Do Music and Language Help? The Influence of Musicianship and Language Background on Tonal Perception

Xinhui (Zhuohan) Zhang

MSc in Applied Linguistics and Second Language Acquisition, 2024

DECLARATION BY THE CANDIDATE AS AUTHOR OF THE DISSERTATION

1. I understand that I am the owner of this dissertation and that the copyright rests with me unless I specifically transfer it to another person.
2. I allow the Department to deposit on my behalf a copy of this dissertation in the Oxford University Research Archive ('ORA') where it shall be freely available online for use in accordance with ORA's Terms and Conditions of Use [https://ora.ox.ac.uk/terms_of_use].
3. I understand that this dissertation should not contain material that can be used to personally identify individuals or specific groups of individuals (unless permission has been obtained from the individuals) and that such material should be removed before this dissertation is deposited in ORA.
4. I agree to be bound by the terms of the ORA Grant of Non-exclusive Licence [https://ora.ox.ac.uk/deposit_agreements] and I warrant that to the best of my knowledge, making my thesis available on the internet will not infringe copyright or any other rights of any other person or party, nor contain defamatory material.
5. I agree that my dissertation shall be available for download in ORA in accordance with paragraphs 2, 3 and 4 above.

Signed [an electronic signature is sufficient]:	Xinhui (Zhuohan) Zhang
Date:	21 August 2024

Table of Contents

List of Tables

List of Figures

Acknowledgement

On an early morning, sitting in the small reading room in Linacre, I have never felt so strongly that this master journey is coming to an end. The past ten months have vanished in the twinkling of an eye. I always remember the first time I met my supervisor, Dr Faidra Faitaki, at her office and I always remember the sweetness of the cake she shared with her supervisees. She was like an angel, always there to support us whenever we have any difficulties. I could feel that she was doing her utmost to use her limited time to clear up as many of our confusions as possible. I have learned a disciplined scholarly attitude from her, which will be a lifelong benefit to me. The academic journey is never easy. I would also like to extend my gratitude to Prof Eva Schultz-Berndt, Dr Martina Faller, and Dr Richard Zimmermann who are from my undergraduate university, the University of Manchester. Even though my academic journey has concluded there, they have always been caring and supportive of me.

I am deeply appreciative of all the people who contributed to the completion of this project. I would like to begin by expressing my love to my long-time best friends 芦俊同 (Lú Jùntóng) and 崔映茜 (Cuī Yìngqiàn, Evelyn), who never hesitate to pick up the phone and respond my messages when I feel low. We have proved that long-distance friendships always work despite time differences. I would like to sincerely thank all the participants. The contribution from every and each one of you is invaluable to the success of this research. My friends, 蒋芊浔 (Jiǎng Qiānxún, Frances) and ก้องภัค ภูผา (Kongpak Phupa), continuously and sincerely supported my audio generation in the tonal discrimination task. There would be no existence of such a wonderful online experiment task without the contribution from you both. Particularly, I would like to thank Kongpak who has been diligently and patiently answering my questions despite the 6-/7-hour time difference. My cousins 石路阳 (Shí Lùyáng, Allen) and 赵婉婷 (Zhào Wǎntíng) helped me spread words about experiments and call for participants. The completion of this research has your names.

Furthermore, I am deeply grateful to all the professors, lecturers and my cohorts in Class of 2024 in ALSLA. The kindness and support from you all have added a vibrant touch to my life as a master student. We are such a huge group. However, I know that the sense of belonging I had whenever I sat in the classroom is not because of the big room size, but since everyone has their unique characteristics that keep shining all way along. I see persistence, intelligence, romance (like someone X put the roses and the chocolate on the morning of Valentine’s Day!), support, warmth, and so on. Since there is no single word to accurately describe you all, I will use ALSLA’24 for us, the wonderful and incomparable us. It is my absolute honour to meet every and each one of you.

Moreover, I have to mention the beautiful souls I met in Linacre. It is a small college, but also a cozy and inclusive one. I love the soft sofa and free coffee (finally it’s free!!), and the nice food which always surprises you no matter there is formal or not. I noticed the matriculation photo at the common room the other day. It shocked me when I realised that we were all total strangers when we firstly arrive, yet now we know each other in one way or another. My life in Oxford is more memorable by being a Linacrite.

Most importantly, I wish to express my heartfelt gratitude to my family members. My mother’s question around ten years ago inspired me to conduct this research. Thank you all for supporting me no matter what happened. Especially,

I would like to mention my late grandparents – grandpa 闫秉孝 (Yán Bǐngxiào) and grandma 潘玉芝 (Pān Yùzhī). I would never be the person I am today without your gentle care and love since my birth. I love you three thousand times.

Very lastly, I would like to write down a few words for myself. Zhuohan, this is the preferred name you chose for yourself and of course, you have rewritten your fate. From Dalian, to Hong Kong, to Manchester, and now to Oxford, no one knows what you have experienced except you. You are a kind girl and also a bold, courageous, and tough one. You have spent so much effort to achieve what you have right now. Yes, sometimes not everything is satisfying. However, you are way too better than you ever were, and you will keep growing and glowing in that way because you mind is stronger than ever. Please be confident and avoid excessive self-doubt. You deserve everything at its best. Like you promised, now this journey is over, get yourself a huge bunch of red roses and enjoy your summer!

Abstract

The interaction between musicianship and language background has attracted attention in the academic research field. The OPERA hypothesis suggests that since music processing requires higher neural sensitivity, together with its overlapping neural network with speech processing and the existence of neural plasticity, the music training enhances speech processing. In education and neurolinguistics, there have been numerous investigations providing neural and behavioural evidence on how musical training could significantly improve sensitivity to pitch procession. The co-existence of the concept of pitch in music and language has poses a question about how musical aptitude and language background (specifically focus on whether one's first language is tonal) influence tonal perception. This study aims to address this question. 167 participants (97 with a tonal L1 and 70 with a non-tonal L1) participated in a musical aptitude test (which included a melody perception and a rhythm perception component) and assessed a tonal discrimination task which included 20 pairs Thai words. The results showed that, for both groups, the participants' overall musical skills, their rhythmic perception, and their melodic perception in particular had a statistically significant and positive correlation with performance in the tonal discrimination task. These correlations did not present significant difference between tonal and non-tonal L1 groups, indicating that regardless of native (non-)tonal language experience, higher musical aptitudes facilitated tonal perception performance. This echoed the OPERA hypothesis by supporting the music-language effect. The musical aptitudes and (non-)tonal L1 backgrounds were both found as significant predictors of tonal discrimination task scores. When musical aptitudes were controlled as covariate, there was significant difference in tonal discrimination task scores between native tonal and non-tonal speakers. The tonal L1 group scored significantly higher than the non-tonal L1 group. This study had several limitations. The uncontrolled environmental factors, lack of break between tests, and unfamiliar of the tests might undermine participants' performance on musical aptitudes. The categorisation of participants ignored their proficiencies in languages other than L1, which might influence the investigation of their relationship with tonal perception. For future studies, it is suggested to facilitate participants in test familiarisation, carefully control and measure between-subject and within-subject factors if possible. Moreover, since this study only utilised discrimination task for tonal perception measurement, future studies can involve tonal identification task to evaluate the ability of tonal perception more comprehensively.

1 Introduction

Musical sounds could be the foundation of language evolution, considering that our human ancestors initially used various tones to communicate before the emergence of language (Darwin, 1874). Music and language commonly exist across all human ethnic groups (James & Leather, 1986; Oudeyer, 2006; Roederer, 2008; Small, 1997; Zatorre, 2024). Numerous studies have been conducted to investigate their complicated relationship (e.g., Asano & Boeckx, 2015; Chang et al., 2024; Feld & Fox, 1994; Hodges & Thaut, 2019; Jackendoff, 2009; Levman, 1992; O'Callaghan, 2018; Oesch, 2019; Patel, 2007; Peretz, 2019; Pino et al., 2023; Rebuschat et al., 2011; Zatorre et al., 2007). These studies have delved into the commonalities and discrepancies between the two constructs in the fields of biology, cognitive sciences, education, philosophy, psychology, sociology, etc. More recent interdisciplinary publications have also introduced the fields of music and linguistics from each other's perceptive (Camp, 2023; Larroque, 2023).

With an increasing emphasis on interdisciplinary research and attention to language education, studies have been focusing on the perceptive processing of tones in music and language. For example, a small-scale review article based on neural evidence has proposed that the perception of music and language prosody predominantly relied on certain regions in the right hemisphere (Asano et al., 2022). On the other hand, left hemisphere is crucial for speech sound learning while more complicated language learning (e.g., pragmatic functions), seems to require the participation of right hemisphere and bilateral brain networks (Qi & Legault, 2020).

Another topic of research concerns how music and language interact with one another (e.g., Bidelman & Alain, 2015; Francois & Schön, 2011; Hallam & Himonides, 2022; Lee et al., 2020; Shen & Froud, 2019; Besson et al., 2017). Musical abilities are suggested to provide supportive transferrable effects for language learning (Besson et al., 2017; Jentschke, 2014). Conversely, the language backgrounds (i.e., tonal and non-tonal languages as L1s) of participants can also potentially impact pitch processing of musical tones (e.g., Bidelman & Chung, 2015). For example, in Bidelman and Chung's (2015) study, Chinese speakers exhibited hemispheric asymmetry when processing different types of pitch types while English speakers did not.

Researchers have also examined how music can be applied to first language (L1) and second language (L2) education (e.g., Green, 2017). Neural evidence also shows that a certain period of musical training could enhance children's performance on language tonal perception (e.g., Nan et al., 2018). Utilising songs in L2 classrooms is suggested to be beneficial for linguistic gains in listening and speaking skills, vocabulary knowledge, grammar use and for boosting intercultural understanding (e.g., Ludke, 2019). There have been suggestions for using songs in primary education for learning foreign language (García Conesa & Juan Rubio, 2015; Ludke & Good, 2020). It is suggested that since second language learning involves negative emotion due to examination and peer pressure, music provides an escape from academic competition (Kumar et al., 2022). It is also reported that teachers may use songs, liking singing together, to motivate students in their language learning process (Oktavia et al., 2022). The application of music is also effective during personal language learning journey (Kao & Oxford, 2014).

In many meticulous studies on tonal perception (e.g., Bidelman & Alain, 2015; Lee et al., 2020; Wang et al., 2022), neuro-imaging data was carefully collected via techniques like electroencephalogram (EEG[1], as in Lee et al., 2020), event-related potentials (ERPs[2], as in Bidelman & Alain, 2015), and mismatch negativity (MMN[3], as in Chen et al., 2018; Shen & Froud, 2019; Wang et al., 2022). Despite the application of advanced techniques, these studies tended to suffer from some serious limitations. Firstly, they often encompassed relatively small sample sizes (e.g., 10 musicians and 10 non-musicians in Bidelman & Alain, 2015; 13 musicians and 11 non-musicians in Lee et al., 2020; 10 participants in each language group in Shen & Froud, 2019; 16 native Chinese and Dutch speakers respectively in Chen et al., 2018; around 20 participants in each language group in Wang et al., 2022). The inadequate sample size indirectly exemplifies the challenges of conducting such studies and consequently increases the difficulty of generalising research results across various contexts (Etz & Arroyo, 2015), which may diminish the persuasiveness of the research.

Secondly, the definitions of musician and non-musician are not consistent across studies. For example, in Bidelman and Alain's (2015) study, musicians were defined as amateurs with five-continuous-year musical training which started before the age of 14 years old, while in Lee et al.'s, (2020) study, musicians should have received professional musical training for more than ten years, but their starting ages were uncontrolled. Rogala et al. (2023) selected students from professional musical schools and labelled them as musicians. However, musicians in Sajjadi et al. (2021) were self-reported to have at least 10 years of formal musical training, should have consistent musical practice per week and start training before the age of 9 years. The inconsistency in selection criteria may make readers confused about the precise definition of musicians in academia and implies the lack of agreements across scholars. The chaotic categorization may indicate the unequal musical performance within the same label across studies. For example, the musician in Bidelman and Alain's (2015) study may not be musicians or non-musicians in Sajjadi et al.'s (2021) study. Such inconsistency in selection criteria may cause a mix of abilities and backgrounds within each category, which is difficult for researchers to synthesize studies in this field. Additionally, novice researchers may struggle to apply consistent criteria for participant classification or encounter difficulties in participant recruitment given the mixed criteria in current studies.

Lastly, although many studies (for a review, see Neves et al., 2022) have provided valuable neural evidence that musical training can increase neural sensitivity when detecting a sound change, it is uncertain whether this sensitivity can be directly extended to vocabulary/phrase/sentence acquisition from a natural language. The systematic review (Neves et al., 2022) mentioned 14 studies investigating the influence of musical training on auditory processing and more than half utilised artificial sounds (including synthesised instrumental sounds) as materials to detect auditory sensitivity (e.g., Alain et al., 2019; Tervaniemi et al., 2022). The results from research employing artificial may not be applicable to studies about natural language learning for several reasons. To begin with, natural languages are inherently complicated. They involve various elements apart from phonetic strings, like semantic, syntactic, and pragmatic components. The physical features of human speech have led to the existence of unique

[1] The EEG captures the neural electrical activity of the brain (de Silva, 2022).
[2] ERPs are a small voltage generated by brain to respond certain events, which are detected in a non-invasive way (Blackwood & Muir, 1990).
[3] The MMN, a type of ERP, is applied to investigate the neural reflection of pitch variations. MMN is a waveform which appears after the averaged time window of 100-200ms of detected differences of acoustic patterns and the response usually peaks at around 5μV maximally shown at about 100-250ms (Garrido et al., 2009).

natural sound patter, which are arguably non-existent for unnatural sounds (Blevins, 2008). Additionally, natural languages involve complex auditory cues like intonations, stress patterns, and phonotactics, which are not typically presented in synthesized sounds but crucial for meaning (Ladd, 2008). Furthermore, it is possible that neural mechanisms may differ for processing natural, musical and artificial sounds (Salvari et al., 2019). As a result, the findings from studies using synthesized sounds should be treated with caution, and it may not be possible to generalise their results to the field of natural language learning. Instead, these studies have paved the way for the studies about the possible impact of musical training on natural language acquisition. Therefore, there is an increasing necessity to conduct studies with audio materials in natural languages rather than artificial ones.

Subsequently, the present study aims to investigate the influence of musicianship and language backgrounds on tonal perception by encompassing a larger number of participants (i.e., 167 people), quantifying musical ability (i.e., measured via a widely used musical test), and shedding more light on how musicianship and (non-)tonal first language background could influence the tonal perception of a natural language (i.e., evaluating tonal discrimination ability via Thai words). There are two research questions:

(1) Do musical abilities influence tonal perception?
(2) Does language background (i.e., tonal vs. non-tonal L1) influence the tonal perception?

In this paper, it should be noted that the terms – musicianship, musical skills, musical abilities, and musical aptitudes – are used interchangeably unless further clarification in that all terms reflect the competency of acquiring musical skills. Regarding the flow of this study, this chapter offered a general introduction of research into tonal perception in music and language, discussed how tonal perception could be potentially influenced by musicianship and language backgrounds (i.e., tonal and non-tonal languages as L1s), and highlighted the necessity of this research by identifying current research gaps. Chapter 2 will introduce the concept of tones in music and language, critically explain hypotheses about neural functions on pitch processing, and review previous research on the impact of musical training and language acquisition on auditory perception in detail. The substantial role of auditory perception in language learning will also be discussed. Chapter 3 will present the methods of this study (including the participants' profile, the pilot study, the experiment materials, procedures, and statistical models). Chapter 4 will present the results of correlation test, linear mixed-effect models and nonparametric ANCOVA. Chapter 5 will discuss the possible reasons of the results and limitations in current study. Finally, chapter 6 will conclude the overall research and provide suggestions for future studies.

2 Literature Review

This chapter aims to provide research background of the influence of musicianship and language background on tonal perception both in depth and width. The first subsection thoroughly introduces the concept of tone, which exists in both music and language. The second subsection presents neural evidence on how the brain processes tonal information, followed by the emphasis on reciprocal interaction across hemispheres in successful speech processing. Then, I present how musical training and language backgrounds (i.e., tonal and non-tonal L1) can influence auditory perception and its importance in second language learning. Next, the theoretical framework, OPERA hypothesis by Patel (2011, 2012) is introduced to explain the music-to-language transfer. The final subsections briefly present the tonal system in Thai language, followed by reiterating the research questions of current study and the corresponding predictions.

2.1 What does 'Tone' Mean in Music and Language? – The Parallelism between Music and Language

Ever since the proposal that language originated from music by utilising different tones was put forth (Darwin, 1876), the discussion about the connections between music and language started to gain traction in the 20th century. Leonard Bernstein delivered a series of speeches about the parallelism between music and language. The musical note, motive, phrase, section, and movement were correspondingly compared with phonemes, morphemes, words, clauses, and sentences in languages. For example, in the lecture of musical syntax, he mentioned that certain notes in Beethoven's 6th symphony aimed to evoke images of birds, streams, and peasants. The elicitation process employed metaphor to establish links between musical and non-musical meanings (1976).

The existence of tones, with pitch being a physical property in both music and language, provides a solid fundamental linkage between the two disciplines, even if the precise definitions of tones are not identical in the two fields (e.g., music: Oxenham, 2013; Rasch & Plomp, 1999; linguistics: Gandour, 1978; Gussenhoven, 2010; Massaro et al., 1985). In music, a tone is 'a periodic sound that elicits a pitch sensation' (Oxenham, 2013, p. 1). A single musical tone can be depicted in three dimensions: pitch, loudness, and timbre (Rasch & Plomp, 1999). For this study, it is worth noting that each musical tone delivers a unique pitch measured by fundamental frequency (F0) in hertz (Hz) and pitch is crucial for melody and harmony in musical composition (Howard & Angus, 2013; Oxenham, 2013). Loudness is the perceived intensity of a sound, measured in decibel (dB) while timbre is the unique quality of a sound, differentiating two sounds of identical pitch and loudness based on sound spectrum and other attributes (Oxenham, 2013).

In the field of linguistics, the concept tone is defined in an alternative way. Tone represents the unique application of pitch (Gandour, 1978; Massaro et al., 1985). Tonal languages require the combination of consonants, vowels, and tones to represent word meanings (Gandour, 1978; Gussenhoven, 2010). There are two types of tone languages: register tone languages and contour tone languages (Pike, 1948). The former requires a specific pitch level of the tone (Bao, 1999). The latter has contour tones, where contour refers to the trajectory of the acoustic correlate of pitch (i.e., F0) over time (de Cheveigné & Kawahara, 2002). In other words, contour tone languages have a dynamic perspective of 'tone', as the syllable should be articulated with a certain pitch movement (Bao, 1999). The changes in pitch contour can alter the meanings of words (Yip, 2002). For example, in Mandarin Chinese, the same *pīnyīn* (the romanisation system of Chinese characters) 'ma' has four different meanings in four different tones: level tone – 'mother', rising tone – 'hemp', dipping tone – 'horse', and falling tone – 'scold' (Zhang, 2018, p. 4). It has been found that a language can have both register tones and contour tones. For example, Kom, a language in Cameroon, has been found to have three level tones (i.e., low, middle, and high) and five contour tones (i.e., high-low, middle-low, low-high, low-middle, and middle-high) (Bernard et al., 2002).

Even though the non-tonal languages also apply tonal changes for communicative purposes, it is important to acknowledge that their applications do not differentiate linguistic meanings (Ladd, 2008). For example, English speakers would employ tune changes in 'five' or 'pounds' to emphasise their points (Ladd, 2008, p. 8) while the linguistic meaning of either word maintained the same regardless of the intonational variances. Contrastively, as previously exemplified by Mandarin Chinese, the tune changes would possibly cause changes in lexical tones depending on specific contexts, therefore delivering different semantic meanings.

2.2 How does the Brain Work? – The Overlapping Resources in Tonal Processing in Language and Music

The interest in revealing the brain functions dates back to 3000 B.C., and Wepfer in the 17th century is arguably the first scholar employing modern neuropsychological approach to investigate cognitive functions in brain with consistent work reporting the lesions and aphasia (Whitaker, 1998). There are several hypotheses about hemispheric functions of processing tonal, stress, and intonation information. To begin with, the functional scale hypothesis was put forward by Van Lancker (1980). The hypothesis proposed that the left hemisphere mainly controls pitch patterns with more linguistic information (i.e., tones) while the right hemisphere mainly controls pitch patterns with less linguistic information (i.e., intonation for showing emotion). A weakness with this hypothesis, however, is that it is difficult to classify the pitch into two distinct categories since the function spans a continuum rather than a dichotomy (Moen & Sundet, 1996). Counter-evidence was also found in later studies (e.g., Proverbio et al., 2016; Vigneau et al., 2011). For example, the left-hemispheric dominance in linguistic tasks was found in healthy adults (for a review, see Vigneau et al., 2011). Moreover, the left hemisphere could also detect musical tones among musicians (Proverbio et al., 2016).

Next, the hypothesis regarding hemispheric preference in temporal and spectral processing proposed by Zatorre and Belin (2001). Their hypothesis suggested that left hemisphere is more specialised than right hemisphere in temporal processing while right hemisphere is more specialised in spectral processing. Considering that the spectro-temporal information of speech is essential for phonology and auditory perception (Gautam & Singh, 2017), if one hemisphere was impaired, the less ability to process one type of information would require the other hemisphere to compensate for the injuries. For example, some studies found that the right hemisphere could substantially compensate for linguistic abilities when patients suffered from left-hemispheric impairment (Kourtidou et al., 2021; for reviews, see Lindell, 2006; Riès et al., 2016). The stimulation of processing linguistic information in the right hemisphere indicated the potential of brain functions. Zatorre and colleagues (2002) also found evidence echoing the functional scale hypothesis and their own hypothesis. They argued that the two hemispheres functioned reciprocally regarding temporal and spectral processing, despite the existence of relative specialisation in each cortical area.

