm) ($n = 14$), and the percentage of musicians ($N = 32$) who play them.

For details of individual musician backgrounds, refer to Appendix C.

4.3 Supplementary tasks

4.3.1 Musical ability test

A test of musical ability was included in the experiment to control for the potential confound of musical aptitude. This confound arises because enhanced pitch discrimination may not necessarily be a result of musical training alone. For instance, it is possible that some who have been classified as non-musicians would nonetheless have a natural aptitude for musical skills such as pitch discrimination. Disregarding this possibility may thus confound the results of the study, as it would be unclear whether musical training or musical ability was the real factor affecting non-native lexical tone discrimination. Given that the tone and non-tone musician groups also differed in their years and recency of musical training, having an objective measure of their musical ability was also important in the event that they showed differences in the non-native tone processing tasks.

To this end, the Musical Ear Test (MET) designed by Wallentin et al. (2010) was used to test participants' musical ability. The MET was specifically chosen because of the following advantages: It can be taken by both musically trained or untrained individuals, is short (< 20 minutes to administer), objective, well-documented and reliable in distinguishing groups of varying musical expertise (e.g., professional musicians vs. amateur musicians vs. non-musicians) (Wallentin et al., 2010).

4.3.1.1 Stimuli

Stimuli in the MET consisted of 52 pairs of melodic phrases and 52 pairs of rhythmic phrases (see Figure 3 for example trials). The melodic phrases formed the Melody subtest, and the rhythmic phrases formed the Rhythm subtest. Half of the

phrases in each subtest were “same” trials, and the other half were “different” trials, i.e., trials with one pitch violation (Melody subtest) or one rhythmic change (Rhythm subtest). Melodic phrases consisted of piano tones, and varied in tonality (25 non-diatonic trials, 20 major key trials, 7 minor key trials). Each melodic phrase contained three to eight tones. Rhythmic phrases consisted of woodblock sounds, and varied in complexity (21 trials with triplets, 31 trials with even subdivisions of the beat; 37 trials beginning on the downbeat, 15 trials beginning later) (Wallentin et al., 2010).

Figure 3 illustrates example trials from the Musical Ear Test (MET). The figure is divided into two sections: "Melody test" and "Rhythm test".

Melody test:

- A:** Labeled "Same". It shows a sequence of notes on a staff.
- B:** Labeled "Different". It shows a sequence of notes on a staff, with two notes circled to indicate pitch violations.

Rhythm test:

- C:** Labeled "Same". It shows a sequence of rhythmic marks on a staff.
- D:** Labeled "Different". It shows a sequence of rhythmic marks on a staff, with two marks circled to indicate rhythmic changes.

Figure 3. Example trials from the Musical Ear Test (MET) (Wallentin et al., 2010). This figure illustrates “same” and “different” trials in each MET subtest (Melody, Rhythm), with differences between musical phrases in the “different” trials circled.

4.3.1.2 Apparatus and procedure

All stimuli from the MET were presented to the participants through a pair of Sennheiser HD 380 Pro headphones, on a desktop computer. The presentation order of the two MET subtests (Melody, Rhythm) was counterbalanced across and within participant groups to control for the confound of difficulty (based on pilot testing, most participants found the Rhythm subtest to be more difficult). Participants were given the

response sheets for each subtest and instructed to check the “Yes” box for musical phrases which they judged to be identical, and the “No” box for musical phrases they judged to be not identical. Two example trials with feedback were given prior to each subtest, and each pair of musical phrases was played only once.

4.3.1.3 Results and conclusion

Mean accuracy scores (in percentage) for musician and non-musician participants are presented in Figure 4. Based on a series of independent-samples *t*-tests, it can be concluded that the MET scores validated the classification of participants into musician and non-musician groups based on their self-reported musical background.

As depicted in Figure 4a, musicians ($M = 77.4\%$, $SD = 6.92$) exhibited higher overall accuracy scores in the MET compared to non-musicians ($M = 69.3\%$, $SD = 8.07$), $t(68) = 4.40$, $p < .001$. Musicians also performed better than non-musicians in each MET subtest. As shown in Figure 4b, musicians had higher accuracy scores in the Melody subtest ($M = 79.0\%$, $SD = 8.89$) compared to non-musicians ($M = 68.6\%$, $SD = 8.31$), $t(68) = 4.91$, $p < .001$, and they also performed better ($M = 75.7\%$, $SD = 7.66$) than non-musicians ($M = 70.1\%$, $SD = 10.5$) in the Rhythm subtest despite most of them being non-percussionists, $t(68) = 2.59$, $p = .012$. These results indicate that participants who had been classified as musicians did have higher musical abilities compared to those who had been classified as non-musicians.

A comparison of the MET scores in the current study with those in Experiment 3 of the study by Wallentin et al. (2010) further validates the classification of musician and non-musician participants in the current study. As shown in Table 4, the non-musicians in the current study had MET scores that were comparable to those tested by Wallentin et al. (2010). The MET scores of musicians in the current study were also

comparable to the scores of the amateur musicians and not the professional musicians tested by Wallentin et al. (2010), which is accurate as none of the musicians in the current study were professional musicians.

Figure 4. Musicians (n = 32) and non-musicians' (n = 38) mean accuracy scores in the Musical Ear Test (MET), with standard error bars. Graph (a) shows participants' combined scores for both subtests. Graph (b) shows participants' scores in each subtest (Melody, Rhythm). (Note: * $p < .05$. *** $p < .001$)

Table 4

Comparison of the Musical Ear Test (MET) scores of musicians and non-musicians in the current study and in Experiment 3 of the study by Wallentin et al. (2010).

Non-musicians			
	Current Study (<i>n</i> = 38)	Wallentin et al. (2010) (<i>n</i> = 21)	
	<i>M</i> (<i>SD</i>)	<i>M</i> (<i>SD</i>)	
MET - Total	69.3 (8.1)	67.8 (6.1)	
MET - Melody	68.6 (8.3)	66.7 (8.4)	
MET - Rhythm	70.1 (10.5)	69.0 (8.7)	
	Musicians	Amateur Musicians	Professional Musicians
	Current Study (<i>n</i> = 32)	Wallentin et al. (2010) (<i>n</i> = 21)	Wallentin et al. (2010) (<i>n</i> = 18)
	<i>M</i> (<i>SD</i>)	<i>M</i> (<i>SD</i>)	<i>M</i> (<i>SD</i>)
MET - Total	77.4 (7.0)	78.0 (6.0)	85.7 (4.6)
MET - Melody	79.0 (8.9)	78.6 (7.5)	89.1 (5.6)
MET - Rhythm	75.7 (7.7)	77.4 (7.1)	82.3 (7.5)

Note. *M* = Mean. *SD* = Standard Deviation. All values are in percentage and reflect mean accuracy. For clearer comparison with data from the study by Wallentin et al. (2010), all values are rounded to one decimal place.

Given that Tm and NTm were found to differ significantly with respect to their years of musical training and years since musical training (see Section 4.2.4.2), separate one-way between subjects ANOVAs were conducted to compare the MET scores of each participant group across the subtests, in order to examine whether there were any differences in the musical abilities of Tm and NTm respectively as a result of the differences in their musical background. Results showed a significant effect of group on MET scores for both subtests (Melody, $F(3, 66) = 7.85, p = .0001$; Rhythm, $F(3, 66) =$

2.87, $p = .043$). For clarity and relevance's sake, only post hoc Bonferroni-corrected comparisons within the musician groups (Tm vs. NTm) and within the non-musician groups (Tnm vs. NTnm) are reported here. Within the musician groups, it was found that Tm and NTm did not have significantly different scores in the Melody subtest, ($M = 79.2$, $SD = 6.93$ and $M = 78.8$, $SD = 11.2$ respectively; $p = 1.000$), nor did they have significantly different scores in the Rhythm subtest ($M = 75.6$, $SD = 7.17$ and $M = 75.8$, $SD = 8.52$ respectively; $p = 1.000$). Similarly, within the non-musician groups, it was found that Tnm and NTnm did not have significantly different scores in the Melody subtest ($M = 69.2$, $SD = 8.50$ and $M = 68.5$, $SD = 8.44$ respectively; $p = 1.000$) and the Rhythm subtest ($M = 67.9$, $SD = 10.65$ and $M = 72.1$, $SD = 10.04$ respectively; $p = 1.000$).

In conclusion, it can be reasonably assumed that those classified as musicians in the current study did differ from those classified as non-musicians in terms of actual musical abilities, as measured via the MET scores, and through comparisons with MET scores of other musicians and non-musicians from the study by Wallentin et al. (2010). Furthermore, the two musician groups do not differ from each other in terms of actual musical abilities despite some differences in their musical background, and likewise for the two non-musician groups. Therefore, for the present study, there is a reduced likelihood for performance in the non-native lexical tone processing tasks to be confounded by musical aptitude.

4.3.2 Absolute pitch test

A test for AP was used to screen for AP ability among musicians, which is a potential confound in the current study. AP refers to the ability to either identify or produce the pitch of a musical note without relying on a reference pitch (Takeuchi &

Hulse, 1993). It is most commonly assessed by pitch identification tasks, where given musical notes in isolation, participants are asked to identify each note by its pitch class (i.e., chroma) and by its pitch height (i.e., octave) (Levitin, 1994). As AP musicians have been found to perform better than non-AP musicians using the same perception task paradigm in the current study (Burnham et al., 2015), an AP screening for musician participants was necessary.

4.3.2.1 Stimuli

A total of 36 pure tones with fundamental frequencies corresponding to musical notes ranging from C3 – B5 (three octaves) were created using the Praat software (Boersma & Weenink, 2017), based on the tuning standard of A4 = 440Hz. Each note lasted 500ms, and the intensity (dB) of each note was adjusted so that each note was generally perceived to be equally loud (see Appendix D for further details).

Pure tones were chosen because they would presumably be equally unfamiliar to all musicians regardless of their instrumental background. This was important to control for the confound of timbre familiarity. For instance, Ward and Burns (1982) cautions against assuming that “absolute piano” performance (i.e., high note identification rates when tested with piano tones) is equivalent to AP ability. This is corroborated by evidence from Miyazaki (1989), which revealed AP possessors trained on the piano to have better rates of note name identification when tested with real piano tones (accuracy = 91.6%) rather than synthetic piano tones (accuracy = 80.4%), with the poorest identification rates observed when they were tested with pure tones (accuracy = 74.4%).

4.3.2.2 Apparatus and procedure

The entire AP test was presented to the participants using the E-Prime 2.0 software (Psychology Software Tools, 2016) on a desktop computer, through a pair of Sennheiser HD 380 Pro headphones. The test consisted of two 36-trial blocks, where each block contained the 36 pure tone notes ranging from C3 to B5. Within and across the blocks, the presentation order of the notes was pseudorandomized, such that every subsequent note was more than an octave apart from the preceding note (e.g., C4 would be followed by any note above C5 or below C3). This was done to discourage the use of relative pitch ability, i.e., the use of a note as a reference to identify subsequent notes by comparing the note intervals. To minimize fatigue, participants received a break after every 12 trials.

The response method was modelled on that used by Bermudez and Zatorre (2009). After listening to each note, participants were instructed to first identify the note by its musical label, and then by the octave range it was from. Participants identified the name of the note by pressing one out of 12 response keys, which were labelled with the 12 note names of the Western musical scale and arranged in a circle-like configuration such that all adjacent keys were one semitone apart (see Figure 5a). Participants then identified the pitch height of the note by pressing keys labelled “1”, “2”, “3”, “4” and “5”, which corresponded to the octave ranges of C2 – B3, C3 – B4, C4 – B5, and C5 – B6 respectively (see Figure 5b). Trials were self-paced, but participants were told to respond as quickly and as accurately as possible. To aid participants in their response, the arrangement of the 12 note response keys and the five octave ranges were presented visually to them on the screen. Four practice trials without feedback were given before the actual test.

Figure 5. Response screens (modified after those used by Bermudez and Zatorre (2009)) presented to the participants during the AP test. Panel (a) shows the note response screen and physical response key arrangement during the note name identification segment of each trial. Panel (b) shows the octave response screen during the octave identification segment of each trial, with each band of colour representing the note range of C to B within each octave.

The purpose of this circle-like configuration for note name identification, which was modified after that used by Bermudez and Zatorre (2009), was twofold. Firstly, it controlled for the potential confound of familiarity. If response keys were arranged either horizontally from left to right, like how notes are arranged on a piano, or vertically from top to bottom, like how notes are arranged on instruments like the violin and the guitar, either arrangement would have introduced bias to some musicians and not others depending on their musical background. Secondly, this arrangement meant that the 12 possible responses were relatively centred with all the keys close by to each other, which was important to collect reliable reaction time data.

4.3.2.3 Data analysis

A total of 35 musician participants completed the AP test. Of these, 31 were included in the final sample of 32 musicians (one non-tone musician did not know Western musical notation and hence was not administered with the test). Three were excluded from the final sample because they were initially classified as non-tone musicians but later reported some formal training or receptive proficiency in Mandarin Chinese. Another was excluded from the final sample because the AP test showed them to be an AP possessor.

Figure 6. How mean absolute deviation (MAD) was computed. This figure presents two examples of how absolute deviation, as opposed to actual semitone deviation, is calculated. For visualization purposes, a piano keyboard with the notes F3, C4, F4, and C5 marked is included.

Three measures were used to analyse the AP test results: percentage accuracy, mean absolute deviation (MAD), and reaction time (RT). Percentage accuracy was the more stringent accuracy measure, as it was calculated for responses which had to be correct in both note name and octave range, i.e., responses with zero semitone deviations from the tested note. In contrast, MAD was computed by disregarding octave information, and taking the absolute value of the shortest semitone deviation between the pitch classes of the tested note and the response note (Bermudez & Zatorre, 2009). Absolute deviation values always ranged between zero to six semitones, and their values across the 72 trials were averaged for each participant to give individual MAD scores. The closer a MAD score is to zero, the closer the pitch identification is to being perfectly correct. To illustrate how MAD was computed, two examples are given in Figure 6.

Traditionally, the percentage accuracy measure has been used to assess AP. In this study, the MAD measure serves to complement the percentage accuracy measure. This is because AP ability is characterized primarily by an enhanced ability in identifying pitch class (i.e., note name) rather than pitch height (i.e., octave range) (Bachem, 1937; Miyazaki, 1988; Miyazaki, 1989). Thus, an assessment of AP ability based strictly on both accurate pitch class and pitch height identification would be less ideal. Besides that, as responses of AP possessors are also fast in addition to being accurate, a third measure, RT was used.

4.3.2.4 Results and conclusion

Figure 7 presents the percentage accuracy, MAD, and reaction times of the 35 musicians who completed the AP test. Based on the test scores, one participant was identified as an AP possessor (see Figure 7). As can be observed from Figures 7a and 7b respectively, participant 19 exhibited high accuracy rates on both accuracy measures,

as well as fast response times (61.1%, MAD = 0.28 semitones, mean RT = 2824ms). This was comparable to the scores of the reference AP possessor (91.7%, MAD = 0.08, mean RT = 1643ms), a known AP possessor administered with the AP test for purposes of reference and comparison. Besides that, participant 19's accuracy rate for note name identification was 77.8%, which was also comparable to the pure tone note name identification rates of 74.4% found in Miyazaki's (1989) study, where 10 AP participants were tested with 72 notes over six octaves (C2 – B7). Furthermore, boxplots analyses corroborate these results by indicating participant 19 as an extreme outlier on both accuracy measures (see Figures 8a and 8b).

For reference, although participants 8 and 15 were also shown to be outliers on both accuracy measures, and participants 2 and 10 as outliers on the MAD measure, they were not considered AP possessors in this study. This was mainly because of their relatively lower pitch class and height identification accuracy rates or slower reaction times compared to participant 19 and the reference AP possessor. Their note name identification rates were also not comparable to the 74.4% mean accuracy rate of Miyazaki's (1989) AP participants: 37.5% (participants 2 and 8), 22.2% (participant 10), and 40.3% (participant 15). Instead, these four participants may be best classified as intermediate AP possessors. Participant 15 best exemplifies this, with a pitch class and height identification accuracy rate (29.2%) which is > 2 standard deviations (SDs) above the mean of non-AP participants ($M = 7.6\%$), and a MAD score (1.08 semitones) which is < 2SDs below the mean of non-AP participants ($M = 2.75$ semitones).

Figure 7. Results of AP test. Scatterplot (a) shows participants' mean percentage accuracy scores plotted against their mean absolute deviation (MAD) scores. Scatterplot (b) shows participants' mean reaction times plotted against their MAD scores. Scores of a known AP possessor on the AP test were included for the purposes of reference and comparison. Participants found to be outliers on either one of the accuracy measures based on boxplot analysis were labelled with their identification number.

Figure 8. Boxplots of AP test scores. Boxplot (a) shows the spread of scores measured in percentage accuracy. Boxplot (b) shows the spread of scores measured in mean absolute deviation (MAD). Outliers with values between one and a half to three times the interquartile (IQ) range, i.e., beyond the whiskers are indicated with a circle. Extreme outliers, i.e., cases with values more than three times the IQ range are indicated with a star.

