Native language, L2 experience, and pitch processing in music

Ao Chen,[1,2] Melis Çetinçelik,[3,4]
M. Paula Roncaglia-Denissen,[5] and Makiko Sadakata[3]
[1] Beijing Language and Culture University | [2] Utrecht University |
[3] University of Amsterdam | [4] Max Planck Institute for Psycholinguistics
| [5] Tilburg University

The current study investigated how the role of pitch in one's native language and L2 experience influenced musical melodic processing by testing Turkish and Mandarin Chinese advanced and beginning learners of English as an L2. Pitch has a lower functional load and shows a simpler pattern in Turkish than in Chinese as the former only contrasts between presence and the absence of pitch elevation, while the latter makes use of four different pitch contours lexically. Using the Musical Ear Test as the tool, we found that the Chinese listeners outperformed the Turkish listeners, and the advanced L2 learners outperformed the beginning learners. The Turkish listeners were further tested on their discrimination of bisyllabic Chinese lexical tones, and again an L2 advantage was observed. No significant difference was found for working memory between the beginning and advanced L2 learners. These results suggest that richness of tonal inventory of the native language is essential for triggering a music processing advantage, and on top of the tone language advantage, the L2 experience yields a further enhancement. Yet, unlike the tone language advantage that seems to relate to pitch expertise, learning an L2 seems to improve sound discrimination in general, and such improvement exhibits in non-native lexical tone discrimination.

Keywords: music advantage, tone language, L2, L2 proficiency

1. Introduction

Speech and music are two unique products of the human brain that serve communicative purposes, and are present across all cultures (Patel, 2008). In both, pitch variation plays a fundamental role. Language makes use of pitch to convey meaning, and musical melodies are formed by structuring pitches. Expertise with

https://doi.org/10.1075/lab.20030.che | Published online: 16 July 2021
Linguistic Approaches to Bilingualism ISSN 1879-9264 | E-ISSN 1879-9272

pitch in language has been found to facilitate musical processing, and vice versa. Lexical tones are commonly used to test the transfer effect between language and music processing. In tone languages, lexical tones are pitch variation that distinguishes lexical meaning. A widely cited example is Mandarin Chinese, in which the same syllable /ma/ means *mom, hemp, horse*, and *scold* when it carries a high-level, a low-rising, a low-dipping, and a high-falling tone respectively. Native speakers of a tone language have been found to outperform non-tone language speakers at pitch discrimination as well as music perception tasks (Bidelman, Hutka, & Moreno, 2013; Chen, Liu, & Kager, 2016; Pfordresher & Brown, 2009), and musicianship facilitates lexical tone perception in return (Lee & Hung, 2008; Schellenberg, 2015; Wong, Skoe, Russo, Dees, & Kraus, 2007).

