Tone language experience modulates the effect of long-term musical training on musical pitch perception

Xiuli Tong, William Choi, and Yuen Ying Man

Citation: The Journal of the Acoustical Society of America **144**, 690 (2018); doi: 10.1121/1.5049365
View online: https://doi.org/10.1121/1.5049365
View Table of Contents: https://asa.scitation.org/toc/jas/144/2
Published by the Acoustical Society of America

Tone language experience modulates the effect of long-term musical training on musical pitch perception

Xiuli Tong[a] and William Choi
Division of Speech and Hearing Sciences, The University of Hong Kong, Hong Kong

Yuen Ying Man
Caritas—Hong Kong, Hong Kong

(Received 10 November 2017; revised 21 June 2018; accepted 19 July 2018; published online 10 August 2018)

Long-term musical training is widely reported to enhance music pitch perception. However, it remains unclear whether tone language experience influences the effect of long-term musical training on musical pitch perception. The present study addressed this question by testing 30 Cantonese and 30 non-tonal language speakers, each divided equally into musician and non-musician groups, on pitch height and pitch interval discrimination. Musicians outperformed non-musicians among non-tonal language speakers, but not among Cantonese speakers on the pitch height discrimination task. However, musicians outperformed non-musicians among Cantonese speakers, but not among non-tonal language speakers on the pitch interval discrimination task. These results suggest that the effect of long-term musical training on musical pitch perception is shaped by tone language experience and varies across different pitch perception tasks.
https://doi.org/10.1121/1.5049365

[TS] Pages: 690–697

I. INTRODUCTION

Two decades of empirical research suggests that long-term musical training enhances listeners' musical pitch acuity, such as the ability to perceive pitch height (e.g., Herholz and Zatorre, 2012; Kishon-Rabin *et al.*, 2001; Tervaniemi *et al.*, 2005) and pitch interval (e.g., Zarate *et al.*, 2012). However, these previous studies have largely focused on English and other non-tonal language speakers (e.g., French and Finnish) whose native languages involve little use of pitch information in spoken word recognition. In contrast, tone language speakers, such as Cantonese, rely heavily on the variation of pitch, known as lexical tones, to distinguish word meanings (e.g., Fok-Chan, 1974). For example, the Cantonese monosyllable /fu/, when pronounced with six different tones,[1] can represent six different meanings: /fu55/ 膚 (*skin*), /fu25/ 虎 (*tiger*), /fu33/ 褲 (*trousers*), /fu21/ 符 (*symbol*), /fu23/ 婦 (*woman*), and /fu22/ 父 (*father*). Emerging evidence also shows the facilitation of tone language experience on musical pitch perception (e.g., Bidelman *et al.*, 2013), implying an inter-connectivity between language and musical domains (Patel, 2011). In the present study, we extended these previous findings by examining the possible interaction between tone language experience and long-term musical training. Specifically, we compared Cantonese-speaking musicians, Cantonese-speaking non-musicians, non-tonal language-speaking musicians, and non-tonal language-speaking non-musicians on pitch height and pitch interval discrimination.

Long-term musical training induces positive changes on auditory processing (e.g., Hannon and Trainor, 2007; Swaminathan *et al.*, 2015) that are traceable to plastic changes in auditory, motor, and sensorimotor integration areas of the human brain (e.g., Hyde *et al.*, 2009; Miendlarzewska and Trost, 2013). For example, Schlaug *et al.* (1995) showed that the anterior half of the corpus callosum was larger among musicians who began their musical training before the age of 7. Similarly, Gaser and Schlaug (2003) demonstrated that musicians had more grey matter volume than non-musicians in auditory regions. In a neurophysiological study, Tervaniemi *et al.* (2005) reported greater amplitudes of N2b and P3 responses to pitch height deviations in musicians relative to non-musicians, reflecting that musicians had higher neuronal sensitivity in conscious pitch discrimination and target detection (Ritter *et al.*, 1992; Näätänen, 1992). Additionally, Bidelman *et al.* (2011b) found that musicians showed enhanced brainstem frequency-following responses compared to non-musicians, suggesting that musicians were faster in neural synchronization, and had more robust brainstem encoding of pitch sequences than non-musicians.

Behaviorally, converging evidence suggests that long-term musical training enhances pitch height and pitch interval discrimination among listeners without tone language background. For example, in behavioral pitch height discrimination, Kishon-Ribon *et al.* (2001) showed that musicians demonstrated lower pitch discrimination thresholds than non-musicians by half, while Micheyl *et al.* (2006) indicated the threshold to be lower by one-sixth, suggesting that musicians were more sensitive to minute differences in pitch than non-musicians. Tervaniemi *et al.* (2005) likewise reported that musicians had significantly higher accuracy and shorter reaction time than non-musicians on a pitch detection task. In addition, the perceptual advantage associated with long-term musical training appears to extend to pitch interval discrimination as well. For example, Zarate

[a]Electronic mail: xltong@hku.hk

 0001-4966/2018/144(2)/690/8/$30.00

et al. (2012) reported that musicians demonstrated higher perceptual sensitivity to 100 cent- and 125 cent-interval differences than non-musicians. Furthermore, in the same study, the perceptual discrimination enhancement associated with increased pitch interval difference (25 cents) was shown among musicians starting at 100 cent-interval difference, but the same perceptual enhancement was evident only beyond 125 cent-interval difference among non-musicians. Collectively, these previous studies suggest that long-term musical training enhances pitch height and pitch interval discrimination among listeners without tone language background.

A growing body of theoretical and empirical research suggests that extensive tone language experience may also facilitate musical pitch perception. The Overlap, Precision, Emotion, Repetition, Attention (OPERA) hypothesis (Patel, 2011) posits a music-to-language transfer in which long-term musical training would enhance the plasticity and increase the neural precision for acoustic encoding in the subcortical circuit through emotional, repetitive, and attentional engagement with music. Moreover, music and speech processing employ overlapping subcortical networks, such that the enhanced plasticity and neural precision resulting from musical training is transferrable to the language domain. Of direct relevance to the current study is that, although the OPERA hypothesis was proposed to account for music-to-language transfer, Patel (2012) formulated the possibility of bi-directional cross-domain transfer, implying that the enhanced neural plasticity for speech encoding associated with language experience might be transferrable to the musical domain.