CHAPTER 5: EXPERIMENT 1

5.1 Overview

The purpose of Experiment 1 was to investigate how musical training and tone language experience affect the perception of non-native lexical tones. A 2 x 2 x (3) mixed factorial design was used. The independent variables consisted of two between-subjects factors, i.e., the musical background (musician or non-musician) and language background (tone or non-tone language speaker), in addition to one within-subjects factor, i.e., stimulus type (speech, low-pass filtered speech, and violin analogues of speech). The dependent variables were the accuracy and reaction times of the participants.

As previously mentioned (see Chapter 3 Section 3.3), tone language experience was hypothesized to have a greater facilitative effect than musical training for non-native lexical tone perception. It was also hypothesized that musical training and tone language experience would not be additive. Based on these hypotheses, as well as past findings from studies using the same experimental paradigm (Burnham et al., 1996; Burnham et al., 2015), a detailed summary of predicted results for Experiment 1 is as follows, with an illustration from Figure 9.

Figure 9. Predicted patterns of non-native Thai lexical tone discrimination for tone musicians (Tm), tone non-musicians (Tnm), non-tone musicians (NTm), and non-tone non-musicians (NTnm) across the three stimulus types (speech, filtered speech, violin). The higher the score on the discrimination index, the better the perceptual discrimination.

Firstly, it was expected that non-tone participants would discriminate the non-native Thai tones poorly in the speech condition due to the phonological irrelevance of pitch to them. The less speech-like the stimuli, however, the better the non-tone participants are expected to perform better as pitch differences will now be perceived to be acoustic rather than linguistic in nature. The non-tone non-musicians (NTnm) would best exemplify this, whereas non-tone musicians (NTm) would conform less to this expected pattern due to the mitigating effects of musical training. It is thus predicted that as found by Burnham et al. (2015), NTnm would show a graduated form of non-native lexical tone discrimination, with improvements in discrimination as becomes less speech-like. In contrast, NTm, who would be analogous to the non-AP musicians in Burnham et al. (2015) would have enhanced discrimination in the speech condition

compared to NTnm, and also have reduced differences in their discrimination of speech and non-speech tone stimuli.

On the other hand, tone participants are expected to show pitch-related transfer benefits from prior experience with pitch as a phonological category. They are thus expected to show equally enhanced performance across the three stimulus types, as found by Burnham et al. (1996). As stronger transfer benefits are expected within the same domain, it is predicted that tone non-musicians (Tnm) would perform better than NTm in the speech (and possibly filtered speech) condition. Lastly, it is predicted that Tm would show the same equally enhanced performance across stimulus types without outperforming Tnm, assuming that no additive effects are observed for musical training and tone language experience.

5.2 Stimuli

The three sets of stimuli used in Burnham et al. (1996) and Burnham et al. (2015) were adapted for use in this study. Each stimulus set consisted of the five Thai lexical tones – mid 32 (M), low 21 (L), falling 51 (F), high 45 (H), and rising 214 (R) (Ladefoged & Johnson, 2001) – but in one of the following forms: Speech, low-pass filtered speech, or violin analogues of speech.

The speech stimulus set comprised five Thai words spoken by a female native Bangkok Thai speaker. The Thai words consisted of the same syllable [pa:] but differed contrastively in lexical tone (see Table 5). For each of the five words, there were three exemplars. The speech stimulus set thus consisted of 15 tokens (5 lexical tones x 3 exemplars). The filtered speech stimulus set was created by applying a low-pass filter on the speech stimulus set, such that all frequencies above 270Hz were removed, and

the phonological cues of the Thai words were no longer discernible. The filtered speech stimuli thus sounded like muffled speech sounds of varying pitch heights and contours.

Table 5

Thai Speech Stimuli with IPA Transcription and Gloss.

Tone	Word	IPA Transcription	Gloss
Mid (32)	ปา	[pā:]	<i>To throw</i>
Low (21)	ป๋า	[pà:]	<i>Forest</i>
Falling (51)	ป้า	[pâ:]	<i>Aunt</i>
High (45)	ป๊า	[pá:]	<i>Father (Chinese origin)</i>
Rising (214)	ป๊า	[pǎ:]	<i>Daddy</i>

To create the violin speech analogues, a professional violinist had listened carefully to the speech stimuli before reproducing them on the violin. Likewise, there were three violin exemplars for each lexical tone in the violin stimulus sets. The violin was chosen as the instrument to recreate musical analogues of the Thai lexical tones because continuously varying pitches can be performed on it rapidly. Thus, unlike non-continuous pitch instruments such as the piano which are constrained to playing pitches in discrete intervals, a violin is able to simulate the rapid pitch changes that occur in lexical tones. For purposes of reference, it should be noted that the violin stimuli had F0 contours that were comparable but not identical to those in the speech (and filtered speech) stimuli. Besides that, the violin stimuli were also in a higher frequency range (293-456Hz) than the speech (and filtered speech) stimuli (138-227Hz).

For this study, all stimuli were normalized in both duration and amplitude. All stimuli were duration normalized to 500ms. The speech and violin stimuli were amplitude normalized to 70dB, whereas filtered speech stimuli were normalized to a

higher amplitude of 80dB to compensate for their dampened, muffled nature, which generally caused them to be perceived as softer (as observed during pilot testing). In addition, three one-minute familiarisation context recordings were also created for each stimulus set. For the speech stimulus set, a female native Thai speaker was recorded reading a Thai short story (see Appendix E). This speech recording was then low-pass filtered to form the context recording for the filtered speech stimulus set. For the violin stimulus set, a violin recording of Canon No. 1 a 2 cancrizans in Bach's Musical Offering BWV 1079 was used. Lastly, two 500ms tones of 600Hz and 1000Hz were also created using the software Praat (Boersma & Weenink, 2017) for use in practice trials.

5.3 Apparatus and procedure

The experiment was presented using the E-Prime 2.0 software (Psychology Software Tools, 2016) on a desktop computer, with the stimuli played to participants through a pair of Sennheiser HD 380 Pro headphones.

Participants first received a short training phase in AX discrimination. Two 500ms tones (600Hz, 1000Hz) were presented in AX pairs to form two "same" and two "different" trials. The inter-stimulus interval was set at 1500ms to induce phonemic processing, which is reliant on long-term memory instead of being purely acoustic or phonetic (Werker and Logan, 1985; Werker and Tees, 1984b). Participants were instructed to judge as quickly and as accurately as they can, within a time limit of 1000ms, whether the two sounds in each sound pair were the same or different. They pressed a key labelled "S" for pairs that sounded the same, and "D" for pairs that sounded different. Feedback was provided in this phase, and participants were required to respond correctly to all four practice trials.

Thereafter, each participant completed three AX discrimination tasks, one in each of the three stimulus types (speech, filtered speech, violin analogues of speech), without feedback. The presentation order of the three AX discrimination tasks was counterbalanced within and across the four participant groups, such that each participant did the three AX discrimination tasks in one of the following permutations: SVF, FSV, or VFS (where S, V, and F refer to the speech, violin and filtered speech stimuli respectively). Each AX discrimination task consisted of its corresponding one-minute familiarisation context recording (a Thai speech recording for the speech condition, which was low-pass filtered for the filtered speech condition, and a violin recording for the violin condition), followed by two stimulus blocks, with a break in between blocks to minimize fatigue. The context recordings were played before each stimulus set to cue the stimulus type. This was to prime a more “music-like” processing mode in the violin context, and a more “speech-like”, linguistic processing mode in the speech and filtered speech contexts.

As there were five Thai tones (M, L, F, H, R), ten AX tone pairs were tested: Three Level-Level tone pairs (M-L, M-H, L-H), one Contour-Contour tone pair (R-F), and six Level-Contour tone pairs (M-R, M-F, L-R, L-F, H-R, H-F). Distribution of the tone pairs in each block was made by taking into consideration the tone pair type: Block 1 consisted of the five tone pairs L-H, R-F, M-R, M-F, and L-R, and Block 2 consisted of the five tone pairs M-L, M-H, L-F, H-R, and H-F. Thus, each block had three Level-Contour tone pairs, with two Level-Level tone pairs in Block 2, and one Level-Level and one Contour-Contour tone pair in Block 1. The distribution of tone pairs was also decided based on the perceptual difficulty of the tone pairs when tested on adult English speakers (Burnham & Francis, 1997); difficulty levels across the two blocks were balanced accordingly.

In each stimulus block, each tone pair (e.g., M paired with L) was presented in four possible order combinations, (e.g., M-M, L-L, M-L, L-M), with each of the four possible order combinations presented twice to give a total of 40 trials (5 contrast pairs x 4 order combinations x 2 trials). Each 40-trial block thus consisted of 20 “same” (AA, BB) and 20 “different” (AB, BA) trials, which were randomized in presentation order. For every tone pair, the tokens for each tone that were presented in each trial were also randomly selected from a pool of three exemplars. For the current study, in the “same” trial condition, there was pseudo-randomization such that the same exemplar was not used (e.g., M-M would consist of two different exemplars of the mid tone). This was done to closely simulate lexical tones heard in real-world contexts, where a balance has to be made between a strictly acoustic processing of the tone stimuli, and a general abstraction of the key features of each tone category.

5.4 Data analysis

Responses in the AX non-native lexical tone discrimination tasks were analysed using the same two measures used in Burnham et al. (2014): d-prime (d') sensitivity index and reaction time (RT). This would allow better comparisons of results in the current study with those in past studies.

To calculate mean d' scores for each participant group, the following formula was used: $d' = Z(\text{Hit rate}) - Z(\text{False positive rate})$. Hit rate was defined as the proportion of correct responses for “different” trials (i.e., judging “different” trials as different), and false positive rate was defined as the proportion of incorrect responses for “same” trials (i.e., judging “same” trials as different). When hit rates were equivalent to the probability of 1, adjustments were made using the formula, $\text{Hit rate} = 1 - \frac{1}{N}$, where N = maximum number of hits. When false alarm rates were equivalent to the probability of

0, adjustments were made using the formula, False alarm rate = $\frac{1}{N}$. A higher d' score indicates better discrimination abilities i.e., greater hit rates and lower false alarms.

Mean RT scores were also calculated for each participant group. As the main purpose of AX discrimination tasks is to measure discrimination of “different” trials, with “same” trials serving as a check against response bias, only RTs of “different” trials were used. Of these, only correct responses were included in the analysis to minimize confounds arising from any additional cognitive processing during error trials. RTs which exceeded ± 2 standard deviations (SDs) from the overall mean (of correct responses to “different” trials) were considered outliers and were further excluded. The final RT dataset used for analysis thus comprised RTs of correct responses to “different” trials which fell within ± 2 SD of the overall mean RT.

Analyses of d' and RT data were conducted using two separate $2 \times 2 \times (3)$ mixed design ANOVAs, with Musical Background (Musicians, Non-Musicians), and Language Background (Tone Speaker, Non-Tone Speaker) as between-subjects factors, and Stimulus Type (Speech, Filtered Speech, Violin) as the within-subjects factor. In addition, planned contrasts were conducted for the d' data. This was in the form of two separate $2 \times 2 \times (2)$ mixed design ANOVAs with the same between-subjects factors of Musical Background and Language Background, but different levels in the within-subjects factor of Stimulus Type. In the first instance, the within-subjects factor is Stimulus Type (Linguistic Tones, Non-linguistic Tones), where the speech and filtered speech conditions were combined into a single level, i.e., “Linguistic Tones”, and contrasted with the violin condition as another level, i.e., “Non-linguistic Tones”. In the second instance, the within-subjects factor is Stimulus Type (Clear Speech, Filtered Speech), with speech vs. filtered speech as its two levels.

5.5 Results

5.5.1 d-prime data

The mean d' scores of the four participant groups (18 Tm, 19 Tnm, 14 NTm, 19 NTnm) for each stimulus type are presented in Figure 10. In general, overall performance was near ceiling, with each participant group having mean discrimination accuracies of $> 85\%$ across the stimulus types (see Table 6).

Figure 10. Mean d' -prime scores of 18 tone musicians (Tm), 19 tone non-musicians (Tnm), 14 non-tone musicians (NTm), and 19 non-tone non-musicians (NTnm) for each of the three stimulus types (speech, filtered speech, violin), with standard error bars.

Table 6

Mean percentage discrimination accuracies of tone musicians (Tm), tone non-musicians (Tnm), non-tone musicians (NTm), and non-tone non-musicians (NTnm) for each stimulus type, with the respective standard deviations in parentheses.

<u>Stimulus Type</u>	<u>Participant Group</u>			
	<u>Tm</u>	<u>Tnm</u>	<u>NTm</u>	<u>NTnm</u>
Speech	89.58 (10.56)	89.61 (9.20)	89.46 (7.41)	83.82 (12.57)
Filtered Speech	86.04 (13.23)	88.95 (8.81)	90.27 (9.38)	85.26 (10.89)
Violin	83.47 (12.28)	90.92 (5.83)	87.68 (9.47)	87.76 (5.51)
Overall	86.37 (12.11)	89.82 (7.99)	89.14 (8.66)	85.61 (10.06)

At a glance, with reference to the predicted patterns of discrimination (see Section 5.1), Tnm and NTm largely followed their predicted discrimination patterns, with pitch discrimination scores that were relatively equivalent for all stimulus types. Meanwhile, NTnm showed a trend towards their predicted discrimination pattern, with discrimination scores improving as the stimuli became less speech-like, i.e., speech < filtered speech < violin. Together, this suggests that musical training and tone language experience may independently enhance non-native lexical tone perception, by reducing the attenuation of discrimination abilities for non-native lexical tones.

Interestingly, unlike predictions that they would also have comparable discrimination across stimulus types, the Tm group displayed a trend of discrimination pattern that was the opposite of NTnm, with discrimination scores decreasing as the stimulus became less speech-like, i.e., speech > filtered speech > violin. This suggests the possibility of an interactional effect of musical training and tone language experience,

whereby the combination of the two factors may result in some hindrance in the processing of musical stimuli in a context less similar to natural music processing.

Analyses revealed no main effect of musical background; overall, musicians ($M = 2.77$) did not perform significantly better than non-musicians ($M = 2.68$), $F(1, 66) = 0.20$, $p = .655$, $\eta_p^2 = .003$. Similarly, there was no main effect of language background; overall, tone speakers ($M = 2.80$) did not perform significantly better than non-tone speakers ($M = 2.63$), $F(1, 66) = 0.47$, $p = .497$, $\eta_p^2 = .007$. No significant interactions between musical background and language background were found, $F(1, 66) = 1.96$, $p = .166$, $\eta_p^2 = .029$.

With respect to stimulus type, no main effect of stimulus type was found, $F(2, 132) = 0.071$, $p = .931$, $\eta_p^2 = .001$. All interactions with stimulus type were also non-significant: Stimulus Type x Musical Background, $F(2, 132) = 1.83$, $p = .165$, $\eta_p^2 = .027$; Stimulus Type x Language Background, $F(2, 132) = 1.62$, $p = .202$, $\eta_p^2 = .024$; Stimulus Type x Musical Background x Language Background, $F(2, 132) = 0.11$, $p = .894$, $\eta_p^2 = .002$. This shows that overall discrimination rates were similar across stimulus types, and that this was true when comparing musicians with non-musicians, tone speakers with non-tone speakers, and the four participant groups.

In addition, no significant linear trends were observed for stimulus type, $F(1, 66) = 0.088$, $p = .768$, $\eta_p^2 = .001$, as well as for all interactions with stimulus type: Stimulus Type x Musical Background, $F(1, 66) = 3.10$, $p = .083$, $\eta_p^2 = .045$; Stimulus Type x Language Background, $F(1, 66) = 1.92$, $p = .170$, $\eta_p^2 = .028$; Stimulus Type x Musical Background x Language Background, $F(1, 66) = 0.12$, $p = .725$, $\eta_p^2 = .002$. This indicates that discrimination patterns across stimulus types were similar when the same

three comparisons were made (between musician and non-musician groups, between tone speakers and non-tone speakers, and between the four participant groups).

5.5.1.1 Planned contrasts: *Linguistic tones vs. non-linguistic tones*

Planned contrasts of participants' performance in the linguistic (speech and filtered speech) vs. non-linguistic (violin) condition found no main effects of musical background, $F(1, 66) = 0.027, p = .869, \eta_p^2 < .0001$, language background, $F(1, 66) = 0.35, p = .555, \eta_p^2 = .005$, and stimulus type, $F(1, 66) = 0.18, p = .672, \eta_p^2 = .003$. Thus, there were no overall effects of musical background, language background, and stimulus type on participants' performances in discriminating linguistic vs. non-linguistic tone stimuli. Musicians generally did not differ from non-musicians, and neither did tone speakers differ from non-tone speakers. Overall performances across the stimulus types were also comparable.