Although the association between tone language and music processing has been supported by a great deal of experimental research, the exact underlying mechanism needs to be further scrutinized. The current study focuses on two questions, first, where does the tone language benefit in music processing originate? And second, how may L2 experience influence musical pitch processing? With regard to the tone language benefit, one hypothesis is that the higher functional load of pitch in tone versus non-tone language may facilitate pitch processing in music (Shaefer & Darcy, 2014). In tone languages, pitch variation alone is sufficient to distinguish lexical meaning while such phonemic function of pitch is absent in non-tone languages. Hence, compared to non-tone language speakers, the need to represent and discriminate the lexical tones may improve sensitivity to certain aspects of pitch (Bradley, 2016), such as interval and contour, which may extend beyond language and exhibit in music processing. Yet interestingly, a positive correlation between the processing of lexical tones and that of musical melodies has only been observed among non-tone language listeners (Chen et al., 2016; Chen, Peter, Wijnen, Schnack, & Burnham, 2018), whereas such correlation was absent for the Chinese listeners. Therefore, it seems that learning the phonemic role of lexical tones enhances processing of musical melody, as the Chinese listeners outperformed the Dutch listeners on music tasks. However, once the tonal categories are well established, the lexical tones are split from general pitch processing due to the phonemic function (Chen, Liu, & Kager, 2016; Mok & Zuo, 2012). The tone language advantage in music processing seems to relate to *learning* the linguistic function of native pitch patterns rather than the *knowledge* of these patterns. The second hypothesis is that phonetic realization rather than phonological contrastivity of pitch patterns *per se* is essential for enhancing music processing. If phonological contrastivity of pitch alone is sufficient for enhancing music processing, then in musical pitch processing tasks, listeners whose native language makes minimal use of pitch to distinguish lexical meaning should exhibit a comparable advantage as observed among native speakers of a canonical tone language where multiple tones are distinguished by different phonetic features such as pitch level and contour (e.g., Cantonese; Bauer & Benedict, 1997; Gandour, Jack, 1983; Li & Thompson, 1989). On the other hand, if multiple tonal contrasts in one's native language are essential for improving musical pitch processing, then speakers whose native language has a rich inventory of lexical tones and make use of multiple phonetic features should outperform those whose native language shows limited lexical pitch variation, and different levels of pitch sensitivity should be observed among native speakers of different tone languages. Similarly, if phonological contrastivity *per se* is sufficient to dissociate the processing of lexical tones and musical pitch, then we would expect a lack of cross-domain correlation as long as the listeners' native language makes use of pitch to distinguish lexical meaning, regardless of whether one or many tonal contrasts are present. On the other hand, if the split can only be achieved when multiple lexical tones are contrasted, then a cross-domain correlation should be observed among speakers whose native language shows limited contrastivity of pitch.

To test whether phonological contrastivity of pitch *per se* is sufficient to trigger advanced processing of music, we compared Mandarin Chinese and Turkish native listeners on their discrimination of short musical phrases. Chinese is a typical tone language with four citation tones and a neutral tone. Lexical tone is obligatory for any Chinese syllable, and most syllables in Chinese can carry any of the four tones to form a real morpheme. The neutral tone is restricted to certain lexical structures and always follows a syllable carrying a citation tone. Turkish makes use of lexical accent when distinguishing word meaning with the default location for primary accent being on the last syllable of the word. In Turkish, lexical meaning can be contrasted by the position of the lexical accent (e.g., final versus non-final accent, [bébek] 'name of a suburb of Istanbul' and [bebék] 'baby', cited from(Levi, 2005)), and importantly, elevated pitch serves as the most prominent cue for marking lexical accent (Levi, 2005). Yet unlike Mandarin Chinese that contrasts between four different pitch patterns lexically for almost every syllable, Turkish lexical accent only contrasts presence versus absence of pitch elevation at a certain position, and it does not distinguish pitch direction or contour differences. In addition, only some words have a minimal pair contrasted by lexical accent. Hence, although Turkish makes use of pitch for distinguishing lexical meaning, it is phonetically simpler than the Chinese lexical tones, and pitch has a lower functional load than in Mandarin Chinese (Shaefer & Darcy, 2014). By comparing these two language groups, we hypothesize that if contrastivity (i.e., lexical meaning can be distinguished by pitch variation) alone is sufficient to enhance pitch perception, then in music processing tasks, Turkish listeners should perform comparably to the Mandarin Chinese listeners, while if a rich tonal inventory (i.e., multiple contrastive pitch patterns) is necessary to trigger such an advantage, Mandarin Chinese listeners should surpass the Turkish listeners.