Another hypothesis, lateralization-carryover hypothesis, was proposed by Gu and colleagues (2013). The hypothesis based on two assumptions: function-dependent brain asymmetry and acoustic-dependent brain asymmetry. The former suggested that for native nontonal language speakers, the left hemisphere preferred speech processing while the right hemisphere favoured musical processing, echoing the function scale hypothesis. The latter suggested that the left hemisphere concentrated on acoustic temporal processing while the right hemisphere focused on spectral/pitch processing, resonating the hypothesis by Zatorre and colleagues (2001, 2002). However, Gu et al. (2013) found different patterns in native tonal speakers who processed the acoustic pitch contrast in the left hemisphere, different from the functional area in nontonal speakers (i.e., the right hemisphere). Therefore, they argued that the inconsistency was caused by the carryover effect from function-dependent brain asymmetry. Contrastively, this finding was not previously supported by previous studies (e.g., Xi et al., 2010) and there were limited research investigating the validity of this hypothesis.

It is shown that these hypotheses do not contradict with each other but intend to explain brain functions from various angles. This implies the complexity of brain structures and functions. More evidence from neuroscience data strongly indicates that linguistic and musical information processing occupies overlapping neural resources (Bidelman et al., 2013; Koelsch et al., 2005). The right and left hemispheres cooperate to reach successful pitch and rhythm memory (Schaal et al., 2017), and music perception (for a review, see Janata, 2015; Rosenthal, 2016). Other noteworthy evidence indicates that the area that mediates syntactical processing is situated within the same regions as for musical harmony and rhythm processing (Herdener et al., 2012; Maess et al., 2001).

With the mixed empirical evidence in aforementioned studies, it is difficult to provide a clear explanation of the precise brain mechanism. Despite this, these studies have suggested the necessity of cooperation between two hemispheres for successful language and music comprehension and the shared resources for musical and linguistic processing.

2.3 The Impact of Musical Training on Auditory Perception

The focused attention on music perception and the relevant talents of musicians is suggested to start from early 20th century with Seashore (1920) as one of the pioneers (Kraus, 2011). Recent studies have shown the importance of musical aptitudes on auditory perception (e.g., Lee et al., 2020; Hennessy et al., 2022). The long-term musical training (e.g., more than 10 years) could enhance neural plasticity in that musicians required lower thresholds for detect frequency discrimination compared to their non-musician peers (Lee et al., 2020). A meta-analysis encompassing 31 studies and 62 effect sizes revealed the significant advantage across young musicians on speech-in-noise perception (Hennessy et al., 2022).

The possibility of the music-to-language transferrable skills have been proposed, given the overlapping neural mechanism of processing music and language (Thaut et al., 2014). Moreover, the aforementioned studies about the beneficial effect of musical training on auditory perception indicate the possibility of generalising the findings to the facilitation of music in learning linguistic components of natural languages. For example, Tang et al. (2016) found that for native Mandarin speakers, the pitch expertise derived from L1 is transferrable to the speech domain. In this study, the musician group (who had continuous musical training for at least 8 years) presented significantly more intense MMN responses than the age- and IQ-matched non-musician groups in discriminating lexical tones. Similar findings were obtained from a study investigating the influence of musical training on categorical perception in Mandarin Chinese (Wu et al., 2015). Categorical perception is a phenomenon when a continuum of sounds tends to be perceived as different categories; discrimination across categories is easier than the same category (Goldstone & Hendrickson, 2010). Wu et al. (2015) generated within-category and cross-category stimuli using speech and non-speech soundwaves imitating two different tones of the Mandarin monosyllable /pa/. The stimuli were heard by Mandarin L1 speakers, who were musicians and non-musicians. It was found that there was significant effect of group (i.e., musicians and non-musicians) on improving within-category discrimination accuracy in their native language. More specifically, the musician group performed significantly better in both speech and non-speech discrimination. Furthermore, the effect of musical training is possible to remain until the older age, since older musicians (at the mean age of around 70 years) were found to have enhanced neural encoding and differentiation of speech, better performance on categorical speech perception, and more robust neural reflections on perception compared with age-matched non-musicians (Bidelman & Alain, 2015).

It Is also worth noting that the combination of language and musical training can enhance sensitivity to lexical pitch as well (Wiener & Bradley, 2023). The study found that a short-term musical training (e.g., 8 weeks) combined with Mandarin language teaching significantly improved tonal sensitivity in participants with English L1. Contrastively, the other two groups (both English L1) solely receiving language teaching or musical training did not show the same level of enhancement. The results may imply the significant combined effect of language and music learning rather than the effect of each predictor on its own. Nevertheless, the study did not further elucidate whether musical training or language learning was the more critical predictor of tonal sensitivity.

2.4 The Impact of Language Backgrounds (tonal vs. non-tonal L1) on Auditory Perception

Musical training has been found to impact tonal perception (e.g., Bidelman & Alain, 2015; Lee et al., 2020; Wu et al., 2015), but different language backgrounds (i.e., tonal vs. non-tonal languages) could possibly influence auditory and tonal processing as well (e.g., Chen et al., 2018; Shen & Froud, 2019; Wang et al., 2022). A study with Chinese-English bilinguals, Spanish-English bilinguals, and English monolingual found that only the first group was more sensitive to environmental sound perception (Wang et al., 2022). During the experiment, when the sound changed from complex tones to environmental sounds, the Chinese-English bilinguals exhibited significantly higher neural reflections than the other two groups, suggesting the advantages of tonal and non-tonal language bilinguals in top-down cognitive mechanism. The researchers of an ERP study (Shen & Froud, 2019) generated Tone 1/4 continuum and Tone 2/3 continuum based on four Mandarin tones, with each continuum containing six deviant types of two tones and with equal distances between all pitch contours. Within each continuum, one pitch contour was chosen as standard sound, one cross-category pair was selected based on the appearance of discrimination peak, and one within-category pair was selected by researchers. The neural responses showed that an effect of deviant types (i.e., within-and cross-category pairs) was only significant for native Mandarin Chinese speakers, but not for advanced learners of Mandarin Chinese nor for native English speakers without knowledge of Mandarin Chinese. The outperformance of native tonal speakers indicated the profound influence of L1 in pitch perception, which was challenging to acquire even for advanced learners of the same tonal language. Therefore, it sheds light on this study to categorise the participants based on their L1 backgrounds.

However, a later study by Chen et al. (2018) encompassing native Dutch and Mandarin speakers may provide an alternative implication. When controlled for musical training, it was native Dutch speakers, rather than native Mandarin speakers that presented significantly earlier detection of lexical tone change. The results suggested that tonal language background does not necessarily lead to increasing sensitivity in pitch perception in language or in music. Moreover, the neural responses to the different soundwaves in Dutch speakers were found to have significant correlates between lexical and musical tones, while this was not exhibited in Mandarin speakers. This outcome may imply that, compared with tonal language speakers, non-tonal language speakers tended to connect lexical tones with musical ones. The differences of neural reflections implied that native tonal language users may use different neural networks for lexical tones and musical tones (Chen et al., 2018). Differences can also be found in hemisphere usage for the specific perception of pitch patterns. For example, compared with native American English speakers, Mandarin Chinese speakers have hemispheric-pitch type interaction, indicating that left hemisphere produced significantly more intense responses to interval pitch while right hemisphere produced significantly more intense responses to contour pitch (Bidelman & Chung, 2015). However, although the American and English participants relied slightly more on left hemisphere in pitch contour, no hemispheric asymmetry was found. This discrepancies in neural responses across native tonal and non-tonal speakers have suggested the mixed influence of languages on pitch perception. Consequently, it is important to consider participants' L1 background in this study.

Moreover, a bidirectional relationship between music and language was discovered in the research including English musicians, English non-musicians, and Cantonese non-musicians (Bidelman et al., 2013). The research found that when the musical pitches in melody discrimination task were similar to Cantonese tones, the Cantonese non-musicians and English musicians both performed significantly better than the English non-musicians. It implied that tonal language background could to some extent compensate for naïve musical abilities. Furthermore, the outperformance of Cantonese speakers than English non-musicians in F0 differentiation and pitch processing speed in the study suggested the facilitated effect of tonal language background on musical pitch perception. These results corroborated the shared neural resources across domains and further indicated the interinfluence between language and music.

2.5 Roles of Auditory Processing in Second Language Acquisition (SLA)

Many previous studies have shown that people with substantial musical exposure tend to exhibit improvements in language learning compared to those lacking musical involvement. A review encompassing 3928 participants (including children, adolescents, and adults) from 62 longitudinal studies published between 1974 and 2022 concluded that musical training can enhance auditory and linguistic processing (Neves et al., 2022). However, it is worth noting that the enhancements were small, and it is questionable if the results were affected by publication bias in that studies with researchers' favourable results and significant effects are more likely to be published (Neves et al., 2022). The review applied trim-and-fill method (Duval & Tweedie, 2000) and the precision-effect test and precision-effect estimate with standard errors (PET-PEESE, Stanley & Doucouliagos, 2014) to investigate the potential existence of publication bias. However, the former method detected bias while the latter did not. After the inclusion of eight missing studies and data correction by the trim-and-fill method, the effect of musical training was found non-significant. The reviewers advocated for further research due to the inconsistent results across methods. While the review focused on the effect of musical training on auditory and linguistic processing, another review (Pino et al., 2023) gathered evidence from language tests. The review suggested that rhythm and melody are important for children in predicting multiple linguistic abilities, such as phonemic, phonological, syntactic, and semantic aspects. Rhythm was found to be positively associated with phonological awareness (Gordon et al., 2015), while the relationship between melody and language development was more complicated (e.g., Cohrdes et al., 2016; Politimou et al., 2019). Children aged between 3-4 years old with better ability to melody perceptual processing could performed better in language grammar (Politimou et al., 2019). For children in the 5 to 7-year age range, melodic discrimination ability was positively correlated with emotional prosody processing ability in languages (Cohrdes et al., 2016).

With the above reviews demonstrating the beneficial effect of musical training on auditory and linguistic gains, it is important to zoom in the impact of auditory processing on SLA among adults, which underscores the empirical implications of present study. It was found that when demographic backgrounds (immersion to L2 environment and length of instruction) and cognitive abilities were controlled for, the auditory processing remained a significant predictor for L2 speech perception proficiency (p = .014) in Japanese L1 young adults (Saito et al., 2022a). The proficiency test in this study was measured by discrimination task using audio materials in general American English produced by a male native speaker and requiring participants to choose the written word they heard in two options. This perception proficiency task inspires the design of tonal discrimination task in this study by inviting a native speaker. However, since this study requires participants to have no prior knowledge of the target tonal language, the orthographical displays of the words will not be presented in this experiment. Another study by Saito and colleagues (2022b) further investigated which type of auditory processing training could enhance L2 speech acquisition in Japanese L1 and English L2 young adults. It was found that the phonetic-only group improved English [æ] and [ʌ] discrimination skills while acoustic-only and acoustic-phonetic groups improved in both discriminating vowels and detecting F2 variations. The comparable results of acoustic-only group to other two groups in discrimination task implied the transferrable skills acquired by acoustic training in SLA (Saito et al., 2022b).

The auditory processing ability of discrimination is critical for SLA according to Auditory Precision Hypothesis-L2 (Saito, 2023). Based on the empirical research presenting the significant link between pitch discrimination and the linguistic rule learning in infancy and adulthood (Mueller et al., 2012), the hypothesis further predicts that the precise auditory processing allows the utilisation of input opportunities at maximum and leads to enhanced L2 speech learning (Saito, 2023). The hypothesis aligns with the findings from previous studies (Saito et al., 2022a; 2022b) and is manifested by the empirical research (Saito, 2023) suggesting a moderate-to-strong correlation (r = .4-.6) between L2 speech proficiency and both biographic and auditory factors.

The relationship between musical training and linguistic gains should be interpreted with caution. After examining 60 studies from 23 countries, a systematic review investigating the linguistic effects of using songs in formal settings (e.g., schools, after-school clubs) argued that the causal links between the use of songs and its linguistic outcomes could not be validly confirmed (Hamilton et al., 2024). The main reason was that among all these intervention studies, some research designs were not meticulous, such as inappropriate measurements for outcomes and lack of control groups. This review also mentioned publication of bias like the review by Neves et al. (2022) which provided neural evidence of the benefits of musical training in enhancing auditory and linguistic processing. The publication of bias possibly reflects the expectation of improving linguistic skills via musical intervention, whereas it also implies the turbulence and complexity in the field of musical and linguistic interdisciplinary research. Therefore, the examination of research results should be performed with caution.

Considering the increasing attention to auditory processing as aptitude framework and its influence on SLA, it is necessary to conduct more studies in this field. It is worth noting that even though numerous studies examined the influence of musicianship and language backgrounds on auditory perception, few encompassed both factors simultaneously. Moreover, few considered the tonal connection between music and language when studying the ability of tonal perception. Building on prior academic research, this study investigating the predictive roles of musicianship and L1 backgrounds (i.e., tonal vs. non-tonal) on tonal perception is essential and meaningful.

2.6 Why does Music Help? – Potential Theoretical Framework

There are a couple of theoretical frameworks to interpret the positive influence of musical training on language skills. The OPERA hypothesis, proposed by Patel (2011, 2012), suggests that musical training benefits neural plasticity in speech processing by fulfilling five conditions: 1) Overlap: the overlapping neural networks in processing acoustic information in music and speech, 2) Precision: higher requirements in precision when the shared neural networks processing music than speech, 3) Emotion: intense positive emotion stimulated by musical engagements within the networks, 4) Repetition: repetitive musical engagements happening within the networks, and 5) Attention: zoomed attention correlated with musical engagements within the networks. To clarify, when these conditions are met, the OPERA hypothesis posits that since neural plasticity allows the precision required for speech processing in shared neural networks to be improved by musical training, the musical training is beneficial for speech processing (Patel, 2011).

By highlighting the connections between speech encoding and musical training, the hypothesis can elaborate music-to-language transferrable effects from a theoretical view. The OPERA hypothesis can explain why in numerous studies musicians performed significantly better than non-musicians in terms of pitch and temporal sound information processing (Zendel & Alain, 2009), detection of speech and non-speech sounds in noisy environments (Hennessy et al., 2022; Rogala et al., 2023), speech perception (Sajjadi et al., 2021), speech production (Götz et al., 2023), and native stress pattern perception (Choi, 2022).

The OPERA hypothesis mentioned speech processing rather than language processing, whereas it is possible to use this framework in current study considering a couple of reasons: the OPERA hypothesis establishes for all the acoustic features shared in music and speech which includes pitch processing (Patel, 2011); and the concept of pitch exists in both speech perception and tonal language acquisition. Furthermore, Patel (2011) applied the hypothesis to discuss the facilitation of musical training in linguistic reading skills by referring to empirical work (e.g., Moreno et al., 2009). The OPERA hypothesis was also adopted to explain L2 learning in kindergarten (Kulset, 2015).

However, the framework may not account for all types of sounds. For example, a study by Miśkiewicz et al. (2018) encompassing 10 musicians and 10 non-musicians did not find significant difference between the groups in terms of the recognition and detection thresholds for environmental sounds (i.e., natural occurrences of sounds except speech and music). The finding seems to be contradictory to the OPERA hypothesis, since the latter suggests that musical training would enhance language perception abilities. One possible explanation is that the experimental materials did not use speech sounds. It is possible that frequent musical training may benefit neural plasticity solely regarding pitch and speech sound processing but not natural sounds.

2.7 Tones in Thai Language

In this study, a general knowledge of Thai language is essential since the words in audio materials were derived from this language. Thai (*ภาษาไทย* 'phaasǎa thay', Jenny, 2019, p. 559) is a member of the Tai language family and is the official national language in Thailand (Hudak, 2009). Thai has five tones and tonal alternations result in changes in semantics. The pitch contour, pitch height, exemplar word (in romanised transcription and tone marker) and its meaning of each tone are presented in Table 1.

Table 1

Presentation of Thai Tones (Hudak, 2009, p. 664)

2.8 Research Questions and Predictions of Current Study

The research questions are as follows:

(1) Do musical abilities influence tonal perception?

(2) Does language background (i.e., tonal vs. non-tonal L1) influence the tonal perception?

The current study investigates how musicianship and first language background influences tonal perception in adults by applying quantitative methods. Participants are grouped into two groups: 1) tonal language as L1, and 2) non-tonal language as L1. Here present two predictions:

(1) Regardless of L1 backgrounds, there is a significantly positive relationship between musicianship and tonal perception.
(2) It is highly likely that the two language groups perform significantly different in tonal discrimination task, indicating different abilities in tonal perception.

3 Methodology

This study employs an experimental design to investigate how musicianship and language background influence tonal perception. This chapter will provide a detailed and comprehensive presentation of the participants, experiment materials, procedures, and statistic models applied throughout the study.

3.1 Ethics

The study (including piloting) was reviewed by the CUREC (Central University Research Ethics Committee) and approved as shown in Figure 2 (see Appendix A). The research began after the CUREC approval. During the study, there was a need to amend the CUREC form to include more participants. The final approval is shown in Figure 3 (see Appendix A). The participant information sheet for each approval is shown in Appendices B1 and B2 respectively. The participant recruitment advert is shown in Appendix C.

3.2 Participants

The target number of participants was obtained via G*Power[4] version 3.1.9.6 (Faul et al., 2007; Faul et al., 2009). The effect size was set based on the meta-analysis by Liu et al. (2023), which examined the effect of languages on three categories of music processing tasks and found the estimated effects in parentheses: melodic discrimination (0.50), pitch discrimination (0.26), and rhythm tasks (-0.01). Since the test used to measure musical ability in the present study only included melodic and rhythmical tasks, the average of melodic discrimination and rhythm tasks was computed. Therefore, the power of language background is 0.255. The meta-analysis (Liu et al., 2023) also considered the power of musical processing tasks (shown in parentheses): melodic discrimination (0.49), fine-grained pitch discrimination (0.29), and rhythm tasks (0.48). For the fine-grained pitch discrimination was not included in the musical aptitude test in this study, the power of musical training is 0.485. Therefore, since the predictors of this study are language background and musical abilities, the final effect size is set as 0.37. Other parameters were set as: α error probability as 0.05 and power (1-β error probability) as 0.8. The a priori power analysis suggested that the sample sizes for Mann-Whitney test would be 244, for MANOVA would be 166, and for ANCOVA would be 60 (see Appendix D).

This study intended to recruit approximately 250 participants representing various language backgrounds. The ideal participant pool would consist of 125 individuals with a tonal L1 and 125 individuals with a non-tonal L1, meeting the total sample size and the sample size of each group indicated by the power analysis at minimum (i.e., 122). Extra spaces would help accommodate any possible withdrawals and invalid data during the process.

However, a total number of 198 participants ended up taking part the experiment. Twenty-eight were excluded after manual screening. Seven of them had provided irrelevant or unreasonable responses to the questions. Nineteen of them skipped questions about musical training experience, which could hinder the interpretation of outcomes. Eleven of them answered questions inconsistently. Seven of them had more than one type of invalid data.

[4] https://www.psychologie.hhu.de/arbeitsgruppen/allgemeine-psychologie-und-arbeitspsychologie/gpower

In the end, there were 168 people from various linguistic and musical backgrounds providing valid data in this study. Among them, 98 participants had tonal languages as their L1. The remaining 70 participants had a non-tonal language as L1(s). With the power analysis indicating the requirement of 64 participants from each language background at minimum, the number of participants failed to exceed the threshold of Mann-Whitney test in this study. Consequently, its results should be interpreted with caution. Contrastively, the number reached the minimum number to apply MANOVA and ANCOVA, ensuring robust statistical power of this statistic model. An overview of 168 participants' profiles is shown in Table 2.

Table 2

Participants' Profile

Linguistic background		
Tonal language as L1	Cantonese	4
	Cantonese, Mandarin	3
	Changshahua, Mandarin	1
	Chinese (unspecified)	65
	Gan Chinese	1
	Hubeinese, Mandarin	1
	Mandarin	20
	Afrikaans, English	1
	Cantonese, English	1
	English, Tsonga	1
Non-tonal language as L1	Bengali	1
	Catalan, Spanish	1
	English	16
	French	1
	German, Lower Saxon	1
	Greek	2
	Hebrew	1
	Hungarian	2
	Ilonggo, Filipino	1
	Indonesian	1
	Italian	4
	Korean	3
	Polish	12
	Portuguese	9
	Romanian	1
	Russian	3

	Spanish	8
	Tegulu	1
	Urdu	1
	Uzbek	1
	Musical background	
Tonal language as L1	Instrumental training	74
	Singing training	35
	Instrumental and singing training	31
Non-tonal language as L1	Instrumental training	41
	Singing training	14
	Instrumental and singing training	13

Participants were all adults (*M* = 26.38, *SD* = 7.15, range = 18-54) with no severe listening impairments. Considering that the impact of gender difference is negligible in musical perception (Bertolo et al., 2023), this study did not collect information on participants' genders. Any participants with relevant knowledge on Thai language were excluded. This criterion guaranteed that the participants did not know the tone rules or phonotactic constraints in Thai and all their judgements were therefore based on their sensitivity to tonal discrimination. Potential participants were contacted via the dissemination of poster advert, posts on social media and the student researcher's personal website. The student researcher's personal contact was included, and participants could get in touch to receive information about the experiment (see Appendix C).