However, a significant Stimulus Type x Musical Background interaction was found, $F(1, 66) = 4.31, p = .042, \eta_p^2 = .061$, indicating different patterns of performance between musicians and non-musicians across linguistic and non-linguistic stimulus types. To investigate the nature of this interaction effect, simple main effects analyses were conducted. In the first analysis, which was on the effect of musical training across stimulus types, no significant effect of musical training was found for the discrimination abilities of linguistic tones, $F(1, 66) = 0.85, p = .359, \eta_p^2 = .013$, and non-linguistic tones, $F(1, 66) = 0.45, p = .503, \eta_p^2 = .007$. Results of the second analysis, which was on the effect of stimulus type across musician and non-musician participants, was similar. No significant effect of stimulus type was found within the musician group, $F(1, 66) = 2.86, p = .095, \eta_p^2 = .042$, as well as the non-musician group, $F(1, 66) = 1.51, p = .224, \eta_p^2 = .022$.

Together, these results indicate a pattern of musicians performing better in the linguistic than in the non-linguistic condition, which was contrasted significantly from non-musicians, who had a pattern of performing better in the non-linguistic than in the linguistic condition (see Figure 11). However, participants' discrimination of linguistic vs. non-linguistic tones were not significantly different regardless of whether they were musicians or non-musicians. Besides that, musicians did not outperform non-musicians in either condition.

Figure 11. Mean d-prime scores of musicians and non-musicians for the linguistic stimulus types (speech, filtered speech) and the non-linguistic (violin) stimulus type, with standard error bars.

5.5.1.2 Planned contrasts: Clear speech vs. filtered speech

Planned contrasts of participants' performance in the speech vs. filtered speech condition found no main effects of musical background, $F(1, 66) = 0.86, p = .357, \eta_p^2 = .013$, language background, $F(1, 66) = 0.66, p = .419, \eta_p^2 = .010$, and stimulus type, $F(1, 66) = 0.001, p = .971, \eta_p^2 < .0001$. All interactions were also non-significant: Musical Background x Language Background, $F(1, 66) = 1.52, p = .223, \eta_p^2 = .022$; Stimulus Type x Musical Background, $F(1, 66) = 0.13, p = .718, \eta_p^2 = .002$; Stimulus Type x Language Background, $F(1, 66) = 2.31, p = .133, \eta_p^2 = .034$; Stimulus Type x Musical Background x Language Background, $F(1, 66) = 0.17, p = .685, \eta_p^2 = .003$.

Thus, there were no independent or interactive effects of musical background, language background, and stimulus type on participants' performances in discriminating non-native lexical tones in speech vs. filtered speech form.

5.5.2 Reaction time data

The mean RT scores (in milliseconds) of the four participant groups (18 Tm, 19 Tnm, 14 NTm, 19 NTnm) for each stimulus type are presented in Figure 12.

Figure 12. Mean reaction times (in milliseconds) of 18 tone musicians (Tm), 19 tone non-musicians (Tnm), 14 non-tone musicians (NTm), and 19 non-tone non-musicians (NTnm) for each of the three stimulus types (speech, filtered speech, violin), with standard error bars. Only reaction times of correct responses to “different” trials within ± 2 standard deviations of the overall mean are shown.

Analyses showed no main effect of musical background; overall, musicians ($M = 867\text{ms}$) did not respond significantly faster than non-musicians ($M = 874\text{ms}$), $F(1, 66) = 0.14$, $p = .711$, $\eta_p^2 = .002$. Similarly, there was no main effect of language background; overall, tone speakers ($M = 873\text{ms}$) did not respond significantly faster than non-tone speakers ($M = 867\text{ms}$), $F(1, 66) = 0.079$, $p = .779$, $\eta_p^2 = .001$. The interaction between musical background and language background was not significant, $F(1, 66) = 0.008$, p

$= .931$, $\eta_p^2 = .0001$, indicating that all four participant groups generally had comparable response speeds.

In contrast to the d' results, a main effect of stimulus type was found here, $F(2, 132) = 35.44$, $p < .001$, $\eta_p^2 = .35$, indicating that RTs significantly differed among the three stimulus types. Bonferroni corrected post hoc tests revealed that responses in the violin context ($M = 834\text{ms}$) were faster compared to those in both speech ($M = 878\text{ms}$), $p < .001$ and filtered speech contexts ($M = 900\text{ms}$), $p < .001$ respectively. Responses in the speech context were also faster than those in the filtered speech context, $p = .045$.

All interactions with stimulus type were non-significant: Stimulus Type x Musical Background, $F(2, 132) = 0.039$, $p = .962$, $\eta_p^2 = .001$; Stimulus Type x Language Background, $F(2, 132) = 0.17$, $p = .846$, $\eta_p^2 = .003$, Stimulus Type x Musical Background x Language Background, $F(2, 132) = 0.48$, $p = .619$, $\eta_p^2 = .007$. This shows that response patterns across the three stimulus types were similar when comparing musicians with non-musicians, tone speakers with non-tone speakers, and the four participant groups.

5.6 Discussion

The purpose of Experiment 1 was to investigate whether musical training interacted with tone language experience in affecting the perception of non-native lexical tones. We hypothesized that while both factors would enhance the perception of non-native lexical tones, with the benefits of musical training also extending to tone language speakers, tone language experience would be more beneficial than musical training due to a nearer transfer effect (from the domain of native to non-native language as opposed to a further, cross-domain transfer effect from music to language). We also hypothesized that there would be no additive effects of the two factors. In effect, it was

predicted that participant groups would show different perceptual discrimination patterns across the different stimulus types, depending on their musical and tone language background.

Contrary to expectations, no effect of musical training or tone language experience was found. Overall, musicians did not perform significantly better than non-musicians, and neither did tone language speakers perform significantly better than non-tone language speakers. This also means that no evidence was found for tone-language-speaking musicians having an additional benefit over tone-language-speaking non-musicians here, unlike in past studies (Dawson et al., 2018; Nan et al., 2018). Besides that, discrimination scores for the non-native lexical tones were generally near ceiling across participant groups, and contrary to the predictions made based on the findings from Burnham et al. (1996) and Burnham et al. (2015), there were no statistically significant group differences in perceptual discrimination patterns across the three stimulus types. Together, these results suggest that all participant groups generally had good pitch discrimination skills and did not appear to be differentially affected by their musical and linguistic backgrounds or the stimulus type, which is also corroborated by the similar patterns in their reaction time scores.

Based on this, one possible explanation for the lack of differences across participants is that the task may have been too easy, and thus could not satisfactorily differentiate the participants by their musical or linguistic backgrounds. However, considering that the same task has been able to differentiate musicians from non-musicians (Burnham et al., 2015), one other possible reason that no effect of musical training was found here could be that the musicians in the current study were amateur musicians. Unlike the musicians in Burnham et al. (2015), who were selected from music institutions, none of the musicians in the current study were majoring, or had ever

majored, in music. In addition, a comparison of musicians in the current study with their non-AP counterparts in the study by Burnham et al. (2015) revealed that they also had fewer number of years of musical training and practice hours per week (see Table 7). Thus, it is possible that the musicians tested by Burnham et al. (2015) generally had higher musical abilities, which enabled a statistically significant effect of musical training to be found.

Table 7

Characteristics of non-absolute-pitch musicians in the current study and in the study of Burnham et al. (2015).

	Current Study	Burnham et al. (2015)
Mean age	21.2 (range: 18 – 27)	25.5 (range: 18.4 – 48.8)
Mean years of musical training	9.9 (range: 3 – 20)	12.2 (range: 4 – 29)
Mean practice hours per week	5.1 (range: 0.5 – 23.0)	9.1 (range: 2 – 20)

Similarly, because an effect of tone language experience has been found using this paradigm in previous studies (Burnham et al., 1996), that no effect of tone language experience was found here may have two other explanations. Firstly, it could be that instead of a facilitative effect, tone language experience hindered subsequent perception of non-native lexical tones instead (which is a possibility based on past literature (Francis et al., 2008; Hao, 2012)). In other words, we would assume that if there had been a facilitative effect, there could have been even higher discrimination scores among tone language speakers. Instead, it could be that the tone participants were disadvantaged by their prior knowledge of Mandarin Chinese tones, which interfered with their perception of the non-native Thai tones, causing them to perform at similar levels as the non-tone participants. From this perspective, it may be that no benefits of

tone language experience were observed because the Thai tonal inventory is slightly more complex than that of Mandarin Chinese. In the study by Burnham et al. (1996), positive transfer benefits of tone language experience were found for Cantonese speakers in the same non-native Thai tone perception paradigm. However, Cantonese has a tonal inventory (6 tones) that is more complex than that of Thai (5 tones). Similarly, in the study by Lee et al. (1996), Mandarin Chinese speakers did not outperform English speakers when discriminating Cantonese tones, which has a more complex tonal inventory than Mandarin Chinese. Although facilitatory effects have been shown for non-native Thai lexical tone perception as a result of prior experience with Mandarin Chinese as a native language (Wayland and Guion, 2004), only two of the five Thai tones were tested in that study, which may be a reason for the different results found presently.

Another possibility, which is not mutually exclusive with the previous explanation, is that there has been an additional confounding factor of tone language exposure. In the study by Burnham et al. (2015), the graded discrimination pattern effect across the speech, filtered speech, and violin stimuli was observed for non-tone non-musicians who not only never had any experience in tone languages, but were also from an environment which was predominantly without ambient exposure to tone languages, i.e., Sydney, Australia. Similarly, Wayland and Guion's (2004) non-tone language speakers were from the Gainesville, Florida, United States, which has a relatively non-tonal linguistic landscape. In contrast, even though all of the non-tone participants in the current study self-reported no native acquisition, formal training or receptive proficiency in any tone languages, most, if not all of them would have been exposed to tone languages for many years from living in multilingual Singapore, which consists of an ethnic Chinese majority who are generally proficient in Mandarin Chinese

and its vernaculars (Department of Statistics of Singapore (2011)). Thus, it is possible that the non-tone participants in this study experienced enhancement in their perception of the non-native Thai lexical tones due to prior long-term exposure to other non-native lexical tones in their ambient surroundings. Subsequently, this could have led to them having comparable discrimination as the tone participants, thereby masking potential facilitatory effects of tone language experience.

Aside from the lack of effects of musical training and language background, what was also interesting from the results was the significant differences in discrimination patterns between musicians and non-musicians when the three stimulus types were re-categorized into linguistic (speech and filtered speech) vs. non-linguistic (violin) categories. In this analysis, musicians had a pattern of better discrimination in linguistic rather than non-linguistic contexts. In contrast, non-musicians had a pattern of better discrimination in non-linguistic rather than linguistic contexts. No such group differences were found when comparing the effects of clear speech vs. filtered speech. Together, these results raise the possibility that some of the musically-trained participants may have been slightly more sensitive to the violin stimuli than the speech and filtered speech stimuli due to their musical background. Consequently, some musicians may have perceived differences among violin tokens to be greater than the differences among the speech and filtered speech tokens respectively, even when the tokens are from the same tone category. If this was the case, discrimination scores would be lowered since within-category tokens meant to be judged as “same” would now be judged as “different”. This possibility is corroborated by how musicians appeared to have greater standard error bars compared to non-musicians in the violin stimuli (non-linguistic) condition (see Figure 11 in Section 5.5.1.1), which suggests more variable performance by the musicians in this condition compared to the non-musicians.

Furthermore, this pattern of heightened sensitivity to pitch differences in the non-linguistic compared to the linguistic condition may have been even more pronounced for musicians who were also tone language speakers. Accordingly, this would explain the trend of discrimination pattern shown by Tm across the speech, filtered speech, and violin stimulus types, i.e., decreasing discrimination abilities in the direction of speech > filtered speech > violin, which is not observed for their non-tone musician counterparts, NTm. Further investigations would be needed, however, to test for this conjecture.

In short, Experiment 1 shows no evidence of any advantages of musicianship or tone language experience. Thus, neither experience outweighed the other, and the combination of both experiences did not result in significant advantages or disadvantages (although musicians had significantly different discrimination patterns for linguistic vs. non-linguistic tones compared to non-musicians, the actual difference in their discrimination of linguistic vs. non-linguistic tones was numerical). Despite this, numerically-speaking, participants with both experiences (Tm) appeared to be more disadvantaged when discriminating musical stimuli in a non-musical context, while participants without any of the two experiences (NTnm) seemed to benefit instead, as seen from the trends of their oppositely-sloped discrimination patterns across the three stimulus types. It is thus possible that some interactive effects may be observed under different conditions, e.g., the trends may have been more pronounced if the Tm group were professional musicians, and if the NTnm group did not have ambient exposure to tone languages.

Overall, we can tentatively conclude that the effects of musical training and linguistic experience may not be as straightforward as it seems. Whether an effect of musical training is observed may depend on finer details, such as the characteristics and musical abilities of the musicians, as well as the difficulty of the perception task. In

particular, the results of Experiment 1 raise the possibility that amateur musicians do not exhibit the same levels of enhancement for non-native lexical tone perceptual discrimination compared to expert musicians. Finally, considering that the non-tone participants in this study largely differed from those tested in previous study in terms of their receptive exposure to tone languages, it remains to be investigated whether tone language exposure itself may yield similar benefits as native experience with tone languages. To answer this question, a secondary data analysis was conducted to compare non-tone participants in the current study, who have been exposed to tone languages, with non-tone participants in Burnham et al. (2015), who have not.

CHAPTER 6: SECONDARY DATA ANALYSIS

6.1 Overview

As the results of non-tone speakers in the tone perception task largely differed from those expected based on the study by Burnham et al. (2015), which utilized the same experimental paradigm, an additional data analysis was conducted. In an attempt to explain why non-tone speakers exhibited equally enhanced discrimination ability as tone speakers, d-prime (d') responses of the non-tone speakers in the present study were compared with those in the study by Burnham et al. (2015).

The basis for this comparison was that non-tone speakers in the present study may have exhibited equivalent discrimination ability as the tone speakers due to an advantage of being in Singapore's tone language exposed environment. Singapore's population consists of an ethnic Chinese majority (74.1% are ethnic Chinese, according to the Department of Statistics of Singapore (2011)), with tone languages such as Mandarin Chinese and other Chinese tone vernaculars such as Hokkien, Teochew, and Cantonese being relatively ubiquitous in the linguistic landscape. In contrast, non-tone speakers from the study by Burnham et al. (2015) were from an environment which was predominantly unexposed to tone languages; they were recruited in Sydney, Australia and reported no experience at all with any tone languages. Therefore, should participants' linguistic environment be an additional factor which influences non-native tone discrimination ability, it was predicted that the tone exposed non-tone (NT^+) speakers from the present study would have higher discrimination scores than the tone unexposed non-tone (NT^-) speakers from the study by Burnham et al. (2015).

6.2 Data analysis

Analysis of the d' data from the present study and from the study by Burnham et al. (2015) was conducted using a 4 x (3) mixed design ANOVAs, with Participant Group (tone exposed/unexposed non-tone musicians/non-musicians) as a between-subjects factor, and Stimulus Type (Speech, Filtered Speech, Violin) as the within-subjects factor. For comparison purposes, only responses from Burnham et al. (2015) which were based on the presentation of the Thai lexical tone stimuli using the same interstimulus interval as in the current study (1500ms) were included in the analysis.

6.3 Results

The mean d' scores of the four non-tone participant groups for each stimulus type are presented in Figure 13. The four groups consisted of 14 tone exposed non-tone musicians (NT^+m), 12 tone unexposed non-tone musicians (NT^-m), 19 tone exposed non-tone non-musicians (NT^+nm), and 13 tone unexposed non-tone non-musicians (NT^-nm).

Figure 13. Mean d-prime scores of 14 tone exposed non-tone musicians (NT⁺m), 12 tone unexposed non-tone musicians (NT⁻m), 19 tone exposed non-tone non-musicians (NT⁺nm), and 13 tone unexposed non-tone non-musicians (NT⁻nm), for each of the three stimulus types (speech, filtered speech, violin), with standard error bars.

Although analyses showed a main effect of stimulus type, $F(2, 108) = 3.40$, $p = .037$, $\eta_p^2 = .059$, post hoc Bonferroni corrected pairwise comparisons revealed no significant differences between all three stimulus type pairs. Discrimination scores in the speech context ($M = 2.51$, $SD = 0.95$) were comparable with those in the filtered speech ($M = 2.63$, $SD = 0.93$) and the violin context ($M = 2.78$, $SD = 0.69$); $p = .873$ and $p = .072$ respectively. Likewise, scores in the filtered speech context were comparable with those in the violin context, $p = .247$. That being said, an overall linear trend was found to be significant for stimulus type, $F(1, 54) = 5.40$, $p = .024$, $\eta_p^2 = .091$, which indicated that discrimination scores generally increased as the stimuli became less speech-like, i.e., speech < filtered speech < violin.

