The Effects of Ethnicity, Musicianship, and Tone Language Experience on Pitch Perception

Yi Zheng [1] and Arthur G. Samuel [1,2,3]

1. Department of Psychology, Stony Brook University
2. Basque Center on Cognition, Brain, and Language
3. Ikerbasque, Basque Foundation for Science

Contact Information:

Yi Zheng

Dept of Psychology

Stony Brook University

Stony Brook, NY 11794-2500

Email: Yi.Zheng.1@stonybrook.edu

Abstract

Language and music are intertwined: music training can facilitate language abilities, and language experiences can also help with some music tasks. Possible language-music transfer effects are explored in two experiments in the present study. In Experiment 1, we tested native Mandarin, Korean and English speakers on a pitch discrimination task with two types of sounds: speech sounds and fundamental frequency (F0) patterns derived from speech sounds. To control for factors that might influence participants' performance, we included cognitive ability tasks testing memory and intelligence. Additionally, two music skill tasks were used to examine general transfer effects from language to music. Prior studies showing that tone language speakers have an advantage on pitch tasks have been taken as support for three alternative hypotheses: specific transfer effects, general transfer effects, and an ethnicity effect. In Experiment 1, musicians outperformed non-musicians on both speech and F0 sounds, suggesting a music to language transfer effect. Korean and Mandarin speakers performed similarly and they both outperformed English speakers, providing some evidence for an ethnicity effect. Alternatively, this could be due to population selection bias. In Experiment 2, we recruited Chinese Americans approximating the native English speakers' language background to further test the ethnicity effect. Chinese Americans, regardless of their tone language experiences, performed similarly to their non-Asian American counterparts in all tasks. Therefore, while the current study provides additional evidence of transfer effects across music and language, it casts doubt on the contribution of ethnicity to differences observed in pitch perception and general music abilities.

Keywords: music & language; specific transfer; general transfer; ethnicity effects; pitch perception

The Effects of Ethnicity, Musicianship, and Tone Language Experience on Pitch Perception

It has been said that music is a universal language. Certainly, language and music share a number of properties: they exist in every culture, and they both convey emotions. In music, a major key or minor key can express happy or sad emotions, respectively. In language, a high pitch can express anger or happiness, and a low pitch can express fear, sadness, tenderness, etc. (Patel, 2010). However, music and language use different building blocks and different means to construct their systems. In general, music uses notes and language uses words; music uses tonal relations whereas language uses semantic and syntactic relations.

Previous research has investigated whether language and music share mechanisms in the brain. Early lesion studies suggested that language and music have separate processing regions in the brain: Patients with impaired language abilities still showed preserved musical abilities (Assal, 1973). There are also cases of patients being able to recognize a song by its lyrics but not by its melody (Peretz et al., 1994). Nevertheless, other studies have provided evidence that music and language are intertwined (Koelsch et al., 2002; Patel, Gibson, Ratner, Besson, & Holcomb, 1998; Schön et al., 2010). For example, Patel et al. (1998) found that music and language share neural resources in structure processing, with a violation of syntax in music (i.e., an out-of-key tone) eliciting the same ERP pattern as a syntactic violation in language. Patel (2003) has proposed the “shared syntactic integration resource hypothesis” (SSIRH), arguing that music and linguistic syntax share neural resources. Some converging evidence comes from Koelsch et al.’s (2002) study showing that Broca’s area is active for syntax violations in both language and Western tonal music.

The similarities between language and music raise the possibility of transfer effects between them (for a review, see Asaridou & McQueen, 2013, or Bradley, 2013). There are also several studies showing that musical training facilitates language abilities. Musicians outperform non-musicians in extracting prosodic information from speech in a familiar language as well as in a foreign language (Thompson, Schellenberg, & Husain, 2003). Moreover, musicians have better abilities in shadowing speech sounds (Pastuszek-Lipińska, 2007, 2008). In addition to music transfer effects to language, there is evidence of transfer to other areas, such as mathematics, visual-spatial abilities, emotion interpretation, and vocal abilities (Amir, Amir, & Kishon-Rabin, 2003; Cheek & Smith, 1999; Nilsonne & Sundberg, 1985; Thompson, Schellenberg, & Husain, 2001, 2004).

Patel (2012) proposed the OPERA theory to explain transfer effects from music to language. OPERA stands for Overlap, Precision, Emotion, Repetition, and Attention. This theory claims that when all five of the following conditions are met then musical training can facilitate speech processing: First, there must be anatomical overlap in the brain regions that process music and speech; second, music must demand greater precision than speech; third, musical activities must elicit strong positive emotion; fourth, musical activities must be frequently repeated; lastly, musical activities must be associated with focused attention.

Compared to the music to language transfer effect, the other direction -- a language to music transfer effect-- is less studied. According to Patel's (2012) OPERA theory, this transfer effect should be weak because language generally demands less precision than music in terms of pitch. However, tone languages may constitute a special case in which language requires high precision, as tone-language speakers use precise and stable pitch patterns when producing words, compared to non-tone-language speakers (Deutsch, Henthorn, & Dolson, 2004a). In Mandarin, for example, there are four lexical tones (see Figure 1): tone 1 is flat, tone 2 is rising, tone 3 is falling-rising, and tone 4 is falling. When the same syllable (e.g., "ma") is pronounced with each of the four different tones, it means different things (Ota, 2003). Given this, tone language use requires attention to lexical pitch changes in order to efficiently understand the meanings of words.

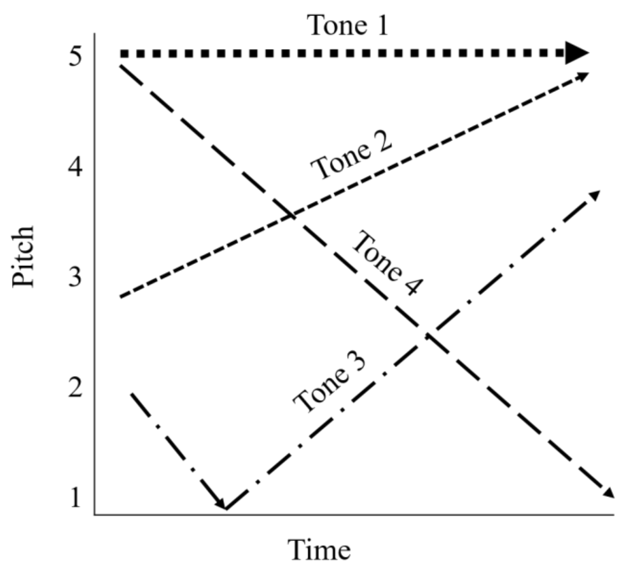

Figure 1. **Four lexical tones in Mandarin: tone 1 is flat, tone 2 is rising, tone 3 is falling-rising, and tone 4 is falling.**

In fact, there is evidence showing that tone-language speakers have an advantage on some pitch tasks over non-tone-language speakers (Bent, Bradlow, & Wright, 2006; Bradley, 2012; Chen, Liu, & Kager; Hove, Sutherland, & Krumhansl, 2010). The evidence for this advantage has been taken as support for three different theories: specific transfer effects, general transfer effects, and what has been called an ethnicity effect (Hove et al., 2010). Although there is no clearcut boundary between these theories, categorizing them helps to organize the substantial literature on this issue. The **specific transfer effect** view is that tone language speakers are only better at detecting lexical pitch changes or changes that are closely related to the participants' first languages. The **general transfer effect** view is that tone language speakers have an advantage on pitch tasks relevant to their native tone languages, but also have enhanced general pitch skills that are not directly linked to their first languages. The **Ethnicity effect** hypothesis is that East Asians (e.g., Mandarin speakers) have better pitch-based auditory abilities due to their Asian identity, independent of whether or not they speak tone languages (Gregersen, Kowalsky, & Li, 2007).

Specific transfer effects.

Ngo, Vu, and Strybel (2016) found that tone language speakers did not differ from non-tone language speakers on a relative pitch task, suggesting that when the pitch task was less similar to language (i.e., more relevant to music), the advantage of speaking a tone language was lost. They argued that the benefit of speaking a tone language might be observed if a language related pitch task was used. Similar evidence was provided for domain-specific transfer effects by showing that Mandarin speakers were more accurate at lexical pitch identification than English speakers, but not on a non-speech pitch discrimination task (Bent et al., 2006). Specifically, in a Mandarin tone identification task, participants were asked to identify if a syllable had a level, rising, low-dipping, or falling tone (i.e., the four Mandarin lexical tones, as shown in Figure 1); in the non-speech pitch discrimination task, participants heard pairs of short sounds differing in pitch and had to identify if the higher frequency was the first or second stimulus. There was also a third task in which participants identified if sine waves were flat, rising or falling, and Mandarin speakers misidentified flat and falling tones partly due to a negative influence from their native language.

Asaridou, Hagoort, and McQueen (2015) found that early Cantonese bilinguals processed sung stimuli more holistically than Dutch late bilinguals, which they attributed mainly to tone language experience. However, this advantage did not transfer to general musical pitch processing. Collectively, these results suggest that the advantage of tone-language speakers does not apply to every aspect of pitch perception -- it is specific to speech or speech-like stimuli.

General transfer effects.