One music-relevant speech coding experience is the use of lexical tones for word identification, given that lexical tones are variations of pitch information, such as pitch height and pitch contour (Fok-Chan, 1974). For example, Cantonese has a rich and complex tonal inventory comprising six tones, including three level tones, i.e., high-level (55, T1), mid-level (33, T3), and low-level (22, T6), and three contour tones, i.e., high rising (25, T2), low falling (21, T4) and low rising (23, T5).[2] To illustrate, words such as /si33/ 試 (*try*) and /si22/ 事 (*event*) differ mainly in terms of relative pitch height, given their corresponding tones, i.e., mid-level and low-level. Similarly, the lexical tones corresponding to /si55/ 師 (*teacher*) and /si25/ 史 (*history*), i.e., high-level and high-rising, differ mainly in terms of pitch contour. Perceptually, a neurophysiological study has suggested that pitch contour and pitch height are critical in distinguishing tonal contrasts among native Cantonese speakers (Tsang *et al.*, 2011). Placing this into the extended view of the OPERA hypothesis (Patel, 2012), long-term experience in attending to pitch in the linguistic dimension may be transferrable to the musical domain.

Indeed, several behavioral and neurophysiological studies have reported language-to-music transfer by consistently showing that tone language experience enhances the perceptual capacity of listeners on musical pitch (e.g., Bidelman *et al.*, 2013; Hutka *et al.*, 2015; Pfordresher and Brown, 2009). For example, Bidelman *et al.* (2013) found that Cantonese-speaking non-musicians outperformed English-speaking non-musicians on pitch memory, musical melody discrimination, and pitch processing speed tasks. Similar results have been reported in a study by Pfordresher and Brown (2009), in which a mixed group of tone language-speaking (Cantonese, Mandarin, and Vietnamese) non-musicians had higher accuracy than English-speaking non-musicians on a pitch interval discrimination task. In a study by Hutka *et al.* (2015), although English-speaking musicians showed a larger mismatch negativity (MMN) amplitude in response to musical pitch changes than Cantonese-speaking non-musicians, they performed similarly on a behavioral musical pitch discrimination task. Also, Cantonese-speaking non-musicians outperformed English-speaking non-musicians the same behavioral task.

The effect of tone language experience on musical pitch perception has also been supported by neurophysiological studies. For example, Bidelman *et al.* (2011a) demonstrated that Mandarin-speaking non-musicians had higher pitch tracking accuracy than English-speaking non-musicians, reflecting that Mandarin-speaking non-musicians were better able to extract pitch contour at the brainstem level. Similarly, in an event-related brain potential (ERP) study that utilized a same/different task to assess sensitivity to pitch height of pure tones, Mandarin-speaking non-musicians detected pitch changes earlier than English-speaking non-musicians (Giuliano *et al.*, 2011).

Although there are two parallel lines of evidence suggesting long-term musical training (e.g., Bidelman *et al.*, 2011b; Kishon-Rabin *et al.*, 2001; Micheyl *et al.*, 2006; Musacchia *et al.*, 2007; Tervaniemi *et al.*, 2005) and tone language experience (e.g., Bidelman *et al.*, 2011a; Bidelman *et al.*, 2013; Chandrasekaran *et al.*, 2009; Hutka *et al.*, 2015) facilitate musical pitch perception, it remains unclear as to how long-term musical training and tone language experience jointly influence musical pitch perception. Given that music and speech processing share at least partially common neural resources (see Patel, 2003, for the shared syntactic integration resource hypothesis), it seems plausible that the effect of long-term musical training on musical pitch perception is dependent on tone language experience.

Empirically, a recent study by Tang *et al.* (2016) provided an important insight regarding the possible interaction between long-term musical training and tone language experience in musical pitch perception. In particular, they reported that Mandarin-speaking musicians and Mandarin-speaking non-musicians showed similar sensitivity on a behavioral pitch height discrimination task, and similar MMN amplitude to pitch height changes, although the former group exhibited larger P3a amplitude than the latter at frontal electrodes. These findings imply that long-term musical training did not enhance listeners' perceptual sensitivity to pitch height despite the presence of tone language experience. However, English-speaking musicians and non-musicians were not included in the above study, which limits the possibility to exclude the possibility that the effect of long-term musical training on pitch height discrimination was absent, *per se*, regardless of tone language experience. For instance, it might be the case that English-speaking musicians would not demonstrate a perceptual advantage relative to English-speaking non-musicians in the same study; or it might be the case that the effect of long-term musical training was absent only under the context in which tone language experience was present. Thus, to further clarify how long-term musical training and tone language experience interacted in different types of musical pitch perception, the present study tested Cantonese-speaking musicians, Cantonese-speaking non-musicians, non-tonal language-speaking musicians, and non-tonal language-speaking non-musicians.

Noticeably, Cooper and Wang (2012) holistically tested four listener groups, i.e., Thai-speaking musicians and non-musicians, and English-speaking musicians and non-musicians, on music aptitude, tone word learning, and tone identification. On the music aptitude task, musicians outperformed non-musicians regardless of their tone language experience. The lack of an interactive effect in the study by Cooper and Wang (2012), and its deviance from the study by Tang *et al.* (2016), might be due to stimuli differences. Specifically, Cooper and Wang (2012) adopted the Advanced Measures of Music Audiation (AMMA) (Gordon, 1989) in which the deviant stimuli differed from the standard either in pitch or rhythm. Without separating the two sets of pitch-deviant and rhythm-deviant stimuli, it was difficult to infer whether the observed perceptual advantage was rooted in pitch height discrimination, rhythmic discrimination, or both. Unlike the study by Cooper and Wang (2012), the study by Tang *et al.* (2016) consisted of a more "pure" measure of pitch height discrimination as they adopted the G3-C4 (196.0–261.6 Hz) pair, which differed only in pitch height (65.6 Hz). Despite the lack of non-tonal language-speaking participants (both musicians and non-musicians) in the study by Tang *et al.* (2016), their preliminary results suggest that tone language experience inhibits the effect of long-term musical training on pitch height discrimination.