Participation was voluntary. If the participants quitted the experiment the halfway, their data was deleted automatically. This information was made explicit in the information sheet. Both the pilot study and the formal experiment began after the CUREC approval and there was no ethical implication for participants.

3.3 Piloting

Five adults (all aged between 22 and 28 years old; 2 males and 3 females) were invited to take part in the online experiment.

The pilot study aimed to find out the estimated time to complete the whole procedure and to determine the feasibility of the experiment. It is found that the average time was around 35 minutes. After the pilot study, the following amendments were made as suggested by pilot participants.

1. The tonal discrimination task was moved to the beginning of the procedure. The reason was proposed by one participant in that if the task came after the musical ability test, the participants may be primed or even trained to perceive tonal differences. Their awareness of different tones may be raised so the result of tonal discrimination task may not be accurate.

2. The instructions of the tonal discrimination task were made clearer in terms of how much time participants had to respond. A reminder about the rest time was also added to each question since it assisted participants in responding within the time limit.

3. The interface of the participant information sheet, tonal discrimination task, and language background questionnaire were made more user-friendly.

4. A rest session was added between tasks. Participants could take a break if they needed to before continuing to the next task as several pilot participants highlighted that the lack of break would hinder their performance on subsequent questions.

3.4 Materials

3.4.1 Questionnaires

The self-report questionnaire measured demographic features (i.e., age), language background (i.e., first language), and musical training experience. There were three questions in total (see Appendix E). The first question recorded all the languages (including first languages) that participants have studied or learned and the time length they have spent using each of them. The two musical training questions recorded instrument and singing training experience, specifically the starting age and years of performance/training. All the questions and their corresponding specific instructions were adapted from Language History Questionnaire version 3.0 (LHQ3) by Li et al. (2020).

3.4.2 Music Ear Test (MET)

The MET is a measurement for musical skills (Swaminathan et al., 2021; Wallentin et al., 2010a, 2010b). Its reliability and validity have been confirmed by a large-scale study containing 604 participants (Correia et al., 2021). The test includes two subtests: a melody subtest and a rhythm subtest. Each subtest contains 52 trials and the MET lasts around 18 minutes in total. Melody is presented by piano sounds. Each trial contains two melodies with 3-8 tones each. The two melodies are separated by empty sound. Within the 52 trials, there are 26 trials, each containing two melodies differentiated in terms of one pitch, and other 26 trials each including two identical melodies. Among the 52 trials, 25 trials contain chromatic tones, and the remaining 27 trials include 20 Major keys and 7 Minor keys.

Rhythm is presented by wood block beats. Each trial contains two rhythms containing 4-11 beats each. The two sounds are separated by metronome at a lower volume. Within the 52 trials, there are also 26 trials, each containing two rhythms differentiated in terms of one rhythmic change maximally, while the other 26 trials including two identical rhythms. Among the 52 trials, 21 contain triplets and 31 contain subdivisions of the beat. The trials of each subtest are randomly allocated and both melody and rhythm are played at 100 beats per minutes.

Before starting each subtest, participants were asked to answer two sample trials to familiarise themselves with the process. In each trial, participants heard two sounds and then pressed ‘Yes’ if they believed the melodies/rhythms in each audio were identical and ‘No’ if they did not. For each trial, participants needed to respond in just over one second. They only received feedback on the two trials and no further feedback was given during the performance of the task. The MET was adapted from Correia et al. (2021) using an available template on Gorilla[5].

In this research the template version was directly applied without further modification. After the completion, this experiment obtained three types of scores after participants' completion: melody subtest score, rhythm subtest score, and MET total score.

3.4.3 Tonal Discrimination Task

This task included 20 pairs of Thai disyllabic words produced by a male native Thai speaker. Each pair contained two words which were identical in terms of consonants and vowels and different in terms of tones (see Appendix F). These words were selected from the book Basic Thai 1 (Liao, 2008), a Thai language learning book for Chinese speakers. The Thai orthography and the corresponding English translation is presented in the 2nd column and their romanised pronunciation is presented in the 4th column. The romanisation in the 6th column is the altered version of the real Thai words. All of the words in the 6th column (except 1B 'ddii^{3}zai^{1}' meaning happy in English) are nonsense words. The superscripts in the two romanisation columns indicate the tone in Thai. Each number from one to five indicates middle tone, low tone, falling tone, high tone, and rising tone respectively. There were 11 pairs of words without any alternation of tones and nine pairs with words in different tones. The forms of alternation included a change between any of the five tones in Thai. For example, the pair 1A and 1B indicates the change between low tone (tone 2) and middle tone (tone 1). It is worth noting that among the nine pairs of alternations, five pairs differed in their first syllables (i.e., 1A and 1B, 3A and 3B, 4A and 4B, 10A and 10B, 20A and 20B) while the remaining four pairs differed in their second syllables. After listening to each pair, participants had around three seconds to answer if the tones of two words were the same or not by clicking the button 'Yes' or 'No'. Same as in the MET, participants answered two trial questions with answers provided (see Appendix G) to familiarise themselves before the formal trial start. The correct judgement of the different/identical tones was awarded one point, with 20 as the full mark.

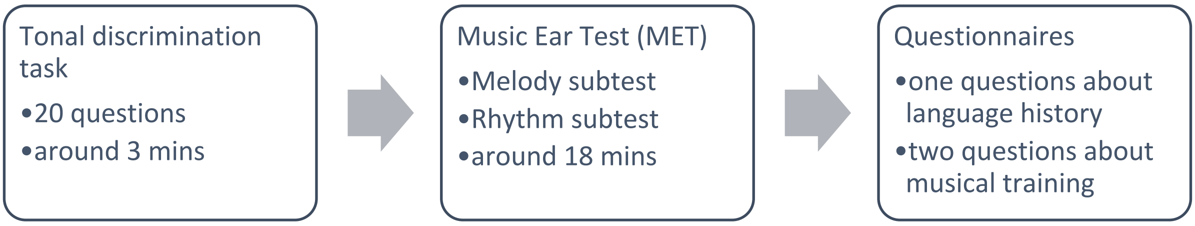
3.5 Procedures

The experiment was programmed on Gorilla. All the participants needed to read and sign the participant information sheet (see Appendices B1 and B2) by ticking the confirmation and agreement square boxes before they started the experiment. The sequence for the experiment tasks after amendments resulting from pilot study became: 1) Tonal discrimination task, 2) Musical Ear Test (MET), and 3) Language Background questionnaire (see Figure 1).

Figure 1

Procedures of Whole Experiment Process

[5] https://app.gorilla.sc/openmaterials/218554



3.6 Statistical Methods

3.6.1 Mann-Whitney U Test

Mann-Whitney U test (Mann & Whitney, 1947) was conducted to detect if there were significant differences across two groups. Since the MET total scores and subtest scores were not normally distributed, Mann-Whitney U test was applied to compare the differences in musical abilities between the two language groups (Hollander et al., 2013). To conduct this test, the null hypothesis (H_0) and alternative hypothesis (H_1) were formulated below.

H_0: There is no difference in mean ranks of MET scores between tonal and non-tonal L1 groups.

H_1: There is a difference in mean ranks of MET scores between tonal and non-tonal L1 groups.

3.6.2 Spearman's Correlation Coefficient

Considering that not all test scores were normally distributed across language groups, Spearman's rho correlation was selected to analyse the inter-relationships among the test scores. Correlation depicts the magnitude of association between two or more variables from a statistical perspective, with bivariate correlation examining the relationship between two variables (Meyers et al., 2009). Coefficient ranges from -1 to 1: if the coefficient is 1, it indicates that the two variables are positively and perfectly correlated; if the coefficient is -1, it indicates that the two variables are negatively and perfectly correlated; if the coefficient is 0, the two variables do not have linear or monotonic relationship (Gogtay & Thatte, 2017). The correlation between two variables can be very strong, strong, moderate, weak, or negligible depending on the value of the coefficient (Schober et al., 2018).

3.6.3 Linear Mixed-effect Model

The linear mixed-effect model (LMM) describes the relationship between a dependent variable and its numerous predictors. The predictors have different levels of effects: the effects that are consistent across levels are called fixed effects, while the ones that are unpredictable and vary across levels are called random effects (Bates, 2010). The model does not require the normality of residuals (Bates, 2010). Therefore, it can be used in this study regardless of the normality of the data. It is worth mentioning that G*Power version 3.1.9.6 (Faul et al., 2007; Faul et al., 2009) does not have an *a priori* condition to calculate the sample size directly. The MANOVA was selected to derive the sample size in G*Power in that LMM is a more advanced method improved from MANOVA (de Melo et al., 2022).

This LMM was conducted in RStudio (Posit Team, 2024), employing the lme4 package (Gałecki & Burzykowski, 2013). The glmer function in lme4 package is suitable for binary outcomes and it can generate precise, consistent, and reliable outcomes (Lee & Grimm, 2018). One notable advantage of using this function is its capacity to address item-level heterogeneity, controlling individual-specific differences and different impact of each test item on participants (Gilbert, 2023). The glmer function was therefore applied in RStudio (Posit Team, 2024) in this study for binary responses (i.e., 'Yes' vs. 'No') to each question in tonal discrimination task for each participant.

3.6.4 Nonparametric ANCOVA

The Analysis of Covariance (ANCOVA), blending the analysis of variances and regression, is a statistical technique to investigate the influence of predictor(s) on dependent variable while controlling the impact of covariate (Field, 2014; Syam et al., 2023). The covariate is a continuous variable that does not serve as the object of the research but could have an impact on the dependent variable. When applying the method, it is important to confirm that the covariate does not account for the variances explained by predictors (Field, 2014). ANCOVA was initially developed to improve the accuracy of experimental data analysis (Fisher, 1932). Nowadays, it is widely used in various fields like psychology, education, chemistry, etc (Gjana & Kosova, 2021; Gupta et al., 2024; Mazlomi Barm Sabz et al., 2021). In this study, the MET score is a continuous variable and will be controlled as covariate when examining the influence of language background on tonal perception. However, considering the nonnormality of the residuals, nonparametric ANCOVA should be conducted.

Regarding the sample size, it is suggested that applying the corresponding nonparametric method of a certain parametric statistic model usually needs extra 15% of the parametric test sample size (Lehmann, 1998). In this study, the ANCOVA requires 60 participants at minimum and therefore, the nonparametric ANCOVA requires at least 69 participants.

This model was performed via RStudio (Posit Team, 2024), utilising npsm package (Kloke & McKean, 2024) and rfit function (Kloke & McKean, 2012). Rfit can be used to compute the rank-based ANCOVA analysis (Hettmansperger & Kloke, 2010; Kloke & McKean, 2014).

4 Results

This chapter will firstly introduce descriptive statistics of participants' MET melody scores, MET rhythm scores, MET total scores, and tonal discrimination task scores. The correlation coefficients between variables and the outcomes of all the statistic models will also be presented.

In this study, the software Microsoft Excel Version 16.86 (Microsoft, 2024), SPSS Version 29.0.1.0 (IBM Corps, 2023), and RStudio Version 2024.04.2.764 (Posit Team, 2024) were used for data management and analysis. The language background variable was dummy coded, with 0 indicating non-tonal L1 and 1 indicating tonal L1. The normal distribution of the MET melody scores, MET rhythm scores, MET total scores, and tonal discrimination task scores was checked according to the skewness and kurtosis values (see Table 3).

Table 3

Skewness and Kurtosis of All Test Results

Group	Test type	Skewness	Kurtosis
All participants	MET melody subtest	-.864	2.728
	MET rhythm subtest	-.749	1.924
	MET	-1.001	3.509
	Tonal discrimination task	-1.181	1.278
Tonal language as L1	MET melody subtest	-1.527	5.515
	MET rhythm subtest	-.984	3.267
	MET	-1.449	5.260
	Tonal discrimination task	-1.664	2.704
Non-tonal language as L1	MET melody subtest	.075	-.613
	MET rhythm subtest	-.336	-.601
	MET	-.128	-.413
	Tonal discrimination task	-.732	.605

According to George and Mallery (2010), the skewness and kurtosis within ±1.5 are reasonable in many cases, (within ±1 if for a small sample). Hatem et al. (2022) also have suggested that the ranges of skewness and kurtosis within ±2 are acceptable as well. Considering the sample size 168, this study considers skewness and kurtosis within ±2 as the benchmarks for normality. In this study, the result of tonal discrimination task is the dependent variable.

One outlier with the tonal discrimination task score as 6 was detected. Their data were deleted as they substantially affected the normal distribution of the data. After the deletion, the descriptive statistics of participants' age remained roughly unchanged (M = 26.40, SD = 7.16). Moreover, the skewness and kurtosis of the tonal discrimination task score indicated that they were normally distributed (see Table 4, and Appendix H for visual illustrations). Another spreadsheet with a detailed information about each participant's response to each item in tonal discrimination task was created as well, as shown in Table 5. Each participant occupied 20 rows, corresponding to 20 questions in tonal discrimination task. Each participant was anonymised using an 8-digit ID number. The Item column indicated the question number in the task and the 1 and 0 in the Response column respectively meant right and wrong answer, respectively.

Table 4

Skewness and Kurtosis of Tonal Discrimination Task Scores According to Different L1 Groups

Data type	Group	Skewness	Kurtosis
Tonal discrimination task score	All participants	-.972	.404
	Tonal language as L1	-1.336	1.042
	Non-tonal language as L1	-.732	.605

Table 5

Participant's Response to Each Item in Tonal Discrimination Task (Partial)

Participants	MET Melody Score	MET Rhythm Score	MET Total Score	L1	Item	Response
11085059	39	44	83	tonal	1	0
11085059	39	44	83	tonal	2	0
11085059	39	44	83	tonal	3	0
11085059	39	44	83	tonal	4	1
11085059	39	44	83	tonal	5	1
11085059	39	44	83	tonal	6	0
11085059	39	44	83	tonal	7	1
11085059	39	44	83	tonal	8	1
11085059	39	44	83	tonal	9	1
11085059	39	44	83	tonal	10	0
11085059	39	44	83	tonal	11	1
11085059	39	44	83	tonal	12	1
11085059	39	44	83	tonal	13	1
11085059	39	44	83	tonal	14	1
11085059	39	44	83	tonal	15	1
11085059	39	44	83	tonal	16	1
11085059	39	44	83	tonal	17	1
11085059	39	44	83	tonal	18	1
11085059	39	44	83	tonal	19	0
11085059	39	44	83	tonal	20	1
11084290	38	34	72	tonal	1	1
11084290	38	34	72	tonal	2	1
11084290	38	34	72	tonal	3	1
11084290	38	34	72	tonal	4	1
11084290	38	34	72	tonal	5	1

11084290	38	34	72	tonal	6	1
11084290	38	34	72	tonal	7	1
11084290	38	34	72	tonal	8	1
11084290	38	34	72	tonal	9	1
11084290	38	34	72	tonal	10	1
11084290	38	34	72	tonal	11	1
11084290	38	34	72	tonal	12	1
11084290	38	34	72	tonal	13	1
11084290	38	34	72	tonal	14	1
11084290	38	34	72	tonal	15	1
11084290	38	34	72	tonal	16	1
11084290	38	34	72	tonal	17	1
11084290	38	34	72	tonal	18	1
11084290	38	34	72	tonal	19	1
11084290	38	34	72	tonal	20	1

The descriptive statistics of MET melody score, MET rhythm score, MET total score, and tonal discrimination task score is shown in Table 6. Despite similar variability of the MET melody and rhythm scores, the mean score of rhythm test was lower than that of melody test. This implies that it is reasonable to separate melody scores from rhythm scores in subsequent data analysis. Further details of both language groups were presented in Table 7.

Table 6

Descriptive Statistics of All Participants

Score type	**Full score**	**Range**	**Minimum**	**Maximum**	***M***	***SD***
MET melody subtest score	52	45	7	52	35.780	6.388
MET rhythm subtest score	52	46	0	46	32.810	6.924
MET total score	104	86	7	93	68.600	11.587
Tonal discrimination task score	20	12	8	20	16.920	2.628

Note. *M* = mean, *SD* = standard deviation

Table 7

Descriptive Statistics of Participants in Tonal and Non-tonal L1 Groups

Language group	**Test type**	**Minimum**	**Maximum**	**Range**	***M***	***SD***
Tonal language as L1	MET melody subtest	7	52	45	37.250	6.354
	MET rhythm subtest	0	46	46	32.600	7.122
	MET	7	93	86	69.850	12.178
	Tonal discrimination task	7	20	14	17.510	2.534
Non-tonal language as L1	MET melody subtest	22	48	26	33.760	5.901
	MET rhythm subtest	19	45	26	33.110	6.680

	MET	47	45	47	66.870	10.558
	Tonal discrimination task	8	20	12	16.100	2.555

4.1 Differences in Musical Aptitudes Across L1 Backgrounds

The Mann-Whitney U test (two-tailed) was conducted to examine the differences of MET total scores across tonal and non-tonal L1 groups. In general, it seemed that tonal L1 group demonstrated slightly higher musical aptitude (mean rank = 90.06, sum of ranks = 8736.00) than their non-tonal counterparts (mean rank: 75.60, sum of ranks = 5292.00). However, there was a marginally significant difference in MET scores between the two language groups (U = 2807.00, p = .056).

The Mann-Whitney U tests were also conducted using MET melody scores and MET rhythm scores as the dependent variable, respectively. The tonal L1 group achieved higher scores in MET melody subtest (mean rank = 96.10, sum of ranks = 9322.00) than non-tonal L1 group (mean rank = 67.23, sum of ranks = 4706.00). The difference in MET melody scores was statistically significant across groups (U = 2221.00, p < .001). By contrast, the non-tonal L1 group (mean rank = 85.84, sum of ranks = 6009.00) outperformed tonal L1 group (mean rank = 82.67, sum of ranks = 8019.00) in the MET rhythm subtest. However, this difference was not statistically significant (U = 3266.00, p = .675). In short, tonal L1 group performed significantly better than non-tonal L1 group in melodic abilities, while there were no significant differences in MET rhythm scores and MET total score across language groups.

However, it should be noted that given the limited number of participants (167 instead of 244 as required according to the power analysis), the results of Mann-Whitney U tests should be interpreted with caution. It is possible that the marginal significance in melody subtest and insignificant cross-group differences in rhythm subtest and MET total score were caused by insufficient number of participants.

4.2 Effects of Musicianship and L1 Backgrounds on Tonal Perception

For all participants regardless of their language backgrounds, the correlations among the variables in were presented in Table 8. The results indicated that all variables were associated with tonal discrimination task scores, with the MET melody scores showing the strongest correlation (r_s = .445). All Spearman's correlation coefficients with tonal discrimination task scores were weak (.10 < r_s < .39) to moderate (.40 < r_s < .69) (Schober et al., 2018). For the MET total scores, on the other hand, they were strongly correlated with melody (r_s = .827) and rhythm scores (r_s = .865), indicating potential collinearity (r > .80 or .90, Field, 2014). These strong associations were intuitive, since the MET total scores consisted of both melody and rhythm scores.

Table 8

Spearman's Correlation Between Variables in All Participants

	Tonal discrimination task score	Language background	MET total score	MET melody score
Language background	.316** (< .001)			
MET total score	.349** (< .001)	.148 (.056)		

MET melody score	.445* (< .001)	.296** (< .001)	.827** (.000)	
MET rhythm score	.191* (.014)	-.033 (.677)	.865** (< .001)	.456** (< .001)

Note. The *p*-value is presented within parentheses. ** indicates significant correlation at the 0.01 level (two-tailed) and * indicates significant correlation at the 0.05 level (two-tailed).

To examine whether the magnitude of these correlations differed across L1 backgrounds, participants were categorised into two groups: 1) tonal languages as L1, and 2) non-tonal languages as L1. The detailed results are presented in Table 9. For both tonal L1 and non-tonal L1 groups, there were weak correlations between overall musicianship, measured by MET total score, and tonal discrimination task scores (r_s = .290 for tonal L1 group and r_s = .372 for non-tonal L1 group), since the coefficients were between .10 and .39 (Schober et al., 2018). It was also found that the MET melody scores had the strongest correlation with tonal discrimination task scores in both groups (r_s = .393 for tonal L1 group and r_s = .384 for non-tonal L1 group). The Fisher's *z*-transformation was conducted (Lowry, 2023) to examine if there were significant differences in magnitude of associations with tonal discrimination task scores across groups. It was shown that for both groups, there were no significant differences in correlations between tonal discrimination task score and MET total score (*z* = -.58, *p* = .562), MET rhythm score (*z* = -.27, *p* = .787), or MET melody score (*z* = .07, *p* = .944).

Table 9

Spearman's Correlation Between Overall Musicianship on Tonal Discrimination Task in Tonal and Non-Tonal L1 Groups

Language group	**Test type**	**Tonal discrimination task score**	**MET total score**	**MET rhythm score**
Tonal language as L1	MET total score	.290** (.004)		
	MET rhythm score	.190 (.062)	.908** (< .001)	
	MET melody score	.393* (< .001)	.842** (< .001)	.568** (< .001)
Non-tonal language as L1	MET total score	.372** (.002)		
	MET rhythm score	.232 (.054)	.388** (< .001)	
	MET melody score	.384* (.001)	.802** (< .001)	.388** (< .001)

It is common to use regression models to investigate the interaction between two continuous variables, however, the residuals of dependent variable (i.e., tonal discrimination task score) was not normally distributed in this study. Therefore, the regression model could not be performed for data analysis. Instead, the linear mixed-effect model was conducted. Considering that the MET includes two subsections, melody subtest and rhythm subtest, three models were performed via RStudio (Posit Team, 2024) to detect the most significant predictor (see Table 10). The displays of software interface are shown in Appendix I.

