

Language, Cognition and Neuroscience

ISSN: 2327-3798 (Print) 2327-3801 (Online) Journal homepage: www.tandfonline.com/journals/plcp21

Cross-domain correlation in pitch perception, the influence of native language

Ao Chen, Liquan Liu & René Kager

To cite this article: Ao Chen, Liquan Liu & René Kager (2016) Cross-domain correlation in pitch perception, the influence of native language, Language, Cognition and Neuroscience, 31:6, 751-760, DOI: 10.1080/23273798.2016.1156715

To link to this article: https://doi.org/10.1080/23273798.2016.1156715

© 2016 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group

Published online: 25 Apr 2016.

Submit your article to this journal

Article views: 2752

View related articles

View Crossmark data

Citing articles: 15 View citing articles

Full Terms & Conditions of access and use can be found at
https://www.tandfonline.com/action/journalInformation?journalCode=plcp21

LANGUAGE, COGNITION AND NEUROSCIENCE, 2016
VOL. 31, NO. 6, 751–760
http://dx.doi.org/10.1080/23273798.2016.1156715

Routledge
Taylor & Francis Group

OPEN ACCESS

Cross-domain correlation in pitch perception, the influence of native language

Ao Chen[a], Liquan Liu[a,b] and René Kager[a]

[a]Utrecht Institute of Linguistics, Utrecht University, Utrecht, The Netherlands; [b]School of Social Sciences and Psychology, Western Sydney University, South Penrith, Australia

ABSTRACT

The current study explores how language experience may shape the correlation between lexical tone and musical pitch perception. A two domains (music and lexical tone) by two languages (tone, Mandarin Chinese and non-tone, Dutch) design is adopted. Participants were tested on their discrimination of Mandarin Chinese lexical tones, Montreal Battery of Evaluation of Amusia (MBEA), and Musical Ear Test (MET). The Chinese listeners outperformed the Dutch listeners on both MBEA and MET, but had comparable accuracies for the lexical tone discrimination. Importantly, a significant cross-domain correlation was only observed for the Dutch listeners but not for the Chinese listeners. For tone language listeners, once lexical tones have been acquired, native listeners perceive them as phonological categories, and split them from other pitch variations that do not play a phonemic role. Non-tone language listeners, on the other hand, perceive both lexical tones and musical pitch on a psycho-acoustical basis, hence exhibit a unified perception of pitch across the two domains.

ARTICLE HISTORY
Received 17 July 2015
Accepted 29 January 2016

KEYWORDS
Pitch perception; lexical tones; musical pitch; cross-domain correlation; language experience

Introduction

Language and music are both high-level cognitive domains that are specific to humans. The issue of whether these two domains coexist in a bimodular fashion or are unified cognitively and neuro-physiologically has received much attention. When examining the relation between music and language, the processing of pitch, which plays a fundamental role in both domains, has been a main focus. Earlier studies that compared pitch processing in language and music present us with a mixed picture. On the one hand, several studies have found evidence suggesting autonomy of pitch processing in the two domains. Individuals with a disorder in one domain may have the other domain largely intact (Hébert, Racette, Gagnon, & Peretz, 2003; Peretz, Champod, & Hyde, 2003; Peretz & Coltheart, 2003; Racette, Bard, & Peretz, 2006; Schlaug, Marchina, & Norton, 2008; Ullman et al., 1997; Wilson, Parsons, & Reutens, 2006). Recent brain studies have pinpointed areas that are involved in linguistic pitch but not in music processing (Fedorenko, Behr, & Kanwisher, 2011).

On the other hand, numerous studies have found that both musicianship and speaking a tone language (i.e. languages in which pitch variations are used to distinguish meaning lexically) are beneficial to pitch processing in the complementary domain. With regard to the enhancement in lexical tone perception brought about by music aptitude, it has been found that, compared to non-tone language speakers without music training, non-tone language musicians are better at detecting lexical tonal variations (Alexander, Wong, & Bradlow, 2005; Delogu, Lampis, & Belardinelli, 2006; Marie, Delogu, Lampis, Belardinelli, & Besson, 2011), and learning to pair pitch patterns to word meaning (which is similar to learning lexical tones) (Wong & Perrachione, 2007). Speaking a tone language has also been found to be beneficial for music perception (Alexander, Bradlow, Ashley, & Wong, 2008; Bidelman, Hutka, & Moreno, 2013; Jiang, Hamm, Lim, Kirk, & Yang, 2010; Pfordresher & Brown, 2009; Stevens, Keller, & Tyler, 2004; Tillmann et al., 2011; Wong et al., 2012). In sum, earlier studies on pitch processing in language and music show mixed results, suggesting autonomy as well as interdependencies between the two domains.

Both speech and music perception are shaped by learning and exposure. In the language domain, as infants grow older, they become more sensitive to native contrasts, but less sensitive to non-native contrasts (e.g. Harrison, 2000; Kuhl, Williams, Lacerda, Stevens, & Lindblom, 1992; Mattock & Burnham, 2006; Mattock, Molnar, Polka, & Burnham, 2008; Werker & Tees, 1984; Yeung, Chen, & Werker, 2013; but see also Liu & Kager, 2014). Similarly, in the music domain, infants become more capable of discriminating melodies

CONTACT Ao Chen a.chen@uu.nl

© 2016 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way.

that conform to the structure in the ambient input (Lynch & Eilers, 1992; Trainor & Trehub, 1992; Trehub, Schellenberg, & Kamenetsky, 1999). So far, however, whether the relation between musical pitch processing and lexical tone processing can also be shaped by experience remains under-investigated. The cross-domain perception of pitch has been studied either among tone language speakers, or non-tone language musicians, and the benefits observed in the opposite domain are taken as evidence for a shared processing between music and language. Such assumption, however, neglects the important fact that lexical tones do not have the same function for tone and non-tone speakers. The absence of lexical tones makes non-tone language listeners perceive them psycho-acoustically rather than phonologically, whereas for tone language speakers, lexical tones are phonological categories rather than pure acoustical pitch variations (Francis, Ciocca, & Ng, 2003; Hallé, Chang, & Best, 2004; Xu, Gandour, & Francis, 2006). Music perception, on the other hand, calls for accurate perception of discrete pitches as small as one semitone, regardless of language background (McDermott & Oxenham, 2008; Trainor & Hannon, 2013). Thus, for non-tone language listeners who are non-musicians, both the perception of musical pitch and the perception of lexical tones are psycho-acoustically driven, and it is reasonable to expect that they show a unified processing. Much evidence for such unification has been found in previous studies (e.g. Delogu et al., 2006; Marie et al., 2011; Tillmann et al., 2011; Wong & Perrachione, 2007). For tone language listeners, however, pitch perception taps into different mechanisms, being used in linguistic contexts to differentiate phonological categories, and in musical contexts as perception of fine acoustical differences. One piece of evidence for such a division is that even amusic native speakers of a tone language are able to produce their native lexical tones accurately and most are able to perceive them correctly (Nan, Sun, & Peretz, 2010).