Turning to the L2 effect on music processing, although many studies reported positive effect of music training on L2 learning, whether L2 learning benefits music processing in return remains largely unknown. For the musicianship induced advantage in speech processing. It has been hypothesized that people with good musical ability are equipped with "good ears" to analyze and discriminate non-native speech sounds (Kempe, Bublitz, & Brooks, 2015; Marie, Delogu, Lampis, Belardinelli, & Besson, 2011; Schellenberg, 2015; Slevc & Miyake, 2006), leading to more accurate L2 perceptual identification and pronunciation. Although the effect of L2 on music processing is under-investigated, it is known that actively learning a second language improves the perception of phonological contrasts in the target language (Bradlow, Akahane-Yamada, Pisoni, & Tohkura, 1999), and experienced learners outperform the inexperienced ones (Flege, Bohn, & Jang, 1997). Physiologically, learning induced neural plasticity has been observed for both the grey and the white matter among adult L2 learners (Callan et al., 2003; Schlegel, Rudelson, & Tse, 2012), and bilingual children and adolescents showed stronger subcortical encoding of speech fundamental frequency than their monolingual counterparts (Krizman, Marian, Shook, Skoe, & Kraus, 2012). Together with the fact that perceptual learning of non-native contrasts show long term retention (Lively, Pisoni, Yamada, Tohkura, & Yamada, 1994), such physiological changes suggest a better wiring of neural resources, which may well facilitate general rather than speech specific auditory performance. Indeed, a second language benefit has been reported in musical rhythm perception. Those who learned a second language (L2) that had different rhythmic properties than their native language outperformed those who learned an L2 with similar rhythmic properties in the discrimination of musical rhythm (Roncaglia-Denissen, Schmidt-Kassow, Heine, Vuust, & Kotz, 2013; Roncaglia-Denissen, Roor, Chen, & Sadakata, 2016), and this discrimination ability seems to be subject to the amount of second language experience (Bhatara, Yeung, & Nazzi 2015). Yet interestingly, Bidelman et al. (2013) found that for native tone language listeners, the later the onset of learning a second language (suggesting more native language input), the better the discrimination between short musical melodies. Hence, experience with a non-tone L2 may hinder native tone language listeners' musical processing. So far however, when testing the tone language advantage in music processing, although native tone language speakers often had knowledge of an L2 (mostly English), whether and how their L2 experience and proficiency may have influenced their performance was largely neglected.

To test the effect of L2 on music processing, we separated the Chinese and the Turkish listeners in to one subgroup with little L2 (English) experience and another highly proficient in English as an L2. With such a manipulation, we aimed to investigate how L2 experience and proficiency may interact with native language in shaping music processing. If L2 enhances auditory acuity in general, then such enhancement is expected to exhibit in the music processing tasks as well, and the advanced L2 learners should outperform the L2 beginners regardless of native language. In contrast, if native language plays a determinant role in music processing, then native Mandarin Chinese listeners should outperform the Turkish listeners irrespective of L2 level. In order to test if L2 advantage (if any) holds true for pitch processing in speech as well, we further tested the Turkish listeners on their discrimination on bisyllabic Chinese lexical tones.

For music processing, we made use of the Musical Ear Test (MET) (Wallentin, Nielsen, Friis-Olivarius, Vuust, & Vuust, 2010) in which participants are tested on their discrimination between pairs of short musical phrases. MET was selected as the tool for two reasons: first, it shows high sensitivity in distinguishing people with different musical ability (e.g., professional versus amateur musicians) which prevents a ceiling effect; second, the melodies retain real life musical structures, melodic as well as rhythmic, which ensures ecological validity. As working memory is involved in these tasks, and as previous studies have shown that tone language listeners outperformed non-tone language listeners on working memory tasks (Bidelman et al., 2013), we further included the a non-word repetition and a backward digit span task to measure phonological and working memory respectively, and we included the performance on these tasks as covariates in the statistical analysis.