Bradley (2012) found that Mandarin speakers were better than English speakers in melodic discrimination tasks on both contour (i.e., the direction of pitch change) and interval (i.e., the slope of pitch change) violation conditions. This is consistent with the findings by Chen et al. (2016) that speaking a tone language enhances pitch perception in general, measured by the music ear test (MET, see Wallentin, Nielsen, Friis-Olivarius, Vuust, & Vuust, 2010) and the Montreal battery of evaluation of amusia (MBEA, see Peretz, Champod, & Hyde, 2003). Mandarin speakers were also better than English speakers at a melody tone discrimination task (Alexander, Bradlow, Ashley, & Wong, 2008) and a melody tone identification task (Chang, Hedberg, & Wang, 2016), and Cantonese (tone language) speakers were better than English speakers at general musical pitch processing (Bidelman, Hutka, & Moreno, 2013; Hutka, Bidelman, & Moreno, 2015). Moreover, Thai (tone language) speakers outperformed Australians at discriminating the pitch contour not only in intact and low-pass filtered language sounds but also in music stimuli (Stevens, Keller, & Tyler, 2004). Interestingly, in a study in which music skills were matched, tone-language speakers (i.e., Cantonese, Vietnamese, and Mandarin) were better than English speakers at discriminating pairs of two-note sounds, but not pairs of single pitch sounds (Pfordresher & Brown, 2009). However, a companion study using exactly the same procedure showed that when the control non-tone language group was changed from monolinguals to bilinguals, the advantage for Mandarin speakers disappeared in both pitch discrimination tasks (Giuliano, Pfordresher, Stanley, Narayana, & Wicha, 2011). In brain imaging studies, compared to non-tone language speaking non-musicians, tone language speaking non-musicians showed a stronger frequency-following response (Bidelman, Gandour, & Krishnan, 2011) and earlier neural responses (Giuliano et al., 2011) on musical pitch discrimination tasks, and a stronger mismatch negativity on pitch contour tasks (Chandrasekaran, Krishnan, & Gandour, 2009).

An ethnicity effect.

Hove et al. (2010) argued that the observed advantage in music ability for Mandarin speakers is an ethnicity effect instead of a language effect. They trained non-musicians to recognize different musical intervals. After training, Mandarin speakers and Korean speakers (non-tone-language) showed no difference in their performance, and both were better than English speakers. Speakers of Hmong (a tone language spoken in Vietnam and adjacent countries that is distinct from Mandarin Chinese) who were living in the US performed similarly to English speakers, and both were significantly worse than Mandarin speakers. Hove et al. (2010) treated the Hmong speakers as culturally and genetically distinct from Chinese, and claimed that the equivalent performance by Hmong and English speakers supports an ethnicity effect. However, given the variety of ethnicities in Asia and the broader definition of "East Asian", the Hmong are not clearly more distant from the Chinese than the Koreans are. Therefore, this research does not provide strong and clear support for an ethnicity effect.

In another study, Vietnamese (tone language) and English speakers were asked to discriminate whether tone pairs were ascending or descending. Vietnamese who arrived in the US as adults significantly outperformed English speakers but did not differ from early arrival Vietnamese who spoke primarily English and were not fluent in Vietnamese (Deutsch, Henthorn, & Dolson, 2004b). This could be taken as potential support for an ethnicity effect. Alternatively, as the authors suggested, the results could reflect the influence of speech experience early in life.

Other studies, primarily focused on musicians, have shown a higher prevalence of absolute pitch (AP-- the ability to name a musical note without a reference note) among East Asians compared to non-Asian populations (Gregersen, Kowalsky, Kohn, & Marvin, 2001; Gregersen et al., 2007; Zatorre, 2003). Deutsch, Dooley, Henthorn, and Head (2009) claimed that the prevalence of AP among Asians was not due to their ethnicity but instead was due to tone language experience. They found that Asian non-fluent tone language speakers were similar to Caucasian non-tone language speakers on AP; within an Asian population, AP depended on the degree of fluency in speaking a tone language. Rather than an ethnicity effect, this could be considered as a general transfer effect of tone language experience on absolute pitch. Schellenberg and Trehub (2008) examined the association between Asian ethnicity and pitch memory, which is considered to be an important determinant of absolute pitch. They tested non-musicians in North America of Asian or European heritage, and found no difference in pitch memory between these two groups, leading them to reject an effect of ethnicity on pitch memory abilities.

The conflicting results in the existing literature could be caused by several factors that vary across studies, as we will explain in detail in the following sections.

Use of control tasks. In many studies, there has been a lack of control between tone-language and non-tone-language groups in terms of factors that might affect participants' performance, such as memory and intelligence. Some studies did try to control for these factors (Asaridou et al., 2015; Bidelman et al., 2013; Hove et al., 2010; Hutka et al., 2015; Pfordresher & Brown, 2009). For example, Hove et al. (2010) had a relative rhythm control task to rule out possible motivational, memory, or general cognitive differences, and Bidelman et al. (2013) tested working memory and general nonverbal intelligence. However, most previous studies have not controlled for those factors in their comparisons between language groups (e.g., Bent et al., 2006; Chandrasekaran et al., 2009; Deutsch et al., 2009; Giuliano et al., 2011; Ngo et al., 2016; Stevens et al., 2004). One version of this concern, with respect to the possibility of an ethnicity effect, is a failure to consider the family history of the native speakers of American English versus the native Mandarin or Korean speakers. For example, a population that is willing and able to move to a new country that is far away may on average differ from other groups in terms of motivation or ability. When these groups are compared in studies conducted in an English-speaking country, the groups are not well matched. This can lead to a selection bias in the populations (e.g., in motivation level).

Differences in stimuli and tasks. As Bent et al. (2006) pointed out, the kinds of stimuli and tasks used may have contributed to the different results. Some studies created pure musical sounds (Bidelman et al., 2011; Deutsch et al., 2009; Hove et al., 2010), while others created sounds that emulate lexical pitches (Bent et al., 2006; Chandrasekaran et al., 2009). Researchers have used different processes to create speech-like stimuli. For example, Stevens et al. (2004) made stimuli that were similar to Thai syllables in duration and pitch range. Alexander et al. (2008) created sequences of short musical notes, imitating rapidly changing lexical pitches in Mandarin. However, these sounds were composed of discrete pitches and thus failed to capture the characteristics of contour tones (such as Mandarin lexical tones 2, 3, and 4) in speech.

In terms of different tasks, studies have used pitch discrimination tasks (i.e., comparing the pitch difference of two sounds), pitch interval size discrimination (i.e. comparing the pitch interval difference of two sounds), pitch identification tasks, pitch contour tasks (i.e., judgments made on pitch direction changes), and absolute pitch tasks (i.e., recognizing the pitch of one sound without a reference note), with the absolute pitch tasks typically used with musicians. Other studies have also used tasks specifically designed to test music pitch processing abilities, such as the Music ear test (MET) and the Montreal battery of evaluation of amusia (MBEA) (Bradley, 2012; Chen et al., 2016; Ngo et al., 2016). In a study in which the tasks involved both tone discrimination and identification, Mandarin non-musicians showed better performance on a melodic identification task, but poorer performance on a melodic discrimination task compared to English non-musicians because of categorical perception of Mandarin tonal patterns (Chang et al., 2016). This indicates a potential influence of task type on the direction and magnitude of transfer effects.

Musicianship and bilingualism. In addition to between-study differences in stimuli and tasks, differences in the tested populations may also have contributed to the inconsistent pattern of results in this literature. For example, an important factor that differs across studies is the musical training of the participants. Some studies had non-musicians (Bent et al., 2006; Giuliano et al., 2011; Hove et al., 2010), while others also had musicians (Bidelman et al., 2011; Chandrasekaran et al., 2009). In some cases, authors did not specify the music background of the participants (e.g., Deutsch et al., 2004b; Stevens et al., 2004). This is problematic because musicians are good at pitch perception regardless of their language experiences. Another potentially important factor is whether or not the non-tone language control group matched the tone language speakers in bilingualism. As we mentioned earlier, the benefits of tone language experience may differ depending on whether the non-tone language speakers are bilinguals or monolinguals (Giuliano et al., 2011). Unfortunately, for most of the existing literature, the control group was monolingual (e.g., Bent et al., 2006; Pfordresher & Brown, 2009); and sometimes this information is not clearly specified (e.g., Alexander et al., 2008; Bidelman et al., 2011; Chen et al., 2016; Hove et al., 2010; Stevens et al., 2004).

Tone language variations. Another issue that may contribute to variability in the findings is the type of tone language. A large proportion of the existing literature is based on Mandarin speakers (e.g., Alexander et al., 2008; Bent et al., 2006; Bidelman et al., 2011; Chang et al., 2016; Chen et al., 2016; Giuliano et al., 2011) or Cantonese speakers (e.g., Bidelman et al., 2013; Hutka et al., 2015). However, an effect observed in one tone language may not appear in another tone language. For example, on an F0 discrimination task, the superior behavioral performance among Cantonese speakers (Bidelman et al., 2013; Hutka et al., 2015) was not found for Mandarin speakers (Bidelman et al., 2011). Cantonese has a larger and thus more complicated set of tones than Mandarin. This speaks to the potential influence of specific tone language type on the transfer effects from language to non-language domain pitch perception. Therefore, caution should be taken in interpreting different findings, with appropriate consideration of any population differences in musicianship, bilingualism, and tone language type.

The current study

The current study aims to test transfer effects between language and music. Our primary focus is on a same-different pitch discrimination task using language stimuli and non-speech sounds derived from them. In addition, we also included two general music skill tasks that measured melody comparison and pitch precision abilities. Importantly, we also had cognitive ability tasks that assessed factors that might influence participants' performance, such as memory and intelligence. Participants also filled out a language and music history questionnaire to provide information about their language and musical training experiences.