Importantly, a significant Participant Group x Stimulus Type interaction was found, $F(2, 108) = 2.33$, $p = .037$, $\eta_p^2 = .115$. To further investigate this, a series of follow-up tests using one-way ANOVA with planned comparisons was conducted. Two comparisons were made, one to compare musicians with non-musicians, and another to compare tone exposed with tone unexposed participants, both on their discrimination scores across the three stimulus types. Planned contrasts revealed no significant differences between tone exposed and tone unexposed participants across all stimulus types. Tone exposed participants had similar discrimination abilities as tone unexposed participants in the speech, ($M_{exp} = 2.52$, $SD = 1.00$; $M_{unexp} = 2.49$, $SD = 0.91$), $t(54) = 0.195$, $p = .846$, filtered speech, ($M_{exp} = 2.70$, $SD = 1.00$; $M_{unexp} = 2.55$, $SD = 0.84$), $t(54) = 0.730$, $p = .470$, and violin contexts ($M_{exp} = 2.66$, $SD = 0.81$; $M_{unexp} = 2.95$, $SD = 0.47$), $t(32.9) = -1.50$, $p = .143$. However, significant differences between musicians and non-musicians were found. Musicians had better discrimination abilities than non-musicians in the speech context ($M_{mus} = 2.87$, $SD = 0.85$; $M_{non-mus} = 2.21$, $SD = 0.94$), $t(54) = 2.82$, $p = .007$, and the filtered speech context ($M_{mus} = 2.92$, $SD = 0.86$; $M_{non-mus} = 2.40$, $SD = 0.93$), $t(54) = 2.24$, $p = .029$, but not in the violin context, where discrimination scores were comparable ($M_{mus} = 2.84$, $SD = 0.77$; $M_{non-mus} = 2.74$, $SD = 0.63$), $t(32.9) = 0.46$, $p = .648$. (Levene's test indicated unequal variances for the d-prime scores in the violin context, $F(3, 54) = 3.56$, $p = .043$, so degrees of freedom were adjusted from 54 to 32.9.)

Interestingly, a significant linear trend effect was also found for the Participant Group x Stimulus Type interaction, $F(1, 54) = 2.93$, $p = .042$, $\eta_p^2 = .140$. This is interesting because it indicates that there are group differences in the linear trend across stimulus types, therefore suggesting that there are different patterns of perceptual discrimination across groups. Through visual inspection of the data in Figure 13, it

becomes apparent that this effect was driven by differences between musicians and non-musicians across the three stimulus types. For both musician groups (NT^+m , NT^-m), no linear trend in discrimination scores was observed across stimuli types. In contrast, a linear trend can be observed for the non-musician groups (NT^+nm , NT^-nm), whereby discrimination scores for each stimulus type increased as the stimuli become more speech-like, i.e., speech < filtered speech < violin. This indicates that unlike non-musicians, who experienced an attenuation in the ability to process non-native linguistic pitch, musicians had pitch discrimination abilities that were less affected by the context of the pitch itself, implying a music-to-speech transfer benefit for non-native pitch perception.

6.3.1 Follow-up analyses

Strikingly, what is also apparent from Figure 13 is that the aforementioned linear trend effect among the non-musicians appears to be much more pronounced for NT^-nm compared to NT^+nm . Based on the pattern of overlap in error bars, NT^-nm shows better discrimination abilities in the violin context than in both the filtered speech and speech contexts. In contrast, NT^+nm shows discrimination abilities that are more similar across the stimulus types. This raises the possibility that tone language exposure had an effect specifically among non-tone speakers who are non-musicians, by inducing more domain-general pitch discrimination abilities which is reflected in a reduced linear trend effect.

To investigate this, a 2 x (3) mixed design ANOVA was conducted on the two non-tone non-musician groups, with Tone Language Exposure (Exposed, Unexposed) as the between-subjects factor, and Stimulus Type (Speech, Filtered Speech, Violin) as the within-subjects factor. The purpose of this mixed ANOVA was to investigate

whether there was a significant interaction of tone language exposure with stimulus type, which would lend statistical support for the differences in discrimination patterns observed between NT⁺nm and NT⁻nm.

Analyses revealed the Stimulus Type x Tone Language Exposure interaction to be extremely close to statistical significance (taken as $p < .05$), with a medium effect size, $F(2, 60) = 3.15, p = .050, \eta_p^2 = .095$. The linear trend effect for this interaction was also close to statistical significance, $F(1, 30) = 3.97, p = .056, \eta_p^2 = .117$. Given that the p -values were this close to statistical significance despite the unequal sample size (19 NT⁺nm, 13 NT⁻nm), as well as the fact that visual inspection of the data showed distinct group differences in the overlap of error bars across stimulus types, a likely assumption is that had there been more NT⁻nm data available, this interaction would have reached significance.

Taking these factors into consideration, a series of post hoc paired-samples t -tests was conducted within each group to compare discrimination scores in the speech vs. filtered speech condition, filtered speech vs. violin condition, and the speech vs. violin condition. As three pairwise comparisons were made in each participant group, a new significance level for all p -values was set at $p < .017$ by dividing the original threshold of $p < .05$ by three, as per Bonferroni corrections. Based on this, results indicated that while NT⁺nm would have equivalent discrimination abilities across the stimulus types, this would not be the case for NT⁻nm. For NT⁺nm, there would be no significant differences in their discrimination abilities in the speech vs. filtered speech contexts, filtered speech vs. violin context, and speech vs. violin context; $t(18) = -0.92, p = .372, t(18) = -0.56, p = .583$, and $t(18) = -1.32, p = .203$ respectively. In contrast, NT⁻nm would have significantly better discrimination abilities in the violin context ($M = 2.94, SD = 0.59$) compared to both the filtered speech ($M = 2.26, SD = 0.83$) and

speech context ($M = 2.06$, $SD = 0.78$); $t(12) = -4.17$, $p = .001$ and $t(12) = -4.17$, $p = .001$ respectively. No significant differences would be found in NT_{nm}'s discrimination of speech and filtered speech stimuli, $t(24) = -1.25$, $p = .235$.

Thus, had the Stimulus Type x Tone Language Exposure interaction reached statistical significance, these results would have been evidence for an effect of tone language exposure among non-tone non-musicians and not non-tone musicians. Subsequently, this would have provided a premise to support the notion that tone language exposure could have been a confounding factor in Experiment 1. As it is, the present set of data is insufficient to allow rejection of the null hypothesis that tone language exposure does not affect the processing of non-native lexical tones.

6.4 Discussion

The purpose of the secondary data analysis was to investigate whether long-term exposure to tone languages in the ambient environment had facilitatory effects for non-native lexical tone perception among non-tone language speakers. This was done in order to better understand why non-tone language speakers had non-native lexical tone discrimination abilities that were on par as those of tone language speakers in Experiment 1. Specifically, that no effect of tone language experience was found in Experiment 1 could at least partially be because long-term ambient exposure to tone languages had enhanced the performance of the non-tone non-musicians. Subsequently, this would have decreased the differences in performance between tone and non-tone participants, especially since musical training would also have enhanced the performance of the non-tone musicians.

It was thus predicted that tone exposed non-tone participants would have better discrimination scores of the non-native Thai tones compared to the tone unexposed non-

tone participants. Results revealed that this was certainly not the case for musicians, but they were inconclusive when it came to the non-musicians. Regardless of their level of tone language exposure, both non-tone musician groups (NT^+m , NT^-m) had the same discrimination pattern whereby the perception of non-native lexical tones was equally good across stimulus types. This indicates that there was no effect of tone language exposure on musicians' non-native lexical tone perception. In contrast, there was a trend of different discrimination patterns within the non-tone non-musician groups (NT^+nm , NT^-nm). While NT^-nm showed an attenuation in linguistic ability to perceive non-native lexical tones, as evident from the gradated discrimination scores which increased as the tone stimuli became less speech-like, this attenuation was much less pronounced among NT^+nm . However, most likely due to unequal sample sizes, this trend of differences within the non-tone non-musicians did not reach statistical significance, and therefore could not be clearly attributed to their differences in tone language exposure.

Thus, what requires further investigation is whether tone language exposure plays a role in mitigating the attenuation in non-native tone perception abilities among non-tone non-musicians. Given that this trend of differences between NT^+nm and NT^-nm was very close to statistical significance despite the unequal sample sizes, it is possible that further research addressing this may find an interaction effect between tone language exposure and musical training. Specifically, it may be that tone language exposure influences pitch discrimination in the absence of musical training, by reducing the gradation effect in the discrimination patterns of pitch from a speech to non-speech context. As it is, the results of the secondary data analysis were not able to conclusively show that tone language exposure had been a confounding factor in Experiment 1 through an enhancement of performance in the $NTnm$ group which may have

subsequently reduced differences across the four participant groups. Nevertheless, the non-significant results are still promising and worthy of further investigations.

Generally, in line with past findings on non-tone populations, the results support the view that musical training has music-to-speech transfer benefits. Given that non-tone musicians did not exhibit the attenuation in linguistic ability to perceive non-native lexical tones which is seen among non-tone non-musicians, it may be inferred that this natural attenuation has been mitigated as an effect of musical training. Furthermore, non-tone musicians were also found to have better discrimination scores than non-tone non-musicians in the speech and filtered speech conditions. Therefore, the implication here is that musical training has cross-domain effects on pitch processing, and this has benefits for the perception of non-native lexical tone in contexts closer to natural speech. For future research, it would be interesting to see whether this facilitatory effect of musical training was found because the sample analysed were from non-tone populations, or if this was driven by the additional presence of expert musicians.

In conclusion, while there was no conclusive support for an effect of tone language exposure on the non-native lexical tone discrimination abilities of non-tone non-musicians, given the inequality in sample sizes, a promising trend was nonetheless observed. What is interesting and adds on to the current literature, however, is that further ambient exposure to tone languages does not have additive benefits to the advantage of musicianship for non-native lexical tone discrimination. To sum it up, just as the characterisation of musical training lies on a continuum (e.g., from non-musician to amateur musician to expert musician), so can the characterisation of tone language experience (e.g., from being receptively exposed to tone languages to being a native tone language speaker). Together with the results of Experiment 1, the present secondary

data analysis raises the possibility that different effects on non-native lexical tone perception may be observed at different points of each continuum.

CHAPTER 7: EXPERIMENT 2

7.1 Overview

The purpose of Experiment 2 was to investigate how musical training and tone language experience affect the production of non-native lexical tones. In this experiment, the same participants from Experiment 1 listened to the Thai stimuli used in Experiment 1 and attempted to imitate it to the best of their abilities. Subsequently, a separate group of native Thai speakers was recruited as informants to evaluate the participants' Thai tone production through a tone identification task and a tone rating task. In the former, informants identified the tone category of each produced token, and in the latter, they rated how well each token had been produced based on their own judgment.

As previously mentioned (see Chapter 3 Section 3.3), it was hypothesized that both musical training and tone language experience would enhance non-native lexical tone production. Whether one factor would have a greater facilitative effect than the other, however, was debatable. In addition, it was also tentatively hypothesized that there would not be additive effects of both factors. Although no effect of either factor was observed in the non-native lexical tone perception task, producing non-native tones is arguably more challenging than perceiving them. Thus, it was predicted that there would be higher identification rates and ratings of the produced tokens by musicians compared to non-musicians, and by tone language speakers compared to non-tone-language speakers. Other than that, it was predicted that tone-musicians (Tm) would be evaluated similarly as tone non-musicians (Tnm) and non-tone musicians (NTm) in both tone identification and tone rating tasks.

7.2 Tone production task

7.2.1 Stimuli

The same five words used in Experiment 1's tone perception task were used to test participants' production of non-native lexical tones. Unlike the tone perception task, only one exemplar per lexical tone was used in the tone production task. Besides that, stimuli recorded by a female native Mandarin Chinese speaker in Tan's (2013) study were also included for use in practice trials. The stimuli were two /da/ syllables carrying the Mandarin Chinese high level (55) tone and high rising (35) tone respectively, i.e., /dǎ//and /dǎ/.

7.2.2 Apparatus and procedure

The tone production task was presented to participants via the E-Prime 2.0 software (Psychology Software Tools, 2016) on a desktop computer, with the stimuli played to participants through a pair of Sennheiser HD 380 Pro headphones. Voice recordings of the participants during the tone production task were made using a Marantz PMD620 handheld audio recorder.

Participants listened to one exemplar set of the five Thai tones in the following order: first, mid (32) [pā:], then low (21) [pà:], followed by falling (51) [pâ:], high (45) [pá:], and finally rising (214) [pǎ:]. After each tone was presented, participants were required to imitate and reproduce the tone verbally. The same exemplar set of the five Thai tones was presented three times to give a total of 15 trials. Before the experimental trials, participants were given two practice trials where they listened to and reproduced two Mandarin Chinese syllables, a high level (55) [tǎ] and a high rising (35) [tǎ].

7.3 Tone evaluation tasks

7.3.1 Informants

A total of 17 native Thai speakers were recruited as informants. They consisted of nine females and eight males aged 21-50 years old ($M_{age} = 28.24$ years, $SD = 7.34$). All informants were from Bangkok, Thailand, and had been in Singapore ranging from 4 months to 17 years at the time of the experiment ($M = 4.53$ years, $SD = 5.79$).

7.3.2 Stimuli and apparatus

Stimuli in the tone identification and tone rating tasks consisted of the Thai tone tokens produced by participants in the tone production task. Only tone tokens from participants in the final sample ($N = 70$) were included, thus making the total number of tone tokens in each task to be 1050 (70 participants x 5 tones x 3 repetitions); 270 tone tokens for Tm ($n = 18$), 285 for Tnm ($n = 19$), 210 for NTm ($n = 14$), and 285 for NTnm ($n = 19$) respectively. Stimuli in both tasks were presented using the E-Prime 2.0 software (Psychology Software Tools, 2016) through a pair of Sennheiser HD 280 Pro headphones, on a laptop.

7.3.3 Procedure

Informants first completed the tone identification task, followed by the tone rating task. In each task, the 1050 tone tokens were presented in two stimulus blocks (35 speakers x 5 tones x 3 tokens x 2 stimulus blocks), which were counterbalanced in order of presentation across listeners. In each stimulus block, the presentation of tone tokens was further blocked by speaker to assist in the listeners' normalization of each speaker's pitch range. Tone tokens within each speaker block were presented in randomized order. All tokens were amplitude normalized to 70dB, with further adjustments when

necessary to ensure that they could be as perceptually similar as possible in loudness, especially across speakers.

To ensure that the two stimulus blocks were comparable, the distribution of speakers was pseudorandomized such that each stimulus block contained a similar number of speakers from each of the four participant groups (Tm, Tnm, NTm, and NTnm).² Lastly, to ease experimental fatigue, self-paced breaks were included between every speaker block, and a short break was given after every stimulus block.

In the tone identification task, listeners were instructed to identify the word which was most closely associated with each tone token based on their own judgment. For example, if a token was perceived as a Thai mid level (32) tone, listeners pressed “1” to select the word “ป้า” [pā:], which is associated with the mid level (32) tone. In a similar manner, listeners pressed “2” to select the word “ป๊า” [pà:] which is associated with the low level (21) tone, “3” for “ป๋้า” [pâ:], which has a falling (51) tone, “4” for “ป๋า” [pá:], which has a high level (45) tone, and “5” for “ป๊้า” [pǎ:], which has a rising (214) tone.

As some tokens had mispronunciations in the consonant segment of the syllable, listeners were informed about this, and specifically instructed to ignore the consonant in each token and instead focus only on the tone when making their decisions. In particular, it was important that the Thai listeners understood the nature of this task, given that tone learning is usually implicit in Thai, and dependent on various rules such

² Both stimulus blocks had 9 Tm speakers and 7 NTm speakers. One block had 11 Tnm speakers and 8 NTnm speakers, while the other had 8 Tnm speakers and 11 NTnm speakers.

as the consonant grapheme and vowel characteristics. To help listeners understand that they were to identify each token based solely on its tone, examples were given by the experimenter: E.g., assuming that a falling (51) [pâ] was mispronounced as a falling (51) [dâ], listeners would then still press the button “3” to identify it as being closest to the word “ป้า” [pâ:], which has the falling (51) tone. All listeners indicated that they understood the instructions, and all passed a practice task in tone identification before completing the experimental trials. The practice task consisted of 15 trials, where the same 15 tone tokens spoken by the female native Bangkok Thai speaker in the tone perception and production tasks were presented.

While the tone identification task assessed the extent to which Thai tones produced by non-Thai speakers could be correctly recognized by native Thai listeners, the tone rating task assessed the extent to which the produced tones were perceived to be good examples of their intended Thai tone categories. In each tone rating trial, listeners were auditorily presented with the tone token, followed by the word associated with its intended tone category. For instance, the word “ป้า” [pā:], which is associated with the mid level (32) tone, was shown after all trials in which mid level tones were auditorily presented. Listeners were then instructed to rate each tone token on a scale of 1 (very poor example) to 7 (very good example) based on their own judgments of how well it had been produced.