Table 10

Summary of Mixed-effect Model Analyses of Tonal Discrimination Task Response to L1 Backgrounds and MET Scores

Models	**Fixed effects**	***β***	***SE***	***z***	***p***	**Random effects**	**Variances**	***SD***

Model 0	Intercept	-.431	.527	-.818	.414	Participants	.668	.818
	L1 background	.533	.169	3.156	.002**	Item	.790	.889
	MET total score	.034	.007	4.717	< .001***			
Model 1	Intercept	-.527	.500	-1.055	.291	Participants	.640	.800
	L1 background	.394	.171	2.302	.021*	Item	.791	.890
	MET melody score	.070	.013	5.426	< .001***			
Model 2	Intercept	.533	.458	1.138	.255	Participants	.746	.864
	L1 background	.651	.174	3.737	< .001***	Item	.789	.888
	MET rhythm score	.039	.012	3.167	.002**			

Note. β = estimate, *SE* = standard error, *z* = z scores, R^2 = *R* squared, Item means each question in tonal discrimination task

* Statistical significance at the $p < .05$

** Statistical significance at the $p < .01$

*** Statistical significance at the $p < .001$

All the three models (Model 0, Model 1, and Model 2) controlled the influence of variances across individuals and questions in the tonal discrimination task. Model 0 indicated that the language background and MET total score respectively had a significant relationship with tonal perception. Since MET total score ($p < .001$) had lower *p*-value than L1 background ($p = .0016$), it meant that MET total score was a more significant predictor than L1 background.

In Model 1, L1 background and MET melody subtest score were selected as two predictors. It was found that the two predictors were both significant, with MET melody score ($p < .001$) as more significant than L1 background ($p = .021$). It is worth underscoring that the *p*-value of L1 background in Model 1 was higher than that in Model 0, possibly indicating that it was MET melody score within MET total score that influenced tonal discrimination more significantly.

In Model 2, the impact of L1 background and MET rhythm score on tonal discrimination task was examined. Unlike the previous two models where musical scores were more significant predictors than L1 background, a more statistically significant effect of L1 background ($p < .001$***) than MET rhythm ($p = .002$**) was discovered. It indicated that compared with L1 background, melodic skills rather than rhythmic skills played a more important role in tonal perception, suggesting the criticality of cultivating melodic abilities in musical training.

It is worth highlighting that the LMM included data from 167 participants, exceeding the minimum sample size required from the initial power analysis (i.e., 166), thereby enhancing the statistical power and robustness of the findings. In other words, the reliability in the observed effect is guaranteed for interpretation.

4.3 Performance on Tonal Discrimination Task by Tonal and Non-tonal L1 Groups

The nonparametric ANCOVA was conducted in RStudio (Posit Team, 2024), controlling MET total scores as covariate and L1 backgrounds as predictor. It was presented in Table 11 (see Appendix J for display of software interface) that for every unit of increase in MET total score, the tonal discrimination task score was expected to improve by .067 when the participant's language backgrounds were identical. If the participants were native tonal language speakers, when they have the same MET performance, their tonal discrimination task scores were expected to be 1.067 point higher than native non-tonal language speakers. Both language backgrounds (p = .005) and MET total scores (p < .001) were significant predictors. The two variables, language backgrounds and MET total scores, could explain 14.3% of variances in tonal discrimination task scores. The dispersion test compared the model with and without the covariate (i.e., MET total scores). Its outcome (p = .000) indicated that including the covariate could significantly improve the model fit. Therefore, the significant effect of musical aptitudes could not be neglected when conducting experiments about language backgrounds and tonal discrimination.

Table 11

Coefficients of nonparametric ANCOVA

Source	β	***SE***	***t***	***p***
Intercept	10.500	1.102	9.532	< .001***
Language backgrounds	1.067	.375	2.843	.005**
MET total scores	.067	.016	4.291	< .001***

Note. Multiple R^2 (Robust) = 0.143; reduction in dispersion test = 13.888, p = .000.

5 Discussion

This chapter will firstly discuss if (non-)tonal language backgrounds influence the performance on musical aptitudes, measured by the MET, and if there is a significant difference across groups. The following two subsections will present the answers to the two research questions and their alignments with the predictions. Whether the interrelation between musicianship and speech processing exists will also be exhibited along with the discussion, linking back to previsou research. The fourth subsection will endeavour to explain why the results were produced and relate back to the OPERA hypothesis. The suitability of this hypothesis for explaining the findings will be considered. The next subsection will provide educational implications for language teaching in classroom. Lastly, several limitations of the current study, their corresponding improvements, and suggestions for future studies will be critically presented.

5.1 Musical Performance across Tonal and Non-tonal L1 Groups

In general, the tonal and non-tonal L1 groups did not present significant differences in results of MET total score and MET rhythm score as suggested by Mann-Whitney test. The significantly better performance by tonal L1 group rather than non-tonal L1 group in MET melody subtest indicated the facilitation of native tonal languages in perceiving new tonal words. This finding resonates with previous neurolinguistic research which indicated that native tonal language speakers were more sensitive in auditory processing compared with native non-tonal speakers even though both groups used English as L2 (Wang et al., 2022). The study by Wang et al. (2022) used environmental sounds as materials, whereas another study (Shen & Froud, 2019) used deviations of lexical tones. The latter is more similar to current study design which also applied audio materials from a natural language (i.e., Thai). The study by Shen and Froud (2017) demonstrated that participants' own L1 background facilitated detection of various pitch contours which are deviated from their L1. Current study further indicated that the transfer of auditory processing was possible from tonal L1 to musical melodies.

The profound effect of native tonal language on musical performance was manifested by other studies (e.g., Tong et al., 2018; Zhang et al., 2020). The native Chinese speakers were found to have better performance in melody rather than rhythm test, while native Japanese speakers performed oppositely, with higher achievement in rhythm than melody test (Zhang et al., 2020). These results were to some extent parallel to current study. The Mann-Whitney U test only observed marginally significant advanced melodic skills in tonal L1 group while the cross-group differences in melodic skills and MET total scores were not significant. The discrepancies in results may be caused by the lack of participants in this study. However, another study found that the modulating effect of native language: for Cantonese speakers, there was no significant difference in pitch height discrimination task between musicians and non-musicians, while the differences were presented among musicians and non-musicians with non-tonal native language (Tong et al., 2018). In other words, the existence of musicianship was not critical for native tonal speakers since their tonal L1 could enhance their sensitivity of pitch detection to a large extent, while for non-tonal native individuals, musicianship was important for pitch height detection. The impact of native tonal language was manifested in that it could possibly substitute the effect of musical training to some degree. It was resonated by Choi (2021) who found that language experience enhanced pitch perception only in individuals without long-term musical training. In the current study, participants' musicianship was measure as a continuous variable. Therefore, even if there were the transferrable effect of language backgrounds, it could not affect participants with or without musical abilities since the musical aptitudes were a continuous variable in this study. The complicated relationships between musicianship and native (non-)tonal languages and the discrepancies of research designs led to the discussion about their impacts on pitch discrimination as follows.

5.2 Musicianship and L1 Backgrounds as Significant Predictors of Tonal Discrimination

Regarding the first research question, it was predicted that tonal perception would exist for both language groups. The findings confirmed this prediction. Furthermore, both musicianship and L1 backgrounds were significant predictors and there was a significantly positive correlation between tonal discrimination task scores and both musicianship and L1 backgrounds. For all the 167 participants, the Spearman's coefficient showed that the MET melody score had the strongest correlation with tonal discrimination task, with the MET total score also having a significant correlation. The same held for language group separately as well. However, the Fisher's *z*-transformation (two-tailed) did not demonstrate significant differences in the correlation between tonal discrimination task and each test across groups. This indicated that the correlation for tonal L1 speakers was the roughly same as the correlation for non-tonal L1 speakers. In other words, the connection between enhanced MET skills and better tonal discrimination held for both language groups. These results to some degree filled the gaps in previous research. On one hand, the investigations of language background and auditory processing solely focused on these two aspects while the influence of musical aptitude was ignored (e.g., Chen et al., 2018; Shen & Froud, 2019; Wang et al., 2022). On the other hand, the studies about musical training and auditory processing were conducted recruiting all the participants from the same language group (e.g., Tang et al., 2016; Wu et al., 2015). Therefore, the research into the impact of L1/musicianship on auditory processing while also considering cross-group comparisons with the other factor was lacked.

The Model 0 of LMM suggested that both musicianship and L1 backgrounds (i.e., tonal vs. non-tonal) were significant predictors with more substantial effect from musicianship. Together in the comparisons with Models 1 and 2, it was proposed that musical aptitudes in general, melodic skills and rhythmic skills were all significant predictors of tonal discrimination competency. Among three, the MET melody score, with its lowest *p*-value, presented the strongest association with tonal discrimination. The results therefore suggested the multifaceted relationship between tonal perception and musical aptitudes with the emphasis on melodic detection. However, the findings could not corroborate previous review (Pino et al., 2023) which put forward the complicated correlation between pitch processing and improvements in language performance (e.g., enhancements in phonology, syntax, semantics, etc.) in that the dependent variable was solely tonal words in Thai. Contrastively, it could shed light on the benefits of melodic training in improving suprasegmental processing ability such as learning tones in SLA. The finding of another review (Neves et al., 2022) was substantiated by the association found in this study, supporting the beneficial impact of musical training on pitch processing. Even if the publication bias mentioned in the review was concerning, the first-hand data collected by this study could possibly inspire further research.

Arguably, this study might partially provide another empirical evidence for bidirectional music-language interaction as suggested by Bidelman et al. (2013). To remind, the musical ability was reflected by MET and the language processing ability was tested by tonal discrimination task in this study. Regarding the bidirectional relationship, on one hand, the impact of musicianship on language processing was manifested by the correlation between higher MET total score and better tonal discrimination performance. Contrastively, on the other hand, the impact of language background on musical abilities was only marginally significant in this study. It could be interpreted that for different L1 groups, no significant difference in musicianship was detected. However, considering the insufficient number of participants in Mann-Whitney test, this finding should be explained with caution. Therefore, the current research results could robustly reaffirm the music-to-language transfer while the manifestation of language-to-music effect should be followed by further studies.

It is also worth noting that in the comparison between findings of this study with those of previous ones, the different designs of participant categorisation should be considered. To begin with, this study solely dichotomously grouped participants based on their L1 while measured their musicianship as a continuous variable, unlike previous studies which classified participants into musicians and non-musicians (e.g., Bidelman & Alain, 2015; Choi, 2021; Lee et al., 2020; Rogala et al., 2023; Sajjadi et al., 2021). The definitions of musicians and non-musicians were justified by each group of researchers in their studies. However, the categorisation of non-musician and musician could be problematic in the study by Choi (2021) since by measuring musical pitch perception via MET, the researcher should be aware that musical aptitudes were not dichotomously categorised. The dichotomous classification could therefore be inappropriate. Next, the selections of (non-)musicians were based on their musical experience. The inconsistency in understanding of concepts emerged in that musicianship, the focus of the study, was defined as musical skills in Cambridge dictionary (n.d.) and also in this study, while the long-term musical experience was a description of training events. The higher musicianship does not necessarily lead to longer term of musical experience considering the existence of absolute pitch and other musical talents. Consequently, the participant groups and the research focus did not meticulously echo with each other in the study by Choi (2021).

Furthermore, the clear categorical classification of participants into four groups of approximately equal size (i.e., Cantonese musicians, English musicians, Cantonese non-musicians, and English non-musicians) in Choi's study (2021) provided convenience of investigating interactive effects. In the musical pitch perception measured by MET, the native Cantonese speakers achieved significantly higher scores than native English speakers in the non-musician group but not musician group. The interactive language x musicianship effect was found in musical pitch perception since only for non-musicians, Cantonese speakers outperformed English speakers, indicating language-to-music transfer. Among musicians, the tonal language experience did not result in significant differences in pitch discrimination. Therefore, the interactive effect between (non-)tonal language background and (non-)musicianship on pitch discrimination was substantiated in Choi's study (2021).

Contrastively, this study did not find the interactive effect of L1 background and musicianship in tonal discrimination task. As shown by Mann-Whitney test, there was a significantly higher achievement by tonal L1 speakers in MET melody subtest. However, the positive correlation between tonal discrimination task scores and both MET total scores and its subtest scores was not significantly different across language groups as suggested by Fisher's *z*-transformation. In other words, regardless of language backgrounds, the impacts of overall musical skills, melodic skills, and rhythmic skills on tonal discrimination task scores remained significant at the same level. It indicated that there was no cross-linguistic difference on the correlation between musicianship on tonal perception.

Additionally, the impacts of language experience and musicianship were both found significant in discriminating unfamiliar tonal words in this study where the pair of words only differentiated in lexical pitch, indicating the importance of both factors in pitch perception. Therefore, the discussion of language experience and musicianship should be paid extra attention, and any studies including either factor as independent variable should consider the possible existent impact of the other factor on dependent variable. It is therefore suggested that in the investigation of how tonal and non-tonal L1 experience would influence tonal discrimination, the control of musicianship would be critical, leading to the response to the next research question.

5.3 Significant Effect of Tonal and Non-tonal L1 Backgrounds on Tonal Discrimination

The nonparametric ANCOVA indicated the significant differences across language groups in tonal discrimination tasks, with significantly better performance by tonal L1 group than non-tonal L1 group when MET total scores were controlled. These findings echoed the prediction. However, it is also worth noting the significant effect of covariate, the MET total scores since the inclusion of musical aptitudes could significantly enhance the model fit. The significant effect indicated that apart from language backgrounds, the improvement of musicianship played a critical role in enhancing tonal discrimination task scores.

In general, this finding confirmed previous discoveries in other studies comparing participants with tonal and non-tonal L1 (e.g., Shen & Froud, 2019; Wang et al., 2022). The neural data from both presented significantly more sensitive responses to auditory information like cross-boundary types of lexical tones (Shen &Froud, 2019) and environmental sounds (Wang et al., 2022) by speakers with native tonal language experience.

The cross-linguistic effect in detecting lexical tones and musical pitches in certain aspects was also found by Chen et al. (2018). In their study, when listening to Chinese rising tone and dipping tone with the same duration and segment, the Chinese L1 speakers presented longer responses time than Dutch native speakers, indicating that native tonal language individuals required more obvious differences between tones for detection. When listening to musical pitch, native Chinese speakers did not outperform native Dutch speakers when the contrast was obvious. However, the Chinese participants exhibited earlier responses to certain notes, implying the competency to faster integrate new pitch experience to their own perception system due to their tonal language experience. Regarding the amplitude of MMN responses to musical notes, Chinese L1 speakers did not present significant advantages than Dutch L1 speakers. This indicated that for sufficiently and obviously different musical pitches, the tonal language experience did not necessarily lead to intense reflections. The current study also found that MET scores were not significantly different across tonal and non-tonal L1 groups. However, the non-significance of language backgrounds in this study may not be convincing due to the lack of participants for Mann-Whitney U test.

Chen et al. (2018) also found that for monosyllabic lexical tones, native Chinese and Dutch speakers presented comparable MMN responses. This might be explained by the obvious differences in materials. Contrastively, the current study showed significant differences across groups in tonal discrimination task scores, against Chen et al.'s (2018) findings. The discrepancies may be caused by the different test items. This study used disyllabic words for tonal discrimination. Even though for each pair, the tonal contrast was presented in only one syllable, for speakers with lower sensitivity to tonal perception, it might be difficult to distinguish. Another reason would be that for Thai words shown in the tonal discrimination task, the tonal differences might be subtle due to tonal features in Thai languages, unlike the different pitches and tones used by Chen et al. (2018). Therefore, for all the participants unfamiliar to this language in current study, their language background would influence them to distinguish different tones. This study also manifested this viewpoint in that tonal L1 group achieved significantly higher scores than non-tonal L1 did in tonal discrimination task when their musicianship was controlled. Therefore, the tonal language experience could facilitate tonal discrimination in the context of interacting with a new language.

5.4 Attempts to Explain the Findings

Whether the influence of musicianship on pitch perception can be explained by theoretical frameworks is worth discussing. It is suggested by this study that the consistent importance of musicianship on tonal perception in both tonal and non-tonal language backgrounds might be a powerful manifestation of OPERA hypothesis (Patel, 2011, 2012). The OPERA hypothesis proposed that based on the neural network shared by musical and linguistic skills and neural plasticity, the enhancement of musical abilities could also facilitate language development. This study found a significantly positive relationship between musicianship and tonal perception, which holds for both tonal and non-tonal L1 groups. However, as indicated by Fisher's *z*-transformation, the amplitude of correlation between musicianship and tonal perception was not found significantly different between language groups. It implied that regardless of speakers' language backgrounds, the significant effect of musical aptitude on tonal discrimination would maintain. The insensitivity to language backgrounds further confirmed the OPERA hypothesis, the establishment of which did not initially take language backgrounds into consideration. More specifically, the OPERA hypothesis argued for benefits in speech processing. Correspondingly, each pair of audios in tonal discrimination task only differentiated in tonal pitch, echoing the focus of the OPERA hypothesis. Considering that the MET measured musical skills, the correlation between higher MET scores and higher tonal discrimination task scores not only reaffirmed the OPERA hypothesis, but also manifested the music-to-language effect consequently in this case.

However, the completion of all five preconditions (i.e., overlap, precision, emotion, repetition, and attention) of OPERA hypothesis when participants performed the experiment remained unknown. The first two requirements, overlap and precision, were supported by numerous evidence mentioned before (e.g., Bidelman et al., 2013; Herdener et al., 2012; Koelsch et al., 2005). Contrastively, the fulfilments of rest three conditions (i.e., the intense positive emotions driven by musical activities, the occurrences of repetitive musical activities, and the focused attention due to musical engagements) were not tested in this study. Even though these preconditions were not examined in this study, the findings still aligned with the hypothesis in that the importance of musical abilities in speech processing, measured by discriminating tonal words in different lexical pitches, was discovered.

Regardless of the completing status of preconditions, it is arguable that the possible existence of absolute pitch among participants could further increase the uncertainty of the degree to support the OPERA hypothesis via this study's outcomes. The reason is that for individuals with absolute pitch, their sensitivity to pitch processing is abnormally higher than others without extra training (Wong et al., 2020). This advantage could enormously felicitate them in both the MET and tonal discrimination task. However, even if there were participants with absolute pitch, their existence should not be a major concern in this study. The main reason is that their responses to MET were not necessarily linked to their talents since MET also included rhythm subtest. It is likely that they achieved astonishingly high score in melody subtest while low score in rhythm subtest. The unsatisfied rhythmic performance did not lead to high scores in MET total score. It was manifested in actual responses where a participant scored 45 out of 52 in melody subtest while 34 out of 52 in rhythm subtest. The MET total score (i.e., 79) was not an outlier. The largest difference in scores between the two subtests was 24, with 39 in melody and 15 in rhythm. The MET total score (i.e., 54) was not an outlier either. Therefore, it is reasonable to conclude that the OPERA hypothesis and the findings of this study resonated with each other. Alternatively, the musicianship was critical in tonal discrimination.

Moreover, the Auditory Precision Hypothesis-L2 (Saito, 2023) was likely to be manifested to some degree. The hypothesis proposed that the more input opportunity, linking to the more enhanced L2 speech proficiency, was associated with higher precision in auditory processing. In this study, the higher level of L2 proficiency could be reflected in better performance in discriminating tonal words in Thai. The more precise auditory processing could be corresponded with higher scores in MET, which examined musical aptitudes, while also required the involvement of auditory processing when discriminating melodies and rhythms. The significantly positive correlation between MET total score and tonal discrimination task shown in the current study reflected the establishment of the Auditory Precision Hypothesis-L2. Furthermore, the correlation did not present significant difference between tonal and non-tonal L1 groups. In other words, for speakers with higher MET total scores, regardless of their (non-)tonal language experience, their better performance in distinguishing tonal words from a new natural language suggests the possibility that the more input opportunity, led by enhanced musical skills, was seized in new tonal language learning. The success in detecting the tonal discrimination would establish the foundation of successful tonal acquisition in an unfamiliar tonal L2 like Thai.

It seemed that the above interpretation was reasonable, while closer analysis at the hypothesis may lead to another conclusion. As suggested by Saito (2023), the hypothesis argued for causal link, which would lead to endless debates. The interpretation of the results in the current study and their connections with the components of the hypothesis were mostly based on correlation rather than causality. Furthermore, the hypothesis explained for L2 speech proficiency. The tonal discrimination task score measured perceptive skills rather than productive ones like speech. The discrimination task score therefore could not represent the speech proficiency. However, if the noticing hypothesis, i.e., linguistic elements could not be learnt without noticing (Schmidt, 1990, 2001), was considered, the success in discrimination would be viewed as the initial step of language input, therefore possibly converting to language learning in the future. As a result, the fitness of these results to Auditory Precision Hypothesis-L2 interpretation was questionable, while the significant correlation between higher musical aptitudes and higher tonal discrimination task scores was robust.

5.5 Educational Implications

The current study has important implications for different educational settings since the data were collected from 167 participants with around 30 different first language backgrounds. The study found that musical skills, especially melodic skills, are crucial to improve learners' tonal perception abilities when acquiring a new tonal language regardless of learners' language backgrounds. Previous studies showed that as little as months of musical training can significantly improve children's musical and speech sensitivities (Moreno et al., 2009; Nan et al., 2018). In school-age Portuguese children, the 9-month longitudinal study presented that compared with control group and painting training group, children with musical training performed significantly better in reading tasks and showed augmentation in ERPs which indicated increasing sensitivity in both music and speech sensitivities (Moreno et al., 2009). In Mandarin-speaking children at the age of 4-5 years, compared with reading training and control settings, the piano training enhanced children's vowel identification abilities and improved their neural sensitivities to lexical and musical pitch tones (Nan et al., 2018). These studies have suggested the benefits of using music in speech and musical processing.