The enhanced musical pitch perception among tone language listeners does not necessarily entail a shared processing of the lexical tones and musical pitch, and the unification observed among non-tone language listeners may not hold true for tone language listeners. Lexical tones play an indispensable role in languages such as Chinese, where every syllable has to carry a tone in speech. For example, the same syllable has different meanings when carrying different lexical tones in Chinese. The syllable /ma/ means “mom”, “hemp”, “horse”, and “scold” when carrying high-level (T1), rising (T2), low-dipping (T3), and high-falling tone (T4), respectively). Inaccurate perception of the lexical tones severely hampers communication. As a result, tone language speakers are expected to have solidly established phonological representations of their native lexical tones. Such representations are not expected to change whether the speakers are sensitive or not to pitch variations in other domains. The hypothesis that we propose in the present study is that, due to the phonemic role that lexical tones play, native tone language listeners may split lexical tones from other types of pitch variation in perception. In other words, pitch processing might be context-specific for tone-language speakers.

To study this, we adopted a two-domain by two-language design, where both tone (Mandarin Chinese) and non-tone language (Dutch) listeners were tested with lexical tone and musical pitch discrimination tasks. Specifically, we selected non-musicians as participants to limit the influence of music training. Lexical tone discrimination was examined with an AX paradigm, including a monosyllabic discrimination (MT) and a bisyllabic discrimination (BT) task. Music perception was measured by two sets of standard batteries, namely the Montreal Battery on Evaluation of Amusia (MBEA) (Peretz et al., 2003), and the Musical Ear Test (MET) (Wallentin, Nielsen, Friis-Olivarius, Vuust, & Vuust, 2010). We ask the question whether the relation between musical pitch and lexical tone discrimination differs as the result of different native languages. We hypothesise that for tone language listeners the processing of native lexical tone and musical pitch are uncorrelated, whereas for non-tone language listeners, pitch perception is more unified across different domains. As the advantage of music processing has been well replicated in multiple studies, we predicted that Chinese participants would outperform the Dutch participants in the two music tasks. Also, Chinese participants were expected to reach higher accuracy in lexical tone discrimination. Considering our focus on pitch perception, only those parts of the MBEA and the MET that target pitch perception will be reported.

Methods

Participants

Forty-two native Dutch listeners and 42 native Mandarin Chinese listeners participated in the study. All the Dutch listeners were raised and educated in the Netherlands. The mean age of the Dutch participants was 23.7 (range 19–47). All the Chinese listeners were raised and educated to at least high school level in China. They were living in the Netherlands to study or work at the time of the experiment. The mean age of the Chinese participants was 26.6 (range 22–31). Among the Dutch participants, 24 had music instruction (mean length 2.56 years, range 0–8). None of the Dutch participants

had learned or had been exposed to tone languages before the experiment. Twenty-four of the Chinese participants had received music instruction (mean 1.8 years, range 0–8). For both groups, none of the participants were practicing music at the time of the experiment, and none of them had been professional musicians. Univariate ANOVA did not show a significant difference between the two groups in terms of years of musical training, $F(1, 82) = 1.88$, n.s.

Stimuli

Stimuli for the lexical tone experiment

A female native speaker of Mandarin Chinese recorded multiple tokens of monosyllables /ba/, /bwɔ/, /bi/, /da/, /dwɔ/, /di/, /la/, /lwɔ/, /li/, /ma/, /mwɔ/, /mi/, /na/, /nwɔ/, /ni/ in isolation, each carrying all possible tones. When these syllables carry T2, the representing characters are: 拔, 博, 鼻, 达, 夺, 迪, 喇, 罗, 梨, 麻, 摩, 迷, 拿, 挪, 尼. All syllables are phonologically legal in Mandarin Chinese. Except /nwɔ/ carrying T1, and /nwɔ/ carrying T3, all other syllables exist in Chinese. To ensure natural production, other syllables and sentences were also produced among these syllables. The recording took place in a sound-proof phonetic lab equipped with a DAT Tascam DA-40 recorder and a Sennheiser ME-64 microphone. After recording, one token for each syllable of all tones were chosen for further manipulation. First, the original syllables with the same segments were set to the same duration (range: 380–604 ms), using overlap-add method in PRAAT (Boersma & Weenink, 2011), so that duration could not be used as a cue for discrimination. Although T3 tends to have a longer duration compared to the other lexical tones (Xu, 1997), duration is not the primary cue for lexical tone perception. For example, although discrimination based on duration was not directly examined, the stimuli in Hallé et al. (2004) and Shen and Lin (1991) differed with about 100 ms, but in both studies, the native listeners reached ceiling accuracy for identifying T2 and T3. Our stimuli were longer than the average syllable duration in running speech (Chen, Wang, & Lee, 2000), which ensured the naturalness of both T2 and T3. Equal duration of T2 and T3 does not render them unnatural or confusable for native listeners. Secondly, to create the monosyllabic stimuli carrying T1 and T4, for each syllable, we extracted the pitch contour of T1 in PRAAT, replaced the original pitch contour of the T4 syllables with the extracted T1 contour, and then resynthesised the T1 syllable using the overlap-add method in PRAAT. To get the stimuli carrying T2 and T3, similarly, for each syllable, we extracted the pitch contours of the original T3, and replaced the original pitch contour of the T2 syllable with the extracted T3 contour, and then resynthesised the T3 syllable. These manipulated syllables were used as stimuli in the MT. Multiple native Mandarin Chinese speakers listened to the stimuli, and agreed that these sounded natural. The participants' task was to discriminate between T2 and T3 and between T1 and T4. For each tonal pair, all the possible four combinations, T2–T2, T2–T3, T3–T2, and T3–T3 were presented to participants randomly, but with equal chance of each. The experiment had a total of 120 trials.