As mentioned previously, the musical advantage associated with musical and tone language experience has not only been shown in pitch height discrimination (e.g., Bidelman *et al.*, 2011b; Hutka *et al.*, 2015; Micheyl *et al.*, 2006; Musacchia *et al.*, 2007; Tervaniemi *et al.*, 2005), but also in pitch interval discrimination (e.g., McDermott *et al.*, 2010; Pfordresher and Brown, 2009; Zarate *et al.*, 2012). Thus, it is essential to extend the previous studies (e.g., Tang *et al.*, 2016) not only in terms of language and musician groups, but also in terms of the inclusion of musical pitch perception in multiple dimensions, i.e., pitch height and pitch interval discrimination. Unlike the Tang *et al.* (2016) study, which tested Mandarin speakers, we tested Cantonese speakers as most Cantonese tones differ in the scale of a semitone (Peng, 2006), resembling the minimum adjacent pitch differences in music (Bidelman *et al.*, 2013). Also, three out of the six Cantonese tones are steady-state pitch variations (Gandour, 1981; Khouw and Ciocca, 2007) that resemble the acoustic nature of musical notes.

In sum, the current study investigated whether the effect of long-term musical training on musical pitch perception is influenced by tone language experience. Specifically, we compared Cantonese-speaking musicians, Cantonese-speaking non-musicians, non-tonal language-speaking musicians, and non-tonal language-speaking non-musicians on a pitch height discrimination task and a pitch interval discrimination task. We hypothesized that the effect of long-term musical training on pitch height and pitch interval discrimination would be observed in non-tonal language-speaking musicians but not in Cantonese-speaking musicians.

II. METHODS

A. Participants

We recruited 30 native Cantonese participants [22 female, mean age = 20.97 years, standard deviation (SD) = 1.38 years] and 30 non-tonal language participants (21 female, mean age = 22.50 years, SD = 4.12 years). According to the participant's self-reported language and music background questionnaire (Choi *et al.*, 2017), among all non-tonal language-speaking participants, 11 were native English speakers, while 19 reported English as their primary second language and none of them had any tone language background.

Both the Cantonese-speaking group and the non-tonal language-speaking group consisted of 15 musicians and 15 non-musicians. According to the established criteria used in previous research (e.g., Alexander *et al.*, 2005; Cooper and Wang, 2012; Wayland *et al.*, 2010), musicians were defined as individuals who had the ability to play their instruments at the time of testing, and who had either a minimum of seven years of continuous musical training or had obtained an official recognition of grade eight or above on their instruments. Non-musicians were those who had received no more than two years of musical training on any instrument or instruments combined, who had no musical training in the past five years, and who had no recent experience playing a musical instrument at the time of testing. All participants were either university students or graduates recruited at The University of Hong Kong who reported to have normal hearing ability and to not possess absolute pitch.

Among the four groups, Cantonese-speaking musicians (13 females, mean age = 20.53 years, SD = 1.64 years) had received an average of 10.60 years of musical training (SD = 3.31 years), while Cantonese-speaking non-musicians (nine females, mean age = 21.40 years, SD = 0.91 years) had received an average of 0.20 years of musical training (SD = 0.56 years). Non-tonal language-speaking musicians (12 females, mean age = 21.80 years, SD = 2.88 years) had received an average of 10.47 years of musical training (SD = 2.23 years), while non-tonal language-speaking non-musicians (nine females, mean age = 23.20 years, SD = 5.07 years) had received an average of 0.52 years of musical training (SD = 0.80 years). The age of onset of musical training was 7.53 years (SD = 2.92 years) for Cantonese-speaking musicians and 12.50 years (SD = 0.71 years) for Cantonese-speaking non-musicians; and 7.20 years (SD = 2.98 years) for non-tonal language-speaking musicians and 10.00 years (SD = 4.20 years) for non-tonal language-speaking non-musicians.

Our analyses confirmed that Cantonese-speaking musicians and non-tonal language-speaking musicians were comparable in terms of the years of musical training, $t(28) = 0.13$, $p = 0.898$, and the age of onset of musical training, $t(28) = 0.31$, $p = 0.759$. Cantonese-speaking non-musicians and non-tonal language-speaking non-musicians were not significantly different in terms of the years of musical training, $t(25.1) = -1.28$, $p = 0.213$, and the age of onset of musical training, $t(28) = 0.80$, $p = 0.456$.

B. Materials and procedure

1. Pitch height discrimination task

Using the AX discrimination paradigm, this task included the standard stimulus C4 (261.60 Hz) followed by either an identical tone or a deviant tone with higher or lower pitch. The pitch changes occurred at six different intervals, i.e., 1, 2, 4, 8, 15, and 30 Hz that corresponded to 7, 13, 25, 50, 100, and 200 cents, in which 100 cents was equal to one semi-tone. The duration of each tone was 600 ms, with 3 ms onset- and offset-ramping applied.

On each trial, two tones were audibly presented via inserted earphones with the inter-stimulus interval of 600 ms. Participants were then asked to judge whether the two tones were the same or different. There were two practice trials and 144 experimental trials (12 same trials × 6 repetitions + 6 types of different trials × 2 types of direction changes × 6 repetitions). The presentation order of the trials was pseudorandomized, with no more than two consecutive identical trials. The Cronbach's alpha reliability was 0.90.

2. Pitch interval discrimination task

Two types of stimuli were used in each trial: standard and target. Each stimulus consisted of two musical notes. For the standard stimuli, the two notes were C4 and G4, with a 700-cent change from the first to the second note. The target stimuli always started with an F#4 followed by another musical note. In the same-interval trials, the following musical note was C#5 (554.4 Hz), corresponding to a 700-cent change from the starting F#4 note. On different-interval trials, the musical note following F#4 was either higher or lower in pitch than C#5 by 7, 13, 25, 50, 100, or 200 cents. The duration of each tone was 600 ms, with 3 ms onset- and offset-ramping applied.