As in the tone identification task, listeners were explicitly instructed to ignore the consonant segment of each token and to focus only on the tone when making their decisions in the tone rating task. Similarly, to help listeners understand that they were to rate each token based solely on its tone, examples were given by the experimenter: E.g., assuming that a falling (51) [pâ:] was the target tone to be produced, a mispronunciation of it as a falling (51) [dâ:] would still receive a high rating, while a

mispronunciation of it as a mid level (32) [pā:] would receive a low/lower rating. Listeners also completed a practice task with the same 15 trials and tone tokens presented in the tone identification task before completing the experimental trials.

7.4 Data analysis and results

Responses in the tone identification task were analysed using percentage accuracy and confusion matrices. Due to experimental error,³ four trials were excluded from the analyses. The analysis of percentage accuracy data was conducted with a 2 x 2 x (5) mixed design ANOVA, with Musical Background (Musicians, Non-Musicians), and Language Background (Tone Speaker, Non-Tone Speaker) as between-subjects factors, and Tone Category (Mid, Low, Falling, High, Rising) as the within-subjects factor.

7.4.1 Tone identification task

7.4.1.1 Accuracy data

Figure 14 illustrates the accuracy in Thai tone production across the four participant groups, based on the mean of tone identification accuracy scores (in percentage) across the 17 native Thai listeners.

³ In the tone identification task, listeners were supposed to choose from five response options labelled “1” to “5”. Data was thus logged whenever the buttons “1” to “5” were pressed. A typo in the coding, however, resulted in data also being logged when the button “6” was pressed. This happened for four trials, most likely from listeners accidentally pressing “6” when “5” was intended. The typo was rectified immediately upon discovery, and the aforementioned four trials discarded from the analyses.

Figure 14. Thai tone production accuracy of 18 tone musicians (Tm), 19 tone non-musicians (Tnm), 14 non-tone musicians (NTm), and 19 non-tone non-musicians (NTnm), based on the percentage of correct tone identification by 17 native Thai speakers. This figure illustrates the extent to which non-native Thai lexical tones (mid 32 (M), low 21 (L), falling 51 (F), high 45 (H), rising 214 (R)) produced by each participant group were correctly identified by native Thai listeners.

As Mauchly's test indicated that the assumption of sphericity had been violated ($\chi^2(9) = 97.5, p < .001$), degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity ($\epsilon = 0.56$). Analyses showed a significant main effect of tone category, $F(2.25, 148.62) = 18.23, p < .001, \eta_p^2 = .22$, but no main effect of musical training, $F(1, 66) = 1.09, p = .300, \eta_p^2 = .016$, or tone language experience, $F(1, 66) = 1.13, p = .291, \eta_p^2 = .017$. This indicated that musicians did not produce more intelligible Thai tones than non-musicians (tone production accuracy: 83.4% vs. 81.4% respectively), and the same for tone language speakers compared to non-tone language speakers (tone production accuracy: 81.3% vs. 83.4% respectively). Post hoc pairwise

comparisons with Bonferroni correction revealed falling (51) tone tokens to have the highest rate of correct identification compared to other tone tokens (95.2%; all $ps < .001$), followed by mid (32), low (21), and rising (214) tone tokens, which had comparable rates of correct identification with each other (83.5%, 84.1%, and 80.4% respectively; all $ps = 1.000$). High (45) tone tokens generally had the lowest rate of correct identification (68.2%; all $ps < .05$ except when compared to rising (214) tone tokens ($p = .133$)).

The interaction between musical background and language background was not found to be significant, $F(1, 66) = 0.002$, $p = .961$, $\eta_p^2 = .00004$, and neither were other interactions: Tone Category x Musical Background, $F(2.25, 148.62) = 0.46$, $p = .657$, $\eta_p^2 = .007$; Tone Category x Language Background, $F(2.25, 148.62) = 0.38$, $p = .712$, $\eta_p^2 = .006$, Tone Category x Musical Background x Language Background, $F(2.25, 148.62) = 0.19$, $p = .850$, $\eta_p^2 = .003$.

7.4.1.2 Confusion matrices

Overall, the tone identification accuracies across participant groups were comparable: 82.4% for Tm, 80.3% for Tnm, 84.8% for NTm, and 82.4% for NTnm. To examine whether the tone confusion patterns across participant groups were also comparable, confusion matrices were constructed for each participant group. Within each participant group, the percentage of misidentifications of their produced tones was calculated by tone category and averaged across the 17 Thai listeners. For reference, the percentages in each tone category were calculated based on a total of 918 trials for Tm,

969 trials for Tnm,⁴ 714 trials for NTm, and 969 trials for NTnm (17 listeners x number of participants in each group x 3 tone tokens).

Table 8 presents the confusion matrices of the 17 native Thai listeners' identification of the five Thai lexical tones produced by Tm, Tnm, NTm, and NTnm respectively. Based on Table 8, it can be observed that the tone confusion patterns for Tm, NTm, and NTnm were largely similar. For example, across the three groups, identification errors for mid (32) tones were mainly due to confusion with low (21) tones, and vice versa. However, for Tnm, identification errors for mid (32) tones were also due to confusion with high (45) tones, while identification errors for low (21) tones were also due to confusion with rising (214) tones. Besides that, identification errors of Tm, NTm, and NTnm for high (45) tones were mainly due to confusion with just mid (32) tones, but for Tnm, it was also due to confusion with rising (214) tones.

To put these differences into perspective, an error threshold was calculated as a basis to determine whether the observed differences were significant or more a result of human error during the experiment (e.g., unintentionally pressing the incorrect response key). This was especially so given that each Thai listener had to respond 1050 times in the identification task. By averaging the error rates of the falling tone, i.e., the most accurately identified tone across the four participant groups, it can be surmised that errors below the threshold of 4.63% were most likely due to human errors during the

⁴ Due to experimental error, four trials produced by the Tnm were excluded; three from the high tone category, and one from the rising tone category. Thus, for the high tone category and the rising tone category, Tnm had 966 and 968 trials respectively instead of the expected 969 trials.

experiment. To aid visualization of the confusion matrix results given this estimated error threshold, Figure 15 presents the tone confusion rates of the five Thai tones across the four participant groups, with data labels included for percentages of tones that were correctly identified, as well as those which exceeded the error threshold.

Taking the error threshold into consideration, most of the previous observations generally hold true. To aid visualization, the patterns of tone confusion across the four participant groups based on the error threshold analysis are summed up in Table 9. The analysis is as follows: Firstly, across the participant groups, errors in the identification of mid (32) tones were due to confusion with low (21) tones, while errors in the identification of low (21) tones were due to confusion with mid (32) tones. For Tm and Tnm, however, mid (32) tones were further confused with high (45) tones (4.68% and 7.02% respectively). For Tnm, low (21) tones were also confused with rising (214) tones (7.64%). Across participant groups, falling (51) tones were rarely confused, but the error threshold analysis further revealed that for Tnm, falling (51) tones were confused with low (21) tones (5.99%). Besides that, errors in the identification of high (45) tones were generally due to confusion with mid (32) tones, but for Tnm, they were also caused by confusion with rising (214) tones (12.50%). For errors in the identification of rising (214) tones, all participant groups were similar in that the errors were due to confusion with high (45) tones. Therefore, the results converged to reveal that Tnm generally produced more confusable tones compared to the other participant groups, followed by Tm. Overall, NTm showed the best production performance; they generally received the highest identification scores across tone categories and produced the least confusable tones.

Table 8

Confusion matrices (in percentages) of Thai tones (mid 32 (M), low 21 (L), falling 51 (F), high 45 (H), rising 214 (R)) produced by (a) 18 tone musicians (Tm), (b) 19 tone non-musicians (Tnm), (c) 14 non-tone musicians (NTm), and (d) 19 non-tone non-musicians (NTnm), as judged by 17 native Thai listeners.

(a) Tm						(c) NTm							
Perceived Tone Category	Intended Tone Category					Perceived Tone Category	Intended Tone Category						
	M	L	F	H	R		M	L	F	H	R		
	M	80.39	8.61	0.98	25.93		0.00	M	83.47	7.84	0.00	22.97	2.38
	L	13.29	85.19	2.51	0.11		0.44	L	12.89	87.11	0.56	0.98	2.52
	F	0.54	1.74	94.99	1.53		0.00	F	2.24	2.10	98.46	2.10	0.42
	H	4.68	0.54	1.20	69.93		18.08	H	1.26	0.42	0.98	71.99	11.90
	R	1.09	3.92	0.33	2.51		81.46	R	0.14	2.52	0.00	1.96	82.77

(b) Tnm						(d) NTnm							
Perceived Tone Category	Intended Tone Category					Perceived Tone Category	Intended Tone Category						
	M	L	F	H	R		M	L	F	H	R		
	M	82.46	8.05	1.86	20.04		0.21	M	87.41	10.32	0.21	25.39	1.65
	L	7.53	80.80	5.99	0.31		0.62	L	10.42	84.31	1.34	0.83	0.93
	F	1.14	1.75	90.71	1.65		0.52	F	1.24	1.75	97.32	2.79	0.62
	H	7.02	1.75	1.44	65.50		16.67	H	0.93	0.31	1.03	66.67	20.33
	R	1.86	7.64	0.00	12.50		81.99	R	0.00	3.30	0.10	4.33	76.47

Figure 15. Tone confusion rates of participants' Thai tone tokens (mid 32 (M), low 21 (L), falling 51 (F), high 45 (H), rising 214 (R)), as judged by 17 native Thai listeners. Graph (a) shows the tone confusion rates for tones produced by 18 tone musicians (Tm). Graph (b) shows the tone confusion rates for tones produced by 19 tone non-musicians (Tnm). Data labels are included for the percentages of tones that were correctly identified, as well as for error rates that were found to exceed the error threshold of 4.63%.

Figure 15 (continued). Tone confusion rates of participants' Thai tone tokens (mid 32 (M), low 21 (L), falling 51 (F), high 45 (H), rising 214 (R)), as judged by 17 native Thai listeners. Graph (c) shows the tone confusion rates for tones produced by 14 non-tone musicians (NTm). Graph (d) shows the tone confusion rates for tones produced by 19 non-tone non-musicians (NTnm). Data labels are included for the percentages of tones that were correctly identified, as well as for error rates that were found to exceed the error threshold of 4.63%.

Table 9

Summary of tone confusion patterns across tone musicians (Tm), tone non-musicians (Tnm), non-tone musicians (NTm), and non-tone non-musicians (NTnm) for each of the Thai lexical tone categories (mid (32), low (21), falling (51), high (45), rising (214)).

Tone Confusion Patterns				
Tone Category	Tm	Tnm	NTm	NTnm
Mid (32)	→ Low (21) → High (45)	→ Low (21) → High (45)	→ Low (21)	→ Low (21)
Low (21)	→ Mid (32)	→ Mid (32) → Rising (214)	→ Mid (32)	→ Mid (32)
Falling (51)		→ Low (21)		
High (45)	→ Mid (32)	→ Mid (32) → Rising (214)	→ Mid (32)	→ Mid (32)
Rising (214)	→ High (45)	→ High (45)	→ High (45)	→ High (45)

Note. The symbol “→” indicates “is confused with”. Tone confusion patterns which are not observed for all participant groups are bolded, and the Tnm column is shaded to highlight the distinctive tone confusion patterns of tone tokens produced by Tnm.

7.4.2 Tone rating task

Figure 16 shows the rating scores of the Thai tones produced by the four participant groups, as judged by the 17 native Thai listeners. Overall, participants' rating scores were comparable: 5.57 for Tm, 5.50 for Tnm, 5.60 for NTm, and 5.54 for NTnm. On average, falling (51) tones received the highest rating scores, while high (45) tones received the lowest rating scores. Thus, the rating scores showed the same patterns as the tone identification scores across tone categories.

Figure 16. Rating scores of Thai tone tokens (mid 32 (M), low 21 (L), falling 51 (F), high 45 (H), rising 214 (R)) produced by 18 tone musicians (Tm), 19 tone non-musicians (Tnm), 14 non-tone musicians (NTm), and 19 non-tone non-musicians (NTnm), as judged by 17 native Thai speakers. Rating scores are on a scale of 1 (very poor example of the intended tone category) to 7 (very good example of the intended tone category).

7.5 Discussion

The purpose of Experiment 2 was to investigate whether the benefits of musical training and tone language experience for non-native lexical tone perception, as detailed by past literature, would also extend to benefits for non-native lexical tone production. This was because perceptual attunement to pitch as a result of either experience would presumably enhance auditory feedback, and subsequently improve pitch production abilities. Thus, it was hypothesized that both musical training and tone language experience would improve non-native lexical tone production. This means that musicians and tone language speakers were predicted to receive more favourable evaluations of their Thai tone production than non-musicians and non-tone language speakers respectively. Whether musicians would outperform tone language speakers or vice versa, however, was an open question. Tentatively, it was also hypothesized that there would not be any additive effects observed, which would mean that Tm would not outperform Tnm and NTm.

When the production results were considered in terms of accuracy, no effect of musical training or tone language experience was found. Across participant groups, the produced tone tokens were generally comparable in intelligibility and well-formedness, as determined by the accuracy scores of Thai listeners in the identification task, and as further corroborated by the comparable ratings of the tone tokens across participant groups. Moreover, participant groups generally had similar production abilities across tone categories, with the falling (51) tone being the easiest to produce, and the high (45) tone being the hardest to produce. In this sense, the lack of group differences observed here mirrored the perception results in Experiment 1.

On the other hand, when the production results were considered in terms of error rate patterns, an effect of musical training, of tone language experience, and an interaction effect between both factors was revealed. The effect of musical training is seen from the fact that the NTm group could be said to have the best performance in production: They had the highest accuracy rates overall and also produced the least confusable tone tokens. This draws parallel with the findings of Cooper and Wang (2012), where non-tone English musicians had the best perception abilities in Cantonese tone identification.

Meanwhile, the effect of tone language experience can be seen in the different error rate patterns for Tnm and Tm compared to the rest of the participant groups. In general, participants had mainly shown difficulties in producing tones of different pitch heights: Across participants, there was more confusion among level tones than contour tones. When misidentified, mid (32) tones and low (21) tones were confused with one another, while high (45) tones were confused with mid (32) tones. Among contour tones, falling (51) tones were almost never misidentified, while rising (214) tones, when misidentified, tended to be confused with high (45) tones, which is most probably due to the fact that Thai high (45) tones tend to end with a slight rise. When coupled with a poor differentiation of pitch height in tone production abilities, it is reasonable to see how rising (214) tone tokens produced by the participants could have been produced at the incorrect pitch height, and subsequently misidentified as high (45) tones by the native Thai listeners.

However, only the two tone-language-speaking groups, Tnm and Tm, showed differences from these general error patterns across tone categories. For Tnm and Tm, mid (32) tones were also confused with high (45) tones. This could be because the Mandarin Chinese tone categories can be differentiated using pitch contour alone – there

is only one level tone in Mandarin Chinese (high level 55). Accordingly, this may have made it harder for tone participants to produce tones of different heights as in Thai. Tone participants may have assimilated the Thai mid (32) tone to the Mandarin Chinese high level (55) tone category, resulting in them producing Thai mid (32) tones which sounded like Thai high (45) tones. In addition, only for Tnm, low (21) tones were also confused with rising (214) tones. This could be due to an assimilation of the Thai low (21) tones into the Mandarin Chinese falling rising (214) tone category, given that the two tone categories share the same initial pitch characteristics. Furthermore, only for Tnm, high (45) tones were also confused with rising (214) tones, suggesting that they may have over-exaggerated the slight rise at the end of Thai high tones, making the pitch slope more similar to that of the rising (35) tone in Mandarin Chinese. In addition, Tnm produced falling (51) tones which were confused with low (21) tones, which is not observed for any of the other participant groups. It could be that there were some assimilation of the Thai falling (51) and low (21) tones into the same category by Tnm based on the fact that both of these tones have a falling slope. Thus, compared to the other participant groups, the tone participants, especially Tnm, produced Thai tokens that were more confusable, which suggests a potential phonological interference effect due to their prior experience with Mandarin Chinese tones.

Strikingly, the results also pointed to an interaction effect of musical training and tone language experience. This was shown by how Tm only had more confusable tones than other participant groups in their mid (32) tone production, as opposed to Tnm, who had more confusable tones in all tone categories except rising (214) tones. Thus, tone tokens produced by Tnm were the most confusable compared to those produced by the other participant groups, and this was followed by those produced by Tm. This implies that for the tone-language speakers, the disadvantages they exhibited in non-native Thai

lexical tone production due to phonological interference from prior experience with Mandarin Chinese tone categories was mitigated in the presence of musical training.

Given that these results contrast with the lack of group differences observed in Experiment 1, it is likely that task difficulty as well as the type of analysis conducted were important factors in manifesting this disadvantage of tone language experience. Considering that non-native production is generally more difficult than non-native perception, this indicates that the effects of tone language experience and musical training could only be manifested in the context of a more difficult task. Furthermore, a more fine-grained analysis with measures other than overall accuracy rates was also needed to show these effects.