Only /ba/, /bwɔ/, /da/, /dwɔ/, /la/, /lwɔ/, /ma/, /mwɔ/, /na/, /nwɔ/ were used as tone bearing syllables to ensure that the BT part could be administered within a reasonable time frame. For the bisyllabic sequences, we reduced the duration of the mentioned syllable used in the monosyllabic tone discrimination task to 225 ms, using the overlap-add method in PRAAT. Then we concatenated the monosyllables to generate T1T4, T4T1, T4T4, and T1T1, as well as T2T3, T3T2, T3T3, and T2T2 sequences. These bisyllabic sequences were used as stimuli. Although each individual syllable is a legal morpheme, none of the bisyllabic sequences are real words.[1] The participants were asked to discriminate between T3T3–T2T3, T3T3–T3T2, T3T3–T2T2, T4T4–T1T4, T4T4–T4T1, and T4T4–T1T1. Multiple native listeners listened to the bisyllabic sequences, and all agreed that they sounded natural. Repetitions of each bisyllabic sequence, such as T2T2–T2T2, T1T1–T1T1 and T2T3–T2T3, were introduced as "same" trials, and the repetition of each sequence occurred once. The experiment had a total number of 180 trials.

The lexical tones were presented either in T2–T3 or in T1–T4 pairs for the following reasons. First, for both tone and non-tone language listeners, the discrimination between T2 and T3 is the most difficult (Chen, 2013; Huang, 2001; Hume & Johnson, 2001) whereas the discrimination between T1 and T4 is relatively easy (Liu & Kager, 2014). By testing both "difficult" and "easy" contrasts together, we had the chance to examine whether the two groups would differ in the discrimination of lexical tones, and if so, whether such differences were specific to the difficult or easy tones. Second, it has been argued that T2 and T3 share a similar pitch register, which is lower than T1 and T4 (Whalen & Xu, 1992). Our manipulation ensured that the natural register of the lexical tones was not distorted, and at the same time pitch contour was the only cue for discrimination.

Stimuli for music experiments

Two widely used standard tests, the MBEA (Peretz et al., 2003) and the MET (Wallentin et al., 2010) were used to evaluate music processing. Whereas the MBEA is used to screen people with amusia, the MET has been claimed to be sensitive enough to reveal the difference

in musical aptitude, even among musicians (for detailed description of the musical stimuli, please refer to Peretz et al. (2003) and Wallentin et al. (2010). We included both tests to have a more complete measurement of the listeners' music perception. In the MBEA, the first three subtests aim at pitch processing (scale, contour, and interval, respectively) whereas the last three target rhythm processing and short-term memory. In the MET, the first subtest tests the discrimination between musical melodies, and the second one tests the discrimination between different rhythms. In both the MBEA and the MET, the listeners were asked to judge whether two pieces of musical melodies were the same or different, and half of the trials were same trials. The two melodies in the different pairs differed in one note. The lexical tones and the musical stimuli were both ecologically valid. The lexical tones were realised on multiple syllables with natural human voice, and the melodies had real musical structure. This ensured that our tasks do tapped into the processing in music and language domain.

Procedure

The evaluations were divided into two sessions, with at least half a day in between. In the first session, the participants finished the lexical tone discrimination experiment followed by the MBEA in a sound-proof booth. For the lexical tone part, the BT always preceded the MT. Such a fixed order was adopted for the following reason. Pilot experiments found that Dutch listeners were quite accurate at discriminating the monosyllabic tones, hence it was very unlikely that the experience with the bisyllabic tones would enhance the MT to an even better level. Yet if the listeners had first participated in the MT, their performance in the BT might have reflected a learning effect. Altogether the first part took about 100 minutes. In both the MT and the BT tasks, the participants were asked to discriminate between two sequences as fast as possible by pressing a button labelled "same" or "different". Within these tasks, the stimuli were randomly presented. The inter-stimulus interval between the two sequences in each trial was 1500 ms, and the maximum response duration was one second. Failing to give a response within one second was considered as an incorrect response. Such a time limit was expected to increase the difficulty of the experiment, and hence better reveal the participants' perceptual sensitivities to the lexical tones. After these, in a second session, the MET was given to the participants to finish at home. Two Dutch participants and five Chinese participants preferred to return to the lab for a second time to complete the MET. Those who did the MET at home were required to finish it in a quiet room without interruption. To make it easier for the participants, the answer sheet for the MET was converted into an Excel spreadsheet, and the participants were asked to type an "x" in the cells corresponding to their answers. We piloted this with both Chinese and Dutch native listeners, and found that the MET can be finished in this way without difficulty. The participants were paid for their participation after the MET answer sheet was received.

Results and discussion

Statistical analysis

SPSS 20 was used for the statistical tests.

Cross-language comparison within each task

Results

Table 1 lists the mean missing and correct hit trials (including accurate response to both the same and the different trials), together with the mean discrimination accuracy. For the lexical tone discrimination task, we separately looked at the discrimination accuracy (number of correct responses divided by number of trials) of the trials that involved T2 and T3, and those involving T1 and T4 for each language group.

A mixed-design ANOVA was carried out taking tone groups (involving T1 and T4 or T2 and T3) and tasks (MT or BT) as within-subject factors and native language as a between-subject factor. Both tone groups, $F(1, 82) = 170.57$, $p < .01$ and tasks, $F(1, 82) = 26.20$, $p < .01$ showed a significant main effect, but no significant main effect of language background was found, $F(1, 82) = 0.17$, n.s. Both the Chinese and the Dutch listeners reached higher accuracy in the MT tasks than in the BT tasks, and in the T1T4 group than in the T2T3 group. Significant interactions were found between tone groups and language background, $F(1, 82) = 14.50$, $p < .01$, between tone groups and tasks, $F(1, 82) = 5.29$, $p < .05$, and between tone groups, language background and tasks, $F(1, 82) = 12.65$, $p < .01$. Bonferroni corrected *post-hoc* tests indicated a significant difference between the Dutch and Chinese listeners in the BT involving T2–T3, with the Dutch outperforming the Chinese ($p < .05$); none of the rest pair-wise comparisons turned out to be significant. Figure 1 plots the accuracy of the Chinese and the Dutch listeners in each of these tasks.[2]

The raw scores of each of the three subtests in the MBEA and the MET and the total accuracy of these tasks of the two language groups are listed in Table 2. For the subtests of the MBEA, we conducted a mixed-design ANOVA with the raw scores of each subtest

Table 1. Mean misses, correct hits and accuracy in each tone combination in each task.