On each trial, standard and target stimuli were audibly presented via inserted earphones with an inter-stimulus interval of 600 ms. Participants were then asked to judge whether the target stimulus had the same interval change as the standard stimulus. There were nine practice trials and 144 experimental trials (12 same-interval trials × 6 repetitions + 6 types of cent changes × 2 types of direction changes × 6 repetitions). The presentation order of the trials was pseudorandomized. The Cronbach's alpha reliability was 0.84.

III. RESULTS

Prior to our analyses, we calculated the hit rates for both the pitch height and pitch interval discrimination tasks by the number of correct responses only on the different-trials, but not on the same-trials. The false alarm rate was calculated by the number of incorrect responses on the same-trials. We obtained the sensitivity index (d′) by subtracting the z-transform of the hit rate by the z-transform of the false alarm rate (MacMillan and Creelman, 2005). The d′ of the pitch height discrimination and pitch interval discrimination tasks were correlated, $r = 0.27$, $p < 0.05$. The d′ by group on pitch height discrimination and pitch interval discrimination tasks are shown in Fig. 1.

To examine whether the effect of long-term musical training on pitch height and pitch interval discrimination was dependent on tone language experience, we submitted the d′ of both tasks to a multivariate analysis of variance (MANOVA) with tone language experience (Cantonese versus non-tonal language) and long-term musical training (musician versus non-musician) as between-subjects factors. There was a significant interaction between long-term musical training and tone language experience, $F(2, 55) = 12.67$, $p < 0.001$, $\eta_p^2 = 0.32$. The main effect of long-term musical training was significant, $F(2, 55) = 7.33$, $p < 0.01$, $\eta_p^2 = 0.21$, but the main effect of tone language experience was not significant, $F(2, 55) = 1.45$, $p = 0.243$.

To further examine the interaction between long-term musical training and tone language experience, we conducted a two-way analysis of variance (ANOVA) for the pitch height discrimination task and for the pitch interval discrimination task. For the pitch height discrimination task, there was a significant interaction between tone language experience and long-term musical training, $F(1, 56) = 6.35$, $p < 0.05$, $\eta_p^2 = 0.10$. However, the main effects were not significant for long-term musical training, $F(1, 56) = 1.36$, $p = 0.249$, and tone language experience, $F(1, 56) = 0.01$, $p = 0.925$. The simple main effect analysis further revealed that non-tonal language-speaking musicians outperformed non-tonal language-speaking non-musicians, $p < 0.05$, but there was no significant difference between Cantonese-speaking musicians and Cantonese-speaking non-musicians, $p = 0.342$. This indicated that long-term musical training enhanced pitch height discrimination in non-tonal language-speaking musicians, but not in Cantonese-speaking musicians.

For the pitch interval discrimination task, the same two-way ANOVA analysis revealed a significant main effect of long-term musical training, $F(1, 56) = 14.75$, $p < 0.001$, $\eta_p^2 = 0.21$, but not of tone language experience, $F(1, 56) = 2.59$, $p = 0.114$. The interaction between tone language experience

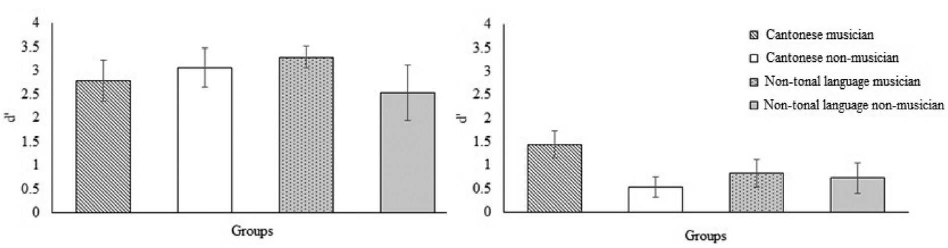

FIG. 1. Sensitivity index (d′) by group on pitch height discrimination (left) and pitch interval discrimination (right). Error bars represent 95% confidence interval.

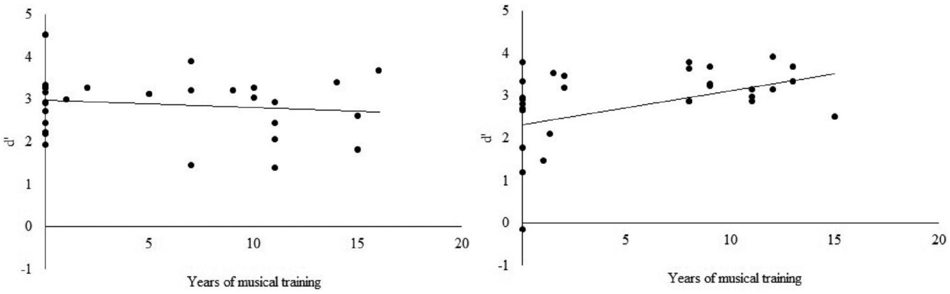

FIG. 2. Pitch height discrimination sensitivity index (d′) by years of musical training in Cantonese (left) and non-tonal language (right) participants.

and long-term musical training was significant, $F(1, 56) = 9.10$, $p < 0.01$, $\eta_p^2 = 0.14$. Simple main effect analysis showed that Cantonese-speaking musicians significantly outperformed Cantonese-speaking non-musicians, $p < 0.001$, but there was no significant difference between non-tonal language-speaking musicians and non-tonal language-speaking non-musicians, $p = 0.563$. In other words, the perceptual advantage associated with long-term musical training is evident only when combined with tone language experience.

To further evaluate the possible inhibitory effect of tone language experience on long term musical training, we conducted correlational analyses between d′ and years of musical training separately on Cantonese and non-tonal language groups for pitch height discrimination. As shown in Fig. 2, the correlation between d′ and years of musical training was significant among non-tonal language participants, $r = 0.42$, $p < 0.05$, but not among Cantonese participants, $r = -0.185$, $p = 0.327$.

The same correlational analyses were also conducted for the pitch interval discrimination task. As shown in Fig. 3, there was a significant association between years of musical training and d′ in Cantonese participants, $r = 0.80$, $p < 0.001$, but not in non-tonal language participants, $p = 0.582$.