2. Materials and methods

2.1 Participants

Fifteen native Turkish speakers with low English proficiency (Turkish beginners, mean age (SD) = 18.93 (1.94) years, 8 women and 7 men), 15 native Turkish speakers with high English proficiency (Turkish advanced learners, mean age (SD) = 24.33 (1.18) years, 9 women and 6 men), 17 native Mandarin speakers with low English proficiency (Chinese beginners, mean age (SD) = 20.59 (1.62) years, 11 women and 6 men), and 16 native Mandarin Chinese speakers with high English proficiency (Chinese advanced learners, mean age (SD) = 25.75 (1.53) years, 13 women and 3 men) participated in the experiment. The Chinese beginners were tested in Beijing, and were either attending a university degree program or had recently graduated. None of the Chinese beginner participants had stayed outside China for more than two weeks and all of them reported not being able to converse in English. The Chinese advanced participants were tested in the Netherlands and were attending or just graduated from English taught bachelor or graduate programs in Dutch universities. All the Chinese speakers were born and grew up in Mandarin in Mainland China. The Turkish L2 beginnning learners were tested in Turkey, and were preparing for university by learning English in a preparatory school at the Middle East Technical University of Ankara, Turkey. Before testing, they had had one or two weeks of English classes and did not have sufficient knowledge of English to converse in it. Fourteen Turkish advanced learners were tested in Turkey; these participants were currently studying in or just finished an English taught program at a university. One Turkish advanced learner was tested in the Netherlands; this participant was attending an English taught program at a Dutch university. Hence, all the advanced L2 learners had abundant experience using English in an academic setting that required oral as well as written proficiency. For all the Chinese and Turkish beginners, courses on English language were obligatory in their curriculum; hence, all of them had some knowledge of English. Yet they had only enrolled in formal education taught solely in Mandarin Chinese/Turkish, and they were not able to communicate in English. All the participants learned English in school, and none of them were simultaneously bilingual. None of the participants reported knowledge of a third language other than English. One Turkish advanced learner failed to participate in the lexical tone discrimination task; hence, analysis of the lexical tone tasks was conducted with the remaining 14 participants, while the analysis of MET and cognitive tests included all Turkish advanced learners. One Chinese beginner participant was excluded due to a technical error, and the data of the remaining 16 participants were analyzed.

2.2 Materials and procedure

Musical Ear Test, the melodic aptitude test

Participants' melodic ability was assessed using the melodic subset of the musical ear test (MET; Wallentin, Nielsen, Friis-Olivarius, Vuust, & Vuust, 2010). The melodic test consisted of 52 pairs of melodic phrases in piano timbre, presenting 3–8 tones, with realistic musical rhythms. The melodies had the duration of one measure and were played at 100 bpm. In different trials (26 pairs), the two phrases differ by one note, and in half of them (13 pairs), such a note change forms pitch contour violation.

Memory tests

The participants' phonological memory was measured by the Mottier test (Mottier, 1951). The Mottier test is a non-word repetition task composed of six sets of six non- words in each, ranging from two to seven syllables for each set. All the syllables in all the words have a CV structure. A female native Turkish and a female native Mandarin Chinese speaker produced the non-words in their native language according to the phonotactic rules of the corresponding language respectively. All the non-words were manipulated afterwards to have a flat tone and comparable intensity.

In the Mottier test, participants heard non-words and were instructed to orally repeat each word as accurately as possible immediately after hearing it. The experimenter recorded whether each syllable was recalled correctly as the participant spoke, and finished the test once the participant failed to recall at least four trials in the same set. The score was computed ad hoc as the total number of correctly recalled non-words with a maximum score of 36 (six sets * six words in each set) correctly recalled non-words per participant.

In the backward digit span (BDS) test, participants heard sequences of numbers and were asked to repeat the sequences orally in reversed order. There were seven sets of two trials, ranging from two to eight digits. The numbers were produced by female native speakers of Turkish and Mandarin Chinese respectively. The experimenter recorded whether each number was correctly recalled as the participant spoke, and finished the test once the participant failed to correctly recall at least one sequence per set. The score was computed ad hoc as the total number of correctly recalled sequences with a maximum score of 14 sequences per participant.