	Chinese misses	Chinese hits	Chinese accuracy	Dutch misses	Dutch hits	Dutch accuracy
MT T2–T3	2.37 (2.85)	51.43 (4.10)	0.86 (0.07)	0.67 (0.98)	51.10 (6.85)	0.85 (0.11)
MT T1–T4	2.91 (2.52)	56.52 (3.48)	0.94 (0.06)	0.98 (1.39)	55.86 (4.86)	0.92 (0.09)
MT overall	4.40 (4.44)	107.95 (5.75)	0.90 (0.05)	1.64 (2.17)	106.95 (11.20)	0.89 (0.09)
BT T2–T3	10.62 (11.43)	68.79 (11.05)	0.77 (0.12)	3.60 (2.60)	74.64 (10.60)	0.82 (0.12)
BT T1–T4	5.19 (6.97)	82.17 (9.83)	0.91 (0.11)	2.83 (2.16)	80.5 (9.22)	0.89 (0.1)
BT overall	15.81 (17.88)	150.95 (19.51)	0.84 (0.11)	6.43 (4.02)	155.14 (18.87)	0.86 (0.10)

being the dependent variable, the subtests being the within-subject factor and language background being the between-subject factor. A significant main effect of subtests was found, $F(2, 164) = 4.32$, $p < .05$; language background also showed a significant main effect, $F(1, 82) = 5.66$, $p < .05$. The interaction between subtests and language background was not significant, $F(2, 164) = 0.04$, n.s. Bonferroni corrected *post-hoc* tests showed a significantly lower score in subtest 3 than in subtest 2, and the Chinese listeners significantly outperformed the Dutch listeners. We computed a total accuracy for the MBEA, namely the total correct answers of the three pitch subtests divided by the total trial numbers of all the three subtests. A mixed-design ANOVA was carried out with the MBEA and the MET accuracy being the dependent variables, taking native language as between-subject variable and task (MBEA or MET) as within-subject variable. The results showed a significant main effect of language background, $F(1, 82) = 14.86$, $p < .01$, where the Chinese listeners outperformed the Dutch listeners. The interaction between language and task was marginally significant, $F(1, 82) = 2.93$, $0.05 < p < .1$.

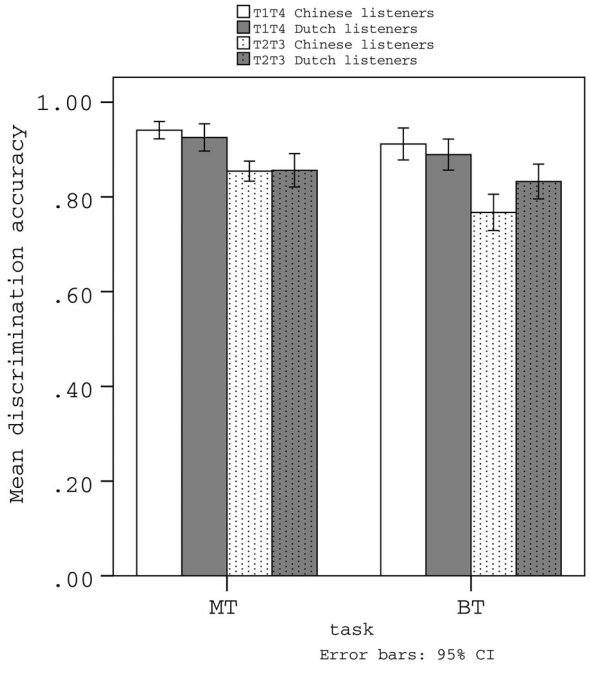


Figure 1. Mean accuracy of the Chinese and the Dutch listeners in the lexical tone discrimination tasks.

Discussion

For the lexical tone tasks, overall, as shown in Table 1, the Dutch and Chinese listeners reached a comparable level of accuracy, regardless of whether the tones were difficult or easy. The observed significant difference in the BT involving T2 and T3 was probably due to T3 sandhi. As T3 sandhi results in assimilation between T2T3 and T3T3, Chinese listeners tend to perceive these two bisyllabic tonal sequences as the same (Chen, Liu, & Kager, 2015; Shen & Lin, 1991). Yet, despite this, Chinese listeners still reached an accuracy rate higher than 70%. These results suggest that non-tone language speakers were as accurate as their tone language peers in these lexical tone discrimination tasks. Although the Dutch listeners were unaware of the contrastive nature of the lexical tones, when presented with monosyllabic and bisyllabic tones, they were nevertheless fairly accurate at discriminating them by listening to the acoustics. It was not expected that Dutch listeners were equally accurate as Chinese listeners. Yet it should be acknowledged that the lexical tone discrimination tasks were quite simple, and the lexical tones were presented in a quiet booth. It is possible that under such a favorable environment, even Dutch listeners, who were naive to lexical tones, could still discriminate these tones accurately. Other studies have also found highly accurate identification of Chinese lexical tones by non-tone language listeners (e.g. Burnham et al., 1996; Burnham et al., 2015; So & Best, 2010). If the tasks had been more complicated and involved more linguistic content (e.g. pairing tones to word meanings), we would expect differences

Table 2. Mean raw scores of each MBEA subtest and the pitch part of the MET, and the MBEA and the MET accuracy of the Chinese and Dutch listeners.

	Dutch listeners	Chinese listeners
MBEA subtest 1 scale	25.29 (2.91)	26.50 (2.92)
MBEA subtest 2 contour	25.69 (2.86)	27.02 (2.28)
MBEA subtest 3 interval	24.81 (3.06)	26.19 (3.58)
MET	36.57 (4.97)	40.48 (3.17)
MBEA accuracy	0.84 (0.08)	0.89 (0.08)
MET accuracy	0.70 (0.1)	0.78 (0.06)

Note: The standard deviations are given in parenthesis.

between the groups to emerge. What our results show is that the lack of phonological representation of the lexical tones does not preclude accurate (phonetic) discrimination.

In both of the music tasks, however, the Chinese listeners outperformed the Dutch listeners. With regard to musical pitch perception, the tone language listeners did show higher accuracy as predicted. In addition, although statistically non-significant, such enhancement seemed to be more evident in the MET, which is a more demanding test when compared with the MBEA (Wallentin et al., 2010). In other words, the more difficult the musical pitch processing is, the more evident the advantage of tone language listeners possibly is. Our findings argue for a benefit of musical pitch processing as the result of growing up with a tonal native language, which is consistent with previous findings (Alexander et al., 2008; Bidelman et al., 2013; Pfordresher & Brown, 2009; Stevens et al., 2004; Wong et al., 2012;). Sensitivity to the pitch differences exhibited in the lexical tones is a prerequisite for acquiring them. Hence, learning to tease apart and build up mental representation of such pitch patterns may have enhanced pitch processing generally.

Cross-task analysis within each language

Results

The overall MBEA accuracy, the MET accuracy, the accuracy of the MT, and the accuracy of the BT were used to examine cross-task correlations. In Tables 3 and 4, we list the correlation coefficients (Pearson's *r*, two-tailed) between the accuracy of each of the two tasks for the Dutch and Chinese listeners respectively. Figure 2 through 5 plot the cross-domain correlations of the Dutch and the Chinese listeners, respectively. For the Dutch listeners, significant correlations were found between each two of these tasks. For the Chinese listeners, however, significant correlations were only observed between tasks of the same domain.