Even so, it should be noted that the overall accuracy rates in identification and ratings of the tone tokens were generally high, which suggests that the disadvantage of tone language experience found here may not actually impact communication detrimentally in real-world contexts. Besides that, as in Experiment 1, there is also a possibility that tone language exposure had a role to play in the results of this experiment, given that the NTnm group had comparable accuracy rates with the other groups and also had fewer tone confusions in their tone tokens than Tnm and Tm. Consequently, this may have caused the disadvantages of tone language experience to appear more prominent than they might have been in the absence of this potential confound.

Overall, Experiment 2 contributes to the lack of literature on potential interaction effects between musical training and tone language experience in non-native lexical tone processing, particularly in the domain of production. The results further support previous findings of a music-to-speech transfer effect in the domain of pitch. More importantly, they extend past findings of the musicianship advantage in non-native

lexical tone processing from the domain of perception to that of production. In addition to this, evidence was found for a disadvantage of tone language experience in non-native lexical tone production, as well as a subsequent mitigating effect of musical training, all of which are novel findings.

CHAPTER 8: GENERAL DISCUSSION

The present study examined two important factors affecting the processing of non-native lexical tones, i.e., musical training and tone language experience, in tandem. Previous studies have mainly focused on the effects of each factor independently, which has limited our understanding of how the two factors may interact with each other. In particular, the present study had the following research questions:

- (1) Do the perceptual benefits of musical training apply to tone language speakers?
- (2) Between musical training and tone language experience, does one bear more weight in non-native lexical tone processing?
- (3) Are the effects of musical training and tone language experience additive?
- (4) In addition to the perceptual benefits, does musical training and tone language experience also benefit the production of non-native lexical tones?

The independent and interactive effects of the two factors on the perception of non-native lexical tones were investigated in Experiment 1, while the same for the production of non-native lexical tones were investigated in Experiment 2.

8.1 Summary of main results and findings

No statistically significant independent or interactive effects of musical training or tone language experience were found in both experiments. Secondary data analysis using perception data from non-tone participants in the present study and from the study by Burnham et al. (2015) found no statistically significant effects of tone language exposure to indicate that this confounded the results of Experiment 1. However, further analysis on the production data based on patterns of error rates revealed that tone participants produced non-native Thai tones which were more confusable compared to

non-tone participants. Notably, the patterns of tone confusion of the tone participants could be explained by the perceptual similarities between their native Mandarin Chinese tones and the non-native Thai tones. Furthermore, tone participants who were musicians produced less confusable tones than their non-musician counterparts, and across groups, the non-tone musicians had the least confusable tones.

Thus, no evidence was found for any perceptual benefits of musical training applying to tone language speakers in this study (because no facilitative effects of musical training on the perception of non-native lexical tones were observed in the first place). Besides that, neither musical training nor tone language experience appeared to bear more weight than the other during non-native lexical tone perception, and there were also no additive effects of the two factors. In contrast, musical training facilitated the production of non-native lexical tones, while tone language experience led to disadvantages. The effects of the two factors on the production of non-native lexical tones were interactive, such that the disadvantages posed by tone language experience were ameliorated in the presence of musical training. Musical training thus appears to be more influential than tone language experience during non-native lexical tone production.

8.2 Influence of task difficulty and sensitivity of analysis used

From the start, what was striking about the results was the lack of statistically significant effects of musical training and tone language experience on non-native lexical tone perception. This ran contrary to expectations based on past studies which have found perceptual benefits of musicianship among tone language speakers and non-tone language speakers (Alexander et al., 2005; Burnham et al., 2015; Dawson et al., 2018; Lee & Hung, 2008; Nan et al., 2018), and either facilitative or inhibitive effects

of tone language experience during non-native lexical tone perception (Burnham et al., 1996; Francis et al., 2008; Hao, 2012; Krishnan, Gandour, & Bidelman, 2010; Lee et al., 1996; Wayland & Guion, 2004). This was especially the case considering that the same experimental paradigm used by Burnham et al. (2015) and Burnham et al. (1996), which found effects of musical training and tone language experience respectively was adopted in this study.

This had led to a few speculations. Firstly, it could be that the perception task was not challenging enough to differentiate the musicians from the non-musicians or the tone language speakers from the non-tone language speakers in this participant sample. The Thai tone stimuli could have simply been acoustically contrastive enough to all participants in this study. This was corroborated by the good discrimination scores across groups.

On the other hand, there were other possible explanations aside from task difficulty. The lack of significant effect of musical training could have been due to the levels of musical expertise among the musicians of this study. Compared to the professional musicians tested by Burnham et al. (2015), the musicians in this study were amateur musicians who had fewer years of musical training as well as fewer practice hours per week. It could be the case that a certain threshold of musical ability had to be met before an effect of musical training could be manifested. Meanwhile, the lack of significant effect of tone language experience could have been caused by tone language participants being hindered by their prior experience with Mandarin Chinese tones, which led to them performing at similar levels as non-tone language speakers. It could also have been caused by non-tone participants having better discrimination of non-native lexical tones as a result of their long-term ambient exposure to tone languages, but this was ruled out based on the results of the secondary data analysis.

Once the results of the production task were taken into consideration, however, this list of possible interpretations could be narrowed down further. Given that perception generally precedes production in the acquisition of both L1 and L2 languages, it would be reasonable to assume that production tasks are generally more difficult than perception tasks. Thus, if the perception task was not challenging enough to differentiate the participant groups, perhaps group differences would be revealed in the context of the more demanding production task instead.

As it is, group differences did emerge in the production task, revealing facilitatory effects of musical training which mitigated the inhibitive effects of tone language experience. That such benefits of musical training could be observed even though the musicians were amateur rather than professional musicians suggests that the level of musical expertise may not have been that crucial in affecting the perception task results after all. It also provides support for the notion that the perception task was relatively too easy for the effects of musical training and tone language experience to be manifested.

Even so, these group differences in tone production were only apparent when a more sensitive and fine-grained measure was used, i.e., error patterns, instead of merely looking at the tone identification accuracies or tone rating scores, which were generally good across groups, just like the discrimination scores in the perception task. Based on this, the most likely inference is that even though the production task was more challenging than the perception task, participants in this study had found the Thai tone stimuli to be sufficiently contrastive and reasonably pronounceable. Thus, the production task can still be considered easy in general. This would explain why no statistically significant effects of musical training or tone language experience were

found in the production task despite their subtle but meaningful effects on participants' tone confusion patterns.

Accordingly, it may be gleaned that if the perception and the production tasks are increased in difficulty level, more apparent differences may be observed among musicians and non-musicians as well as tone language speakers and non-tone language speakers in their non-native lexical tone processing.

8.3 The double-edged sword of tone language experience

Assuming that task difficulty was the key to explaining the lack of statistically significant results in Experiment 1, it may now be surmised that Mandarin Chinese tone language experience did not hinder non-native Thai lexical tone perception in Experiment 1. Instead, tone language speakers were hindered by their tone language background only when it came to the production of these tones. It may thus be posited that the perceptual similarities between Mandarin Chinese tones and Thai tones posed no inhibitive effects on non-native Thai tone perception but was disadvantageous when it came to non-native Thai tone production.

In the perception task, sensitivity to the contrasts in Thai lexical tones in the dimensions of pitch height and pitch contour is needed to discriminate the Thai tones well. Given that this is similar to the way Mandarin Chinese tones can be contrasted, it is conceivable that tone participants would not have faced too much difficulties when having to discriminate the non-native Thai tones. As established previously, similar to Thai tones, Mandarin Chinese tones can also be contrasted in pitch height and pitch contour (although Mandarin Chinese tones can arguably be distinguished by pitch contour alone since there is only one level tone (high level 55), and research has shown native Mandarin Chinese listeners to give more weight to pitch contour than pitch height

(Gandour, 1984; Massaro, Cohen, & Tseng, 1985)). Furthermore, Mandarin Chinese has tones which are similar to some Thai tones in acoustic characteristics, e.g., both languages have high falling (51) tones and low falling rising (214) tones (termed rising tones for the Thai tones in the present study, based on the terminology of Burnham et al. (2015)). It thus follows that an assimilation of some Thai tones to Mandarin Chinese categories based on perceptual similarity in pitch height and/or pitch contour could have even been helpful to tone participants in their Thai tone discrimination.

However, in the production task, an ability to closely imitate the pitch height and pitch contour of each Thai tone category was needed to achieve distinct differentiation between the tones, and thereby produce them well. In this case, an assimilation of Thai tones to Mandarin Chinese categories appeared to be less desirable, as it led to the tone participants producing the Thai tones in ways that were influenced by their attunement to Mandarin Chinese tone categories. Perhaps the clearest example of this is in the Thai falling (51) tones production. Across all participant groups, the Thai falling (51) tone was the most accurately produced tone. Strikingly, only the tone non-musicians (Tnm) had their produced Thai falling (51) tones misidentified by the native Thai listeners as Thai low (21) tones to an extent that was above the error threshold. This indicates a phonological influence of Mandarin Chinese tones, as the Thai falling (51) tone is similar to the Mandarin Chinese (21) tone in pitch contour.

Such a phonological influence of Mandarin Chinese tones could also be observed in the tone participants' production of the Thai level tones. Across the participants, only tone participants produced Thai mid (32) tone tokens that were further misidentified as Thai high (45) tones in addition to Thai low (21) tones. This suggests that tone participants had assimilated the Thai mid (32) tones to the Mandarin Chinese high level (55) tone category because of their similarity in pitch contour. Subsequently,

this influenced the height of the tone participants' Thai mid (32) tone tokens, leading to them being misidentified as Thai high (45) tones. Across groups, the Tnm group also produced the most confusable Thai low (21) tone tokens and Thai high (45) tone tokens. Together, these indicate that relative to the non-tone participants, tone participants faced more difficulty in pronouncing the Thai level tones distinctly. This is most likely attributable to their attunement to Mandarin Chinese tone categories, which do not require them to contrast level tones of different pitch heights.

These patterns of L2 to L1 tone assimilation among tone participants may have mattered less in the perception task because of the increased saliency in contrasts when tones are presented in pairs. For instance, even if Thai low (21) tones were similar to Thai mid (32) tones in pitch contour, tone participants, being sensitive to pitch, may have found it easier to also perceive the differences in pitch height between the two tones in the perception task, where the Thai tones are presented in pairs for discrimination. Conversely, in the production task, the Thai tones were presented in isolation, which may have subjected the tone participants to more phonological influences from their native Mandarin Chinese background.

Thus, possibly due to the nature of the perception and the production tasks, L2 to L1 tone assimilation patterns may be a strength when discriminating L2 tones but a weakness instead when producing them especially in isolation.

Nonetheless, this interpretation has to be contextualized in real-world communication. Overall, native Thai listeners were still able to identify the tones produced by all participants quite accurately regardless of their linguistic background, which implies that the disadvantages of perceptual similarity between Mandarin Chinese

tones and Thai tones during non-native Thai tone production may not be as impactful in real-life communication settings.

8.4 The cross-domain interaction effect of musical training with linguistic experience

Most importantly, the findings from the present study show that musical training interacts with linguistic experience in the processing of non-native lexical tones. Specifically, musical training benefits the production of non-native lexical tones, and mitigates the disadvantages of tone language experience in this process.

This finding is significant because firstly, it extends our knowledge of the benefits of musical training on the perception aspect of non-native lexical tone processing to the production aspect. Generally, past studies have mainly shown a benefit of musical training for segmental production (Milovanov et al., 2010; Slevc & Miyake, 2006) and for non-native lexical tone perception (Alexander et al., 2005; Burnham et al., 2015; Lee & Hung, 2008), but few have investigated whether this is the case as well for non-native lexical tone production. Thus, the present study provides evidence that musical training may not only enhance linguistic pitch perception but also enhance linguistic pitch production.

Besides that, this finding is significant as it contributes to our knowledge of how musical training may interact with linguistic experience in non-native lexical tone processing. For instance, Cooper and Wang (2012) showed that the combined effects of musical training and tone language experience were not beneficial in non-native Cantonese tone word learning. Instead, having just one of the experiences, i.e., either musical training or tone language experience, led to the best outcomes when Thai speakers and English speakers were learning non-native Cantonese tone words. In

contrast, the present study shows otherwise in the context of non-native Thai tone production, as musical training but not tone language experience led to the best outcomes when producing non-native Thai tone words, with the combination of both factors being better than having only tone language experience. This highlights a need for further studies to be conducted using different experimental paradigms and different combinations of L1 and L2 tone languages, so that we can better understand how and when musical training interacts with linguistic experience in specific aspects of non-native lexical tone processing.

8.5 Limitations and future areas of research

8.5.1 Related to the scope

One of the limitations of the present study is that the effects of musical training and tone language experience on the processing of non-native lexical tones have been measured by focusing on the F0, the primary cue of lexical tones. However, lexical tones may also be characterized by other cues such as duration and phonation. For instance, Mandarin Chinese tones also differ in duration, with its low falling rising (214) tone having the longest duration, and its high falling (51) tone having the shortest (Lin, 1965; Tseng, 1990). In contrast, Thai tones appear to be relatively similar in duration (Tilman et al., 2011), and its falling (51) and high (45) tones have also been described to be glottalized in phonation (Hudak, 2009). Nonetheless, the non-speech stimuli (filtered speech, violin) in Experiment 1 would not have captured the glottalized aspects of the Thai falling and high tones. Similarly, the current analyses for Experiment 2 do not reveal whether native tone language experience in Mandarin Chinese would also affect the duration of one's produced Thai tone tokens. Essentially, due to the constraints of the experimental design and the scope of the study, which is focused on pitch processing,

the findings of the study related to the perception and production of lexical tones are thus limited to the aspect of F0 and are not generalizable to other acoustic cues of lexical tones such as duration and phonation.

The contributions of the study in disentangling the effects of musical training from tone language experience in the processing of non-native lexical tones are also limited by the combination of native and non-native tone languages chosen in the study (Mandarin Chinese and Thai respectively). Importantly, in future investigations of both musical training and tone language experience in tandem, different combinations of native and non-native tone languages will need to be considered as the kind of perceptual assimilation between L1 and L2 tones will most probably differ depending on the similarities between the L1 and L2 tonal inventories.

8.5.2 Related to the tasks

Besides this, given that the lack of statistically significant effects of musical training and tone language experience on non-native lexical tone processing in this study appeared to be largely influenced by task difficulty, a possible future area of research would be to test the same four groups of participants using perception and production tasks of greater difficulty. For instance, the perception and production of the non-native Thai lexical tones could be tested by using Thai word stimuli with greater segmental variation, e.g., using different consonants and vowels to form syllables other than just [pa:] or including bi-syllabic words. Besides that, testing perception and production abilities using Thai word stimuli recorded from multiple native Thai speakers of different genders would also increase the task difficulty.

8.5.3 Related to the sample

There are also several limitations related to the sample in the present study.

Firstly, the musicians were all amateur musicians and were mostly non-percussionist instrumentalists. In reality, musicians are considerably non-homogenous, and may differ in the type of musical training received, number of years of musical training, instruments studied, level of expertise, and even in their level of interest in music and self-identity as a musician. Thus, the lack of statistically significant effects of musical training observed in this study may not be generalizable to other contexts involving musicians with different levels of experience, expertise, or type of musical training. For example, we do not know whether a statistically significant effect of musical training would have emerged in Experiment 1 had the musicians been professional musicians instead of amateur musicians.

Therefore, to achieve more comprehensive understanding of the effects of musicianship on non-native lexical tone processing, future research could endeavour to test musicians of different backgrounds, e.g., those who vary in their amount of musical experience, or differ in the type of musical training. In particular, one possible avenue for future research would be to compare the non-native lexical tone production of musicians who are singers vs. musicians who are instrumentalists. A study by Christiner & Reiterer (2015) has shown singers being rated more favourably than instrumentalists in their imitation of non-native Hindi sentences, even though both groups received better ratings compared to non-musicians/non-singers. Based on this, it is possible that singers may also produce non-native lexical tones better than instrumentalists, which would be a noteworthy finding.

Secondly, although care had been taken to ensure that the non-tone participants had no native acquisition, formal training or receptive proficiency in any tone languages, most, if not all of the participants in the current study had had long-term ambient exposure to tone languages. This is unlike the context of the studies by Burnham et al. (1996) and Burnham et al. (2015), where the non-tone English participants were recruited in a relatively tone unexposed environment. Notably, the non-tone participants in the current study had comparable perception of the non-native Thai tones as their tone-language-speaking counterparts. They also produced less confusable Thai tones even in the absence of a musical background. In addition, the secondary data analysis raises the possibility that the NT⁺nm group in the current study may experience less perceptual attenuation in the discrimination of non-native Thai lexical tones compared to the NT⁻nm group tested by Burnham et al. (2015). Due to limitations with unequal sample sizes, however, this remains to be ascertained in future research. In short, given the characteristics of the non-tone sample in this study, the possibility that the NTnm group may have had enhanced non-native Thai lexical tone perception and production due to their exposure to tone languages could not be ruled out entirely.