We expected musicians to be better than non-musicians on speech and other pitch-related tasks. Further, if there is an advantage for tone-language speakers in a pitch discrimination task, we want to know if this is a language transfer effect or an ethnicity effect. If there is a language to music transfer effect, an additional goal is to determine whether the transfer is general or specific. In order to achieve these goals, in Experiment 1 we tested Mandarin (i.e., Asian tone-language), Korean (i.e., Asian non-tone-language), and English (i.e., non-Asian non-tone-language) speaking musicians and nonmusicians. We chose Koreans as an Asian non-tone language group because of their genetic similarity (i.e., East Asian) to Mandarin speakers and because the most relevant previous study also used Koreans as the Asian non-tone language group (Hove et al., 2010). In Experiment 2, we recruited Mandarin native speakers, non-Asian English native speakers, and Chinese Americans with different levels of tone language proficiency to further test the potential ethnicity and tone language effects.

Experiment 1

In this experiment, we test for a language to music transfer effect and/or an ethnicity effect by testing three subject groups: Mandarin, Korean, and English native speakers. In addition, we look for a music to language transfer effect by testing both musicians and non-musicians for each of the three languages above.

Method

Materials

Forty Mandarin short phrases were chosen from proverbs, Chinese ancient poems, and contemporary short sentences (Appendix A). They were recorded by a female native speaker of Mandarin Chinese in a sound attenuated booth and were stored on a PC. The sounds were digitized (16-bit) at a 48-kHz sample rate. They were then edited in Goldwave software. Each Mandarin phrase was composed of five Chinese characters, and each Chinese character was a one syllable word. The length of each Mandarin phrase was about 1.3 sec.

A second set of stimuli, which we will call F0 (fundamental frequency) sounds, was constructed from the Mandarin phrases. We used Praat software to extract the fundamental frequency contour of each Mandarin phrase (ceiling pitch: 400 Hz, floor pitch: 75 Hz). The resulting stimuli had no phonetic content, but matched the phrases in their tone patterns and duration.

A third set of stimuli was created from the speech contours by using them to energize a set of harmonics designed to have the timbre of a violin. More specifically, we created a script running in Praat to extract the F0 information from the spoken Chinese phrases. In SuperCollider, we loaded the data into a buffer and a violin sample sound (retrieved from freesound.org) into another buffer. Reading from the F0 data buffer, we mapped the buffer's output into the Warp1 function, which played the violin buffer using arguments that allowed us to change the pitch of the violin without changing the speed. The resulting stimuli had the same fundamental frequency patterns as the speech and the F0 stimuli, but sounded more musical. Due to an error in the program that presented these stimuli to the subjects the data were not usable in Experiment 1; for Experiment 2 the error was corrected, providing usable data.

In the original Mandarin phrases (and therefore, in the stimuli derived from them), each phrase had a flat tone (Tone 1) in the second, third, or fourth syllable, with each such flat tone preceded and followed by non-flat tones. We constructed a "lure" version of half of the stimuli by shifting the pitch of the flat tone up by approximately 10% (about 27 Hz). Figure 2 shows an example of this pitch change. The shifted/non-shifted pitch patterns were paired with corresponding original pitch sounds. This provided a relative pitch discrimination task in which half of the trials were "same" (no pitch change), and half were different (pitch change). We chose to manipulate a "flat tone" instead of contour changing tones (e.g., tone2, tone3, and tone4, see Figure 1) because flat tones are similar to the discrete tones in music. We chose a 10% shift of the pitch because pilot testing showed that when the original and shifted patterns were played consecutively, this degree of change produced detection rates that avoided floor and ceiling effects.

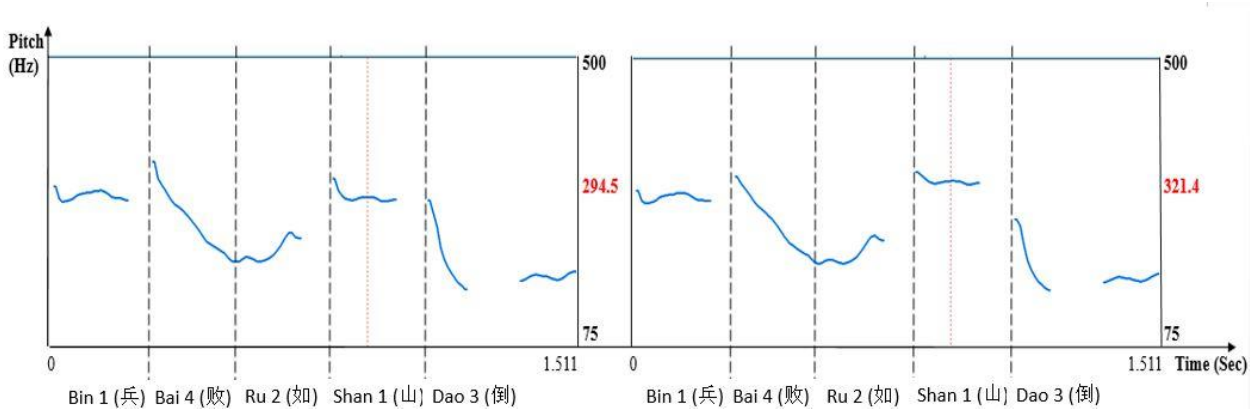

Figure 2. **The pitch manipulation procedure, as illustrated for one Mandarin phrase. The left graph is the original 5-syllable Mandarin speech (e.g., *bin 1* indicates that bin was produced with tone 1). The right graph represents the pitch changed sound. In this example, the pitch height of the flat tone in the fourth syllable was changed from 294.5 Hz to 321.4 Hz, an approximately 10% change.**

Procedures

Participants were tested individually in a sound-attenuating booth. The experiment included several tasks. On the primary pitch tasks, the participants made same-different judgments on the sounds described above. The Mandarin sounds were presented last in order to prevent Mandarin speakers from using a speech-pitch processing strategy on non-speech sounds. A similar stimulus presentation order was also used in Bent et al. (2006). Each task took about 5 minutes, and between tasks, participants were offered a break. After completing the primary pitch tasks, participants did two cognitive ability tasks and two music skills tasks that assessed abilities such as memory capacity. Each task is described in the following paragraphs:

Primary pitch tasks. For the primary pitch tasks, the participants did a same-different discrimination task. There were 80 trials (40 speech, 40 non-speech), half of which were same (20 trials) and half different (20 trials) within each of the two usable stimulus types. On each trial, a sound sequence was presented first, followed by another sequence one second later that either matched or mismatched the first one in terms of pitch pattern. Specifically, on a Different trial, the pitch of one part of the second sequence (located in the 2nd, 3rd, or 4th syllable for speech, and in the corresponding location for the F0 and music stimuli) was higher (about a 10% pitch change) than the corresponding part of the first sequence. Participants were told that they should only listen for a possible pitch difference because everything else would be the same. Participants responded by pushing one of two labeled buttons (Same or Different). A practice trial was presented first, to make sure that participants fully understood the instructions. For each stimulus type, stimuli were presented in a random order, with half of the trials consisting of a stimulus presented twice (Same), and half including the 10% pitch difference on one tone (Different).

Cognitive ability tasks. After completing the primary pitch tasks, participants did tasks that were designed to provide information about their memory and problem solving abilities.

Memory. An online version of the children's game "Simon" was used to test short-term memory (retrieved from http://www.freegames.ws/games/kidsgames/simon/mysimon.htm). For this game the display includes four fan-shaped colors forming a ring; the colors are red, green, yellow, and blue. Each round of the game consists of an increasingly long series of colors lighting up in sequence. A round begins with one color lighting up. The participant's task is to reproduce the sequence by clicking on the colored panels in the correct order. For the first display this is trivial because only one color is presented. After the participant clicks on the correct color, a sequence of two colors lighting up is presented, with the first one being a repetition of the single-item display. If the participant correctly reproduces the two-color sequence, the game moves up to three, then four, etc. A round ends when the participant fails to reproduce the correct order. The Simon game was played with the sound turned off to make this measure independent of auditory short-term memory (in the usual game, each color is associated with a particular tone). Participants played six rounds of this game, and their median final sequence length across the six rounds was used as a measure of their short-term memory abilities.

Problem solving. Participants were given nonverbal multiple choice questions to answer (see Appendix C). They were told that each question needed to be completed within 30 seconds —the PowerPoint slide would automatically move on to the next question at that point. If they finished a question in less than 30 seconds, they could press a button to go to the next question. Participants were told that they could not go back to previous questions once they moved on. There were 14 questions in total, arranged from the easiest to the hardest, based on pilot testing of the items. All questions were chosen from two free intelligence tasks on line (retrieved from http://www.iq-test.com/free-iq-test/, and http://www.quickiqtest.net/). This task was designed to provide an estimate of each person's problem-solving abilities under time pressure. The estimate was simply the number of correct answers.

Music skill tasks. We then used two on-line music tasks created by the Music and Brain Imaging lab at Harvard Medical School (retrieved from http://jakemandell.com/tonedeaf/, and http://tonometric.com/adaptivepitch/) to measure participants' general music skills.

Melody comparison. The first music task was somewhat similar to the primary pitch tasks of the current study. On each trial, participants listened to two melodies that could either be identical or slightly different. In the latter case, the second melody could differ from the first melody in one tone or several tones. Melodies were played by various instruments across trials. Participants made a “same” or “different” judgment, and were scored on the number of correct judgments.