Discussion

As shown in Table 3, among the Dutch listeners, significant correlations were found between each two of these tasks. First and most importantly, the accuracy for both the MT and BT tasks correlated significantly with both the MBEA and the MET. In other words, those listeners who were more accurate in discriminating the lexical tones, were also more accurate in discriminating the musical melodies. Such correlations favor a unified processing of musical pitch and lexical tones, and were likely to be due the psycho-acoustical perception of the lexical tones by the Dutch listeners. Being naive to the linguistic function of the lexical tones, their pitch discrimination possibly resided on the detection of acoustical differences in both the linguistic and the music tasks. Second, within each domain, the accuracy of different tasks correlated with each other significantly, which indicates the listeners' consistency within the same domain. In addition, the significant correlation

Table 3. Correlation between multiple tasks performed by Dutch listeners.

	MBEA	MET	MT	BT
MBEA		$r = .44^{**}$	$r = .63^{**}$	$r = .6^{**}$
MET			$r = .45^{**}$	$r = .42^{**}$
MT				$r = .86^{**}$

*Significance at .05 or lower.
**Significance at .01 or lower.

Table 4. Correlation between multiple tasks performed by Chinese listeners.

	MBEA	MET	MT	BT
MBEA		$r = .56^{**}$	$r = .07$	$r = -.17$
MET			$r = -.15$	$r = -.26$
MT				$r = .38^{a,*}$

[a]As can be seen on Figure 4 and Figure 5, three Chinese participants had an accuracy lower than 60% in BT, which was somewhat unexpected. Yet these participants had an accuracy above 80% in MT. Hence, they were included in the correlation analysis. If these three participants were excluded, then Pearson's *r* between MT and BT for the Chinese participants was 0.58, $p < .01$. The cross-domain correlations remained non-significant.
*Significance at .05 or lower.
**Significance at .01 or lower.

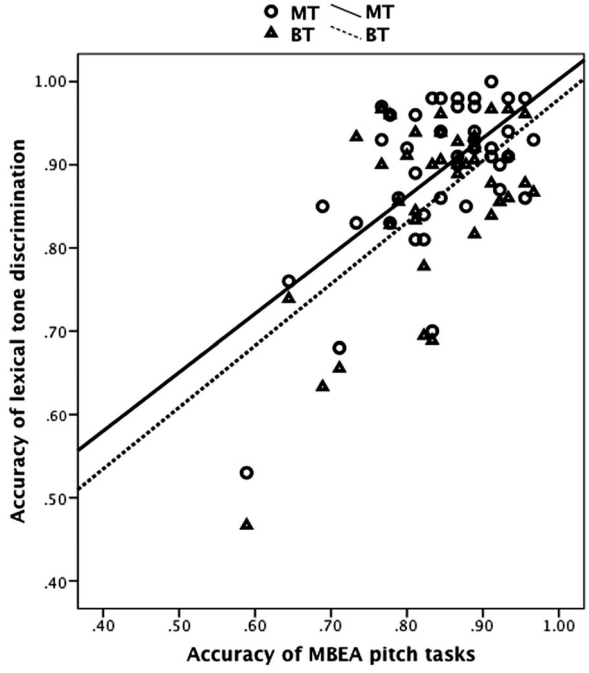


Figure 2. Correlation between the lexical tone discrimination tasks and the MBEA pitch tasks for the Dutch listeners. Linear R^2 MT = 0.4 and linear R^2 BT = 0.36.

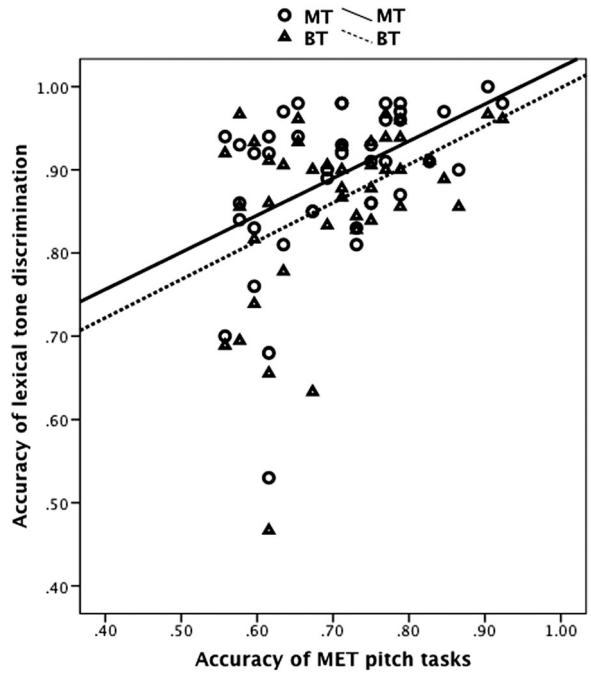


Figure 3. Correlation between the lexical tone discrimination tasks and the MET pitch task for the Dutch listeners. Linear R^2 MT = 0.21 and linear R^2 BT = 0.18.

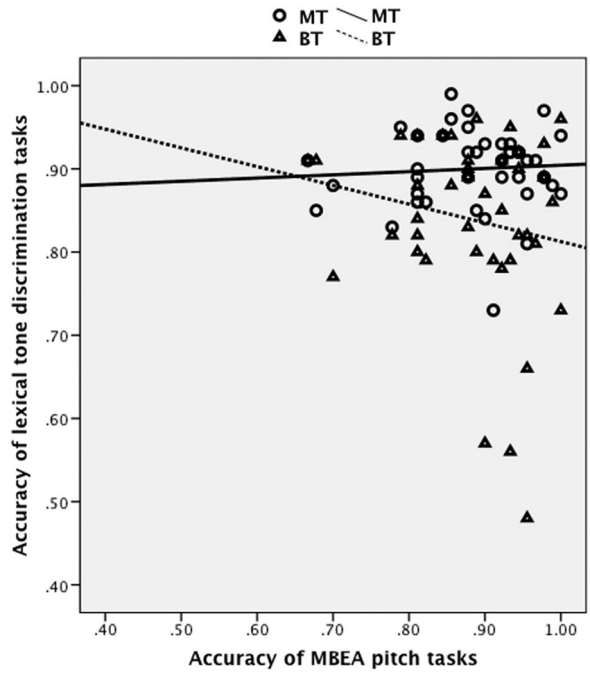


Figure 4. Correlation between the lexical tone discrimination tasks and the MBEA pitch tasks for the Chinese listeners. Linear R^2 MT = 0.01 and linear R^2 BT = 0.03.

between the MBEA and the MET corroborated the validity of the two tests in revealing musical pitch processing abilities.