Future research could possibly explore this issue further with a comparison of participants with varying experiences of tone languages, in order to shed further light on the effects of different extents of tone language experience on non-native tone language processing. In particular, it would be noteworthy to investigate whether there are interactions between musical training and tone language exposure, given that non-tone musicians did not differ in their discrimination patterns across the three stimulus types regardless of their tone language exposure, whereas non-tone non-musicians had a trend of showing more pronounced attenuation in their perception of non-native lexical

tones in the direction of speech < filtered speech < violin if they did not have any tone language exposure.

Thirdly, although the recruitment of both tone and non-tone participants from the same country and linguistic environment minimizes potential confounds of geographical and cultural differences, another limitation is presented. Due to socio-political reasons, both tone and non-tone participants in this study were not exactly homogenous in their respective linguistic profiles. While all of them were bilingual in English and either a tone or non-tone language, most of them had, at a minimum, some receptive proficiency in another tone or non-tone language. It is difficult to predict whether this variation in linguistic background could have had any additional effects, especially among tone language speakers who had receptive proficiencies in other tone languages.

8.5.4 Related to the analyses

With respect to the production data in Experiment 2, one limitation is that no acoustic analyses were conducted. Thus, the evaluation of the production data is based on subjective judgments of native Thai speakers, without objective comparison to the acoustic characteristics of Thai tones produced by native Thai speakers. The reason that no acoustic analyses were conducted was that the tone identification accuracies and rating scores for the produced Thai tones were largely similar across groups, which made it less likely that the acoustic characteristics of the produced Thai tones across groups would differ greatly. Nonetheless, even if the acoustic characteristics differed significantly across groups, it is of greater import that their produced tones are intelligible to native speakers, since the purpose of L2 acquisition is ultimately for

effective communication. In this sense, the identification and rating analyses conducted in Experiment 2 would suffice.

8.5.5 Related to the findings

In addition to the main findings, there were a few observations that were interesting and deserve further investigation. Specifically, different trends of Thai tone discrimination across the three stimulus types (speech, filtered speech, and violin) were observed across the four participant groups in Experiment 1. While tone non-musicians (Tnm) and non-tone musicians (NTm) had equivalent discrimination performance regardless of the stimulus type (speech, filtered speech, and violin), non-tone non-musicians (NTnm) had a trend towards a graded discrimination pattern, with discrimination scores increasing as the stimulus became less speech-like (i.e., speech < filtered speech < violin). Tone musicians (Tm), on the other hand, had a trend towards the opposite discrimination pattern as NTnm, with discrimination scores increasing as the stimulus became more speech-like (i.e., violin < filtered speech < speech).

Although these patterns of discrimination were not statistically significant, it is notable that the trends observed for Tnm, NTm, and NTnm are in line with those found for the NTm, NTm, and NTnm from studies by Burnham et al. (1996) and Burnham et al. (2015). In this light, it would be informative for future research to explore whether these different trends in discrimination patterns for Tm vs. Tnm and NTm vs. NTnm would be significantly different from each other under experimental conditions which are more similar to those of the studies conducted by Burnham et al. (1996) and Burnham et al. (2015), i.e., if the musicians had levels of musical expertise similar to the (non-AP) musicians tested in the study by Burnham et al. (2015), and if the non-tone speakers had no ambient exposure to tone languages just like those tested by Burnham

et al. (1996). A re-examination of these trends would allow us to find out, firstly, whether the past findings by Burnham et al. (1996) and Burnham et al. (2015) can be replicated.

More importantly, had the trends in discrimination patterns for Tm, Tnm and NTm, vs. NTnm been found to be significantly different from each other in this study, it would have indicated that the Tnm and NTm groups benefited from their respective experiences of musical training and tone language background respectively, such that they experienced transfer benefits to non-native lexical tone processing across speech and music mediums. Moreover, it would suggest that the combination of musical training and tone language experience may not be as beneficial. Presumably, the Tm group would be so sensitive to pitch differences as a result of both musical training and tone language experience, they would be able to discriminate even between tokens of the same category. While this would have less of an effect on their discrimination of speech tones given their experience and familiarity with lexical tones, it would lower their discrimination performance in the violin context, where heightened sensitivity to within-category tokens is less desirable. As for the NTnm group, had their trend in discrimination pattern been statistically significant, it would have suggested that they would be at the most disadvantage in non-native lexical tone perception, because of attenuation in their perception of non-native lexical tones in speech form.

A re-examination of the discrimination patterns of these four participant groups under experimental conditions more comparable to past studies conducted by Burnham et al. (1996) and Burnham et al. (2015) would thus allow us to make such inferences about the independent and interactive effects of musical training and tone language experience on non-native lexical tone processing across a continuum of speech to non-speech, should the trends be found to be statistically significant.

CHAPTER 9: CONCLUSION

In tone languages, which form a major subset (70%) of the world's languages, pitch is used as a perceptual cue to contrast word meaning. When pitch takes on this function, it is termed a lexical tone. For speakers of non-tone languages, where the function of pitch is intonational and not lexical, such that pitch rarely affects meaning at the word level, non-native lexical tones can be challenging to perceive and produce.

Notably, musicians have been found to have enhanced perception of non-native lexical tones, which has been attributed to a cross-domain transfer effect between music and language due to their shared domain of pitch. Musical training thus appears to be beneficial to the learning of non-native tone languages. Nonetheless, whether this effect of musical training is influenced by prior linguistic experience and whether musical training may also enhance the production of non-native lexical tones is not well-studied. There are thus some gaps in our knowledge of the extent of influence musical training has on the learning of non-native tone languages.

Through an investigation of the non-native lexical tone perception and production of musicians and non-musicians both with and without prior native experience of tone languages, the current study addressed these research gaps. Contrary to expectations, no statistically significant independent or interactive effects of musical training with prior linguistic experience of tone languages as a native language were found. Secondary data analysis revealed that this was not due to potentially confounding effects of long-term ambient exposure to tone languages among the non-tone language speakers. However, additional analyses using confusion matrices demonstrated an interaction between musical training and tone language experience in the production of non-native lexical tones, revealing a potential of musical training to mitigate

disadvantages brought about by the phonological influences of one's native tone language during non-native lexical tone production.

Overall, the current study suggests that the manifestation of the cross-domain transfer benefits of musical training during non-native lexical tone processing is not as invariable as we may think but rather dependent on various factors. These factors range from the task itself (e.g., perception vs. production), the sensitivity of the analysis used (e.g., accuracy rates vs. confusion matrices), the task difficulty, as well as first language (L1) phonological influences among native tone language speakers. The study also presents new evidence that the cross-domain transfer benefits of musical training to non-native lexical tone processing go beyond perception to affect production, which adds to our knowledge about the extent of influence musical training has in second language (L2) tone learning. Lastly, by demonstrating that musical training can mitigate undesirable L1 phonological influences among native tone language speakers during non-native lexical tone production, this study is in support of the notion that musical training is also beneficial to tone language speakers in their learning of non-native tone languages. The findings of this study thus contribute to our understanding of the domain-generalty of pitch processing and help to inform predicting factors of the L2 acquisition of tone languages.

REFERENCES

- Alexander, J. A., Wong, P. C., & Bradlow, A. R. (2005, September). Lexical tone perception in musicians and non-musicians. In *Interspeech* (pp. 397-400).
- Ambrose, S. E., Berry, L. M. U., Walker, E. A., Harrison, M., Oleson, J., & Moeller, M. P. (2014). Speech sound production in 2-year-olds who are hard of hearing. *American Journal of Speech-Language Pathology*, 23(2), 91-104.
- Bachem, A. (1937). Various types of absolute pitch. *The Journal of the Acoustical Society of America*, 9(2), 146-151.
- Bermudez, P., & Zatorre, R. J. (2009). A distribution of absolute pitch ability as revealed by computerized testing. *Music Perception: An Interdisciplinary Journal*, 27(2), 89-101.
- Best, C. T. (1995). A direct realist view of cross-language speech perception. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 171–204). Baltimore: York Press.
- Bidelman, G. M., Gandour, J. T., & Krishnan, A. (2011a). Cross-domain effects of music and language experience on the representation of pitch in the human auditory brainstem. *Journal of Cognitive Neuroscience*, 23(2), 425-434.
- Bidelman, G. M., Gandour, J. T., & Krishnan, A. (2011b). Musicians and tone-language speakers share enhanced brainstem encoding but not perceptual benefits for musical pitch. *Brain and Cognition*, 77(1), 1-10.
- Boersma, Paul & Weenink, David (2017). Praat: doing phonetics by computer [Computer program]. Version 6.0.29, retrieved from <http://www.praat.org/>

- Burnham, D., Brooker, R., & Reid, A. (2015). The effects of absolute pitch ability and musical training on lexical tone perception. *Psychology of Music*, 43(6), 881-897.
- Burnham, D., Francis, E., Webster, D., Luksaneeyanawin, S., Attapaiboon, C., Lacerda, F., & Keller, P. (1996, October). Perception of lexical tone across languages: Evidence for a linguistic mode of processing. In H. T. Bunnell, and W. Idsardi (Eds.), *Proceedings of the Fourth International Conference on Spoken Language Processing* (Vol. 4, pp. 2514-2517). University of Virginia: IEEE.
- Burnham, D., & Francis, E. (1997). The role of linguistic experience in the perception of Thai tones. In A. S. Abramson (Ed.) *Southeast Asian Linguistic Studies in Honour of Vichin Panupong (Science of Language)*, 8, 29-47. Bangkok: Chulalongkorn University Press.
- Burnham, D., Kim, J., Davis, C., Ciocca, V., Schoknecht, C., Kasisopa, B., & Luksaneeyanawin, S. (2011). Are tones phones? *Journal of Experimental Child Psychology*, 108(4), 693-712.
- Chandrasekaran, B., Krishnan, A., & Gandour, J. T. (2009). Relative influence of musical and linguistic experience on early cortical processing of pitch contours. *Brain and Language*, 108(1), 1-9.
- Chao, Y. R. (1948). *Mandarin primer: An intensive course in spoken Chinese*. Harvard University Press.

- Cooper, A., & Wang, Y. (2012). The influence of linguistic and musical experience on Cantonese word learning. *The Journal of the Acoustical Society of America*, 131(6), 4756-4769.
- Creel, S. C., Weng, M., Fu, G., Heyman, G. D., & Lee, K. (2018). Speaking a tone language enhances musical pitch perception in 3–5-year-olds. *Developmental Science*, 21(1), 1-7. <https://doi.org/10.1111/desc.12503>
- Dawson, C., Tervaniemi, M., & Aalto, D. (2018). Behavioral and subcortical signatures of musical expertise in Mandarin Chinese speakers. *PloS One*, 13(1), 1-14. <https://doi.org/10.1371/journal.pone.0190793>
- Delogu, F., Lampis, G., & Belardinelli, M. O. (2010). From melody to lexical tone: Musical ability enhances specific aspects of foreign language perception. *European Journal of Cognitive Psychology*, 22(1), 46-61.
- Department of Statistics of Singapore (2011). *Singapore Census of Population 2010, Statistical Release 1: Demographic Characteristics, Education, Language and Religion*. Retrieved from https://www.singstat.gov.sg/publications/cop2010/census10_stat_release1
- Flege, J. (1995). Second-language speech learning: Theory, findings, and problems. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 229–273). Timonium, MD: York Press.
- Francis, A. L., Ciocca, V., Ma, L., & Fenn, K. (2008). Perceptual learning of Cantonese lexical tones by tone and non-tone language speakers. *Journal of Phonetics*, 36(2), 268-294.
- Fromkin, V. A. (Ed.). (2014). *Tone: A linguistic survey*. Academic Press.

- Gandour, J. (1983). Tone perception in Far Eastern languages. *Journal of Phonetics*, 11, 149-175.
- Gandour, J. (1984). Tone dissimilarity judgments by Chinese listeners. *Journal of Chinese Linguistics*, 12, 235–261.
- Gandour, J., Wong, D., Lowe, M., Dzemidzic, M., Satthamnuwong, N., Tong, Y., & Li, X. (2002). A cross-linguistic fMRI study of spectral and temporal cues underlying phonological processing. *Journal of Cognitive Neuroscience*, 14(7), 1076-1087.
- Hall, D. A., & Plack, C. J. (2008). Pitch processing sites in the human auditory brain. *Cerebral Cortex*, 19(3), 576-585.
- Hao, Y. C. (2012). Second language acquisition of Mandarin Chinese tones by tonal and non-tonal language speakers. *Journal of Phonetics*, 40(2), 269-279.
- Hudak, T. J. (2009). Thai. In B. Comrie (Ed.), *The world's major languages* (pp. 660-676). London: Routledge.
- Hutka, S., Bidelman, G. M., & Moreno, S. (2015). Pitch expertise is not created equal: cross-domain effects of musicianship and tone language experience on neural and behavioural discrimination of speech and music. *Neuropsychologia*, 71, 52-63.
- Jiang, C., Hamm, J. P., Lim, V. K., Kirk, I. J., & Yang, Y. (2010). Processing melodic contour and speech intonation in congenital amusics with Mandarin Chinese. *Neuropsychologia*, 48(9), 2630-2639.

- Krishnan, A., Gandour, J. T., & Bidelman, G. M. (2010). The effects of tone language experience on pitch processing in the brainstem. *Journal of Neurolinguistics*, 23(1), 81-95.
- Kuhl, P. K., Williams, K. A., Lacerda, F., Stevens, K. N., & Lindblom, B. (1992). Linguistic experience alters phonetic perception in infants by 6 months of age. *Science*, 255(5044), 606-608.
- Ladefoged, P., & Johnson, K. (2015). *A course in phonetics*. Stamford, CT: Cengage Learning.
- Lee, C. Y., & Hung, T. H. (2008). Identification of Mandarin tones by English-speaking musicians and nonmusicians. *The Journal of the Acoustical Society of America*, 124(5), 3235-3248.
- Lee, Y. S., Vakoch, D. A., & Wurm, L. H. (1996). Tone perception in Cantonese and Mandarin: A cross-linguistic comparison. *Journal of Psycholinguistic Research*, 25(5), 527-542.
- Levitin, D. J. (1994). Absolute memory for musical pitch: Evidence from the production of learned melodies. *Perception & Psychophysics*, 56(4), 414-423.
- Lin, M. C. (1965). The pitch indicator and the pitch characteristics of tones in Standard Chinese. *Acta Acoustica (China)*, 2, 8 –15.
- Liu, F., Chan, A. H., Ciocca, V., Roquet, C., Peretz, I., & Wong, P. C. (2016). Pitch perception and production in congenital amusia: Evidence from Cantonese speakers. *The Journal of the Acoustical Society of America*, 140(1), 563-575.

- Liu, F., Patel, A. D., Fourcin, A., & Stewart, L. (2010). Intonation processing in congenital amusia: discrimination, identification and imitation. *Brain*, 133(6), 1682-1693.
- Massaro, D. W., Cohen, M. M., & Tseng, C. (1985). The evaluation and integration of pitch height and pitch contour in lexical tone perception in Mandarin Chinese. *Journal of Chinese Linguistics*, 13, 267–290.
- Mattock, K., & Burnham, D. (2006). Chinese and English infants' tone perception: Evidence for perceptual reorganization. *Infancy*, 10(3), 241-265.
- Milovanov, R., Pietilä, P., Tervaniemi, M., & Esquef, P. A. (2010). Foreign language pronunciation skills and musical aptitude: A study of Finnish adults with higher education. *Learning and Individual Differences*, 20(1), 56-60.
- Miyazaki, K. I. (1988). Musical pitch identification by absolute pitch possessors. *Perception & psychophysics*, 44(6), 501-512.
- Miyazaki, K. I. (1989). Absolute pitch identification: Effects of timbre and pitch region. *Music Perception: An Interdisciplinary Journal*, 7(1), 1-14.
- Nan, Y., Liu, L., Geiser, E., Shu, H., Gong, C. C., Dong, Q., ... & Desimone, R. (2018). Piano training enhances the neural processing of pitch and improves speech perception in Mandarin-speaking children. *Proceedings of the National Academy of Sciences*, 115(28), 6630-6639.
- Patel, A. D., Wong, M., Foxton, J., Lochy, A., & Peretz, I. (2008). Speech intonation perception deficits in musical tone deafness (congenital amusia). *Music Perception: An Interdisciplinary Journal*, 25(4), 357-368.