This study, by investigating the correlation between tonal discrimination and both language background and musicianship, found that the musicianship as the significant predictor of tonal discrimination task in both tonal and non-tonal L1. It indicates that regardless of L1 backgrounds, educational practitioners can provide some musical training to boost tonal discrimination. While short-term musical intervention studies have provided abundant evidence of linguistic enhancement, the studies about long-term (e.g., more than one year) musical intervention and its linguistic effect specifically in tonal perception are comparably limited.

One long-term study was found in terms of informal teaching environment and its integration of L2 learning to choir practice shed light on its practice (Lehtinen-Schnabel & Levänen, 2024). The 12-month track with interviews and phonology tests found that two groups of participants showed significant improvements in phonetic awareness and production in their L2 Finnish. The implication of this study is limited in that the two groups of participants were members of choir, and they differentiated in the arrival time of Finnish environment but not the intervention of musical training. Moreover, the main focus during the process was choir practice while language was an addition. However, this may serve as the evidence for implicit learning since the participants did not receive explicit language instruction. The L2 stimuli were activated during choir rehearsals in a playful and encouraging approach by utilising senses through body, ears, mouth and other ways. The increase in sound discrimination suggested the success of learning L2 through music.

Regarding the acquisition of tonal language, this study implied that the involvement of music training, leading to increase in musical aptitudes, could be beneficial for enhanced tonal discrimination of a new language. This has been confirmed by previous research. For example, in their first exposure to Thai language, individuals with more than five years of formal musical training outperformed individuals with less than two years of formal musical training in tonal identification (Götz et al., 2023). However, during the further research screening about the application of music in language acquisition, it was found that a majority of studies focus on the influence of music on phonological awareness rather than on tonal perception (e.g., Eccles et al., 2021; Vidal et al., 2020). Some studies considered pitch and musical perception (e.g., Wang et al., 2020; Yang et al., 2024), while the focus was on reading developments rather than lexical tone learning. The influence of musical training on pitch perception remained undetected.

Additionally, the sustainability of musical intervention could be questionable. Even if the musical intervention in these studies could last a few weeks (Wiener & Bradley, 2023) or even months (Nan et al., 2018), it is uncertain if the intervention will be continued after the experiment. Since the importance of musical intervention in acquiring different languages has been supported by studies (e.g., English: Islami, 2019; Finnish: Lehtinen-Schnabel & Levänen, 2024; Cantonese: Maggu et al., 2021), it is suggested that educational practitioners will consider the implementation of involving music in tonal language learning. Prior to the execution of this plan, more longitudinal evidence about involving music in tonal language and its effect on tonal perception in the classroom setting is required.

Despite the benefits of musical training on L2 acquisition and multiple viable approaches, the challenges may appear in practical learning contexts (Zhang, 2019). It was found that while musical tones provided a shortcut for learning linguistic tones (Zhang, 2019), Chinese L2 learners might adhere to guidelines or transfer musical tones to language learning, leading to overly melodious speech (Ho & Ho, 2007). In formal education setting, the discrepancies in musical pitch and linguistic tones could also undermine learning effect in Chinese L2 preschoolers (To-Chan et al., 2023). Considering that in preschool classrooms, children were attracted to language learning via musical engagements, further research in ameliorating language teaching methods through music is worth exploring.

5.6 Limitations and Further Improvements

The current study has several limitations which can be improved in future studies. To begin with, this study did not include sufficient number of participants to investigate the influence of language backgrounds on musicianship which was conducted via Mann-Whitney U test. Even though this was not the main research focus of this study, more participants are required for future research to enhance statistical power.

Regarding the experiment setting, the environmental factors should be further controlled. The current study required all the participants to wear earphones and perform the whole procedures in a quiet environment. However, according to feedback from several participants, such requirements were not necessarily met. The researcher also reminded participants to take a rest between each test to avoid fatigue while it was possible that they continued to perform the test when feeling exhausted. These inconsistent behaviours with suggestions could impact participants' performance. In other words, the experiment scores might not reflect their actual musical and linguistic abilities, but the ones diminished due to uncontrolled environmental factors. The future research is suggested to carefully control these elements by reiterating the importance of quiet and undisturbed environment for the research. If the researchers intend to conduct the experiment face-to-face, it can be achieved by inviting small groups of participants (e.g., 10 people per group) to the same venue (e.g., classroom, lecture hall) in turn and requiring performing the task on site. Organising participants into small groups can limit the impact on each other and performing at the same quiet place can diminish the influence of various surroundings.

Secondly, even though the MET is a reliable examination tool for musical aptitudes, some participants were not familiar with its setting which might hinder their performance. For example, according to several feedback, the sound volume of metronome in the MET rhythm subtest was disturbing and participants were distracted in several questions. These responses, therefore, may not accurately reflect participants' actual rhythmic skills. Due to the anonymisation of the data, it was not possible for researchers to separate remove them. However, this disturbance was not a major concern in this study in that only a few participants complained about this design. Moreover, there were two trial questions with correct answers before the formal test started for participants to familiarise with. Despite these, the accuracy of MET results could still be further improved in future studies. For example, participants can be clearly instructed to perform the test by being informed of the volume and duration of metronome beforehand so that they will prepare themselves.

Thirdly, even though this study did not categorise participants' musicianship dichotomously, the categorisation of language groups can be questionable in a way that all the participants were bi-/multilingual. However, they were grouped solely based on their L1 backgrounds without considering their foreign language abilities. They might know limited knowledge of tonal languages which was insufficient to be qualified as proficiency in their perspective and therefore was not captured by the questionnaire. It was also possible that some non-tonal L1 speakers had received years of training in a tonal language and have thus established tonal perception abilities which are as sensitive as those of tonal L1 speakers. Additionally, even though participants were asked to indicate what languages they knew, their language proficiency for each language was unknown. This might have led to an inaccurate grouping into tonal vs non-tonal languages. Moreover, it was possible that the L1 proficiency of some participants deteriorated if they had not frequently spoken their L1 for years. However, the absence of proficiency data should not have been detrimental to the current study. The reason was that the application of linear mixed-effect model considered the random effect of individual factor, i.e., participants' different responses to tonal discrimination task due to personal traits, and it was presented non-significant. Nevertheless, the dichotomous grouping (into tonal L1 vs non-tonal L1) could still be improved by altering it into a continuous variable. The variation can be achieved through testing participants language usage frequency and proficiency, as assessed in a language history questionnaire version 3 (LHQ3.0, Li et al., 2020). The more detailed measurement may produce more comprehensive results, benefitting future research. Nevertheless, the LHQ3.0 contains 27 questions which can be time-consuming to complete all. Therefore, researchers should select their target questions carefully according to their research needs.

Fourthly, the design of tonal discrimination task could be ameliorated. Unlike the MET, in this task, the numbers of yes and no answers were not equal (i.e., 9 vs. 11). This was due to phonotactic constraints. In other words, as reported by the native Thai speaker who recorded the stimuli, it was difficult to pronounce the words in an altered tone which does not exist in the Thai language. For example, the second syllable of word 12B 'ma^{4}kea^{5}' was initially changed into Tone 1 to construct a nonsense word 'ma^{4}kea^{1}'. However, the speaker could not pronounce the syllable in Tone 1 and the tones of this pair were therefore identical. The influence of the uneven split may not be critical to the outcome. However, considering that the answers of MET subtests were equally split into yes and no, the even distribution in tonal discrimination task may further benefit data analysis and discussion. Considering that the random effect of tonal discrimination task was addressed in the linear mixed-effect model and was not marked as significant, the result of the current study is reliable.

Regarding the measurements of tonal perception, this study only concerned the discrimination task. However, tonal identification could also be used as an approach for testing musicians' perception abilities (Ong et al., 2020). Apart from distinguishing two tones, the tonal identification task also examines if participants can match the auditory input to the target audio material. Therefore, for future research, it is suggested to conduct both discrimination and identification tasks for participants to derive more comprehensive results.

6 Conclusion

In conclusion, this study found both musicianship and language backgrounds as significant predictors of tonal discrimination ability. Moreover, the significance of musical aptitudes held for both tonal and non-tonal L1 groups. However, the significant association between tonal discrimination task scores and musicianship did not present significant difference between two language groups. When musical aptitudes served as the covariate, the language backgrounds remained a significant predictor, indicating the enhanced tonal discrimination task performance in tonal rather than non-tonal L1 group. It is worth noting that the inclusion of musical aptitudes resulted in significant enhancement of model fit. The bidirectional relationship between musicianship and language background was not manifested in this study due to insufficient number of participants.

It is worth mentioning that this study had advantages compared with previous ones. Firstly, it used a consistent definition of (non-)musicians by measuring musical ability via a test. This decision has turned the musicianship from a categorical variable to a continuous variable, allowing the statistical investigation of the influence of musical skills in detail. Secondly, the current study involved 167 participants with around 30 different L1s, suggesting that its results are sufficiently inclusive and comprehensive. The findings may help language instructors reflect and improve their teaching methods in linguistically diverse classroom settings by the introduction of music in language learning.

Moreover, this study contributed to the field of music and language interdisciplinary studies and had empirical implications. The research applied real words from a natural language rather than the synthesized pitch. The results are therefore of greater value for language learners and researchers and practitioners in language education. This study also concerned the influence of both musicianship and first language backgrounds on tonal perception. The significant correlations between musicianship and tonal discrimination in both language groups and the non-significant difference in such correlations between language groups indicate that in multilingual education contexts, it is possible for teachers to use music as facilitative methods for linguistic gains.

This research is replicable for further studies regarding musical aptitudes and second language learning with improvements. Firstly, the researchers should prepare participants with the trials in the musical aptitude test and remind them about the importance of quiet environment during the performance of each task. Secondly, participants should be highly recommended to take a break between tests to diminish the potential influence of fatigue on test results. Thirdly, the participants' language backgrounds should be considered comprehensively, i.e., not limited to their first languages. Lastly, the design of tonal discrimination task is suggested to include equal number of questions with identical and different pairs of tonal words.

References

Alain, C., Moussard, A., Singer, J., Lee, Y., Bidelman, G. M., & Moreno, S. (2019). Music and visual art training modulate brain activity in older adults. *Frontiers in Neuroscience*, *13*, 182. https://doi.org/10.3389/fnins.2019.00182

Asano, R., & Boeckx, C. (2015). Syntax in language and music: what is the right level of comparison?. *Frontiers in Psychology*, *6*, 942. https://doi.org/10.3389/fpsyg.2015.00942

Asano, R., Lo, V., & Brown, S. (2022). The neural basis of tonal processing in music: An ALE meta-analysis. *Music & Science*, *5*, 1-15. https://doi.org/10.1177/20592043221109958

Bao, Z. (1999). *The structure of tone*. Oxford University Press.

Bates, D. M. (2010). *Lme4: Mixed-effects modeling with R*. Springer. https://www.researchgate.net/profile/Dimitris-Kavroudakis/post/What_is_the_appropriate_package_to_use_for_performing_NML_using_R/attachment/59d62f0dc49f478072e9f5e7/AS%3A272534858076166%401441988781179/download/lrgprt.pdf

Bernard, H. R., Mbeh, G. N., & Handwerker, W. P. (2002). Does marking tone make tone languages easier to read?. *Human Organization*, *61*(4), 339-349. https://doi.org/10.17730/humo.61.4.3u6lumuux1fn3q3y

Besson, M., Barbaroux, M., & Dittinger, E. (2017). Music in the brain: Music and language processing. In R. Ashley & R. Timmers (Eds.), *The Routledge companion to music cognition*, (pp. 37-48). https://doi.org/10.4324/9781315194738

Bernstein, L. (1976). *The Unanswered Question: Six Talks at Harvard.* Harvard University Press. https://www.speakersacademy.com/wp-content/uploads/2018/06/Leonard-Bernstein-The-Unanswered-Question_web.pdf

Bertolo, M., Müllensiefen, D., Peretz, I., Woolley, S. C., Sakata, J. T., & Mehr, S. A. (2023). Sex differences in music perception are negligible. *BioRxiv*, 1-18. https://doi.org/10.1101/2023.05.23.541970

Bidelman, G. M., & Alain, C. (2015). Musical training orchestrates coordinated neuroplasticity in auditory brainstem and cortex to counteract age-related declines in categorical vowel perception. *Journal of Neuroscience*, *35*(3), 1240-1249. https://doi.org/10.1523/JNEUROSCI.3292-14.2015

Bidelman, G. M., & Chung, W. L. (2015). Tone-language speakers show hemispheric specialization and differential cortical processing of contour and interval cues for pitch. *Neuroscience*, *305*, 384-392. https://doi.org/10.1016/j.neuroscience.2015.08.010

Bidelman, G. M., Hutka, S., & Moreno, S. (2013). Tone language speakers and musicians share enhanced perceptual and cognitive abilities for musical pitch: evidence for bidirectionality between the domains of language and music. *PloS One*, *8*(4), e60676. https://doi.org/10.1371/journal.pone.0060676

Blackwood, D. H. R., & Muir, W. J. (1990). Cognitive brain potentials and their application. *The British Journal of Psychiatry*, *157*(S9), 96-101. https://doi.org/10.1192/S0007125000291897

Blevins, J. (2008). Natural and unnatural sound patterns. In K. Willems & L. De Cuypere (Eds.), *Naturalness and iconicity in language* (pp. 121-148). John Benjamins Publishing Company.

Cambridge Dictionary. (n.d.). Musicianship. In *Dictionary.cambridge.org*. Retrieved August 6, 2024, from https://dictionary.cambridge.org/dictionary/english/musicianship

Camp, G. (2023). *Linguistics for Singers: An Introduction*. Routledge. https://doi.org/10.4324/9781003320753

Chang, A., Teng, X., Assaneo, M. F., & Poeppel, D. (2024). The human auditory system uses amplitude modulation to distinguish music from speech. *PLoS Biology*, *22*(5), e3002631. https://doi.org/10.1371/journal.pbio.3002631

Chen, A., Peter, V., Wijnen, F., Schnack, H., & Burnham, D. (2018). Are lexical tones musical? Native language's influence on neural response to pitch in different domains. *Brain and Language*, *180*, 31-41. https://doi.org/10.1016/j.bandl.2018.04.006

Choi, W. (2021). Musicianship influences language effect on musical pitch perception. *Frontiers in Psychology*, *12*, 712753. https://doi.org/10.3389/fpsyg.2021.712753

Choi, W. (2022). Towards a native OPERA hypothesis: Musicianship and English stress perception. *Language and Speech*, *65*(3), 697-712. https://doi.org/10.1177/00238309211049458

Cohrdes, C., Grolig, L., & Schroeder, S. (2016). Relating language and music skills in young children: A first approach to systemize and compare distinct competencies on different levels. *Frontiers in Psychology*, *7*, 1616. https://doi.org/10.3389/fpsyg.2016.01616

Correia, A. I., Vincenzi, M., Vanzella, P., Pinheiro, A. P., Lima, C. F., & Schellenberg, E. G. (2022). Can musical ability be tested online?. *Behavior Research Methods*, *54*(2), 955-969. https://doi.org/10.3758/s13428-021-01641-2

Darwin, C. (1874). *The descent of man and selection in relation to sex*. A.L. Burt Company. https://archive.org/details/descentofmansele1874darw/page/n11/mode/2up

de Cheveigné, A., & Kawahara, H. (2002). YIN, a fundamental frequency estimator for speech and music. *The Journal of the Acoustical Society of America*, *111*(4), 1917-1930. https://doi.org/10.1121/1.1458024

de Silva, F. L. (2022). EEG: Origin and measurement. In C. Mulert & L. Lemieux (Eds.), *EEG-fMRI: Physiological basis, technique, and applications* (2nd ed., pp. 23-48). Springer, Cham. https://doi.org/10.1007/978-3-031-07121-8_2

de Melo, M. B., Daldegan-Bueno, D., Menezes Oliveira, M. G., & de Souza, A. L. (2022). Beyond ANOVA and MANOVA for repeated measures: Advantages of generalized estimated equations and generalized linear mixed models and its use in neuroscience research. *European Journal of Neuroscience*, *56*(12), 6089-6098. https://doi.org/10.1111/ejn.15858

Duval, S., & Tweedie, R. (2000). Trim and fill: A simple funnel-plot–based method of testing and adjusting for publication bias in meta-analysis. *Biometrics*, *56*(2), 455-463. https://doi.org/10.1111/j.0006-341X.2000.00455.x

Eccles, R., Van der Linde, J., le Roux, M., Holloway, J., MacCutcheon, D., Ljung, R., & Swanepoel, D. W. (2021). Is phonological awareness related to pitch, rhythm, and speech-in-noise discrimination in young children?. *Language, Speech, and Hearing Services in Schools*, *52*(1), 383-395. https://doi.org/10.1044/2020_LSHSS-20-00032

Etz, K. E., & Arroyo, J. A. (2015). Small sample research: Considerations beyond statistical power. *Prevention Science*, *16*, 1033-1036. https://doi.org/10.1007/s11121-015-0585-4

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175-191. https://doi.org/10.3758/BF03193146

Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*(4), 1149-1160. https://doi.org/10.3758/BRM.41.4.1149

Feld, S., & Fox, A. A. (1994). Music and Language. *Annual Review of Anthropology*, *23*, 25-53. http://www.jstor.org/stable/2156005

Field, A. (2014). *Discovering Statistics Using IBM SPSS Statistics* (4th ed.). Sage Publications.

Fisher, R. A. (1932). *Statistical methods for research workers*. Oliver and Boyd.

Francois, C., & Schön, D. (2011). Musical expertise boosts implicit learning of both musical and linguistic structures. *Cerebral Cortex*, *21*(10), 2357-2365. https://doi.org/10.1093/cercor/bhr022

Gałecki, A., & Burzykowski, T. (2013). *Linear mixed-effects models using R: A step-by-step approach*. Springer. https://doi.org/10.1007/978-1-4614-3900-4_13

Gandour, J. T. (1978). The perception of tone. In V. A. Fromkin (Ed.), *Tone: A Linguistics Survey* (pp. 41-76). Academic Press.

García Conesa, I. M., & Juan Rubio, A. D. (2015). The use of rhymes and songs in the Teaching of English in Primary Education. *Docencia e Investigacíon: Revisita Científica de Educacíon*, *25*(2), 83-101. https://hdl.handle.net/10481/90058

Garrido, M. I., Kilner, J. M., Stephan, K. E., & Friston, K. J. (2009). The mismatch negativity: a review of underlying mechanisms. *Clinical neurophysiology*, *120*(3), 453-463. https://doi.org/10.1016/j.clinph.2008.11.029

Gautam, S., & Singh, L. (2017). Development of spectro-temporal features of speech in children. *International Journal of Speech Technology*, *20*(3), 543-551. https://doi.org/10.1007/s10772-017-9424-2

George, D., & Mallery, P. (2011). *SPSS for windows step by step: A simple study guide and reference, 18.0 update*. Allyn & Bacon/Pearson.

Gilbert, J. B. (2023). Modeling item-level heterogeneous treatment effects: A tutorial with the glmer function from the lme4 package in R. *Behavior Research Methods*, 1-13. https://doi.org/10.3758/s13428-023-02245-8

Gjana, A., & Kosova, R. (2021). Traditional class, and online class teaching. Comparing the students performance using ANCOVA. *Journal of Multidisciplinary Engineering Science and Technology (JMEST)*, 14806-14811. https://www.jmest.org/wp-content/uploads/JMESTN42353932.pdf

Gogtay, N. J., & Thatte, U. M. (2017). Principles of correlation analysis. *Journal of the Association of Physicians of India*, *65*(3), 78-81. https://www.kem.edu/wp-content/uploads/2012/06/9-Principles_of_correlation-1.pdf

Goldstone, R. L., & Hendrickson, A. T. (2010). Categorical perception. *Wiley Interdisciplinary Reviews: Cognitive Science*, *1*(1), 69-78. https://doi.org/10.1002/wcs.26

Gordon, R. L., Shivers, C. M., Wieland, E. A., Kotz, S. A., Yoder, P. J., & Devin McAuley, J. (2015). Musical rhythm discrimination explains individual differences in grammar skills in children. *Developmental Science*, *18*(4), 635-644. https://doi.org/10.1111/desc.12230

Götz, A., Liu, L., Nash, B., & Burnham, D. (2023). Does Musicality Assist Foreign Language Learning? Perception and Production of Thai Vowels, Consonants and Lexical Tones by Musicians and Non-Musicians. *Brain Sciences*, *13*(5), 810. https://doi.org/10.3390/brainsci13050810

Green, L. (2017). *Music, informal learning and the school: A new classroom pedagogy*. Routledge. https://doi.org/10.4324/9781315248523

Gu, F., Zhang, C., Hu, A., & Zhao, G. (2013). Left hemisphere lateralization for lexical and acoustic pitch processing in Cantonese speakers as revealed by mismatch negativity. *Neuroimage*, *83*, 637-645. https://doi.org/10.1016/j.neuroimage.2013.02.080

Gupta, R., Singh, V. P., & Agarwal, M. (2024). Metallurgical and interfacial co-relation of HA-ZnO-Fe3O4 composite for implant application with properties: experimental interaction study with ANOVA and ANCOVA model. *Phase Transitions*, *97*(4-5), 322-336. https://doi.org/10.1080/01411594.2024.2304802

Gussenhoven, C. (2010). *The Phonology of Tone and Intonation*. Cambridge University Press. https://doi.org/10.1017/CBO9780511616983

Hallam, S., & Himonides, E. (2022). *The power of music: An exploration of the evidence*. Open Book Publishers. https://doi.org/10.11647/OBP.0292

Hamilton, C., Schulz, J., Chalmers, H., & Murphy, V. A. (2024). Investigating the substantive linguistic effects of using songs for teaching second or foreign languages to preschool, primary and secondary school learners: a systematic review of intervention research. *System*, 103350. https://doi.org/10.1016/j.system.2024.103350

Hatem, G., Zeidan, J., Goossens, M., & Moreira, C. (2022). Normality testing methods and the importance of skewness and kurtosis in statistical analysis. *BAU Journal-Science and Technology*, *3*(2), 1-5. https://digitalcommons.bau.edu.lb/cgi/viewcontent.cgi?article=1072&context=stjournal

Hennessy, S., Mack, W. J., & Habibi, A. (2022). Speech-in-noise perception in musicians and non-musicians: A multi-level meta-analysis. *Hearing Research*, *416*, 108442. https://doi.org/10.1016/j.heares.2022.108442

Herdener, M., Humbel, T., Esposito, F., Habermeyer, B., Cattapan-Ludewig, K., & Seifritz, E. (2014). Jazz drummers recruit language-specific areas for the processing of rhythmic structure. *Cerebral Cortex*, *24*(3), 836-843. https://doi.org/10.1093/cercor/bhs367

Hettmansperger, T. P., & McKean, J. W. (2010). *Robust nonparametric statistical methods*. CRC press. https://doi.org/10.1201/b10451

Ho, R. S. C., & Ho, S. F. (2007). Musical scales and Cantonese level tones. In E. Schubert, K. Buckley, R. Eliott, B. Koboroff, J. Chen & C. Stevens (Eds.), *Proceedings the Inaugural Conference on Music Communication Science* (pp. 64-67).