Pitch Precision. The second music task tested participants’ pitch perception abilities using a staircase procedure. On each trial, participants listened to two short tones and had to judge whether the pitch of the second tone was higher or lower than the pitch of the first tone. The frequency separation of the two tones was increased when participants made errors, and decreased when they were accurate, until the staircase converged on a value that served as an estimate of the participant’s frequency sensitivity. Thus, for the pitch precision task, unlike the other tasks, lower values indicated better performance. The adaptive pitch procedure measured sensitivity centered on a 500 Hz tone.

Music and language questionnaire. A music and language history questionnaire was created based on the Language Experience and Proficiency Questionnaire (LEAP-Q; Marian, Blumenfeld, & Kaushanskaya, 2007) and Bradley’s language and music profile (Bradley, 2012). Appendix B shows the questionnaire.

The cognitive ability tasks and music skill tasks were used for three purposes. First, they potentially provide information about what kind of cognitive abilities support processing of a particular kind of stimulus. Second, they provide information about how to interpret any possible differences found on the primary tests. Third, the two general music skills tests provide an assessment of possible general transfer effects and/or ethnicity effects

The entire experiment, including all the tasks described above, took around 45 minutes.

Participants

Stony Brook undergraduate students with self-reported normal hearing participated in this experiment. All participants were at least 18 years old. 48 Mandarin native speakers, 48 Korean native speakers, and 48 English native speakers (non-Asians) each received one credit to fulfill a research requirement in Psychology courses, or payment of $10.

Given previous work showing transfer effects from music to language, it is important to characterize the three subject groups in terms of musical training. Prior studies have used various criteria for categorizing participants' musical abilities. Here, to be considered a musician, a participant had to have more than three years of formal musical training. For participants who played more than one musical instrument, only the instrument that corresponded to the longest training was counted. Using these criteria, we had 24 musicians and 24 non-musicians in each of the three language groups – native speakers of American English, Mandarin, or Korean. The categorization of each participant was used in the statistical analyses using a factor of "musicianship" (musicians versus non-musicians). The number of years of musical training was matched for the three language groups, both for non-musicians (Mandarin, M=1.04, SD=1.27; Korean, M=1.42, SD=1.21; English, M=1.04, SD=1.37) and for musicians (Mandarin, M=7.96, SD=3.59; Korean, M=7.58, SD=3.36; English, M=8.54, SD=3.24). A one-way ANOVA showed no differences across the three language groups in terms of the number of years of music training, F (2, 141) = 0.07, p=.929.

All participants knew at least two languages. The majority of native English speakers had Spanish as their second language, with others being able to speak other non-tone languages other than English. Korean and Mandarin speakers reported speaking English as their second language.

The second language proficiencies, on a scale of 0-10 (0=cannot speak, 10=native speaker), were matched across language groups both for non-musicians (Mandarin, M=7.14, SD=1.28; Korean, M=7.43, SD=1.9; English, M=7.21, SD=2.06) and for musicians (Mandarin, M=7.83, SD=1.66; Korean, M=7.92, SD=1.84; English, M=7.41, SD=2.04). A one-way ANOVA showed no differences across the three language groups in terms second language proficiency, F (2, 136) = 0.5, p=.608.

Results and Discussion

On the primary pitch tasks, participants received an equal number of stimulus pairs that were identical and ones in which a flat tone had been shifted up in frequency by around 10%. Errors on "same" trials were treated as false alarms, and errors on "different" trials were treated as misses. For each participant, these proportions were used to compute a *d'* score as the measurement of performance for each primary pitch task. If a hit rate or false alarm rate was 0 or 1, 0.01or 0.99 was used instead to calculate the d' score.

An important feature of the current study was the inclusion of tasks that assessed basic cognitive ability. As the primary pitch tasks involved remembering pairs of five-syllable sounds and making comparisons between them, short term memory could possibly affect participants' performance. In addition, we measured problem solving abilities because our participants came from various cultural and language backgrounds. If intelligence affects participants' performance on pitch-related tasks, it is important to have some measure of this. To look for cognitive functions that might support performance on the primary pitch tasks and the two music skill tasks, and to examine possible similarities among task requirements, Pearson correlations (two-tailed) were conducted. Given the substantial literature showing that musical training can affect performance on a wide range of abilities, we conducted these correlations separately for the musicians and the non-musicians.

Table 1 and Table 2 show that for both musicians and non-musicians, participants' performance on the Mandarin speech sounds was significantly correlated with that on F0 sounds. This strong relationship between the two primary pitch tasks confirmed that the two tasks were closely related, and that our measurements were stable enough to produce a strong correlation. Another feature shared by musicians and non-musicians was the significant correlation between the two measures of cognitive ability – the memory and problem-solving tasks. In addition, melody comparison was significantly correlated with the two primary tasks – judgments made on the Mandarin tones and F0 stimuli. These correlations presumably reflect the fact that the melody comparison task, like the Mandarin tone and F0 tasks, involves listening for pitch changes in a sound sequence. Importantly, for both musicians and non-musicians, performance on problem-solving was not correlated with either primary pitch task (or with the music skill tasks), showing that the primary pitch discrimination tasks do not depend very much on whether a person is intelligent or not. Comparing the results across Tables 1 and 2, we see that performance across tasks tends to be more correlated for the musicians (note, in particular, the patterns shown for Memory, and for Pitch Precision), consistent with the idea that years of musical training can broadly affect perceptual and cognitive processing. The different patterns for musicians and non-musicians confirm that the two groups should be treated separately. In fact, the results suggest that it may be difficult to isolate effects of tone language experience in musicians because musical training has the potential to influence performance across a wide range of tasks.

Table 1

Correlations between tasks among musicians in Experiment 1.

	F0	Speech	Memory	Problem solving	Melody comparison	Pitch precision
F0	--					
Speech	.711**	--				
Memory	.225	.284*	--			
Problem solving	.092	.109	.361**	--		
Melody comparison	.265*	.260*	.265*	.163		
Pitch precision	-.316**	.372**	-.341**	-.147	-.100	--

*$p < .05$, **$p < .01$.

Table 2

Correlations between tasks among non-musicians in Experiment 1.

	F0	Speech	Memory	Problem solving	Melody comparison	Pitch precision
F0	--					
Speech	.736**	--				
Memory	.108	.063	--			
Problem solving	.206	.135	.264**	--		
Melody comparison	.471**	.327**	.164	.098		
Pitch precision	-.066	.013	-.113	-.104	-.081	--

*$p < .05$, **$p < .01$.

Table 3 shows the average performance of Korean, Mandarin, and English speakers, broken down by musicians versus non-musicians, on the cognitive ability and music skill tasks. We conducted two-way ANOVAs with Musicianship and language Group as factors for each of the four tasks. The results revealed that musicians significantly outperformed non-musicians on the melody comparison task, F (1, 138) =5.287, p=.023, η2=.037, and on the pitch precision task, F (1, 138) =6.91, p=.01, η2=.048. This result shows that musical training is associated with music skills more generally, which is consistent with previous findings that musical training facilitates auditory acuity (Bidelman et al., 2013). It is also consistent with prior studies showing that musicians' frequency discrimination thresholds are smaller than non-musicians' (Bideman et al, 2013; Kishon-Rabin, Amir, Vexler, &Zaltz, 2001; Spiegel & Watson, 1984). Although musicians generally showed numerical advantages over their non-musician counterparts, no significant main effect of Musicianship was found for memory F (1, 138) =2.66, p=.105, η2=.019, or for problem solving F (1, 138) =.208, p=.649, η2=.002. Williamson, Baddeley, and Hitch (2010) found that musical training benefited short-term memory, but they measured auditory verbal short-term memory whereas the current study measured non-verbal memory.

Looking at the effect of language Group, there was a significant effect for melody comparison, F (2, 138) =4.81, p=.01, η2=.078, and for problem solving, F (2, 138) =5.85, p=.004, η2=.078, but not for the pitch precision or memory tasks, *p's*>.05. The means in Table 3 indicate that the effect for melody comparison is due to somewhat better performance by the Korean group than the other two groups. For problem solving, the English speakers did not do as well as the other two groups. The lack of correlation between problem solving and any of the pitch/music tasks, as shown in Tables 1 and 2, suggests that the problem-solving difference is unlikely to contaminate the pitch-relevant tasks. The fact that there were no differences across the three language groups on memory indicates that any differences on our primary pitch tasks are unlikely to be driven by confounded memory abilities across groups. The interaction between Musicianship and Group was not significant for any of the four tasks, *p's*>.05.

Table 3

Performance of musicians and non-musicians on memory, problem-solving, and two music skill tasks. Note that a lower value on the pitch precision task indicates better performance (standard deviations are shown in parentheses).

		Memory (#)	Problem-Solving (#: 0-14)	Melody comparison (% correct)	Pitch precision (Hz)
Musicians	Mandarin	8.56 (2.37)	10.46 (2)	74.53 (8.27)	9.46 (10.92)
	English	7.69 (1.67)	9.21 (1.89)	74.08 (8.3)	11.88 (9.89)
	Korean	8.42 (1.69)	10.21 (1.93)	78.35 (6.74)	8.27 (11.47)
Non-musicians	Mandarin	7.31 (2.1)	10.33 (2.2)	71.65 (9.17)	14.46 (14.07)
	English	7.92 (1.88)	8.96 (2.07)	70 (9.67)	17.19 (14.15)
	Korean	7.79 (2.3)	10.13 (1.94)	75.73 (7.5)	15.67 (18.55)

Having addressed the relationship between the cognitive abilities and music skill tasks, we return to the primary pitch tasks. A three-way ANOVA was conducted with Musicianship (musician, non-musician) and language Group (English, Mandarin, and Korean) as between-subject factors, and Sound (F0, Speech) as a within-subject factor. Figure 3 shows the overall performance of musicians and non-musicians in the three language groups on the two primary tasks. It is immediately apparent, looking at Figure 3, that musical training is a dominant factor in performance. The main effect of Musicianship was significant, F (1, 138) =17.48, p<.001, $\eta 2$=.11, with musicians consistently outperforming non-musicians. The significant main effect of Musicianship on the two primary tasks is consistent with previous research showing that musical training enhances auditory skills (Kraus & Chandrasekaran, 2010; Parbery-Clark, Skoe, & Kraus, 2009).