Figure 4 shows the hit rates on 7, 13, 25, 50, 100, and 200 cents of changes by group for the pitch interval discrimination task. To examine whether the hit rate increased as pitch interval change increased, we first conducted a Page's trend test (Page, 1963). The results indicated that increased pitch interval change facilitated discrimination in Cantonese-speaking musicians, $L = 1321$, $\chi^2 = 51.96$, $p < 0.001$, Cantonese-speaking non-musicians, $L = 1235$, $\chi^2 = 19.11$, $p < 0.001$, non-tonal language-speaking musicians, $L = 1287$, $\chi^2 = 37.05$, $p < 0.001$, and non-tonal language-speaking non-musicians, $L = 1267$, $\chi^2 = 29.63$, $p < 0.001$.

Next, we examined whether the interdependency between tone language experience and long-term musical training was evident across all pitch interval changes or confined only to certain pitch interval changes. A three-way ANOVA on hit rates was conducted with pitch interval changes (7, 13, 25, 50, 100, and 200 cents) as the within-subjects factor, and tone language experience (Cantonese versus non-tonal language) and long-term musical training (musician versus non-musician) as between-subjects factors. Greenhouse-Geisser adjustment to the degree of freedom was applied to correct for violation of sphericity whenever necessary. There was a marginally significant three-way interaction of pitch interval change, long-term musical training, and tone language experience, $F(2.89, 161.61) = 2.24$, $p = 0.088$, and a significant two-way interaction between pitch interval change and long-term musical training, $F(2.89, 161.61) = 4.71$, $p < 0.01$. The main effects of pitch interval changes and long-term musical training were both significant, $ps < 0.05$.

To further examine the three-way interaction, there were significant and marginally significant interactions between tone language experience and long-term musical training under 50-cents change, $p < 0.05$, and 200-cents change, $p = 0.063$. The main effect of long-term musical training was significant in 50-, 100-, and 200-cents changes, $ps < 0.05$. In further examining the two-way interactions, simple main effect analysis showed that Cantonese-speaking musicians significantly outperformed Cantonese-speaking non-musicians, $ps < 0.01$, but non-tonal language-speaking musicians did not significantly outperform non-tonal language-speaking non-musicians, $ps > 0.264$, under 50- and 200-cents changes. These results suggest that the interdependency between tone language experience and long-term musical training was evident only under 50- and 200-cents changes.

IV. DISCUSSION

The present study set out to explore the interaction between long-term musical training and tone language experience in musical pitch perception. On the pitch height discrimination task, musicians outperformed non-musicians

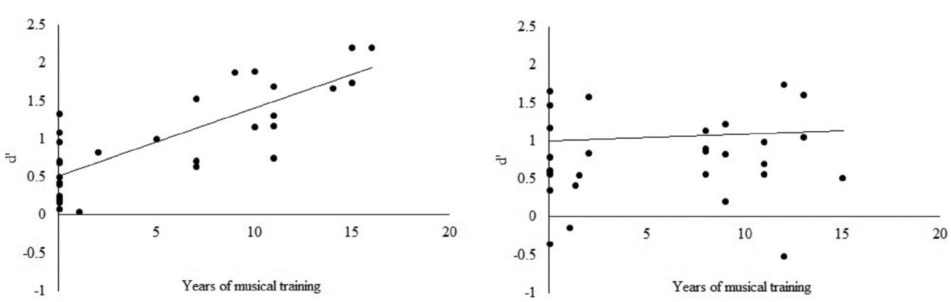

FIG. 3. Pitch interval discrimination sensitivity index (d′) by years of musical training in Cantonese (left) and non-tonal language (right) participants.

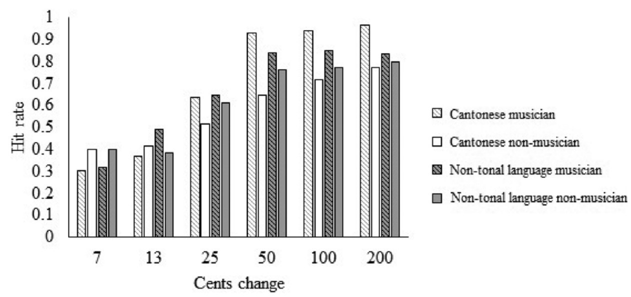

FIG. 4. Hit rates on different cents changes by group on pitch interval discrimination.

only among non-tonal language speakers, but not among Cantonese speakers, whereas on the pitch interval discrimination task, musicians outperformed non-musicians only among Cantonese speakers, but not among non-tonal language speakers. These results suggest that the effect of long-term musical training on musical pitch perception is modulated by tone language experience. The modulation occurred in two forms: as an inhibitory effect on the pitch height discrimination task, and as an interdependency on the pitch interval discrimination task. These results were further discussed in terms of the different natures of the pitch height discrimination and pitch interval discrimination tasks.

One striking finding in the current study is the inhibitory effect of tone language experience on the effect of long-term musical training on pitch height discrimination, which is indicated by the observation that musical advantage was shown only among non-tonal language speakers, but not among Cantonese speakers. This clearly shows that the perceptual advantage associated with long-term musical training, as exhibited among non-tonal language speakers, is inhibited by Cantonese tone language experience. One possible account could be the assimilation of musical pitch into native tonal categories. Unlike non-tonal languages, Cantonese is a tone language that involves a prominent use of pitch for lexical distinctions (Cutler and Chen, 1997), and Cantonese speakers have formed rigid tonal representations in their phonological system (So and Best, 2010). In the present study, it may be possible that the Cantonese speakers assimilated musical pitch to their native tonal categories. Our hypothesis is theoretically grounded on the perceptual assimilation model for suprasegmentals (So and Best, 2008), which posits that listeners assimilate non-native pitch information into their native suprasegmental categories, and that the effect of native language experience on perceptual discrimination depends on how the contrasts are assimilated (Best *et al.*, 2001). For example, if the contrasts are assimilated into two separate native suprasegmental categories, i.e., two category assimilation, discrimination of the contrasts among listeners is expected to be good. However, if the contrasts are assimilated into one single category, i.e., single category assimilation, discrimination is expected to be poor.