Self-reported language skills and history questionnaire

Self-reported language skills have been shown to correlate highly with objective measures of language skills (Marian, Blumenfeld, & Kaushanskaya, 2007) and were successfully used in previous research to assess individuals' language skills (e.g., Garbin et al., 2011; Roncaglia-Denissen, Schmidt-Kassow, Heine, & Kotz, 2013). The language skills and history questionnaire in the current study is the same as in Roncaglia-Denissen et al. (2015). In this questionnaire, participants rate their second language listening, writing, reading, and speaking skills on a ten-point scale, with 1 being very difficult and 10 being very easy to use the language in that particular mode. Participants also report their age of first and second language first exposure, situations of acquisition, and current use. They were also asked to indicate any other language they knew and rate their proficiency on the same ten-point scales, as well as how often they used that particular language in daily life.

Music background questionnaire

Participants were given a music background questionnaire to assess information about their formal musical training (number of years) and daily exposure to music (hours) (Roncaglia-Denissen et al., 2018). Formal musical training was assessed for each participant in terms of the number of years they attended private lessons to learn an instrument or to learn how to sing. Previous research has shown that both singing and instrument playing shapes auditory perception (e.g. Kraus & Chandrasekaran, 2010), and as we aimed to exclude musicians rather than differentiate musicians with different learning experience, whether the participants learned one or multiple instruments or whether an instrument was learned concurrently with singing lessons was disregarded.

Lexical tone discrimination task

A native female speaker of Mandarin Chinese recorded individual CV syllables carrying the four different lexical tones. Ten different syllables were used as tone bearing units in the bisyllabic discrimination task. For each syllable separately, after normalizing the duration of the syllables, the pitch contours of the dipping tones (T3) were extracted and were used to replace the pitch contour of the rising tone (T2) syllables, so that the T2-T3 pair solely differed in lexical tones. Similarly, the pitch contour of falling tone (T4) was extracted and used to replace the high-level tone (T1), so that T1-T4 pair solely differed in the lexical tones. These monosyllables were concatenated to form the bisyllabic sequences. The two sounds in one trial always had the same CV segments. Both "different" (e.g., /ma_{T1}//ma_{T1}/-/ma_{T1}//ma_{T4}/) and "same" (e.g., /ma_{T1}//ma_{T1}/-/ma_{T1}//ma_{T1}/) trials were included in the experiment. For details, see Chen, Liu, & Kager (2016). After the pair of stimuli was presented auditorily, the participants had one second to give their response by pressing keys labeled as "same" or "different" on a computer keyboard. If they failed to respond within one second, the trial was considered as an incorrect response. Before the experiment started, there were eight practice trials with feedback for the participants to acclimate to the procedure. No feedback was given in the experiment. There were 180 trials in total. It should be emphasized that all the four lexical tones can be produced with large duration variation in real life, and duration normalization of the lexical tones does not lead to unnatural stimuli. Native Chinese listeners had no difficulty at identifying the lexical tones of individual syllables.

For the Turkish speakers, the order of the lexical tone task and the MET was counterbalanced, and the working memory tasks and the questionnaires followed these tasks. For the Chinese speakers, the MET preceded the working memory tasks and the questionnaires.

3. Statistical analysis

First, to test whether the advanced L2 learners were indeed more proficient in English than the beginner groups, we compared the language skills of the participants with separate Kruskal-Wallis tests, using each skill as the dependent variable and language group (Chinese beginners and advanced learners, Turkish beginners and advanced learners) as a between-subjects factor. Next, an ANCOVA was conducted with the MET score as dependent variable, with L2 proficiency and native languages as between-subject variables and the scores of the Mottier and BDS as well as years of formal music training as covariates. For the Turkish listeners, an additional ANOVA was performed with the accuracy of lexical tone discrimination task (correct responses divided by total trial number), with L2 proficiency as a between-subject variable, and with Mottiers and BDS scores as covariates. To determine whether and how native language and L2 experience may have influenced working memory, for the Mottier test and the BDS, a MANOVA were conducted with the scores in both tasks being the dependent variable, and L2 proficiency (beginners versus advanced learners), and native languages (Chinese versus Turkish) as between-subject factors. It should be acknowledged that the advanced L2 learners were older than their beginner counterparts, yet seeing all the participants were in their twenties, it is unlikely that age would significantly influence their performance in the current experiments (Salthouse, 2009); hence, we did not include age as a variable in the analysis.