Turning to Table 4, the Chinese listeners, on the other hand, showed very different results. Significant correlations were only found between tasks within the same domain. The significant correlation between the MBEA and the MET demonstrates the consistency in pitch perception within the music domain of tone language listeners. Those who were more accurate in one musical task were also more accurate in the other one. The same holds true for the discrimination of native lexical tones. The absence of cross-domain correlations suggests that native lexical tones and musical pitch might be processed independently. As we predicted, the discrimination of lexical tones did not seem to be influenced by whether the participants showed high or low sensitivity to musical pitch. Comparing Figures 4 and 5 to Figures 2 and 3, the distribution of the accuracy of the music discrimination tasks was somewhat more dispersed among the Dutch listeners than among the Chinese listeners. Seeing the higher consistency of the Chinese listeners in the music task, one plausible explanation of the enhancement in musical pitch perception might be that, growing up with a native tone language helps to improve pitch perception of those listeners with low pitch sensitivity.

General discussion

In the current study, we tested Dutch and Chinese listeners on two musical pitch and two lexical tone discrimination tasks. For the discrimination of lexical tones, the Chinese and the Dutch listeners reached a comparable level of accuracy whereas the Chinese listeners outperformed their Dutch counterparts in the MBEA and the MET. With regard to the relationship between musical pitch and lexical tone perception, significant cross-domain correlations were only observed among the Dutch but not among the Chinese listeners.

Our findings of improved accuracy in musical pitch processing among tone language listeners are consistent with previous studies (Alexander et al., 2008; Bidelman et al., 2013; Pfordresher & Brown, 2009; Stevens et al., 2004; Wong et al., 2012). The enhanced music perception is likely to be the result of growing up with a native tone language. After all, lexical tones are mainly pitch variations, and acquiring the tonal categories requires building up separate representations for each tonal pattern. Learning to tease apart the tones may have enhanced tone language listeners' perception of pitch in general. Establishing native phonology occurs early in life when the brain is still highly malleable. Learning the contrastive function of pitch in one's native language may bring a "bonus" to other cognitive abilities (Bidelman et al., 2013; Wong et al., 2012), and the benefit brought by learning early in life may well stay until adulthood.

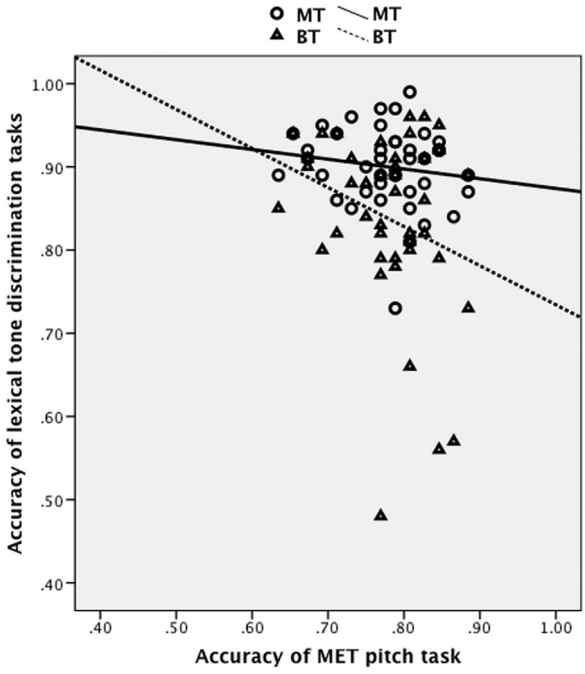


Figure 5. Correlation between the lexical tone discrimination tasks and the MET pitch tasks for the Chinese listeners. Linear R^2 MT = 0.02 and linear R^2 BT = 0.07.

Somewhat unexpectedly, we did not observe a better performance on lexical tone tasks among the Chinese listeners. Yet comparable levels of accuracy do not equal same way of processing, nor do they imply equal sensitivity to lexical tones. Several studies have shown that lexical tones are processed in different ways by native and non-native listeners (Francis et al., 2003; Hallé et al., 2004). In light of these previous findings, presumably, in our tasks, Chinese listeners perceived their native lexical tones as phonological contrasts, whereas Dutch listeners paid attention to the acoustical differences. The simplicity of the task may be another reason for the high accuracy of the Dutch listeners. If the tasks had been conducted in a less favorable environment, the advantage of perceiving native lexical tones would have possibly been visible.

One crucial finding of the current study is that only for the non-tone language listeners did the discrimination of musical pitch and lexical tones correlate, whereas no cross-domain correlation was observed for the tone language listeners. It is not a new idea that the human perception system can be shaped by input. Within the domains of language and music, perceptual attunement has been found in numerous studies (e.g. Mattock & Burnham, 2006; Mattock et al., 2008; Trainor & Trehub, 1992; Trehub et al., 1999; Werker & Tees, 1984). What we find here is that not only within-domain perception can be shaped by experience, but also that the connection between language and music does not seem to be hard-wired. Having learned the linguistic function of lexical tones in their native language, Chinese listeners no longer perceive the lexical tones as simple acoustical events, but treat them as phonological categories. Hence, the perception of lexical tones seems to be split from the perception of other phonologically irrelevant pitches. Acquiring a non-tone language, on the contrary, does not require establishing phonological representations of pitch variations in a lexically contrastive way, which may explain why non-tone language listeners showed stronger cross-domain pitch processing in the present study. In sum, language experience can shape how linguistic and musical pitch are correlated.

This lack of correlation among the Chinese listeners may seem surprising at first glance. Previous studies with amusics have found that tone language amusic listeners were impaired at perceiving their native intonation patterns (Jiang et al., 2010), and non-tone language amusics were less accurate compared to controls when discriminating lexical tones (Tillmann et al., 2011). These studies seem to argue for a shared processing between speech pitch and musical pitch regardless of language background. Such findings, however, do not conflict with our results. The crucial difference between these studies and ours is whether these linguistic pitches are phonologically contrastive, or in other words, phonemic. Neither intonation, nor the lexical tones for non-tone language listeners, serves the phonological contrastiveness that the native lexical tones serve to Chinese listeners. Another piece of evidence supporting the split hypothesis is that even amusics who speak a tone language natively are able to produce their native lexical tones correctly and most are also able to discriminate them accurately (Liu et al., 2012; Nan et al., 2010).

For future research, it would be interesting to test native tone language listeners with non-native lexical tones or intonations, and examine whether the discrimination of these linguistic stimuli correlate with the discrimination of musical pitch. A significant correlation would strengthen the evidence in support of the split hypothesis, as neither non-native lexical tones nor intonations are lexically contrastive.