Although the perceptual assimilation model was originally proposed to account for the perception of non-native lexical tones, it seems very likely that Cantonese speakers may assimilate musical pitch into their L1 tonal categories. In fact, this claim is supported by a pitch-MMN lateralization study on Cantonese speakers (Gu *et al.*, 2013). While non-tonal language speakers showed left- and right-lateralization for speech and musical processing, respectively (e.g., Zatorre *et al.*, 1992; Tramo, 2001), Cantonese speakers demonstrated left hemispheric lateralization for their MMN response to both lexical tone and harmonic pitch variations, indicating that linguistic processing occurs during both lexical tone and harmonic pitch perception. Given that the three level tones in the Cantonese lexical tone system are steady state pitch variations resembling the time-variant nature of musical pitch (Gu *et al.*, 2013), it is reasonable to hypothesize that Cantonese speakers assimilate musical pitch into their tonal categories.

However, it remains uncertain whether the assimilation facilitates pitch height discrimination or depends on the specific contrasts as posited by the perceptual assimilation model. Nonetheless, we believe that perceptual assimilation has constrained the effect of long-term musical training on pitch height discrimination, or, more specifically, has inhibited the musical advantage induced by long-term musical training. Given that the templates that the contrasts are assimilated to, i.e., native phonological categories, are established by extensive native language experience (Best *et al.*, 2001; So and Best, 2008), long-term musical training which enhances listeners' sensitivity to fine-grained psychoacoustic features of musical pitch, *per se*, may not be directly relevant to linguistic perceptual processing. Consistent with this claim, the relationship between years of musical training and perceptual sensitivity only existed among non-tonal language speakers but not among Cantonese speakers in the present study.

Another important finding is the interdependency between long-term musical training and tone language experience on pitch interval discrimination, in which tone language experience was necessary to induce musical advantage in discriminating pitch intervals. One possible account is that tone language experience and long-term musical training enhanced separate perceptual operations that facilitated pitch interval discrimination when both were present but not in isolation. In a model for pitch interval recognition (Deutsch, 1975), incoming complex acoustic waveforms are first subject to low level sensory processing in which spectral information is extracted and encoded, thereby allowing for the identification of individual musical notes (e.g., C4 or G4). In this regard, there is ample evidence suggesting that long-term musical training enhances the sensory operations underlying this early stage of pitch detection/identification (e.g., Micheyl *et al.*, 2006; Näätänen, 1992). According to the Deutsch (1975) model, the later stages of pitch interval recognition entail cognitive operations for extracting and comparing abstract information obtained from an array of musical notes. Specifically, listeners had to evaluate the pitch difference of the second note (e.g., G4) relative to the first note (e.g., C4), and the fourth note (e.g., C#5) to the third note (e.g., F#4), then compare whether the interval differences were the same. Thus, unlike pitch height discrimination, listeners had to ignore local differences in absolute pitch and attend only to the abstract representations of pitch intervals. This is similar to the perceptual normalization of level tones on which Cantonese speakers rely to overcome talker variation in absolute pitch (Wong and Diehl, 2003). With reference to the Deutsch (1975) model, the perceptual normalization may be domain-general (Peretz, 2006; Wong *et al.*, 2012), thereby enhancing Cantonese-speaking musicians' ability to extract pitch intervals, a cognitive operation that functions interdependently with the sensory enhancement associated with long-term musical training to facilitate pitch interval discrimination.

On the other hand, despite the enhanced sensory processing relevant to the first stage of pitch interval discrimination, non-tonal language-speaking musicians' lack of tone language experience may have limited the potential of musical advantage induced by long-term musical training. At a granular level, it remains unclear why this interdependency is evident at 50- and 200-cent intervals, but not at other pitch intervals. Also, it has been suggested that music context influences pitch interval discrimination (e.g., Graves and Oxenham, 2017; Wapnick *et al.*, 1982). For example, Wapnick *et al.* (1982) demonstrated that musicians performed less accurately on pitch interval discrimination when the pitch information was carried in isolation than when the information was presented in a melodic context. Thus, it is worthwhile in future studies to explore how the interdependency between tone language experience and long-term musical training varies with specific pitch intervals and the musical context.

On the pitch interval discrimination task, the pitch height of the first three musical notes was fixed, and it might be possible that listeners performed the pitch interval discrimination task in the same way as the pitch height discrimination task: by establishing an imaginary reference pitch, i.e., C#5, and focusing only on the fourth note. However, the current results provide little support for this possibility. In particular, the correlation between the pitch interval discrimination task and the pitch height discrimination task was very small, which suggests that the two tasks may not measure the same construct. Also, adopting a 70% hit rate criterion, all four groups of listeners showed higher thresholds on the pitch interval discrimination task than on the pitch height discrimination task. If it was indeed the case that the listeners performed the pitch interval discrimination task in the same way as the pitch height discrimination task, the listeners should have shown the same threshold on the two discrimination tasks. Nevertheless, future studies may include both fixed interval and roving pitch interval discrimination tasks.

The present study showed for the first time that the effect of long-term musical training on musical pitch perception is not universal across languages. Specifically, there is a need to consider long-term musical training's interaction with tone language experience, which may appear in different forms depending on the nature of the task. However, it should be noted that the present study focused on Cantonese tone language only. Thus, future studies may consider testing another tonal population, such as Mandarin or Thai, in order to evaluate whether the interaction between long-term musical training and tone language experience depends on the complexity of the tonal system.

It should also be noted that although the non-tonal language speakers tested in the present study had no tone language experience and spoke English as their second language, they had different first languages. Future studies may consider having the non-tonal language group composed of participants who speak the same first language, which will eliminate any possible effect of native language background.

Another promising way to extend the current study is to incorporate both musical pitch and lexical tone perception among the listener groups. Moreover, future studies may evaluate multiple dimensions of musical aptitude: not only pitch, but rhythmic and dynamic changes as well (Kania, 2014), thus extending the present results to a holistic set of musical perceptual sensitivity. Taking this further, a recent study has reported that pitch sensitivity training improved linguistic and musical pitch perception in tonal amusics (Liu *et al.*, 2017). Given the bidirectional transfer between language and musical domains (Patel, 2011), future studies may examine whether and how tone language experience can be employed to facilitate pitch interval discrimination among speakers with amusia. Critically, there is a need for future studies to include nonverbal intelligence as a control measure, given its positive correlation with pitch discrimination (Tang *et al.*, 2016).