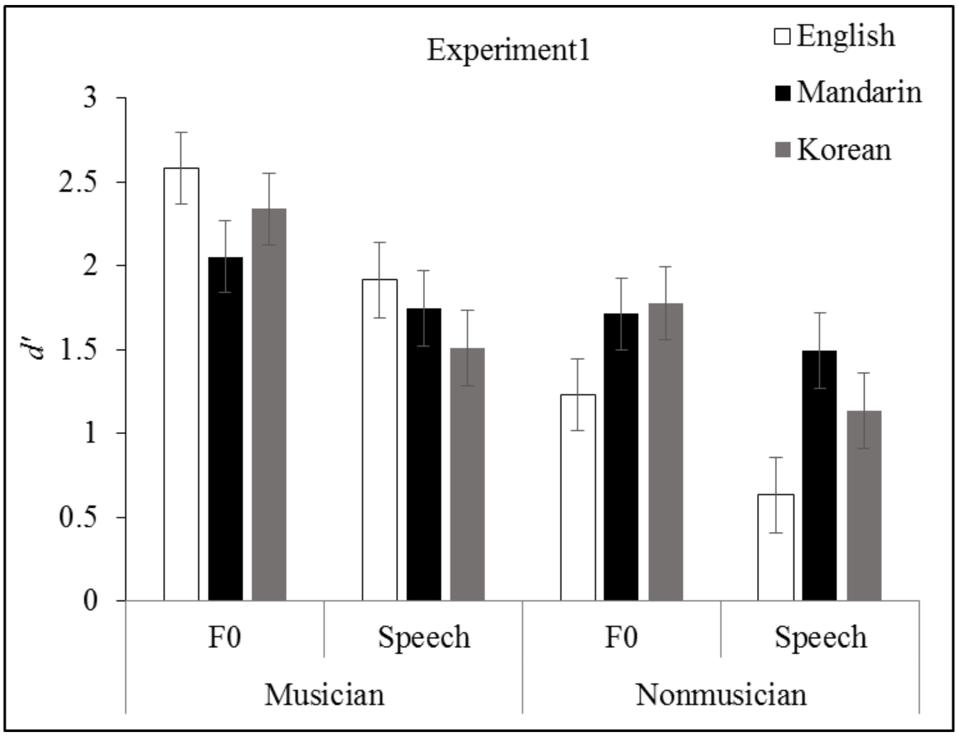

Figure 3. **The overall performance of musicians and non-musicians on F0 and speech sounds, broken down by the participant's native language. Error bars represent the standard error of the mean.**

Not only did musical training improve performance overall on the primary tasks, it also clearly moderated the effect of native language: While there was no significant main effect of language Group, F (2, 138) =.32, p=.73, $\eta 2$=.005, the interaction between language Group and

Musicianship was significant, F (2, 138) =3.61, p=.03, $\eta 2$=.05. If we consider the non-musicians' results to be an estimate of the "pure" effect of language background, we see that the English speakers lag behind the Mandarin and Korean groups. In contrast, for musicians this difference has been eliminated because of greatly improved performance by English speaking musicians. Pairwise comparisons confirmed that English speaking musicians significantly outperformed English speaking non-musicians (mean difference=1.32, p<.001), whereas there was no significant difference between musicians and non-musicians for the Korean or Mandarin speakers, both *p's*>.05. Thus, the benefit of musical training on speech and F0 sound processing was stronger for the English speakers than for the Korean and Mandarin speakers. As Figure 3 shows, the weaker musical training effect for the Mandarin and Korean speaking groups seems to be due to better performance by the non-musicians in these two groups: Korean non-musicians and Mandarin non-musicians were similar to the musicians in all three language groups on the primary tasks. Collectively, the results of the Musicianship and language Group factors support two tentative conclusions. First, the better performance by the "pure" (non-musician) Mandarin and Korean speakers than their English-speaking counterparts suggests a potential ethnicity effect. Second, the disappearance of this difference for the musicians indicates that years of musical training can provide non-tone-language speakers with the same perceptual advantage that tone-language speakers enjoy.

The main effect of Sound was significant, F (1, 138) =65.65, p<.001, $\eta 2$=.32, with better discrimination for the F0 stimuli than for the full Mandarin syllables. Pairwise comparisons showed that F0 sounds yield better performance than speech sounds among English speakers, p<.001, Mandarin speakers, p=.025, and Korean speakers, p<.001. This was expected because the removal of complex lexical context in the F0 condition can make the pitch discrimination task easier. Furthermore, the interaction between Sound and language Group was significant, F (2, 138) =4.57, p=.012, η2=.06, because the difference between the F0 and Speech stimuli was bigger for the participants who did not know Mandarin. The interaction between Sound and Musicianship was not significant, F (1, 138) =.75, p=.389, η2=.005, nor was the three-way interaction, F (2, 138) =.07, p=.931, η2=.001.

We noted above that the pattern of between-task correlation was different for musicians than for non-musicians, with musicians showing more highly correlated performance across a range of tasks than non-musicians. The results in Figure 3 also clearly demonstrate that differences on our primary tasks among the language Groups are apparent for the non-musicians, but are gone for the musicians. These finding confirm a well-established finding in the literature that musical training can have strong and broad effects on perceptual and cognitive processing. With respect to our questions of interest – are there transfer effects from language to music, and/or is there an ethnicity effect – the effect of musical training can mask any existing differences. Thus, as we examine these questions below, we will restrict our analyses to the "pure" non-musicians.

A specific transfer effect?

Using data from the non-musicians, a two-way repeated measures ANOVA was conducted on the *d'* scores for the primary pitch tasks, with Sound (speech, F0) as a within-subject factor and language Group (Mandarin, Korean, English) as a between-subject factor. There was a significant main effect of Sound (F (1, 69) = 30.53, p <.001, η2=.31) and Group (F (2, 69) =3.4, p=.039, η2=.09). The interaction between Group and Sound was not significant (F (2, 69) =2.32, p =.106, η2=.06).

To index the differential performance of the three language groups on the two primary tasks (shown in the right side of Figure 3), we calculated Cohen's *d*. For F0 sounds, the difference between Korean and English speakers was medium-large, *d*=.54, as was the difference between Mandarin and English speakers, *d* =.48. In contrast, the difference between Mandarin and Korean speakers was small, *d* =.06. For speech sounds, the difference between Korean and English speakers was medium-large, *d* =.57, the difference between Mandarin and English speakers was large, *d* =.85, and the difference between Mandarin and Korean speakers was medium, *d* =.35. The results suggest a native language advantage for Mandarin speakers for the Speech stimuli, and a small specific transfer effect for the F0 stimuli, compared to the English speakers. The advantage of Korean speakers over English speakers, and Korean speakers performing similarly to Mandarin speakers, suggests a potential ethnicity effect.

A general transfer effect?

Performance on the melody comparison and the pitch precision tasks can provide information about whether there was a general transfer effect from tone-language experience to these musical tasks. For the non-musicians, the difference between Korean and English speakers was large, *d* =.66, the difference between Mandarin and English speakers was small, *d* =.17, and the difference between Mandarin and Korean speakers was medium-large, *d* = .49. On the pitch precision task, the difference between Korean and English speakers was small, *d*=.09, the difference between Mandarin and English speakers was small, *d* =.19, and the difference between Mandarin and Korean speakers was also quite small, *d* =.07. As no clear advantage was found for Mandarin speakers over the other two language groups on the two general music skill tasks, there is no strong evidence to support general transfer effects from language to music. The advantage of Korean speakers over English speakers on the melody comparison task again provides some preliminary evidence for an ethnicity effect.

An ethnicity effect?

Recall that Hove et al. (2010) suggested that the tone language advantage shown for music processing could potentially be an ethnicity effect, rather than a language transfer effect. To examine this issue, the Korean and Mandarin speakers were combined to form an Asian group, with the English speakers constituting a non-Asian group. Looking first at our two primary tasks, a two way repeated measures ANOVA (Sound × Ethnicity) showed a main effect of Ethnicity, F (1, 70) = 6.57, p=.013, η2=.09, with no interaction of Sound and Ethnicity, F (1, 70) = .74, p=.392, η2=.01. The effect size of Ethnicity for speech sounds was d =0.72; for F0 sounds, it was d=0.51. However, neither the melody comparison (F (1, 70) = 2.73) nor the pitch precision task (F (1, 70) = 0.3) produced any evidence for an ethnicity effect, both *p's* >.05.

It might be argued that the possible ethnicity effect on the primary tasks was mainly due to the characteristics of the Korean group. Koreans are geographically and culturally closer to Chinese than to Americans, and they therefore might have more exposure to Chinese speakers. In the current study, self-report information on questionnaires showed that none of the Korean speakers had any experience of living in China, learning Mandarin, or being able to speak or understand any tone languages. A potentially more substantial concern is that some Korean dialects use pitch accent, such as Northern Kyungsang Korean (Jun, Kim, Lee, & Jun, 2006). Even though pitch accent languages impose weaker pitch processing demands than true tone languages, such experience could potentially have some effect on the pitch discrimination task being tested. This does not appear to be an issue in the current study because most of the Korean non-musicians (17 of the 24) reported that they came from Seoul. Five participants did not provide specific information on their Korean dialects, and only two reported that they were from the regions where the dialects have pitch accents (Daejeon and Busan, respectively). Therefore, it is unlikely that experience with a pitch accent language produced the good performance by Korean speakers.