Hodges, D., & Thaut, M. (Eds.). (2019). *The Oxford handbook of music and the brain*. Oxford University Press.

Hollander, M., Wolfe, D. A., & Chicken, E. (2013). *Nonparametric statistical methods*. John Wiley & Sons.

Howard, D., & Angus, J. (2013). *Acoustics and psychoacoustics* (4th ed.). Routledge. https://doi.org/10.4324/9780080961873

Hudak, T. J. (2009). Thai. In B. Comrie (Ed.), *The world's major languages*, (2nd ed., pp. 660-676). https://doi.org/10.4324/9780203301524

Islami, F. T. (2019). Using song as a media in teaching vocabulary to young learners based on total physical response (TPR) method. *Universitas Pendidikan Indonesia*. https://core.ac.uk/download/pdf/289220943.pdf

IBM Corps. (2023). *IBM SPSS Statistics* (Version 29.0.1.0) [Computer software]. https://www.ibm.com/products/spss-statistics?lot=5&mhsrc=ibmsearch_a&mhq=spss

Jackendoff, R. (2009). Parallels and Nonparallels between Language and Music. *Music Perception: An Interdisciplinary Journal*, *26*(3), 195–204. https://doi.org/10.1525/mp.2009.26.3.195

James, A. & Leather, J. (1986). *Sound patterns in second language acquisition*. De Gruyter Mouton. https://doi.org.ezproxy-prd.bodleian.ox.ac.uk/10.1515/9783110878486

Janata, P. (2015). Neural basis of music perception. *Handbook of Clinical Neurology*, *129*, 187-205. https://doi.org/10.1016/B978-0-444-62630-1.00011-1

Jenny, M. (2019). Thai. In V. Alice & W. Justin (Eds.), *The mainland southeast Asia linguistic area* (pp. 559-608). De Gruyter Mouton. https://doi.org/10.1515/9783110401981-013

Jentschke, S. (2014). The relationship between music and language. In S. Hallam, I. Cross & M. H. Thaut (Eds.), *The Oxford handbook of music psychology* (2nd ed., pp. 343-355). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780198722946.001.0001

Kao, T. A., & Oxford, R. L. (2014). Learning language through music: A strategy for building inspiration and motivation. *System*, *43*, 114-120. https://doi.org/10.1016/j.system.2014.01.003

Kim, T. K. (2015). T test as a parametric statistic. *Korean Journal of Anesthesiology*, *68*(6), 540-546. https://doi.org/10.4097/kjae.2015.68.6.540

Kloke, J. D., & McKean, J. W. (2012). Rfit: Rank-based estimation for linear models, *The R Journal*, *4*(2), 57–64. https://svn.r-project.org/Rjournal/html/archive/2012/RJ-2012-014/RJ-2012-014.pdf

Kloke, J. D., & McKean, J. W. (2014). *Nonparametric statistical methods using R*. Chapman & Hall/CRC. https://doi.org/10.1201/b17501

Kloke, J. D., & McKean, J. W. (2024). *Nonparametric statistical methods using R* (2nd ed). Chapman & Hall.

Koelsch, S., Gunter, T. C., Wittfoth, M., & Sammler, D. (2005). Interaction between syntax processing in language and in music: An ERP study. *Journal of cognitive neuroscience*, *17*(10), 1565-1577. https://doi.org/10.1162/089892905774597290

Kourtidou, E., Kasselimis, D., Angelopoulou, G., Karavasilis, E., Velonakis, G., Kelekis, N., Zalonis, I., Evdokimidis, I., Potagas, C., & Petrides, M. (2021). The role of the right hemisphere white matter tracts in chronic aphasic patients after damage of the language tracts in the left hemisphere. *Frontiers in Human Neuroscience*, *15*, 635750. https://doi.org/10.3389/fnhum.2021.635750

Kraus, N. (2011). Musical training gives edge in auditory processing. *The Hearing Journal*, *64*(2), 10,12,14,16. https://doi.org/10.1097/01.HJ.0000394539.18079.14

Kulset, N. B. (2015). Singing in the multicultural kindergarten leads to second language acquisition and social inclusion: Explained by the OPERA hypothesis. 1-13. https://www.researchgate.net/profile/Nora-Kulset/publication/314756155_Singing_in_the_multicultural_kindergarten_leads_to_second_language_acquisition_and_social_inclusion_Explained_by_the_OPERA_hypothesis/links/58c58746a6fdcce648e8b11f/Singing-in-the-multicultural-kindergarten-leads-to-second-language-acquisition-and-social-inclusion-Explained-by-the-OPERA-hypothesis.pdf

Kumar, T., Akhter, S., Yunus, M. M., & Shamsy, A. (2022). Use of music and songs as pedagogical tools in teaching English as foreign language contexts. *Education Research International*, *2022*(1), 3384067. https://doi.org/10.1155/2022/3384067

Ladd, D. R. (2008). *Intonational phonology* (2nd ed.). Cambridge University Press. https://doi.org/10.1017/CBO9780511808814

Larroque, P. (2023). *An introduction to linguistics through popular music*. Springer Nature. https://doi.org/10.1007/978-3-031-24703-3

Lee, W., & Grimm, K. J. (2018). Generalized linear mixed-effects modeling programs in R for binary outcomes. *Structural Equation Modeling: A Multidisciplinary Journal*, *25*(5), 824-828. https://doi.org/10.1080/10705511.2018.1500141

Lee, J., Han, J. H., & Lee, H. J. (2020). Long-term musical training alters auditory cortical activity to the frequency change. *Frontiers in Human Neuroscience*, *14*, 329. https://doi.org/10.3389/fnhum.2020.00329

Lehmann, E. L. (1998). *Nonparametrics: Statistical Methods Based on Ranks*. Prentice Hall.

Lehtinen-Schnabel, J., & Levänen, S. (2024). Singing Language–Integrating Second Language Learning into Choir Practice. *Music & Science*, *7*, 1-25. https://doi.org/10.1177/20592043241244542

Levman, B. G. (1992). The Genesis of Music and Language. *Ethnomusicology*, *36*(2), 147–170. https://doi.org/10.2307/851912

Li, Y., Tang, C., Lu, J., Wu, J., & Chang, E. F. (2021). Human cortical encoding of pitch in tonal and non-tonal languages. *Nature communications*, *12*(1), 1161. https://doi.org/10.1038/s41467-021-21430-x

Li, P., Zhang, F., Yu, A., & Zhao, X. (2020). Language History Questionnaire (LHQ3): An enhanced tool for assessing multilingual experience. *Bilingualism: Language and Cognition*, *23*(5), 938-944. https://doi.org/10.1017/S1366728918001153

Liao, Y. F. (2008). *Jichu Taiyu (1)* [Basic Thai Language 1]. World Publishing Corporation. https://fliphtml5.com/hdnkn/xpzz/

Lindell, A. K. (2006). In your right mind: Right hemisphere contributions to language processing and production. *Neuropsychology Review*, *16*(3), 131-148. https://doi.org/10.1007/s11065-006-9011-9

Liu, J., Hilton, C. B., Bergelson, E., & Mehr, S. A. (2023). Language experience predicts music processing in a half-million speakers of fifty-four languages. *Current Biology*, *33*(10), 1916-1925. https://doi.org/10.1016/j.cub.2023.03.067

Lowry, R. (2024, July 20). *Significance of the Difference Between Two Correlation Coefficients*. VassarStats. http://vassarstats.net/rdiff.html

Ludke, K. M. (2019). Songs and music. In M. Dressman & R. W. Sadler (Eds.), *The handbook of informal language learning*, (pp. 203-213). https://doi.org/10.1002/9781119472384.ch13

Ludke, K. M., & Good, A. (2020). Singing to support foreign language learning: Examples from two cultural and developmental contexts. In H. R. Gudmundsdottir, C, Beynon, K. Ludke & A. J. Cohen (Eds.), *The Routledge companion to interdisciplinary studies in singing, Volume II: Education* (pp. 403-414). Routledge.

Maess, B., Koelsch, S., Gunter, T. C., & Friederici, A. D. (2001). Musical syntax is processed in Broca's area: an MEG study. *Nature Neuroscience*, *4*(5), 540-545. https://doi.org/10.1038/87502

Maggu, A. R., Lau, J. C., Waye, M. M., & Wong, P. C. (2021). Combination of absolute pitch and tone language experience enhances lexical tone perception. *Scientific reports*, *11*(1), 1485. https://doi.org/10.1038/s41598-020-80260-x

Mann, H. B., & Whitney, D. R. (1947). On a test of whether one of two random variables is stochastically larger than the other. *The Annals of Mathematical Statistics*, *18*(1), 50–60. http://www.jstor.org/stable/2236101

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*(2), 267–289. http://www.jstor.org/stable/23767517

Mazlomi Barm Sabz, A., Asgari, P., Makvandi, B., Ehteshamzadeh, P., & Bakhtiyar Pour, S. (2021). Comparison of the effectiveness of positive psychology and emotion regulation training interventions in promoting the psychological well-being in nar-anon group. *International Journal of Mental Health and Addiction*, *19*, 1909-1918. https://doi.org/10.1007/s11469-020-00284-2

Meyers, L. S., Gamst, G., & Guarino, A. J. (2009). Bivariate correlation: Pearson product-moment and Spearman rho correlations. In *Data Analysis Using SAS Enterprise Guide* (pp. 155–161). Cambridge University Press. https://doi.org/10.1017/CBO9780511804786

Microsoft Corporation. (2018). *Microsoft Excel* (Version 16.86) [Computer software]. https://office.microsoft.com/excel

Miśkiewicz, A., Rościszewska, T., Żera, J., Majer, J., & Okoń-Makowska, B. (2018). Detection and recognition of environmental sounds by musicians and non-musicians. *Archives of Acoustics*, *43*(4), 581-592. http://dx.doi.org/10.24425/aoa.2018.125152

Moen, I., & Sundet, K. (1996). Production and perception of word tones (pitch accents) in patients with left and right hemisphere damage. *Brain and Language*, *53*(2), 267-281. https://doi.org/10.1006/brln.1996.0048

Moreno, S., Marques, C., Santos, A., Santos, M., Castro, S. L., & Besson, M. (2009). Musical training influences linguistic abilities in 8-year-old children: More evidence for brain plasticity. *Cerebral Cortex*, *19*(3), 712-723. https://doi.org/10.1093/cercor/bhn120

Mueller, J. L., Friederici, A. D., & Männel, C. (2012). Auditory perception at the root of language learning. *Proceedings of the National Academy of Sciences*, *109*(39), 15953-15958. https://doi.org/10.1073/pnas.1204319109

Nan, Y., Liu, L., Geiser, E., Shu, H., Gong, C. C., Dong, Q., Gabrieli, J. D. E., & 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), E6630-E6639. https://doi.org/10.1073/pnas.1808412115

Neves, L., Correia, A. I., Castro, S. L., Martins, D., & Lima, C. F. (2022). Does music training enhance auditory and linguistic processing? A systematic review and meta-analysis of behavioral and brain evidence. *Neuroscience & Biobehavioral Reviews*, *140*, 104777. https://doi.org/10.1016/j.neubiorev.2022.104777

O'Callaghan, K. (2018). *Essays on music and language in modernist literature*. https://doi.org/10.4324/9781315231716

Oesch, N. (2019). Music and language in social interaction: Synchrony, antiphony, and functional origins. *Frontiers in Psychology*, *10*, 1514. https://doi.org/10.3389/fpsyg.2019.01514

Oktavia, D., Mukminin, A., Marzulina, L., Harto, K., Erlina, D., & Holandyah, M. (2022). Challenges and strategies used by English teachers in teaching English language skills to young learners. *Theory and Practice in Language Studies*, *12*(2), 382-387. https://doi.org/10.17507/tpls.1202.22

Ong, J. H., Wong, P., & Liu, F. (2020). Musicians show enhanced perception, but not production, of native lexical tones. *The Journal of the Acoustical Society of America*, *148*(6), 3443-3454. https://doi.org/10.1121/10.0002776

Oudeyer, P. (2006). *Self-organization in the evolution of speech*. Oxford University Press, Incorporated. https://doi.org/10.1093/acprof:oso/9780199289158.001.0001

Oxenham, A. J. (2013). The perception of musical tones. In D. Deutsch (Ed.), *The psychology of music* (3rd ed., pp. 1-33). Academic Press. https://doi.org/10.1016/B978-0-12-381460-9.00001-8

Patel, A. D. (2007). *Music, language, and the brain*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195123753.001.0001

Patel, A. D. (2011). Why would musical training benefit the neural encoding of speech? The OPERA hypothesis. *Frontiers in psychology*, *2*, 142. https://doi.org/10.3389/fpsyg.2011.00142

Patel, A. D. (2012). The OPERA hypothesis: assumptions and clarifications. *Annals of the New York Academy of Sciences*, *1252*(1), 124-128. https://doi.org/10.1111/j.1749-6632.2011.06426.x

Peretz, I. (2019). *How music sculpts our brain*. Odile Jacob.

Pike, K. L. (1948). *Tone Languages: A Technique for Determining the Number and Type of Pitch Contrasts in a Language, with Studies in Tonemic Substitution and Fusion.* University of Michigan Press.

Pino, M. C., Giancola, M., & D'Amico, S. (2023). The association between music and language in children: A state-of-the-art review. *Children*, *10*(5), 801. https://doi.org/10.3390/children10050801

Politimou, N., Dalla Bella, S., Farrugia, N., & Franco, F. (2019). Born to speak and sing: Musical predictors of language development in pre-schoolers. *Frontiers in Psychology*, *10*, 948. https://doi.org/10.3389/fpsyg.2019.00948

Posit team. (2024). *RStudio: Integrated development environment for R* (Version 2024.04.2.764) [Computer software]. https://www.posit.co/

Proverbio, A. M., Orlandi, A., & Pisanu, F. (2016). Brain processing of consonance/dissonance in musicians and controls: a hemispheric asymmetry revisited. *European Journal of Neuroscience*, *44*(6), 2340-2356. https://doi.org/10.1111/ejn.13330

Qi, Z., & Legault, J. (2020). Neural hemispheric organization in successful adult language learning: Is left always right?. *Psychology of Learning and Motivation*, *72*, 119-163. https://doi.org/10.1016/bs.plm.2020.02.004

Rasch, R., & Plomp, R. (1999). The perception of music tones. In D. Deutsch (Ed.), *Cognition and perception: The psychology of music* (2nd ed., pp. 89-112). Academic Press. https://doi.org/10.1016/B978-012213564-4/50005-6

Rebuschat, P., Rohrmeier, M., Hawkins, J. A., & Cross, I. (Eds.). (2011). *Language and music as cognitive systems*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199553426.001.0001

Ries, S. K., Dronkers, N. F., & Knight, R. T. (2016). Choosing words: Left hemisphere, right hemisphere, or both? Perspective on the lateralization of word retrieval. *Annals of the New York Academy of Sciences*, *1369*(1), 111-131. https://doi.org/10.1111/nyas.12993

Roederer, J. G. (2008). *The physics and psychophysics of music: An introduction* (4th ed.). Springers New York, NY. https://doi.org/10.1007/978-0-387-09474-8

Rogala, T., Szczepańska-Antosik, J., Miśkiewicz, A., Żera, J., & Majer, J. (2023). Identification of environmental sounds: Auditory abilities of musicians and non-musicians. *Professional Acoustics*, *2*, 135-156. http://wydawnictwo.uwm.edu.pl/uploads/documents/czytelnia/monografie/Rosinski-A_Przestrzenie-akustyki-2.pdf#page=135

Rosenthal, M. A. (2016). Hemispheric asymmetry in the hierarchical perception of music and speech. *Psychological Bulletin*, *142*(11), 1165. http://dx.doi.org/10.1037/bul0000076

Saito, K. (2023). How does having a good ear promote successful second language speech acquisition in adulthood? Introducing auditory precision hypothesis-L2. *Language Teaching*, *56*(4), 522-538. https://doi.org/10.1017/S0261444822000453

Saito, K., Cui, H., Suzukida, Y., Dardon, D. E., Suzuki, Y., Jeong, H., Révész A., Sugiura M., & Tierney, A. (2022a). Does domain-general auditory processing uniquely explain the outcomes of second language speech acquisition, even once cognitive and demographic variables are accounted for?. *Bilingualism: Language and Cognition*, *25*(5), 856-868. https://doi.org/10.1017/S1366728922000153

Saito, K., Petrova, K., Suzukida, Y., Kachlicka, M., & Tierney, A. (2022b). Training auditory processing promotes second language speech acquisition. *Journal of Experimental Psychology: Human Perception and Performance*, *48*(12), 1410-1426. https://psycnet.apa.org/doi/10.1037/xhp0001042

Sajjadi, E., Mohammadzadeh, A., Sayadi, N., Nazeri, A., & Tabatabai, S. M. (2021). Comparison of the recognition scores of stop and fricative consonants in babble noise between musicians and non-musicians. *Psychology of Music*, *49*(5), 1344-1355. https://doi.org/10.1177/0305735620953616

Salvari, V., Paraskevopoulos, E., Chalas, N., Müller, K., Wollbrink, A., Dobel, C., Korth, D., & Pantev, C. (2019). Auditory categorization of man-made sounds versus natural sounds by means of MEG functional brain connectivity. *Frontiers in Neuroscience*, *13*, 1052. https://doi.org/10.3389/fnins.2019.01052

Schaal, N. K., Pollok, B., & Banissy, M. J. (2017). Hemispheric differences between left and right supramarginal gyrus for pitch and rhythm memory. *Scientific Reports*, *7*(1), 42456. https://doi.org/10.1038/srep42456

Schober, P., Boer, C., & Schwarte, L. A. (2018). Correlation coefficients: Appropriate use and interpretation. *Anesthesia & analgesia*, *126*(5), 1763-1768. https://doi.org/10.1213/ANE.0000000000002864

Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics, 11*, 129-158. https://doi.org/10.1093/applin/11.2.129

Schmidt, R. (2001). Attention. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 3-32). Cambridge University Press. https://doi.org/10.1017/CBO9781139524780

Seashore, C. E. (1920). The inheritance of musical talent. *The Musical Quarterly*, *6*(4), 586–598. http://www.jstor.org/stable/737981

Shen, G., & Froud, K. (2019). Electrophysiological correlates of categorical perception of lexical tones by English learners of Mandarin Chinese: an ERP study. *Bilingualism: Language and Cognition*, *22*(2), 253–265. https://doi.org/10.1017/S136672891800038X

Stanley, T. D., & Doucouliagos, H. (2014). Meta-regression approximations to reduce publication selection bias. *Research Synthesis Methods*, *5*(1), 60-78. https://doi.org/10.1002/jrsm.1095

Student. (1908). The probable error of a mean. *Biometrika*, *6*(1), 1–25. https://doi.org/10.2307/2331554

Small, C. (1997). *Music, society, education*. Wesleyan University Press.

Swaminathan, S., Kragness, H. E., & Schellenberg, E. G. (2021). The Musical Ear Test: Norms and correlates from a large sample of Canadian undergraduates. *Behavior Research Methods*, *53*(5), 2007-2024. https://doi.org/10.3758/s13428-020-01528-8

Syam, T., Syam, M. M., Khan, A., Syam, M. I., & Syam, M. I. (2023). Statistical analysis of car data using Analysis of Covariance (ANCOVA). In Yilmaz, F., Queiruga-Dios, A., Martín Vaquero, J., Mierluş-Mazilu, I., Rasteiro, D., Gayoso Martínez, V. (Eds.), *Mathematical Methods for Engineering Applications. ICMASE 2022. Springer Proceedings in Mathematics & Statistics: Vol. 414* (pp. 1-11). Springer. https://doi.org/10.1007/978-3-031-21700-5_1

Tang, W., Xiong, W., Zhang, Y. X., Dong, Q., & Nan, Y. (2016). Musical experience facilitates lexical tone processing among Mandarin speakers: Behavioral and neural evidence. *Neuropsychologia*, *91*, 247-253. https://doi.org/10.1016/j.neuropsychologia.2016.08.003

Tervaniemi, M., Putkinen, V., Nie, P., Wang, C., Du, B., Lu, J., Li, S., Cowley, B. U., Tammi, T., & Tao, S. (2022). Improved auditory function caused by music versus foreign language training at school age: is there a difference?. *Cerebral Cortex*, *32*(1), 63-75. https://doi.org/10.1093/cercor/bhab194

Thaut, M. H., McIntosh, G. C., & Hoemberg, V. (2014). Neurologic music therapy: From social science to neuroscience. In M. H, Thaut & V. Hoemberg (Eds.), *Handbook of neurologic music therapy* (pp. 1-6). Oxford University Press.