In conclusion, the facilitative effect of long-term musical training on musical pitch perception is not necessarily universal. It appears to interact with tone language experience in a complex way. Specifically, the interaction can appear in two forms: as an inhibitory effect in the pitch height discrimination task and as interdependency in the pitch interval discrimination task.

ACKNOWLEDGMENTS

The two co-first authors X.T. and W.C. have contributed equally to this work. Material in this article derives from a dissertation submitted by the third author in partial fulfillment of the requirements for the Bachelor of Science in the Division of Speech and Hearing Sciences, The University of Hong Kong. This research was supported, in part, by the Early Career Scheme (ECS) and General Research Fund (17673216) from the Hong Kong Government Research Council (27402514), and the HKU Seed Fund for Basic Research (201611159052) to X.T. We also thank Veronica Ka Wai Lai and Joseph Cheng for their helpful comments on statistical methods.

[1]The six lexical tones were first transcribed by Chao (1930) in a numerical notational system using five levels (from lowest 1 to highest 5) to describe relative height, shape and duration of pitch contour.

[2]Tonal inventory varies across different tone languages, such as Cantonese, Mandarin and Thai. In terms of inventory size, Cantonese (with six lexical tones) has more lexical tones than Thai (five) and Mandarin (four). In terms of pitch height, Cantonese has three level tones (high level, mid level, and low level), while Mandarin has only one level (high-level) tone. In terms of pitch contour, Mandarin has a unique dipping (falling-rising) tone that is absent in Cantonese and Thai.

Alexander, J. A., Wang, P. C. M., and Bradlow, A. R. (**2005**). "Lexical tone perception in musicians and non-musicians," in *Proceedings of the 9th ECSCT*, September 1–8, Lisbon, Portugal, pp. 397–400.

Best, C. T., McRoberts, G. W., and Goodell, E. (**2001**). "Discrimination of non-native consonant contrasts varying in perceptual assimilation to the listener's native phonological system," J. Acoust. Soc. Am. **109**(2), 775–794.