Another possible concern is that since most of our Mandarin and Korean native speakers were foreigners, population selection might play a role in explaining any differences observed between them and the native English speakers who were born in the US. Almost 90% of the Mandarin and Korean non-musicians reported that they had come to the US in recent years. A potential population selection bias extends to all studies in which tone language speakers are tested in Western countries, as the subjects are those who were willing and able to move across the globe. To control for a possible effect of population selection bias, in Experiment 2 we recruited Asian participants who were all from the United States.

Experiment 2

Experiment 2 was designed to test the potential ethnicity effect and/or tone language effect with better control for population selection. Consistent with the very potent effect of musicianship that we found in Experiment 1, we screened potential participants on their formal musical training and only included participants with <= 3years of musical training. We had four groups of participants: Mandarin native speakers (Asian tone-language), English native speakers (non-Asian non-tone language), Chinese Americans who speak tone languages (Asian tone-language), Chinese Americans who had limited experience with any tone language (Asian non-tone language). Critically, the participants in the two Chinese American groups, though ethnically Chinese, all were US-born native English speakers. These participants had upbringing experiences that were similar to those of the English speaking (non-Asian) group, allowing us to tease apart tone language experience and ethnicity effects using the three American groups. If an ethnicity effect exists, then the performance of the two Chinese American groups should be similar, and they should outperform the English speaking non-Asian group. If there is a tone language effect, then the Chinese Americans who speak tone languages should outperform the English speaking group and the Chinese American non-tone-language group.

Method

Participants

One hundred twenty students from Stony Brook University participated in this experiment. They all had self-reported normal hearing and vision and were at least 18 years old. Each participant received one credit to fulfill a research requirement in Psychology courses, or payment of $10. There were 30 Mandarin native speakers (14 male, 16 female; all right-handed; mean age: 23.47), 30 English native speakers (all non-Asians, 20 male, 10 female; 4 left-handed; mean age: 21.67), 30 Chinese American tone language speakers (14 male, 16 female; 1 left-handed, mean age: 19.7), and 30 Chinese Americans non-tone language speakers (23 male, 7 female; 2 left-handed; mean age: 20.23).

The number of years of musical training was matched for the four groups (Mandarin, M=1.1, SD=1.32; English, M=1.53, SD=1.66; Chinese American_Nontone, M=0.9, SD=1.21, Chinese American_Tone, M=1.7, SD=1.66). A one-way ANOVA showed no differences among the four groups in terms of the number of years of music training, F (3, 116) = 1.9, p=.134.

All Chinese Americans spoke English as their native language. The Chinese American non-tone group knew little or no Mandarin (or any other tone language), but all spoke non-tone second languages (e.g., Spanish, French.). Specifically, on a ten-point scale, their average self-rated speaking ability in any tone language was 2.38; for understanding, it was 3.28. In comparison, in the Chinese American tone language group, participants spoke second languages that were all tone languages (i.e., Mandarin, Cantonese, Fuzhounese, etc.). Their average self-rated speaking ability was 7.41 and their average understanding ability was 7.79 in any tone language. Of all the Chinese Americans (30 tone-language, and 30 non-tone-language), three were adopted as babies by American families, and the rest were second or third generation immigrants.

Materials and Procedure

The materials were the same as Experiment 1, except that the questionnaire included additional questions about the participants' tone language background (see Appendix B). We fixed the error in the program that ran the music timbre condition, so that the data from this condition would be usable. The procedure was the same as that in Experiment 1.

Results and Discussion

As in Experiment 1, we first examined the correlations among the different tasks that the participants did (see Table 4). These correlations were strikingly similar to those of the non-musicians in Experiment 1 (compare Table 4 to Table 2, skipping the first column, which pertains to the Music Timbre stimuli that are not in Table 2). Recall that this pattern systematically differed from the correlation pattern of the musicians (see Table 1). The results again indicate that perceptual processing in musicians is different than such processing in non-musicians. Table 4 also shows that the Musical Timbre stimuli produced a correlation pattern very similar to that of the F0 stimuli that were used to produce the Musical Timbre items.

Table 4

Correlations between tasks in Experiment 2.

	Music	F0	Speech	Memory	Problem solving	Melody comparison	Pitch precision
Music	--						
F0	.672**	--					
Speech	.644**	.741**	--				
Memory	.139	.108	.071	--			
Problem solving	.045	.101	.047	.309**	--		
Melody comparison	.426**	.420**	.307**	.079	.141		
Pitch precision	-.121	-.18**	-.145	-.089	-.017	-.210*	--

*p < .05, **p < .01.

Participants' performance on the cognitive abilities tasks and the two general music tasks are presented in Table 5. One-way ANOVAs showed that there were no differences among the four language groups on the memory, pitch precision, and melody comparison tasks, all p*'s*>.05. There was a significant difference on the problem-solving task, F (3, 116) =4.96, p=.003, and as in Experiment 1 this appears to be due to lower performance by the English (non-Asian) group. Also as in Experiment 1, there was no significant correlation between the problem solving and the primary pitch tasks or the music skill tasks, indicating that the difference in problem solving ability is unlikely to contaminate participants' performance on pitch-related tasks.

Table 5

Participants' performance on memory, problem-solving, and two music skill tasks. Note that a lower value on the pitch precision task indicates better performance (standard deviations are shown in parentheses).

	Memory (#)	Problem-Solving (#: 0-14)	Melody comparison (% correct)	Pitch precision (Hz)
Mandarin	8.65 (2.1)	10.23 (2.31)	72.11 (10.01)	17.07 (17.57)
English	7.65 (2.16)	8.6 (2.43)	68.34 (8.81)	13.68 (12.75)
Chinese American_Nontone	7.92 (2.15)	10.5 (1.68)	71.66 (6.5)	16.79 (14.25)
Chinese American_Tone	8.55 (1.55)	10.07 (1.87)	70.7 (7.33)	14.82 (12.2)

We begin our analysis of the main results with the portion of Experiment 2 that overlaps with Experiment 1: performance on the F0 and Speech sounds by the Mandarin and (non-Asian) English speakers. We conducted a three-way ANOVA using data from non-musicians only: language Group (English, Mandarin) × Sound (F0, Speech) × Experiment (1 vs. 2). Figure 4 shows the *d'* scores for the Mandarin and English speaking groups on the tone-shift judgment in Mandarin phrases and in the only-F0 versions. The main effect of Sound was significant, F (1, 104) =24.89, p<.001, η2=.19, as was the main effect of language Group, F (1, 104) =5.47, p=.021, η2=.05. The main effect of Experiment was not significant, F (1, 104) =.36, p=.549, η2=.003, nor was its interaction with Sound (F (1, 104) =0.65, p=.422, η2=.006), Group (F (1, 104) =1.88, p=.173, η2=.018), or both (F (1, 104) =0.4, p=.529, η2=.004). As Figure 4 shows, the differences between Mandarin and English speakers were smaller in Experiment 2 than in Experiment 1, but the pattern is the same.

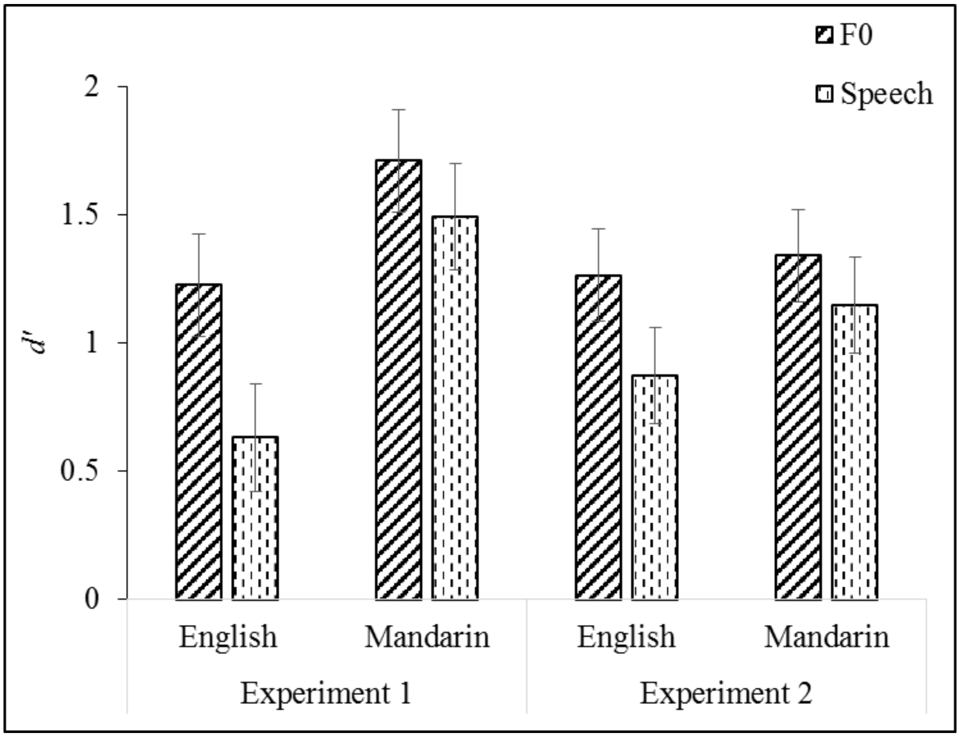

Figure 4. **The performance of Mandarin and English non-musicians on F0 and speech sounds in Experiments 1 and 2. Error bars represent the standard error of the mean.**

Having established that for the overlapping portions of the two experiments the results are similar, we turn to the full data set of Experiment 2. A two-way ANOVA was conducted with language Group (English, Mandarin, Chinese American_tone, and Chinese American_nontone) as a between-subject factor and Sound (Music, F0, and Speech) as a within-subject factor. Table 6 shows the means and standard deviations of the four groups on the primary tasks. The main effect of Sound was significant, F (2, 232) =28.42, p<.001, η2=19.7. Participants were significantly better at discriminating pitch differences in Music and in F0 sounds, compared to Speech sounds, *p's* <.001. The difference between F0 and Music sounds was marginal, p=.067. As we noted in our discussion of the correlations, the Music and F0 sounds were both non-speech sounds without complex lexical context, with the Music stimuli constructed from the F0 sounds. The main effect of Group was not significant, F (3, 116) =1.19, p=.316, η2=.03. The interaction between Sound and Group was non-significant, F (6, 232) =.51, p=.801, η2=.013, suggesting that how well the participants did on the three primary tasks was not influenced by what group they belong to. As we discuss next, this pattern is largely driven by the similarity in performance for all of the native English speakers.