Learning the phonological function of lexical tones involves perception of different phonetic dimensions of pitch (e.g. slope, turning point, and pitch level). It has been shown that only for those non-speech homologues that exhibit the naturally occurring pitch dimensions of their native lexical tones (e.g. curvilinear pitch slope as opposed to linear pitch slope), tone language speakers showed enhanced discrimination (Chandrasekaran, Krishnan, & Gandour, 2007). When native tone language listeners learn a new tone language, their native tonal categories interfere with learning the new ones (Francis, Ciocca, Ma, & Fenn, 2008). It would be interesting for future study to investigate whether the perception of

certain phonetic dimensions might correlate across different domains among tone language listeners.

The findings that the Dutch listeners showed significant cross-domain correlation whereas the Chinese did not, lead to new questions regarding the plasticity of human cognition. We speculate that infants, regardless of language background, start by perceiving lexical tones acoustically when their ability to discriminate both native and non-native sound categories is still intact (Harrison, 2000; Mattock & Burnham, 2006; Mattock et al., 2008; Yeung et al., 2013). Tone language learning infants then learn to incorporate tonal pitch into their phonology as they gradually get attuned to their native language. In this acquisition process, the need to build up representations of lexical tones possibly sharpens their general pitch perception. Such enhanced sensitivity is retained into adulthood and extends beyond speech processing. In other words, although learning native lexical tones occurs early in life, the benefit brought by such early tuning of the brain stays until adulthood. For now, no clear answer can be given to the question of whether an improvement in pitch perception brought by learning lexical tones, once exhibited, can undergo attrition. For future studies, it would be interesting to compare tone language attritors and non-tone language listeners on their pitch perception. Another interesting question is at what age tone language learners start to split their native lexical tones as opposed to pitch in musical contexts. It seems that the acquisition of lexical tones continues until at least early childhood (Hua & Dodd, 2000). Within the early years of life, when and how lexical tones are split from other types of pitch needs further investigation.

Notes

1. /dwɔ-dwɔ/ both carrying a dipping tone can be used as quantifier in lyrics, and toddlers may use /ba-ba/ both carrying a falling tone to refer to father, and /ma-ma/ both carrying a level tone to refer to mother.
2. We also conducted the same mixed-effect ANOVA with accuracy calculated as correct hits divided by the sum of trials where the participants gave a response (i.e. excluding the missing trials). The significant effects remained the same except that task was marginally significant $F(1, 82) = 3.06$, $0.05 < p < .1$, and the interaction between tone and task was non-significant, $F(1, 82) = 1.96$, n.s.

Acknowledgement

We thank Denis Burnham, Mike Olson, and Jessie Nixon for their helpful suggestions.

Disclosure statement

No potential conflict of interest was reported by the authors.

References

Alexander, J., Bradlow, A., Ashley, R., & Wong, P. (2008). Music melody perception in tone-language and non-tone language speakers. *JASA*, 124, 2495. Retrieved from. http://scitation.aip.org/content/asa/journal/jasa/124/4/10.1121/1.4782815

Alexander, J., Wong, P. C. M., & Bradlow, A. (2005, September). *Lexical tone perception in musicians and nonmusicians*. Paper presented at the Interspeech 2005 (Eurospeech) 9th European Conference on Speech Communication and Technology, Lisbon.

Bidelman, G., 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, e60676.

Boersma, P., & Weenink, D. (2011). Praat: Doing phonetics by computer (Version 5.3.25).

Burnham, D., Francis, E., Webster, D., Luksaneeyanawin, S., Attapaiboon, C., Lacerda, F., & Keller, P. (1996). Perception of lexical tone across languages: Evidence for a linguistic mode of processing. *Proceedings of the Fourth International Conference on Spoken Language*, *4*, 2514–2517.

Burnham, D., Kasisopa, B., Reid, A., Luksaneeyanawin, S., Lacerda, F., Attina, V., … Webster, D. (2015). Universality and language-specific experience in the perception of lexical tone and pitch. *Applied Psycholinguistics, FirstView*, 1–33. doi:10.1017/S0142716414000496

Chandrasekaran, B., Krishnan, A., & Gandour, J. T. (2007). Experience-dependent neural plasticity is sensitive to shape of pitch contours. *Neuroreport*, 18(18), 1963–1967. doi:10.1097/WNR.0b013e3282f213c5

Chen, A., Liu, L., & Kager, R. (2015). Cross-linguistic perception of Mandarin tone sandhi. *Language Sciences*, *48*, 62–69.

Chen, B., Wang, H., & Lee, L. (2000). Retrieval of broadcast news speech in mandarin Chinese collected in Taiwan using syllable-level statistical characteristics. *Proceedings of Acoustics, Speech, and Signal Processing*, *3*(3), 1771–1774. doi:10.1109/ICASSP.2000.862096

Delogu, F., Lampis, G., & Belardinelli, M. O. (2006). Music-to-language transfer effect: May melodic ability improve learning of tonal languages by native nontonal speakers? *Cognitive Processing*, 7(3), 203–207.

Fedorenko, E., Behr, M. K., & Kanwisher, N. (2011). Functional specificity for high-level linguistic processing in the human brain. *Proceedings of the National Academy of Sciences*, 108 (39), 16428–16433. doi:10.1073/pnas.1112937108

Francis, A. L., Ciocca, V., Ma, L., & Fenn, K. (2008). Perceptual learning of Cantonese lexical tones by tone and non-tone language speakers. *Journal of Phonetics*, 36(2), 268–294. doi:10.1016/j.wocn.2007.06.005

Francis, A. L., Ciocca, V., & Ng, B. K. C. (2003). On the (non)categorical perception of lexical tones. *Attention, Perception, & Psychophysics*, 65(7), 1029–1044.

Hallé, P. A., Chang, Y., & Best, C. T. (2004). Identification and discrimination of Mandarin Chinese tones by Mandarin Chinese vs. French listeners. *Journal of Phonetics*, 32(3), 395–421. doi:10.1016/S0095-4470(03)00016-0

Harrison, P. (2000). Acquiring the phonology of lexical tone in infancy. *Lingua*, 110(8), 581–616. doi:10.1016/S0024-3841(00)00003-6

Hébert, S., Racette, A., Gagnon, L., & Peretz, I. (2003). Revisiting the dissociation between singing and speaking in expressive aphasia. *Brain*, 126(8), 1838–1850.

Hua, Z., & Dodd, B. (2000). The phonological acquisition of Putonghua (modern standard Chinese). *Journal of Child Language*, 27(01), 3. Retrieved from doi:10.1017/S030500099900402X

Huang, T. (2001). The interplay of perception and phonology in tone 3 sandhi in Chinese Putonghua. *Working Papers in Linguistics Ohio State University*, 55, 23–42.