- Perrachione, T. K., Fedorenko, E. G., Vinke, L., Gibson, E., & Dilley, L. C. (2013). Evidence for shared cognitive processing of pitch in music and language. *PLoS One*, 8(8), 1-9. <https://doi.org/10.1371/journal.pone.0073372>
- Pfordresher, P. Q., & Brown, S. (2009). Enhanced production and perception of musical pitch in tone language speakers. *Attention, Perception, & Psychophysics*, 71(6), 1385-1398.
- Plack, C. J., Barker, D., & Hall, D. A. (2014). Pitch coding and pitch processing in the human brain. *Hearing Research*, 307, 53-64.
- Polka, L., & Werker, J. F. (1994). Developmental changes in perception of nonnative vowel contrasts. *Journal of Experimental Psychology: Human Perception and Performance*, 20(2), 421.
- Psychology Software Tools (2016). E-prime 2.0 [Computer program]. Version 2.0.10.356, retrieved from <http://www.pstnet.com>
- Slevc, L. R., & Miyake, A. (2006). Individual differences in second-language proficiency: Does musical ability matter? *Psychological Science*, 17(8), 675-681.
- Takeuchi, A. H., & Hulse, S. H. (1993). Absolute pitch. *Psychological Bulletin*, 113(2), 345.
- Tan, L. Y. C. (2013). *Effects of musical and linguistic experience on the perception and production of non-native tones* (Unpublished final year project). Nanyang Technological University, Singapore.

- Tillmann, B., Burnham, D., Nguyen, S., Grimault, N., Gosselin, N., & Peretz, I. (2011). Congenital amusia (or tone-deafness) interferes with pitch processing in tone languages. *Frontiers in Psychology*, 2, 120.
- Tossa, W., & MacDonald, M. R. (1996). Pig and Dog [Text with audio]. Retrieved from: <http://www.seasite.niu.edu/thai/storytelling/pig&dog/pig&dogtext.htm>
- Tseng, C.-Y. (1990). *An acoustic phonetic study on tones in Mandarin Chinese*. Taipei, Taiwan: Institute of History and Philology, Academia Science.
- Van Lancker, D., & Fromkin, V. A. (1973). Hemispheric specialization for pitch and “tone”: Evidence from Thai. *Journal of Phonetics*, 1(2), 101-109.
- Van Lancker, D., & Fromkin, V. A. (1978). Cerebral dominance for pitch contrasts in tone language speakers and in musically untrained and trained English speakers. *Journal of Phonetics*, 6(1), 19-23.
- Wallentin, M., Nielsen, A. H., Friis-Olivarius, M., Vuust, C., & Vuust, P. (2010). The Musical Ear Test, a new reliable test for measuring musical competence. *Learning and Individual Differences*, 20(3), 188-196.
- Wang, Y., Jongman, A., & Sereno, J. A. (2001). Dichotic perception of Mandarin tones by Chinese and American listeners. *Brain and Language*, 78(3), 332-348.
- Wang, Y., Jongman, A., & Sereno, J. A. (2003). Acoustic and perceptual evaluation of Mandarin tone productions before and after perceptual training. *The Journal of the Acoustical Society of America*, 113(2), 1033-1043.
- Ward, W. D., & Burns, E. M. (1982). Absolute pitch. *The Psychology of Music*, 431-451.

- Werker, J. F., Gilbert, J. H., Humphrey, K., & Tees, R. C. (1981). Developmental aspects of cross-language speech perception. *Child Development*, 349-355.
- Werker, J. F., & Logan, J. S. (1985). Cross-language evidence for three factors in speech perception. *Perception & Psychophysics*, 37(1), 35-44.
- Werker, J. F., & Tees, R. C. (1983). Developmental changes across childhood in the perception of non-native speech sounds. *Canadian Journal of Psychology/Revue Canadienne de Psychologie*, 37(2), 278.
- Werker, J. F., & Tees, R. C. (1984a). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development*, 7(1), 49-63.
- Werker, J. F., & Tees, R. C. (1984b). Phonemic and phonetic factors in adult cross-language speech perception. *The Journal of the Acoustical Society of America*, 75(6), 1866-1878.
- Wong, P. C., & Perrachione, T. K. (2007). Learning pitch patterns in lexical identification by native English-speaking adults. *Applied Psycholinguistics*, 28(4), 565-585.
- Wong, P. C., Skoe, E., Russo, N. M., Dees, T., & Kraus, N. (2007). Musical experience shapes human brainstem encoding of linguistic pitch patterns. *Nature Neuroscience*, 10(4), 420.
- Wu, X., Munro, M. J., & Wang, Y. (2014). Tone assimilation by Mandarin and Thai listeners with and without L2 experience. *Journal of Phonetics*, 46, 86-100.
- Yip, M. (2002). *Tone*. Cambridge University Press.

APPENDIX A: Laboratory questionnaire

Relevant sections from the laboratory questionnaire which were used to cross-check and supplement participants' responses in the screening questionnaire are shown here.

Language Background

Q1) Please list all language(s)/dialect(s) including English that you can LISTEN to, the level of proficiency on a scale from 1 (not proficient) to 7 (very proficient), and the age at which you began learning to LISTEN to the language:

e.g., English 7 (proficiency), 0 (age); Mandarin Chinese 6, 0; Cantonese 5, 3

Music Background

Q2) Have you ever had any music lessons (e.g., private or group music lessons) and/or do you practice an instrument on your own?)

☐ Yes

☐ No

Q3) If you answered "Yes" to Question 2, please state the type of lesson you had, the instrument studied, and the year(s) you were enrolled in the lessons, your proficiency on a scale from 1 (not proficient) to 7 (extremely proficient/professional), the number of hours, on average, you play the instrument each week in the past 25 months, and the age at which you started playing the instrument. Please be as specific as possible:

e.g., Private; Violin; 1991-1997; 4 (proficiency); 2 hours/week; 10 years old.

Q4) Please list the style of music you studied (e.g., jazz piano; classical voice).

How many years of training did you complete in that particular style?

e.g., Piano; Jazz; 1991-1992

Q5) Can you read music?

☐ Yes

☐ No

☐ Somewhat

Q6) Do you have professional musicians in your family?

☐ Yes

☐ No

Q7) If you answered "Yes" to Question 6, please specify the occupations of those family members who are professional musicians.

Family member; Occupation

Q8) Do you have/are working on a degree in music?

☐ Yes

☐ No

Q9) If you answered "Yes" to Question 8, please list the title of the degree(s), the institution(s) you studied at, and what year it was/will be awarded.

Degree; Institution; Year

Q10) Please state any additional information about your music background:

APPENDIX B: Screening questionnaire

Thank you for your interest in this paid research study. Please complete this screening questionnaire to sign up.

For a start, let us know how to contact you:

Q1) Your name: _____

Q2) Preferred email address: _____

Q3) Handphone number: _____

Demographic Information

For profiling purposes, please let us know a bit more about yourself. Your responses will be kept private and confidential.

Q4) Your nationality:

☐ Singaporean

☐ Other: _____

Q5) Your ethnicity/race:

☐ Chinese

☐ Malay

☐ Indian

☐ Other: _____

Q6) Gender you identify as:

☐ Male

☐ Female

Q7) Your age: _____

Language Background

Q8) Do you know any of the following languages? Please select all those that you know.

- ☐ Thai
- ☐ Hokkien
- ☐ Cantonese
- ☐ Teochew
- ☐ Hakka
- ☐ Mandarin Chinese
- ☐ Punjabi
- ☐ Burmese
- ☐ Lao
- ☐ Vietnamese
- ☐ No, I do not know any of these languages.

Music Background

Q9) Do you know how to play any musical instruments?

- ☐ Yes
- ☐ No

If “Yes”, continue the questionnaire and answer Q10. If “No”, go to the end of the questionnaire.

Q10) Have you ever taken private music lessons before?

(i.e., outside of school-related music activities)

- ☐ Yes
- ☐ No

If “Yes”, continue the questionnaire and answer Q11. If “No”, go to Q12.

You have indicated having taken private music lessons before. Please let us know a little more about it by answering the following questions:

Q11a) Age you first started taking private music lessons: _____

b) Number of years you have taken private music lessons: _____

c) Number of years you have stopped taking private music lessons: _____

Q12a) How did you learn to play your musical instrument(s)?

☐ Through co-curricular activities (CCA) in school, e.g., band, orchestra

☐ Self-taught

☐ Other: _____

b) Age you first started learning to play your musical instrument(s): _____

c) Number of years you learned how to play your musical instrument(s): _____

Continue to Q13.

Q13) What musical instruments do you know how to play?

List them in order of the most proficient to the least proficient.

e.g., Piano, Violin

Q14) How frequently do you play your musical instrument(s)?

☐ Almost everyday

☐ More than once a week

☐ About once a week

☐ About once a month

☐ A few times a year

☐ About once a year

☐ Never

Q15) How many years have you stopped having an established practice schedule?

Established practice schedule = regular practice times whereby you focused on improving your mastery of a musical instrument. If you have never stopped having an established practice schedule, answer "0". If you have never had an established practice schedule, please answer "NA".

-----End of screening questionnaire-----

APPENDIX C: Individual musician backgrounds

Table C1

Musical Background of Tone and Non-Tone Musicians.

Musician group	Age started musical training (in years)	Years of musical training	Years since musical training	Average playing hours per week	Years since established practice schedule	Instruments played (in order of most proficient)
Tone	7	13	0.0	8.0	0.0	Piano, cello
	12	13	0.0	15.0	0.0	Sheng, piano
	11	6	0.0	4.5	0.0	Piano, saxophone, flute
	5	20	0.5	7.5	0.0	Piano, cello
	8	10	0.5	3.0	0.5	Flute, recorder, piano
	10	9	0.8	6.0	NA	Violin, piano, voice
	6	16	1.0	3.3	1.0	Piano, French horn
	5	14	1.0	8.0	6.0	Piano, tuba, flute, drums, ukulele
	10	8	1.0	4.0	NA	Voice, piano
	6	12	2.0	0.5	2.0	Piano, voice, erhu
	7	14	2.5	1.5	0.0	Piano, violin, recorder
	6	12	3.0	5.0	4.0	Violin
	4	18	4.0	2.0	4.0	Erhu, piano
	5	15	4.0	3.0	4.0	Piano
	6	8	5.0	6.0	0.0	Guitar, piano
	7	8	5.0	4.0	0.0	Piano, mallet percussion instruments
	7	9	6.0	2.5	6.0	Piano, double bass, bass guitar
	5	9	7.0	3.0	7.0	Piano, guitar

Musician group	Age started musical training (in years)	Years of musical training	Years since musical training	Average playing hours per week	Years since established practice schedule	Instruments played (in order of most proficient)
Non-tone	4	16	0.0	8.0	0.0	Piano, clarinet, ukulele
	10	11	1.0	6.0	1.0	Veenai
	9	5	1.0	2.0	1.0	Piano, angklung, kolintang, gamelan
	6	10	3.0	2.5	NA	Piano, tuba, ukulele, guitar
	14	7	4.0	23.0	5.0	Voice, ukulele
	10	6	4.0	3.0	NA	Piano
	7	4	4.0	6.0	4.0	Saxophone, piano
	7	10	5.0	6.0	6.5	Voice, piano
	9	7	5.0	2.0	2.0	Voice
	6	6	6.0	2.0	6.0	Piano, voice
	9	11	6.0	2.5	6.0	Voice, keyboard (organ)
	7	4	8.0	7.0	9.0	Voice
	9	4	10.0	3.0	NA	Dholak, mridangam, tabla, dhol, djembe
	13	3	11.0	2.0	0.0	Guitar, drums

Note. Average playing hours per week reported are for the past 25 months. Established practice schedule was defined as regular practice times when one focused on improving the mastery of a musical instrument (i.e., a purposeful and self-aware form of playing which is not for leisure). NA (i.e., not applicable) refers to cases where participants self-report never having had an established practice schedule.

APPENDIX D: Absolute pitch test stimuli

Table D1 shows the frequencies (in Hz) and intensity levels (in dB) of the 36 pure tones used in the absolute pitch (AP) test.

Table D1

Intensity Levels (in dB) and Frequencies (in Hz) of Notes Used in the Absolute Pitch Test.

Note	Hz	dB	Note	Hz	dB	Note	Hz	dB
C3	130.81	81	C4	261.63	80	C5	523.25	70
C#3/Db3	138.59	83	C#4/Db4	277.18	80	C#5/Db5	554.37	70
D3	146.83	83	D4	293.66	80	D5	587.33	70
D#3/Eb3	155.56	83	D#4/Eb4	311.13	78	D#5/Eb5	622.25	70
E3	164.81	83	E4	329.63	78	E5	659.25	67
F3	174.61	85	F4	349.23	78	F5	698.46	67
F#3/Gb3	185.00	85	F#4/Gb4	369.99	78	F#5/Gb5	739.99	67
G3	196.00	87	G4	392.00	73	G5	783.99	67
G#3/Ab3	207.65	87	G#4/Ab4	415.30	73	G#5/Ab5	830.61	67
A3	220.00	85	A4	440.00	73	A5	880.00	65
A#3/Bb3	233.08	85	A#4/Bb4	466.16	73	A#5/Bb5	932.33	65
B3	246.94	85	B4	493.88	73	B5	987.77	65

As loudness perception varies with sound frequency and intensity based on equal-loudness contours, the pure tones used in the AP test had to be adjusted in their intensity levels so that they would be perceived to be having the same loudness despite having different frequencies. These values were determined firstly by referencing the equal-loudness contour at 70 phons as a general guide (see Figure D2), before fine-tuning the values by comparing the perceptual loudness of notes in the same pitch chroma (e.g., C3 vs. C4 vs. C5) and within the same octave range (e.g., C3 – B4).

APPENDIX E: Thai short story for speech and filtered speech familiarisation contexts

The following Thai story, “Pig and Dog” (Tossa & MacDonald, 1996) was read by a female native Thai speaker and the subsequent recording used as the 1-minute familiarization context recording for speech and filtered speech stimulus sets. The English translation of the story is included as follows.

<p style="text-align: center;"><u>หมูกับหมา</u></p> <p>ครั้งหนึ่งนานมาแล้วมี ชาวนาคนหนึ่งเลี้ยงหมูกับหมาไว้ช่วยทำงาน วันหนึ่งชาวนาก็ขอให้หมากับหมูไปช่วยกันไถนา ให้เสร็จก่อนที่ตนเองจะกลับมาจากธุระ</p>	<p style="text-align: center;"><u>Pig and Dog</u></p> <p>Once upon a time, a farmer raised a pig and a dog to help him work. One day he had to go to town, so he told his pig and dog to go plough the field before he returned home in the evening.</p>
<p>หมาและหมูก็พากันไปนา หมาก็บอกให้หมูไถนาก่อนตนเองจะนอนพักผ่อน น พอหมูเหนื่อย ตนเองก็จะลุกมาไถนาต่อแล้วหมูจะได้พักผ่อนพอ อดกลังกันอย่างนั้น หมาก็ตรดตัวลงไถ่ต้นไม้นอนหลับสบาย ส่วนหมูก็ลงมือไถนาอย่างขมกเขม่นจนกระทั่งเ สรีจหมดก็เลยปลุกหมาขึ้นมามากกว่าตอนนี้ไถนา าเสร็จหมดแล้วจะนอนพักผ่อนบ้าง พอหมูนอนหมาก็ลุกขึ้นมาวิ่งเล่นในทุ่งนาที่หมูไ ด้ไถไว้แล้ว</p>	<p>Pig and Dog went to the field. Once they arrived at the field, the dog told the pig to plough first and the dog would take the rest. When the pig got tired, the dog could help. Dog went to sleep, and Pig ploughed the field until the entire field was ploughed. Pig told Dog that he had finished and would take a rest. So Dog got up and ran around, playing in the rice field.</p>

<p>พอค่ำทั้งหมูและหมาก็พากัน กลับบ้านชาวเขาก็ถามว่า ใครเป็นคนไถนา หมาก็รีบบอกว่าตัวเองเป็นคนไถ ส่วนหมูก็บอกว่าตัวเองเป็นคนไถ ทั้งสองก็เถียงกันเป็นเวลานานจนชาวเขาก็รำคาญ เลยบอกว่าผู้ที่ไถนาจะต้องมีรอยตีนอยู่บนทุ่งนา แน่นอนจึงได้ไปยังทุ่งนาเพื่อพิสูจน์ความจริง ปรากฏว่าในทุ่งนาเต็มไปด้วยรอยตีนหมา</p>	<p>In the evening, they returned home to tell the farmer about their work. Dog said he was the one who ploughed the field and Pig said the same. The farmer did not know who to believe, so he went to take a look at the field and found Dog's footprints on the field.</p>
<p>ชาวนาจึงได้ตัดสินใจว่าหมาเป็นสัตว์ ที่ขยันเอาใจใส่ต่องานและเป็นคนไถนา และหมูเป็นสัตว์ขี้เกียจ เอาแต่นอนไม่รู้จักทำงาน ชาวนาจึงตัดสินใจว่า ต่อไปนี้จะให้หมูกินแต่รำและให้อยู่ในโคลนตมอยู่ใ ต่ถุนบ้าน ส่วนหมาให้กินข้าว และอยู่บนเรือนที่สะอาดสะอาด</p>	<p>He believed that dog had done the work and pig was lazy. So he chased the pig to stay under the house in the mud, eating only rice bran and the dog could stay anywhere, eating rice for food.</p>
<p>นี่เองเป็นสาเหตุให้หมูกินรำ และหมากินข้าวตลอดมาจนกระทั่งทุกวันนี้</p>	<p>That is why dogs eat rice and pigs eat bran.</p>