To-Chan, T. S., Loh, E. K., Tam, L. C., Woo, J. P., Chow, R. L., Fung, R. W., & Sun, N. Z. L. (2023). Teaching Chinese to L2 Preschoolers Through Children's Songs: The Cases of Mandarin and Cantonese. In J. F. Hong & C. M. SI (Eds.), *Teaching Chinese language in the international school context* (pp. 161-183). Springer. https://doi.org/10.1007/978-981-99-6376-8_10

Tong, X., Choi, W., & Man, Y. Y. (2018). Tone language experience modulates the effect of long-term musical training on musical pitch perception. *The Journal of the Acoustical Society of America*, *144*(2), 690-697. https://doi.org/10.1121/1.5049365

Van Lancker, D. (1980). Cerebral lateralization of pitch cues in the linguistic signal. *Research on Language & Social Interaction*, *13*(2), 201-277. https://doi.org/10.1080/08351818009370498

Vidal, M. M., Lousada, M., & Vigário, M. (2020). Music effects on phonological awareness development in 3-year-old children. *Applied Psycholinguistics*, *41*(2), 299-318. https://doi.org/10.1017/S0142716419000535

Vigneau, M., Beaucousin, V., Hervé, P. Y., Jobard, G., Petit, L., Crivello, F., Mellet, E., Zago, L., Mazoyer, B., & Tzourio-Mazoyer, N. (2011). What is right-hemisphere contribution to phonological, lexico-semantic, and sentence processing?: Insights from a meta-analysis. *Neuroimage*, *54*(1), 577-593. https://doi.org/10.1016/j.neuroimage.2010.07.036

Wallentin, M., Nielsen, A. H., Friis-Olivarius, M., Vuust, C., & Vuust, P. (2010a). The Musical Ear Test, a new reliable test for measuring musical competence. *Learning and Individual Differences*, *20*(3), 188-196. https://doi.org/10.1016/j.lindif.2010.02.004

Wallentin, M., Nielsen, A. H., Friis-Olivarius, M., Vuust, C., & Vuust, P. (2010b). The Musical Ear Test, a new reliable test for measuring musical competence: Corrigendum. *Learning and Individual Differences*, *20*(6), 705. https://doi.org/10.1016/j.lindif.2010.10.001

Wang, C., Flemming, K., Cortiana, G., Putkinen, V., Lammert, J., Rafat, Y., Tao, S., & Joanisse, M. F. (2022). Chinese-English bilinguals are more sensitive to environmental sound perception than Spanish-English bilinguals through top-down cognitive mechanism. *Biological Psychology*, *175*, 108449. https://doi.org/10.1016/j.biopsycho.2022.108449

Wang, C., Tao, S., Tao, Q., Tervaniemi, M., Li, F., & Xu, P. (2020). Musical experience may help the brain respond to second language reading. *Neuropsychologia*, *148*, 107655. https://doi.org/10.1016/j.neuropsychologia.2020.107655

Whitaker, H. A. (1998). Neurolinguistics from the middle ages to the pre-modern era: Historical vignettes. In H. A. Whitaker & Stemmer, B. (Eds.), *Handbook of neurolinguistics* (pp. 27-54). Elsevier Science & Technology.

Wiener, S., & Bradley, E. D. (2023). Harnessing the musician advantage: Short-term musical training affects non-native cue weighting of linguistic pitch. *Language Teaching Research*, *27*(4), 1016-1031. https://doi.org/10.1177/1362168820971791

Wong, Y. K., Ngan, V. S., Cheung, L. Y., & Wong, A. C. (2020). Absolute pitch learning in adults speaking non-tonal languages. *Quarterly Journal of Experimental Psychology*, *73*(11), 1908-1920. https://doi.org/10.1177/1747021820935776

Wu, H., Ma, X., Zhang, L., Liu, Y., Zhang, Y., & Shu, H. (2015). Musical experience modulates categorical perception of lexical tones in native Chinese speakers. *Frontiers in Psychology*, *6*, 436. https://doi.org/10.3389/fpsyg.2015.00436

Xi, J., Zhang, L., Shu, H., Zhang, Y., & Li, P. (2010). Categorical perception of lexical tones in Chinese revealed by mismatch negativity. *Neuroscience*, *170*(1), 223-231. https://doi.org/10.1016/j.neuroscience.2010.06.077

Yang, S., Tang, L., Liu, L., Dong, Q., Georgiou, G. K., & Nan, Y. (2024). Musical pitch processing predicts reading development in Chinese school-age children. *Journal of Neurolinguistics*, *71*, 101199. https://doi.org/10.1016/j.jneuroling.2024.101199

Yip, M. J. W. (2002). *Tone.* Cambridge University Press.

Zatorre, R. J. (2024). *From perception to pleasure: The neuroscience of music and why we love it*. Oxford University Press. https://doi.org/10.1093/oso/9780197558287.001.0001

Zatorre, R. J., & Belin, P. (2001). Spectral and temporal processing in human auditory cortex. *Cerebral cortex*, *11*(10), 946-953. https://doi.org/10.1093/cercor/11.10.946

Zatorre, R. J., Belin, P., & Penhune, V. B. (2002). Structure and function of auditory cortex: music and speech. *Trends in Cognitive Sciences*, *6*(1), 37-46. https://doi.org/10.1016/S1364-6613(00)01816-7

Zatorre, R. J., Chen, J. L., & Penhune, V. B. (2007). When the brain plays music: auditory–motor interactions in music perception and production. *Nature Reviews Neuroscience*, *8*(7), 547-558. https://doi.org/10.1038/nrn2152

Zendel, B. R., & Alain, C. (2009). Concurrent sound segregation is enhanced in musicians. *Journal of Cognitive Neuroscience*, *21*(8), 1488-1498. https://doi.org/10.1162/jocn.2009.21140

Zhang, H. (2018). *Second Language Acquisition of Mandarin Chinese Tones: Beyond First-Language Transfer*. Brill. https://doi.org/10.1163/9789004364790

Zhang, L. (2019). Tone features of Chinese and teaching methods for second language learners. *International Journal of Chinese Language Education*, *5*, 45-66. https://ijcle.eduhk.hk/uploads/issues/376/IJCLE_No.5_Jun19_03.pdf

Zhang, L., Xie, S., Li, Y., Shu, H., & Zhang, Y. (2020). Perception of musical melody and rhythm as influenced by native language experience. *The Journal of the Acoustical Society of America*, *147*(5), EL385-EL390. https://doi.org/10.1121/10.0001179

Appendix A. Ethic Approvals

The ethic approval for participant recruitment before and after differences of experiment needs are shown in Figures 2-3.

Figure 2

Ethic Approval (Version 1)

CUREC 1B - The influence of musicianship and language background on tonal perception

SC **Student CUREC <student.curec@education.ox.ac.uk>** Monday 11 March 2024 at 14:20

To: Xinhui Zhang

Cc: Faidra Faitaki; Student CUREC

Dear Xinhui,

Thank you for addressing my comments and for providing additional information. Apologies for the delay in getting back to you.

I am writing to acknowledge receipt of your CUREC 1B application entitled '*The influence of musicianship and language background on tonal perception*'. The application was reviewed and approved by Dr Faidra Faitaki, your supervisor. No further approval from the Education DREC is required for applications reviewed under the CUREC 1B process. As such, the project will not receive a formal letter of ethical approval from the SSH IDREC.

The ethics reference for your application is C1B-24HT-Educ-003. Please add this reference to your CUREC 1B form and include it on documents for the research participants such as the participant information sheet. The guidance notes from the footer of the participant information sheet are for you rather than your participants, so I would remove them from there.

Please note that this is contingent on the research project adhering to the criteria set out in the CUREC 1B guidance. Please ensure, therefore, that you comply with the conditions of this process and, should anything change in the course of the project, you should discuss this with your supervisor to determine whether this requires further review and approval by the Education DREC.

Please don't hesitate to get in touch if you have any questions.

All the best for your research – we hope it goes well.

Irina

Irina Lepadatu
Research Manager

Department of Education, University of Oxford
15 Norham Gardens, Oxford, OX2 6PY

Research SharePoint site

LinkedIn I X I YouTube

DEPARTMENT OF EDUCATION UNIVERSITY OF OXFORD

Figure 3

Ethic Approval (Version 2)

RE: CUREC 1B Amendment Process?

SC

Student CUREC <student.curec@education.ox.ac.uk>

Wednesday 5 June 2024 at 17:26

To: Xinhui Zhang; Faidra Faitaki; **Cc:** Student CUREC

Dear Zhuohan,

Thank you for amending that.

Please use this reference in your documents: **C1B-24HT-Educ-003_Amendment_01**.

Good luck with your project!

Best wishes,

Irina

Irina Lepadatu
Research Manager

Department of Education, University of Oxford
15 Norham Gardens, Oxford, OX2 6PY

Research SharePoint site

LinkedIn | X | YouTube

DEPARTMENT OF EDUCATION

Dr Faidra Faitaki

University Direct Line: +44 (0) 1865 274024

University e-mail: faidra.faitaki@education.ox.ac.uk

Participant information for online surveys or tasks

The influence of musicianship and language background on tonal perception

CUREC Approval Reference: C1B-24HT-Educ-003

General Information

The aim of this research is to investigate the influence of musical aptitude and language background on tonal perception.

We appreciate your interest in participating in this online task. Please read through this information before agreeing to participate (if you wish to) by ticking the 'yes' box below.

You may ask any questions before deciding to take part by contacting the researcher (details below).

The Principal Researcher is Xinhui Zhang, who is attached to the Department of Education at the University of Oxford. This research is being completed under the supervision of Dr Faidra Faitaki.

You will be asked to answer the questions about your language background, perform the Music Ear Test, during which you will be asked if two sounds are identical, and complete a tonal discrimination task, during which you will be asked if two words have the same tone. This should take about 25 minutes. No background knowledge is required. The data will only be accessed to the principal researcher and her supervisor. The data is needed for research purposes and will form part of the dissertation of the student researcher and might also be used to inform academic presentations/publications. The data collection process is voluntary and does not require participants to provide any identifying information or contact details.

Do I have to take part?

No. Please note that participation is voluntary. If you do decide to take part, you may withdraw at any point for any reason before submitting your answers by simply closing the browser.

We have included a 'Prefer not to say' option for each set of questions should you prefer not to answer a particular question.

How will my data be used?

The data that we will collect are anonymous. None of personal information that can identify a person will be collected.

Your IP address will not be stored[1]. We will take all reasonable measures to ensure that data remain confidential.

The responses you provide will be stored in a password-protected electronic file on University of Oxford secure servers and may be used in the student's academic assessment, publications in a peer reviewed journal, conference presentations, reports for external organisations, and websites. Research data will be stored for 3 years after publication or public release of the work of the research.

Who will have access to my data[2]?

The University of Oxford is the data controller with respect to your personal data and, as such, will determine how your personal data is used in the research. The University will process your personal data for the purpose of the research outlined above. Research is a task that we perform in the public interest. Further information about your rights with respect to your personal data is available from https://compliance.admin.ox.ac.uk/individual-rights.

Who has reviewed this research?

This research has been reviewed and approved by my supervisor on behalf of the Department of Education's Research Ethics Committee [C1B-24HT-Educ-003].

Who do I contact if I have a concern or I wish to complain?

If you have a concern about any aspect of this research, please speak to Xinhui Zhang, xinhui.zhang@education.ox.ac.uk, or her supervisor Dr Faidra Faitaki, faidra.faitaki@education.ox.ac.uk, and we will do our best to answer your query. I/ We will acknowledge your concern within 10 working days and give you an indication of how it will be dealt with. If you remain unhappy or wish to make a formal complaint, please contact the Chair of the Research Ethics Committee at the University of Oxford who will seek to resolve the matter as soon as possible:

The Chair, Education Departmental Research Ethics Committee

[1] Guidance is available within CUREC's Best Practice Guidance (06) on Internet-mediated research

[2] Guidance is available within CUREC's Best Practice Guidance (09) on Data collection, protection and management

Email: student.curec@education.ox.ac.uk

Address: 15 Norham Gardens, Oxford, OX2 6PY

Please note that you may only participate in this survey if you are 18 years of age or over.

☐ I certify that I am 18 years of age or over

If you have read the information above and agree to participate with the understanding that the data (including my personal data) you submit will be processed accordingly, please tick the box below to start.

☐ Yes, I agree to take part

Appendix B2. Participant Information Sheet (Version 2)

DEPARTMENT OF EDUCATION

Principal investigator: Dr Faidra Faitaki

University Direct Line: +44 (0) 1865 274024

University e-mail: faidra.faitaki@education.ox.ac.uk

Student researcher: Xinhui (Zhuohan) Zhang

University e-mail: xinhui.zhang@education.ox.ac.uk

Participant information for online surveys or tasks

The influence of musicianship and language background on tonal perception

CUREC Approval Reference: C1B-24HT-Educ-003_Amendment_01

General Information

The aim of this research is to investigate the influence of musical aptitude and language background on tonal perception.

We appreciate your interest in participating in this online task. Please read through this information before agreeing to participate (if you wish to) by ticking the 'yes' box below.

You may ask any questions before deciding to take part by contacting the researcher (details below).

The Principal Researcher is Xinhui Zhang, who is attached to the Department of Education at the University of Oxford. This research is being completed under the supervision of Dr Faidra Faitaki.

You will be asked to answer the questions about your language background, perform the Music Ear Test, during which you will be asked if two sounds are identical, and complete a tonal discrimination task, during which you will be asked if two words have the same tone. This should take about 25 minutes. No background knowledge is required. The data will only be accessed to the principal researcher and her supervisor. The data is needed for research purposes and will form part of the dissertation of the student researcher and might also be used to inform academic presentations/publications. The data collection process is voluntary and does not require participants to provide any identifying information or contact details.

Do I have to take part?

No. Please note that participation is voluntary. If you do decide to take part, you may withdraw at any point for any reason before submitting your answers by simply closing the browser.

We have included a 'Prefer not to say' option for each set of questions should you prefer not to answer a particular question.

Expenses and payments

You will be paid at the rate of £8.00 per hour if your data meets the screening criteria: 1) all your responses to questions are in English and 2) relevant to the questions. If your responses do not meet the criteria aforementioned, your data will be deleted permanently and you will not receive any form of reimbursement.

How will my data be used?

The data that we will collect are anonymous. None of personal information that can identify a person will be collected.

Your IP address will not be stored[3]. We will take all reasonable measures to ensure that data remain confidential.

The responses you provide will be stored in a password-protected electronic file on University of Oxford secure servers and may be used in the student's academic assessment, publications in a peer reviewed journal, conference presentations, reports for external organisations, and websites. Research data will be stored for 3 years after publication or public release of the work of the research.

Who will have access to my data[4]?

The University of Oxford is the data controller with respect to your personal data and, as such, will determine how your personal data is used in the research. The University will process your personal data for the purpose of the research outlined above. Research is a task that we perform in the public interest. Further information about your rights with respect to your personal data is available from https://compliance.admin.ox.ac.uk/individual-rights.

Who has reviewed this research?

This research has been reviewed and approved by my supervisor on behalf of the Department of Education's Research Ethics Committee [C1B-24HT-Educ-003_Amendment_01].

Who do I contact if I have a concern or I wish to complain?

---

[3] Guidance is available within CUREC's Best Practice Guidance (06) on Internet-mediated research

[4] Guidance is available within CUREC's Best Practice Guidance (09) on Data collection, protection and management

If you have a concern about any aspect of this research, please speak to Xinhui Zhang, xinhui.zhang@education.ox.ac.uk, or her supervisor Dr Faidra Faitaki, faidra.faitaki@education.ox.ac.uk, and we will do our best to answer your query. I/ We will acknowledge your concern within 10 working days and give you an indication of how it will be dealt with. If you remain unhappy or wish to make a formal complaint, please contact the Chair of the Research Ethics Committee at the University of Oxford who will seek to resolve the matter as soon as possible:

The Chair, Education Departmental Research Ethics Committee

Email: student.curec@education.ox.ac.uk

Address: 15 Norham Gardens, Oxford, OX2 6PY

Please note that you may only participate in this survey if you are 18 years of age or over.

☐ I certify that I am 18 years of age or over

If you have read the information above and agree to participate with the understanding that the data (including my personal data) you submit will be processed accordingly, please tick the box below to start.

☐ Yes, I agree to take part

Appendix C. Advertising

DEPARTMENT OF EDUCATION

University of Oxford

15 Norham Gardens

Oxford, OX2 6PY

Student researcher: Xinhui (Zhuohan) Zhang

xinhui.zhang@education.ox.ac.uk

The influence of musicianship and language background on tonal perception

Ethics Approval Reference: C1B-24HT-Educ-003

VOLUNTEERS NEEDED FOR EXPERIMENT

Music and language are not two isolated entities. The research aims to investigate if musical experience and language background influence people's sensitivity on tonal perception. The participants will listen to 15 pairs of words and answer if there is a tonal difference within each pair. Accurate perception can improve language learners' performance, so finding out what factors influence it can help learners improve their acquisition of tonal languages.

We are looking for volunteers without severe listening impairments, aged at least 18 years old, to take part in an online study, involving a single session. The session will take about 25 minutes of your time. You would be asked to answer questions about your language and music background and perform two short tasks assessing your musical ability and tonal discrimination.

If you are interested and would like more information, please contact Xinhui (Zhuohan) Zhang at xinhui.zhang@education.ox.ac.uk. Zhuohan is based at the Department of Education, 15 Norham Gardens, OX2 6PY, Oxford. There is no obligation to take part.

Thank you!

Appendix D. Power Analysis

The *a priori* power analyses were conducted for Mann-Whitney Test, MANOVA (substitute for LMM), and ANCOVA (to estimate the sample size of nonparametric ANCOVA) to derive minimum sample sizes. Their displays in G*Power are shown in Figures 4-6.