Hume, E., & Johnson, K. (2001). A model of the interplay of speech perception and phonology. In E. Hume & K. Johnson (Eds.), *The role of speech perception in phonology* (pp. 3–26). San Diego: Academic press.

Jiang, C., Hamm, J. P., Lim, V. K., Kirk, I. J., & Yang, Y. (2010). Processing melodic contour and speech intonation in congenital amusics with Mandarin Chinese. *Neuropsychologia*, 48 (9), 2630–2639. doi:10.1016/j.neuropsychologia.2010.05.009

Kuhl, P. K., Williams, K. A., Lacerda, F., Stevens, K. N., & Lindblom, B. (1992). Linguistic experience alters phonetic perception in infants by 6 months of age. *Science*, *255*, 606–608.

Liu, F., Jiang, C., Thompson, W., Xu, Y., Yang, Y., & Stewart, L. (2012). The mechanism of speech processing in congenital amusia: Evidence from Mandarin speakers. *Plos One*, 7, e30374.

Liu, L., & Kager, R. (2014). Perception of tones by infants learning a non-tone language. *Cognition*, 133(2), 385–394. doi:10.1016/j.cognition.2014.06.004

Lynch, M., & Eilers, R. (1992). A study of perceptual development for musical tuning. *Perception and Psychophysics*, 52, 599–608.

Marie, C., Delogu, F., Lampis, G., Belardinelli, M., & Besson, M. (2011). Influence of musical expertise on segmental and tonal processing in Mandarin Chinese. *Journal of Cognitive Neuroscience*, 23, 2701–2715.

Mattock, K., & Burnham, D. (2006). Chinese and English infants' tone perception: Evidence for perceptual reorganization. *Infancy*, 10(3), 241–265. doi:10.1207/s15327078in1003_3

Mattock, K., Molnar, M., Polka, L., & Burnham, D. (2008). The developmental course of lexical tone perception in the first year of life. *Cognition*, 106(3), 1367–1381. doi:10.1016/j.cognition.2007.07.002

McDermott, J. H., & Oxenham, A. J. (2008). Music perception, pitch, and the auditory system. *Current Opinion in Neurobiology*, 18(4), 452–463. doi:10.1016/j.conb.2008.09.005

Nan, Y., Sun, Y., & Peretz, I. (2010). Congenital amusia in speakers of a tone language: Association with lexical tone agnosia. *Brain*, 133(9), 2635–2642.

Peretz, I., Champod, A., & Hyde, K. (2003). Varieties of musical disorders. The Montreal battery of evaluation of Amusia. *Annals of the New York Academy of Sciences*, 999(1), 58–75.

Peretz, I., & Coltheart, M. (2003). Modularity of music processing. *Nature Neuroscience*, *6*, 688–691.

Pfordresher, P., & Brown, S. (2009). Enhanced production and perception of musical pitch in tone language speakers. *Attention, Perception and Psychophysics*, 71(6), 1385–1398. doi:10.3758/APP.71.6.1385

Racette, A., Bard, C., & Peretz, I. (2006). Making non-fluent aphasics speak: Sing along! *Brain*, 129(10), 2571–2584. doi:10.1093/brain/awl250

Schlaug, G., Marchina, S., & Norton, A. (2008). From singing to speaking: Why singing may lead to recovery of expressive language function in patients with Broca's aphasia. *Music Perception*, 25, 315–323.

Shen, X. S., & Lin, M. (1991). A perceptual study of Mandarin tones 2 and 3. *Language and Speech*, 34(2), 145–156. doi:10.1177/002383099103400202

So, C. K., & Best, C. T. (2010). Cross-language perception of non-native tonal contrasts: Effects of native phonological and phonetic influences. *Language and Speech*, *53*(2), 273–293. doi:10.1177/0023830909357156

Stevens, C., Keller, P., & Tyler, M. (2004). Language tonality and its effects on the perception of contour in short spoken and musical items. *Proceedings of the 8th International Conference on Music perception and Cognition*, 713–716.

Tillmann, B., Burnham, D., Nguyen, S., Grimault, N., Gosselin, N., & Peretz, I. (2011). Congenital amusia (or tone-deafness) interferes with pitch processing in tone languages. *Frontiers in Psychology*, 2, 120. doi:10.3389/fpsyg.2011.00120

Trainor, L., & Hannon, E. E. (2013). Musical development. In D. Deutsch (Ed.), *The psychology of music* (pp. 425–497). New York: Elsevier.

Trainor, L. J., & Trehub, S. (1992). A comparison of infants' and adults' sensitivity to western musical structure. *Journal of Experimental Psychology: Human Perception and Performance*, 18, 394–402.

Trehub, S. E., Schellenberg, E. G., & Kamenetsky, S. B. (1999). Infants' and adults' perception of music structure. *Journal of Experimental Psychology: Human Perception and Performance*, 25(4), 965–975.

Ullman, M. T., Corkin, S., Coppola, M., Hickok, G., Growdon, J. H., & Koroshetz, W. J. (1997). Neural dissociation within language: Evidence that the mental dictionary is part of declarative memory, and that grammatical rules are processed by the procedural system. *Journal of Cognitive Neuroscience*, 9, 266–276.

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. doi:10.1016/j.lindif.2010.02.004

Werker, J. F., & Tees, R. C. (1984). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development*, 7(1), 49–63. doi:10.1016/S0163-6383(84)80022-3

Whalen, D. H., & Xu, Y. (1992). Information for mandarin tones in the amplitude contour and in brief segments. *Phonetica*, 49 (1), 25–47.

Wilson, S., Parsons, K., & Reutens, D. (2006). Preserved singing in aphasia: A case study of the efficacy of melodic intonation therapy. *Music Perception*, *24*, 23–36.

Wong, P., Ciocca, V., Chan, A., Ha, L., Tan, L., & Peretz, I. (2012). Effects of culture on musical pitch perception. *Plos One*, 7, e33424.

Wong, P. C. M., & Perrachione, T. K. (2007). Learning pitch patterns in lexical identification by native English-speaking adults. *Applied Psycholinguistics*, 28(4), 565–585.

Xu, Y. (1997). Contextual tonal variations in Mandarin. *Journal of Phonetics*, 25(1), 61–83.

Xu, Y., Gandour, J. T., & Francis, A. L. (2006). Effects of language experience and stimulus complexity on the categorical perception of pitch direction. *Journal of the Acoustical Society of America*, 120(2), 1063–1074. Retrieved from http://link.aip.org/link/?JAS/120/1063/1

Yeung, H. H., Chen, K. H., & Werker, J. F. (2013). When does native language input affect phonetic perception? The precocious case of lexical tone. *Journal of Memory and Language*, 68(2), 123–139. doi:10.1016/j.jml.2012.09.004