Table 6

Participants' performance on the three primary pitch tasks (standard deviations are shown in parentheses).

	Music	F0	Speech
Mandarin	1.67 (.82)	1.34 (1.06)	1.15 (1.11)
English	1.39 (1.03)	1.27 (.84)	.87 (0.97)
Chinese American_Nontone	1.27 (.68)	1.24 (.99)	.77 (.79)
Chinese American_Tone	1.22 (.97)	1.04 (1.15)	.7 (.97)

The core purpose of Experiment 2 was to test participants who are all native English speakers, born and raised in the US, who vary in ethnicity (Asian versus non-Asian) and tone language experience. This combination allows us to control for a possible population selection bias. Thus, we focus on the two Chinese American groups and their fellow (non-Asian) native English speakers to test ethnicity and tone language effects. Figure 5 shows the *d'* scores of these participants on the Music, F0, and Speech sounds. As Figure 5 shows, performance was very similar for the three groups, for all three types of stimuli. All of the effect sizes for English vs. Chinese American_tone, English vs. Chinese American_nontone, and Chinese American_tone vs Chinese American_nontone were quite small: .16, .13, and .06 for the Music condition, .22, .03, and .18 for the F0 condition, and .18, .12, and .08 for the Speech condition. A post-hoc power analysis (one-tailed) showed if there was a medium effect size (*d* =0.5, α=.05) between the non-Asian English speaking group and the Chinese American non-tone language group, the study would have a 61% chance of finding it.

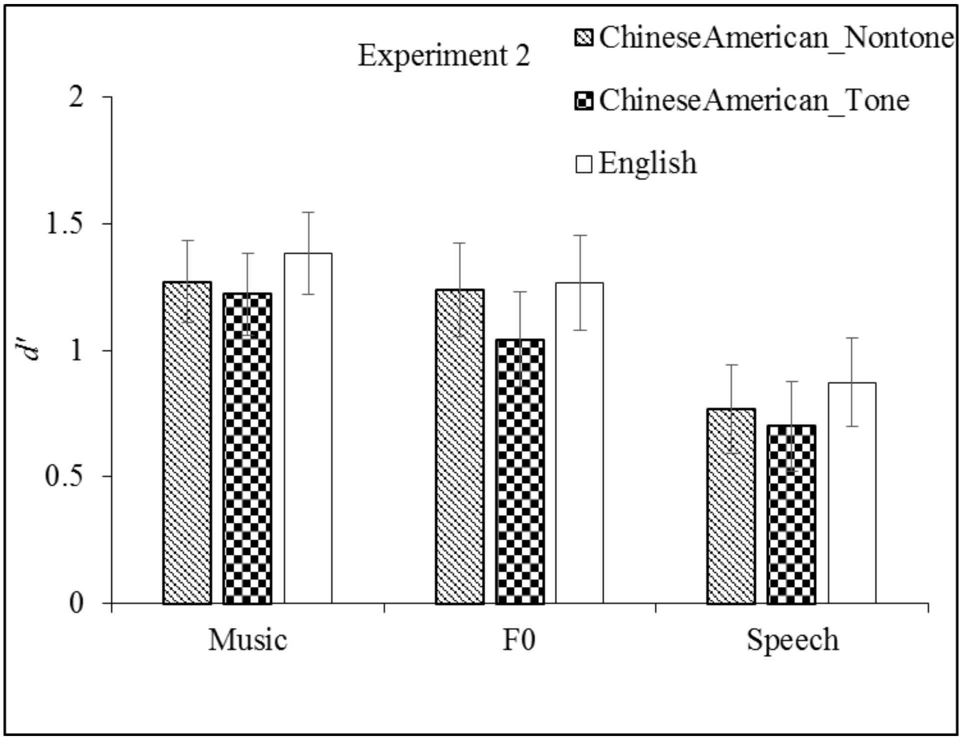

***Figure 5.* The performance of the three language groups on Music, F0, and speech sounds. Error bars represent the standard error of the mean.**

On the two music skills tasks, the effects sizes were slightly larger, but still relatively small. Specifically, the English vs. Chinese American_tone, English vs. Chinese American_nontone, and Chinese American_tone vs Chinese American_nontone comparisons yielded effect sizes were .29, .43, and .14 for the melody comparison task, and.09, .23, and .15 for the pitch precision task. Taken together with the results above, when population selection was carefully controlled, ethnicity and tone language experience did not play a measurable role in relative pitch discrimination or in general music skills. Given the null effect of ethnicity, the observed advantage of native Mandarin speakers (and, in some cases, native Korean speakers) over English speakers found in Experiment 1 is potentially related to a population selection bias.

General Discussion and Conclusions

In the current study, we examined bidirectional transfer effects between language and music, and how ethnicity might play a role in these effects. We investigated the music to language transfer effect by comparing musicians and non-musicians on discriminating pitch differences in pairs of speech sequences and pairs of F0 sequences. The language to music transfer effect was tested by comparing tone and non-tone language speakers. We examined whether there was a specific transfer effect, a general transfer effect, and/or an ethnicity effect. It is worth noting that the three effect types do not necessarily exclude each other.

Consistent with previous studies, Experiment 1 provided strong evidence for a music to language transfer effect. Interestingly, the effect of musicianship was more pronounced among English speakers than among Korean or Mandarin speakers, with musical training eliminating a disadvantage that English non-musicians had relative to Mandarin and Korean non-musicians. Experiment 1 also showed a small advantage of Mandarin speakers over English speakers on the F0 and Speech sounds, suggesting a domain-specific transfer effect. In addition, Mandarin non-musicians showed comparable performance to the Korean and English non-musicians on the two music skill tasks, providing no support for a general language to music transfer effect. Our results are at odds with Bidelman et al. (2013), who reported a general transfer effect from language to music by showing that tone language speaking non-musicians showed better memory and auditory pitch acuity compared to English speaking non-musicians. It is unclear, though, whether the superior performance in auditory pitch tasks was driven by cognitive differences, population selection bias, or other factors.

The advantage of the two Asian groups on the primary tasks in Experiment 1 could potentially be due to an ethnicity effect or to a population selection effect, as the Korean and Mandarin speakers were almost all foreigners or late immigrants. In Experiment 2, we changed from Korean speakers to Chinese Americans who had limited experience with any tone language, in order to control for a possible population selection bias. This switch provides a cleaner test for the existence and robustness of an ethnicity effect. In addition, we also recruited Chinese Americans who were raised in the US but spoke tone languages as second languages to investigate a tone language effect. The Chinese American participants were raised in an environment similar to that for the non-Asian English native speakers. A similar approach was taken by Schellenberg and Trehub (2008) in a study in which they compared Asian-heritage to European Americans on a pitch memory task. However, their Asian-heritage group contained an unknown mixture of people who spoke or did not speak tone languages, potentially confounding the tone language and ethnicity factors. With this caveat, they found no clear Asian advantage. In the current study, the Chinese American non-tone language group and non-Asian Americans did not differ on any tasks.

Dediu and Ladd (2007) suggested that tone language use and genetic profiles are intertwined. If that is the case, there should be an advantage even among those of Asian heritage who did not speak any tone language. The results of Experiment 2 clearly argue against an ethnicity effect, and are at odds with Hove et al.'s (2010) report of an ethnicity effect on pitch perception. It is possible that the conflicting findings are due to task differences, as Hove et al. (2010) used a pitch identification task whereas we used a pitch discrimination task. In the current study, Asian Americans were comparable to non-Asian Americans, indicating that ethnicity and tone language use are separable.

The current study highlights two methodological issues in the investigation of language-to-music transfer effects: the importance of musical training and of population selection. Experiment 1 clearly demonstrated that musicians and non-musicians performed very differently, and the correlation patterns among tasks were also dissimilar for musicians versus non-musicians. The interaction between the participants' musical training background and language groups showed that musical training can eliminate existing differences between speakers of tone language and non-tone languages. Collectively, our results make it clear that musicians and non-musicians should be treated differently in investigations of a possible tone language effect. More specifically, musical training can mask these effects. Some previous studies have paid attention to participants' musical training background (Bent el al., 2006; Giuliano et al., 2011; Hove et al., 2010), but others have not (Deutsch et al., 2004b; Stevens et al., 2004).