Figure 4

*Display in G*Power: Power Analysis for Mann-Whitney Test*

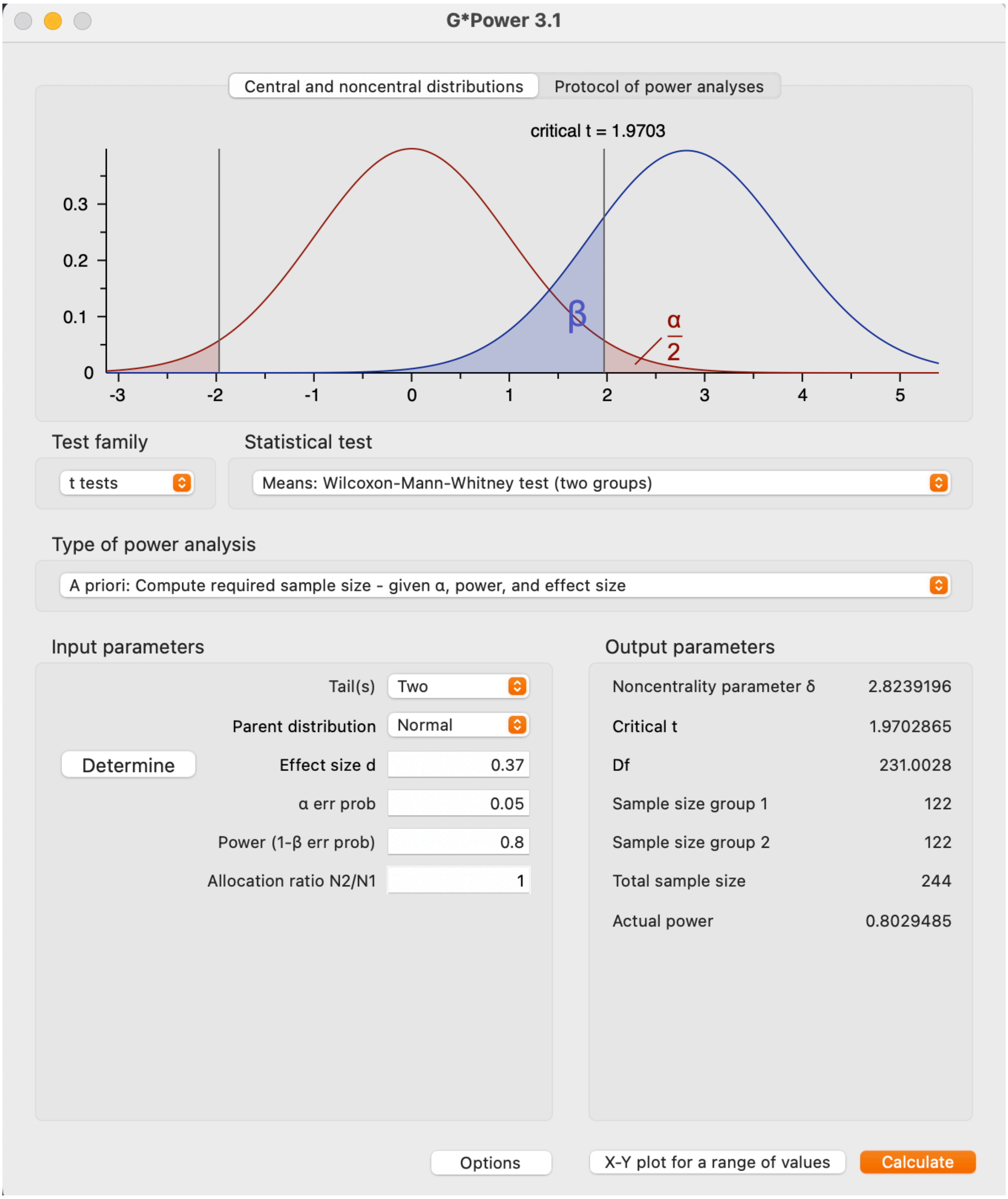


Figure 5

*Display in G*Power: Power Analysis for MANOVA*

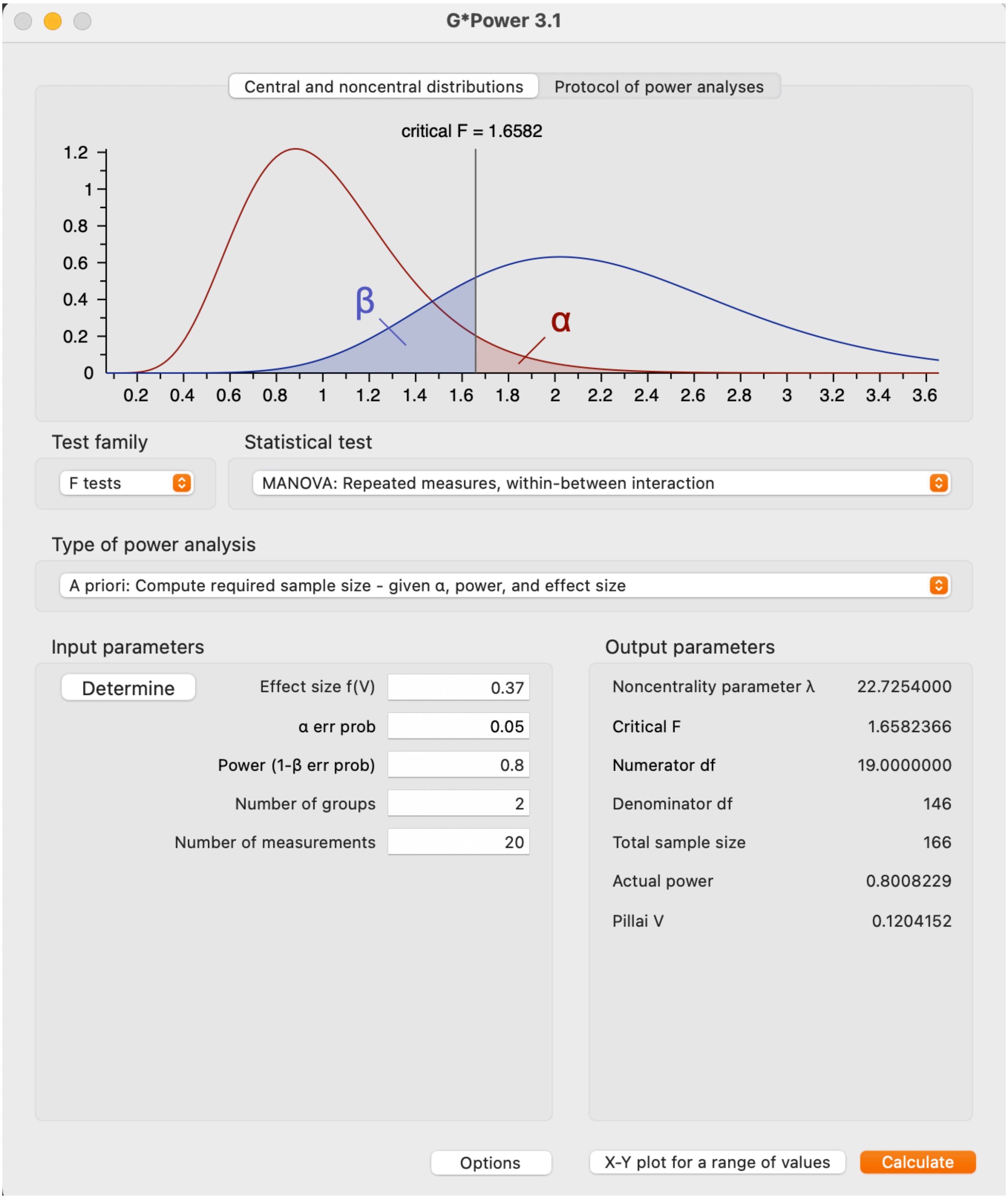
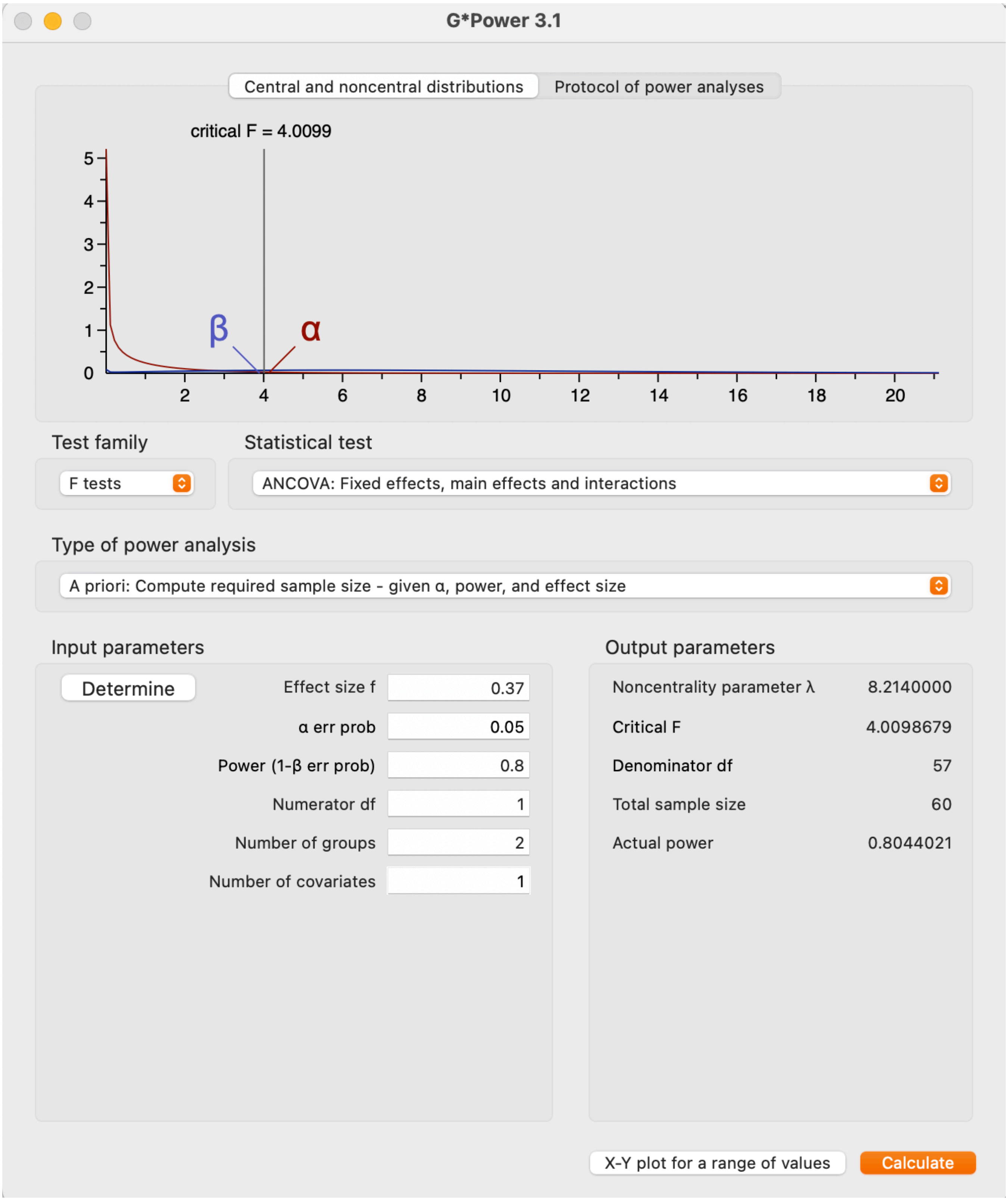


Figure 6

*Display in G*Power: Power Analysis for ANCOVA*



Appendix E. The Language Background and Music Training Experience Questionnaire in Experiment

Language and Music Background Questionnaire

Welcome to the last part!

You will be asked to answer questions about your language backgrounds.

This part will take around 2-3 minutes.

Please feel free to take a short break before the start!

Please try to answer all the questions in this section that are suitable to you and provide all the answers, otherwise your answers in previous sections do not count!

Please firstly enter your current age in years: ____________

Please enter your first language (i.e., the language you were firstly taught since birth). If there is more than one language, please list each one of them: ____________

1. Indicate your native language(s) and any other languages you have studied or learned, and the total number of years you have spent using each language.

*Notes: For "Years of use", you may have learned a language, stopped using it, and then started using it again. Please give the total number of years.

You can type decimals or select integers in all the number entries.

Language	Years of use*

2. List the number of instrument(s) you have been learnt/played, the age of starting to learn/play the instrument, and how many years you have been played since then.

*Notes: For "Years of Training", you may have learned/played an instrument, stopped learning/playing it, and then started to learn/play it again. Please give the total number of years.

If you do not have relevant experience, please type "0" in all three entries.

Number of instrument(s)	
Age of Starting to learn/play (Years)	
Years of Training*	

3. Have you ever received any singing training?

*Notes: For "Years of Training", you may have received training, stopped singing, and then started to sing again. Please give the total number of years.

If you do not have relevant experience, please type "0" in both entries.

Age of Starting to learn (Years)	
Years of Training*	

Appendix F. Thai Word List for Audio Materials

The list of Thai words selected from Liao (2008) is presented in Table 12 below, with their English translation, romanised transcription with original and deviant tones.

Table 12

Thai Word List for Audio Materials in Thai Characters with English Translations and Romanised Transcriptions

No.	Thai words	English translation		Romanisation with original tones		Romanisation with modified tones
1	*ดีใจ*	happy	1A	‘ddii^{1}zai^{1}’	1B	‘ddii^{3}zai^{1}’
2	*เก้าอี้*	chair	2A	‘gao^{3}i^{3}’	2B	‘gao^{3}i^{3}’
3	*โกโก้*	coco	3A	‘goo^{1}goo^{3}’	3B	‘goo^{4}goo^{3}’
4	*ไข่ไก่*	egg	4A	‘kai^{2}gai^{2}’	4B	‘kai^{4}gai^{2}’
5	*เข้าใจ*	understand	5A	‘kao^{3}zai^{1}’	5B	‘kao^{3}zai^{1}’
6	*กะปิ*	shrimp sauce	6A	‘ga^{2}bi^{2}’	6B	‘ga^{2}bi^{5}’
7	*ไฟฟ้า*	electricity	7A	‘fai^{1}faa^{4}’	7B	‘fai^{1}faa^{5}’
8	*น้ำชา*	tea	8A	‘nam^{4}chaa^{1}’	8B	‘nam^{4}chaa^{1}’
9	*ไม่ใช่*	no	9A	‘mai^{3}chai^{3}’	9B	‘mai^{3}chai^{5}’
10	*ที่นี่*	here	10A	‘ti^{3}ni^{3}’	10B	‘ti^{2}ni^{3}’
11	*เสื้อผ้า*	clothes	11A	‘sea^{3}pa^{3}’	11B	‘sea^{3}pa^{3}’
12	*มะเขือ*	eggplant	12A	‘ma^{4}kea^{5}’	12B	‘ma^{4}kea^{5}’
13	*ผัวเมีย*	husband and wife	13A	‘pua^{5}mia^{1}’	13B	‘pua^{5}mia^{1}’
14	*เมื่อเช้า*	this morning	14A	‘mea^{3}chao^{4}’	14B	‘mea^{3}chao^{4}’
15	*เยอะแยะ*	many	15A	‘ye^{4}ye^{4}’	15B	‘ye^{4}ye^{4}’
16	*สีเทา*	grey	16A	‘sii^{5}tao^{1}’	16B	‘sii^{5}tao^{1}’
17	*ราคา*	price	17A	‘laa^{1}kaa^{1}’	17B	‘laa^{1}kaa^{1}’
18	*สาลี่*	pear	18A	‘saa^{5}lii^{3}’	18B	‘saa^{5}lii^{3}’
19	*เชื่อใจ*	believe	19A	‘chea^{3}zai^{1}’	19B	‘chea^{3}zai^{2}’
20	*แน่ใจ*	be sure	20A	‘nee^{3}zai^{1}’	20B	‘nee^{4}zai^{1}’

Appendix G. Illustrations of Trial Questions in Tonal Discrimination Task

Before the start of tonal discrimination task, participants will answer two trial question to familiarise themselves. After the completion of each trial question, its correct answer will be shown immediately (see Figures 7-10).

Figure 7

Online Display of the Tonal Discrimination Task in Gorilla (Trial 1 Question)

Figure 8

Online Display of the Tonal Discrimination Task in Gorilla (Trial 1 Answer)

Figure 9

Online Display of the Tonal Discrimination Task in Gorilla (Trial 2 Question)

Figure 10

Online Display of the Tonal Discrimination Task in Gorilla (Trial 2 Answer)

The figure originally presented here cannot be made freely available via ORA because of copyright.
The figure was sourced at Gorilla Experiment Builder. (n.d.). Gorilla.sc Platform. Retrieved from https://gorilla.sc

Appendix H. Figures of Normal Distribution of Tonal Discrimination Task Scores across Groups

The illustrations of normal distribution of tonal discrimination task scores of all participants (see Figure 11), tonal L1 group (see Figure 12), and non-tonal L1 group (see Figure 13) are presented below.

Figure 11

Histogram of Tonal Discrimination Task Score of All Participants in SPSS

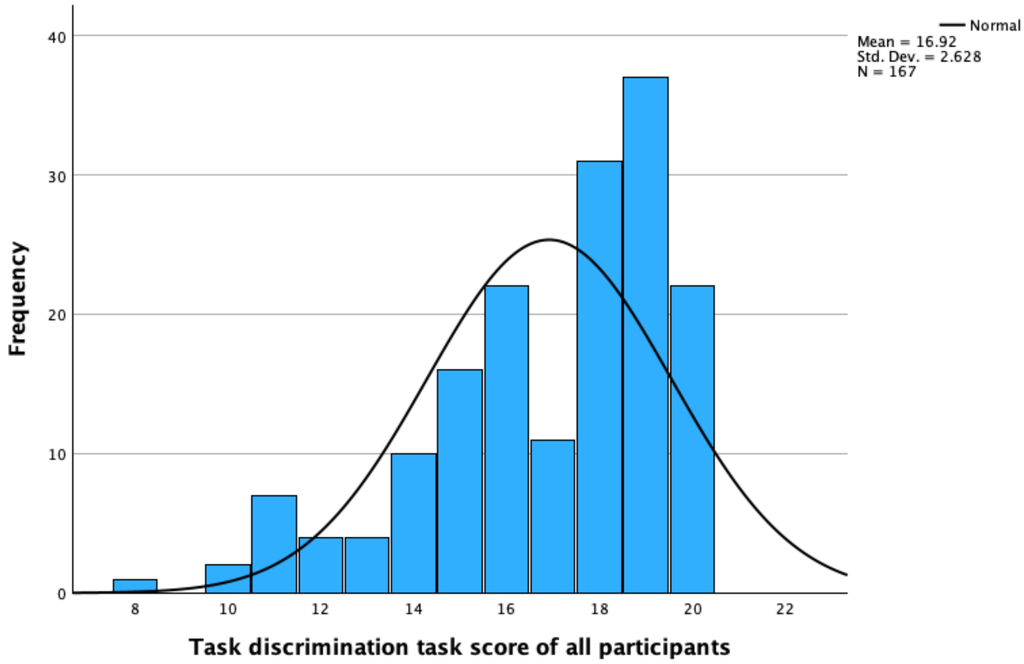


Figure 12

Histogram of Tonal Discrimination Task Score of Tonal L1 Group in SPSS

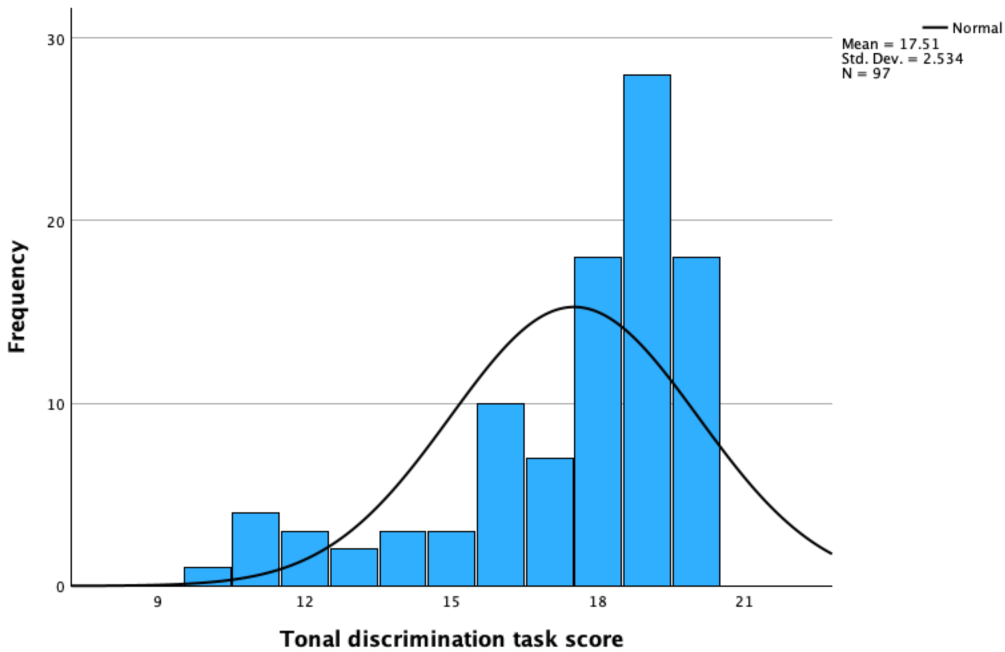
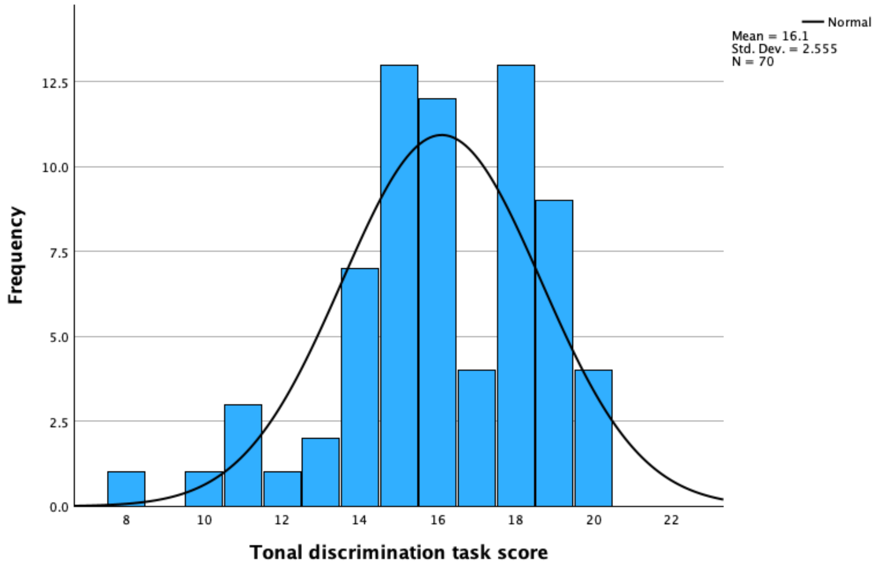


Figure 13

Histogram of Tonal Discrimination Task Score of Non-tonal L1 Group in SPSS



Appendix I. Displays of Linear Mixed-effect Model in RStudio

Three models were conducted via RStudio (Posit Team, 2024), with language background and MET total scores as fixed effect (see Figure 14), language background and MET melody scores as fixed effect (see Figure 15), and language background and MET rhythm scores as fixed effect (see Figure 16). The random effects for all the three models were factors of each participant and each question in tonal discrimination task. The coding and outcomes are shown in these three Figures. The dependent variable, Response, to the left of ~ was predicted by the sum of the variables at the right. The fixed effects were L1, indicating language background, and MET total scores. The random effects were specified with | symbol and, in this case, it was assumed that they had the same random effects across responses.

Figure 14

Display in RStudio: Output of Linear Mixed-Effect Model in RStudio with Language Background and MET Total Scores as Fixed Effect (Model 0)

Figure 15

Display in RStudio: Output of Linear Mixed-Effect Model in RStudio with Language Background and MET Melody Scorse as Fixed Effect (Model 1)

Figure 16

Display in RStudio: Output of Linear Mixed-Effect Model in RStudio with Language Background and MET Rhythm Scores as Fixed Effect (Model 2)

Appendix J. Display of Nonparametric ANCOVA in RStudio

The nonparametric ANCOVA was conducted via RStudio (Posit Team, 2024). The L1 background (i.e., tonal vs. non-tonal), a categorical variable, was predictor, and MET total score was covariate.

Figure 17

Display in RStudio: Output of Nonparametric ANCOVA in RStudio with Language Background as Predictor and MET Total Scores as Covariate

The figure originally presented here cannot be made freely available via ORA because of copyright.
The figure was sourced at RStudio Team. (2024). RStudio: Integrated Development Environment for R. Posit Software.