Bidelman, G. M., Gandour, J. T., and Krishnan, A. (**2011a**). "Cross-domain effects of music and language experience on the representation of pitch in the human auditory brainstem," J. Cog. Neurosci. **23**(2), 425–434.
Bidelman, G. M., Hutka, S., and Moreno, S. (**2013**). "Tone language speakers and musicians share enhanced perceptual and cognitive abilities for musical pitch: Evidence for bidirectionality between the domains of language and music," PLoS One **8**(4), e60676.
Bidelman, G. M., Krishnan, A., and Gandour, J. T. (**2011b**). "Enhanced brainstem encoding predicts musicians' perceptual advantages with pitch," Euro. J. Neurosci. **33**(3), 530–538.
Chandrasekaran, B., Krishnan, A., and Gandour, J. T. (**2009**). "Relative influence of musical and linguistic experience on early cortical processing of pitch contours," Brain Lang. **108**, 1–9.
Chao, Y. R. (**1930**). "A system of tone-letters," Le Maître Phon. **45**, 24–27.
Choi, W., Tong, X., Gu, F., Tong, X., and Wong, L. (**2017**). "On the early neural perceptual integrality of tones and vowels," J. Neuroling. **41**, 11–23.
Cooper, A., and Wang, Y. (**2012**). "The influence of linguistic and musical experience on Cantonese word learning," J. Acoust. Soc. Am. **131**(6), 4756–4768.
Cutler, A., and Chen, H. (**1997**). "Lexical tone in Cantonese spoken-word processing," Percept. Psychophys. **59**(2), 165–179.
Deutsch, D. (**1975**). "The organization of short-term memory for a single acoustic attribute," in *Short Term Memory*, edited by D. Deutsch and J. A. Deutsch (Academic Press, New York).
Fok-Chan, Y. (**1974**). *A Perceptual Study of Tones in Cantonese* (Hong Kong University Press, Hong Kong).
Gandour, J. (**1981**). "Perceptual dimensions of tone: Evidence from Cantonese," J. Chin. Ling. **9**, 20–36.
Gaser, C., and Schlaug, G. (**2003**). "Brain structures differ between musicians and non-musicians," J. Neurosci. **23**(27), 9240–9245.
Giuliano, R. J., Pfordresher, P. Q., Stanley, E. M., Shalini, N., and Wicha, N. Y. Y. (**2011**). "Native experience with a tone language enhances pitch discrimination and the timing of neural responses to pitch change," Front. Psychol. **2**, 146.
Gordon, E. (**1989**). *Advanced Measures of Music Audiation* (Gia Publications, Chicago, IL).
Graves, J. E., and Oxenham, A. J. (**2017**). "Familiar tonal context improves accuracy of pitch interval perception," Front. Psychol. **8**, 1753.
Gu, F., Zhang, C., Hu, A., and Zhao, G. (**2013**). "Left hemisphere lateralization for lexical and acoustic pitch processing in Cantonese speakers as revealed by mismatch negativity," Neuroimage **83**, 637–645.
Hannon, E. E., and Trainor, L. J. (**2007**). "Music acquisition: Effects of enculturation and formal training on development," Trends Cogn. Sci. **11**(11), 466–472.
Herholz, S. C., and Zatorre, R. J. (**2012**). "Musical training as a framework for brain plasticity: Behavior, function, and structure," Neuron **76**(3), 486–502.
Hutka, S., Bidelman, G. M., and 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," Neuropsycholigia **71**, 52–63.
Hyde, K. L., Lerch, J., Norton, A., Forgeard, M., Winner, E., Evans, A. C., and Schlaug, G. (**2009**). "The effects of musical training on structural brain development: A longitudinal study," Ann. N.Y. Acad. Sci. **1169**, 182–186.
Kania, A. (**2014**). "The philosophy of music," in *The Stanford Encyclopedia of Philosophy*, edited by E. N. Zalta, https://plato.stanford.edu/archives/spr2014/ (Last viewed June 1, 2017).
Khouw, E., and Ciocca, V. (**2007**). "Perceptual correlates of Cantonese tones," J. Phon. **35**, 104–117.
Kishon-Rabin, L., Amir, O., Vexler, Y., and Zaltz, Y. (**2001**). "Pitch discrimination: Are professional musicians better than non-musicians?," J. Basic Clin. Physiol. Pharmacol. **12**(2), 125–143.
Liu, F., Jiang, C., Francart, T., Chan, A. H. D., and Wong, P. C. M. (**2017**). "Perceptual learning of pitch direction in congenital amusia: Evidence from Chinese speakers," Music Percept. **34**(3), 335–351.
Macmillan, N. A., and Creelman, C. D. (**2005**). *Detection Theory: A User's Guide* (Lawrence Erlbaum Associates, Mahwah, NJ).
McDermott, J. H., Keebler, M. V., Micheyl, C., and Oxenham, A. J. (**2010**). "Musical intervals and relative pitch: Frequency resolution, not interval resolution, is special," J. Acoust. Soc. Am. **128**(4), 1943–1951.
Micheyl, C., Delhommeau, K., Perrot, X., and Oxenham, A. J. (**2006**). "Influence of musical and psychoacoustical training on pitch discrimination," Hear. Res. **219**, 36–47.
Miendlarzewska, E. A., and Trost, W. J. (**2013**). "How musical training affects cognitive development: Rhythm, reward and other modulating variables," Front. Neurosci. **7**, 279.
Musacchia, G., Sams, M., Skoe, E., and Kraus, N. (**2007**). "Musicians have enhanced subcortical auditory and audiovisual processing of speech and music," Proc. Natl. Acad. Sci. U.S.A. **104**(40), 15894–15898.
Näätänen, R. (**1992**). *Attention and Brain Function* (Lawrence Erlbaum Associates, Mahwah, NJ).
Page, E. B. (**1963**). "Ordered hypotheses for multiple treatments: A significance test for linear ranks," J. Am. Stat. Assoc. **58**(301), 216–230.
Patel, A. D. (**2003**). "Language, music, syntax and the brain," Nat. Neurosci. **6**, 674–681.
Patel, A. D. (**2011**). "Why would musical training benefit the neural encoding of speech? The OPERA hypothesis," Front. Psychol. **2**, 142.
Patel, A. D. (**2012**). "The OPERA hypothesis: Assumptions and clarifications," Ann. N.Y. Acad. Sci. **1252**, 124–128.
Peng, G. (**2006**). "Temporal and tonal aspects of Chinese syllables: A corpus-based comparative study of Mandarin and Cantonese," J. Chi. Ling. **9**, 20–36.
Peretz, I. (**2006**). "The nature of music from a biological perspective," Cognition **100**, 1–32.
Pfordresher, P. Q., and Brown, S. (**2009**). "Enhanced production and perception of musical pitch in tone language speakers," Atten. Percept. Psychophys. **71**(6), 1385–1398.
Ritter, W., Paavilainen, P., Lavikainen, J., Reinikainen, K., Alho, K., Sams, M., and Näätänen, R. (**1992**). "Event-related potentials to repetition and change to auditory stimuli," Electroencephalogr. Clin. Neurophysiol. **83**, 306–321.
Schlaug, G., Jancke, L., Huang, Y., Staiger, J. F., and Steinmetz, H. (**1995**). "Increased corpus callosum size in musicians," Neuropsychologia **33**(8), 1047–1055.
So, C. K., and Best, C. T. (**2008**). "Do English speakers assimilate Mandarin tones to English prosodic categories," in *Proceedings of the 9th Annual Conference of the International Speech Communication Association (Interspeech 2008), incorporating the 12th Australasian International Conference on Speech Science and Technology (SST 2008)*, September 22–26, Brisbane, Australia, p. 1120.
So, C. K., and Best, C. T. (**2010**). "Cross-language perception of non-native tonal contrasts: Effects of native phonological and phonetic influences," Lang. Speech **53**(2), 273–293.
Swaminathan, J., Mason, C. R., Streeter, T. M., Best, V., Kidd, G., and Patel, A. D. (**2015**). "Musical training, individual differences and the cocktail party problem," Sci. Rep. **5**, 11628.
Tang, W., Xiong, W., Zhang, Y., Dong, Q., and Nan, Y. (**2016**). "Musical experience facilitates lexical tone processing among Mandarin speakers: Behavioral and neural evidence," Neuropsychologia **91**, 247–253.
Tervaniemi, M., Just, V., Koelsch, S., Widmann, A., and Schroger, E. (**2005**). "Pitch discrimination accuracy in musicians vs nonmusicians: An event-related potential and behavioral study," Exp. Brain Res. **161**, 1–10.
Tramo, M. J. (**2001**). "Biology and music. Music of the hemispheres," Science **291**(5501), 54–56.
Tsang, Y., Jia, S., Huang, J., and Chen, H. (**2011**). "ERP correlates of pre-attentive processing of Cantonese lexical tones: The effects of pitch contour and pitch height," Neurosci. Lett. **487**, 268–272.
Wapnick, J., Bourassa, G., and Sampson, J. (**1982**). "The perception of tonal intervals in isolation and in melodic context," Psychomusicol.: J. Res. Music Cogn. **2**, 21–37.
Wayland, R. P., Herrera, E., and Kaan, E. (**2010**). "Effects of musical experience and training on pitch contour perception," J. Phon. **38**, 654–662.
Wong, P. C. M., Ciocca, V., Chan, A. H. D., Ha, L. Y. Y., Tan, L-H., and Peretz, I. (**2012**). "Effects of culture on musical pitch perception," PLoS One **7**(4), e33424.
Wong, P. C. M., and Diehl, R. L. (**2003**). "Perceptual normalization of inter- and intra-talker variation in Cantonese level tones," J. Speech Lang. Hear. Res. **46**, 413–421.
Zarate, J. M., Ritson, C. R., and Poeppel, D. (**2012**). "Pitch-interval discrimination and musical expertise: Is the semitone a perceptual boundary," J. Acoust. Soc. Am. **132**(2), 984–993.
Zatorre, R. J., Evans, A. C., Meyer, E., and Gjedde, A. (**1992**). "Lateralization of phonetic and pitch discrimination in speech processing," Science **256**(5058), 846–849.