The second methodological issue that the current study raises is a concern in the selection of populations to be tested. Participants who speak tone languages, when tested in the West, often have quite different personal histories than "typical" Westerners. This can easily create a population selection bias, potentially affecting performance. This is an inherently difficult problem that may only partially be offset by control tasks. We included such control tasks in our two experiments, and in addition tried to address the issue by recruiting participants with similar backgrounds in Experiment 2.

We believe that the current study provides new insights into language-music transfer effects. Tone language experience facilitated domain-specific pitch discrimination, but not general pitch processing skill. Musical training was associated with improved performance on pitch discrimination, both within speech sequences and within F0 tone sequences, and on the melody comparison and pitch precision tasks. By using a carefully selected population of ethnically Chinese Americans, we found that ethnicity (i.e., being Asian) did not enhance linguistic pitch perception or general music skills. Our results expand and clarify the understanding of the relationship between language and music.

Acknowledgements

We thank Susan Brennan for her valuable feedback on this project. We also appreciate the constructive suggestions of three anonymous reviewers. Support was provided by Ministerio de Ciencia E Innovacion, Grant PSI2014-53277, Centro de Excelencia Severo Ochoa, Grant SEV-2015-0490, and by the National Science Foundation under Grant IBSS-1519908.

References

Alexander, J. A., Bradlow, A. R., Ashley, R. D., & Wong, P. C. (2008). Music-melody perception in tone-language and non-tone-language speakers. In *Poster presented at the 156th Meeting of Acoustical Society of America, Miami, FL*.

Amir, O., Amir, N., & Kishon-Rabin, L. (2003). The effect of superior auditory skills on vocal accuracy. *The Journal of the Acoustical Society of America*, *113*(2), 1102-1108.

Asaridou, S. S., & McQueen, J. M. (2013). Speech and music shape the listening brain: evidence for shared domain-general mechanisms.*Frontiers in psychology*, *4*, 321.

Asaridou, S. S., Hagoort, P., & McQueen, J. M. (2015). Effects of Early Bilingual Experience with a Tone and a Non-Tone Language on Speech-Music Integration. PloS one, 10(12), e0144225.

Assal, G. (1973). Aphasie de Wernicke chez un pianist (Wernicke aphasia in a pianist). *Revue Neurologique*, *29*, 251-255.

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

Bidelman, G. M., Gandour, J. T., & Krishnan, A. (2011). Musicians and tone-language speakers share enhanced brainstem encoding but not perceptual benefits for musical pitch. *Brain and Cognition*, *77*(1), 1-10.

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.

Bradley, E. D. (2012). *Crosslinguistic Perception of Pitch in Language and Music* (Doctoral dissertation, University of Delaware).

Bradley, E. D. (2013). Pitch perception in lexical tone and melody. *Reviews of Research in Human Learning and Music*.

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

Chang, D., Hedberg, N., & Wang, Y. (2016). Effects of musical and linguistic experience on categorization of lexical and melodic tones. The Journal of the Acoustical Society of America, 139(5), 2432-2447.

Cheek, J. M., & Smith, L. R. (1999). Music training and mathematics achievement. *Adolescence*, *34*(136), 759-761.

Chen, A., Liu, L., & Kager, R. (2016). Cross-domain correlation in pitch perception, the influence of native language. Language, Cognition and Neuroscience, 31(6), 751-760.

Dediu, D., & Ladd, D. R. (2007). Linguistic tone is related to the population frequency of the adaptive haplogroups of two brain size genes, ASPM and Microcephalin. *Proceedings of the National Academy of Sciences*, *104*(26), 10944-10949.

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

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

Deutsch, D., Henthorn, T., & Dolson, M. (2004b). Speech patterns heard early in life influence later perception of the tritone paradox. *Music Perception: An Interdisciplinary Journal*, *21*(3), 357-372.

Giuliano, R. J., Pfordresher, P. Q., Stanley, E. M., Narayana, S., & Wicha, N. Y. (2011). Native experience with a tone language enhances pitch discrimination and the timing of neural responses to pitch change. *The relationship between music and language*, *14*.

Gregersen, P. K., Kowalsky, E., & Li, W. (2007). Reply to Henthorn and Deutsch: Ethnicity versus early environment: Comment on 'Early Childhood Music Education and Predisposition to Absolute Pitch: Teasing Apart Genes and Environment'by Peter K. Gregersen, Elena Kowalsky, Nina Kohn, and Elizabeth West Marvin [2000]. *American Journal of Medical Genetics Part A*, *143*(1), 104-105.

Gregersen, P. K., Kowalsky, E., Kohn, N., & Marvin, E. W. (2001). Early childhood music education and predisposition to absolute pitch: Teasing apart genes and environment. *American journal of medical genetics*,*98*(3), 280-282.

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

Hutka, S., Bidelman, G. M., & Moreno, S. (2015). Pitch expertise is not created equal: cross-domain effects of musicianship and tone language experience on neural and behavioural discrimination of speech and music. Neuropsychologia, 71, 52-63.

Jun, J., Kim, J., Lee, H., & Jun, S. A. (2006). The prosodic structure and pitch accent of Northern Kyungsang Korean. *Journal of East Asian Linguistics*, *15*(4), 289-317.

Kishon-Rabin, L., Amir, O., Vexler, Y., & Zaltz, Y. (2001). Pitch discrimination: Are professional musicians better than non-musicians?. *Journal of basic and clinical physiology and pharmacology*, *12*(2), 125-144.

Koelsch, S., Gunter, T. C., Cramon, D. Y. V., Zysset, S., Lohmann, G., & Friederici, A. D. (2002). Bach speaks: a cortical "language-network" serves the processing of music. *Neuroimage*, *17*(2), 956-966.

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

Marian, V., Blumenfeld, H. K., & Kaushanskaya, M. (2007). The Language Experience and Proficiency Questionnaire (LEAP-Q): Assessing language profiles in bilinguals and multilinguals. *Journal of Speech, Language, and Hearing Research*, *50*(4), 940-967.

Ngo, M. K., Vu, K. P. L., & Strybel, T. Z. (2016). Effects of Music and Tonal Language Experience on Relative Pitch Performance. *The American Journal of Psychology*, *129*(2), 125-134.

Nilsonne, Å., & Sundberg, J. (1985). Differences in ability of musicians and nonmusicians to judge emotional state from the fundamental frequency of voice samples. *Music Perception*, 507-516.

Ota, M. (2003). The development of lexical pitch accent systems: An autosegmental analysis. *The Canadian Journal of Linguistics/La revue canadienne de linguistique*, *48*(2), 357-383.

Parbery-Clark, A., Skoe, E., & Kraus, N. (2009). Musical experience limits the degradative effects of background noise on the neural processing of sound. *J.* Neurosci. 29, 14100–14107.

Pastuszek-Lipińska, B. (2008). The influence of music education and training on SLA. *Experimental Linguistics ExLing, 2008*, 185.

Pastuszek-Lipińska, B. (2007). Musicians outperform nonmusicians in a study with shadowing speech. *Proceedings of ICPhS XVI, Saarbrücken*,*6*(10), 821-824.

Patel, A. D. (2012). The OPERA hypothesis: assumptions and clarifications. *Annals of the New York Academy of Sciences*, *1252*(1), 124-128.

Patel, A. D. (2003). Language, music, syntax and the brain. *Nature neuroscience*, *6*(7), 674-681.

Patel, A. D. (2010). Music, language, and the brain: Oxford University Press.

Patel, A. D., Gibson, E., Ratner, J., Besson, M., & Holcomb, P. J. (1998). Processing syntactic relations in language and music: An event-related potential study. *Journal of cognitive neuroscience*, *10*(6), 717-733.

Peretz, I., Champod, A. S., & Hyde, K. (2003). Varieties of musical disorders. *Annals of the New York Academy of Sciences*, *999*(1), 58-75.

Peretz, I., Kolinsky, R., Tramo, M., Labreque, R., Hublet, C., Demeurisse, G., & Belleville, S. (1994). Functional dissociation following bilateral lesions of auditory cortex. *Brain*, *117*, 1283–1302.

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

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

Schön, D., Gordon, R., Campagne, A., Magne, C., Astésano, C., Anton, J. L., & Besson, M. (2010). Similar cerebral networks in language, music and song perception. *Neuroimage*, *51*(1), 450-461.

Spiegel, M. F., & Watson, C. S. (1984). Performance on frequency-discrimination tasks by musicians and nonmusicians. *The Journal of the Acoustical Society of America*, *76*(6), 1690-1695.

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

Thompson, W. F., Schellenberg, E. G., & Husain, G. (2001). Arousal, mood, and the Mozart effect. *Psychological Science*, *12*(3), 248-251.

Thompson, W. F., Schellenberg, E. G., & Husain, G. (2003). Perceiving prosody in speech: Effect of music lessons. *Annals of the New York Academy of Sciences*, *999*(1), 530-532.

Thompson, W. F., Schellenberg, E. G., & Husain, G. (2004). Decoding speech prosody: do music lessons help? *Emotion*, *4*(1), 46.

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

Williamson, V. J., Baddeley, A. D., & Hitch, G. J. (2010). Musicians' and nonmusicians' short-term memory for verbal and musical sequences: Comparing phonological similarity and pitch proximity. *Memory & Cognition*, *38*(2), 163-175.

Zatorre, R. J. (2003). Absolute pitch: A model for understanding the influence of genes and development on neural and cognitive function.*Nature neuroscience*, *6*(7), 692-695.