4. Results and discussion

Table 1 lists the mean MET score, mean Mottier and BDS score, mean years of music training, self-reported levels of English reading, writing, understanding, speaking, and grammar of each language group. None of the participants reported knowing a third language besides the native language and English with skills above 2, and none of them used a third language in their daily life.

For all the measurements of English proficiency, the language group showed a significant effect as expected. Bonferroni corrected post-hoc pairwise comparisons indicated that Chinese and Turkish beginners reported significantly lower English levels than their high proficiency counterparts (all ps < .05), while no difference was found between the two beginner groups or between the advanced

Table 1. Mean MET, Mottier and BDS scores, mean years of music training, and self-reported English level of Turkish beginning L2 learners, Turkish advanced L2 learners, Chinese advanced L2 learners, and Chinese beginning L2 learners

	Turkish beginners	Turkish advanced learners	Chinese beginners	Chinese advanced learners
MET score	28.07(5.15)	35.40(4.21)	37.00(2.42)	40.25(4.06)
Mottier	23.40(4.31)	26.13(5.14)	30.06(3.31)	30.25(4.71)
BDS	6.60(2.10)	8.67(2.35)	12.25(1.95)	11.81(3.02)
Music training (years)	1.20(1.57)	1.67(1.23)	0.19(0.54)	0.48(0.95)
L2 age	11.75(4.83)	8.8(2.51)	8.94(2.62)	10.81(2.69)
L2 reading	2.53(1.73)	7.87(1.30)	4.88(2.16)	8.31(1.35)
L2 writing	2.20(1.52)	8.13(1.55)	5.12(2.39)	7.69(1.25)
L2 understanding	3.07(2.22)	8.13(1.60)	3.94(2.29)	8.00(1.54)
L2 speaking	2.47(1.81)	7.87(1.64)	4.31(2.02)	7.81(1.22)
L2 grammar	3.60(2.95)	8.00(1.77)	4.75(2.11)	8.19(1.38)

groups $H_{\text{English reading}}(3) = 42.71$, (results). $p < .001$, $H_{\text{English writing}}(3) = 39.41$, $p < .001$, $H_{\text{English understanding}}(3) = 33.45$, $p < .001$, $H_{\text{English speaking}}(3) = 37.94$, $p < .001$, $H_{\text{English grammar}}(3) = 27.27$, $p < .001$. Thus, the advanced learners were more proficient in English than the beginners, while the Chinese- and Turkish advanced learners were comparable for their English levels, as well as the beginner groups.

For MET, the ANCOVA showed a significant main effect of native language $F(1, 55) = 15.60$, $p < .001$, partial $\eta^2 = 0.22$ as well as a main effect of L2 proficiency $F(1, 55) = 19.69$, $p < .001$, partial $\eta^2 = 0.26$. The interaction between native language and L2 proficiency was not significant, $F(1, 55) = 1.95$, $p = .17$. Neither Mottier nor BDS score showed a significant effect, $F_{\text{Mottier}}\ (1, 55) = 0.94$, $p = .34$, $F_{\text{BDS}}(1, 55) = 1.84$, $p = .18$. Years of musical training showed a significant effect, $F(1, 55) = 4.32$, $p = .04$, partial $\eta^2 = 0.07$. Figure 1 plots the mean score of the MET melodic test of the Chinese and Turkish advanced and beginning learners. As can be seen from the figure, the native Chinese speakers outperformed the Turkish speakers, and the advanced L2 learners outperformed the beginners irrespective of native language. Hence, although pitch signals lexical accent in Turkish, the Turkish listeners were not as accurate as the Chinese listeners in discriminating the melodies in the MET. In fact, compared to the Dutch listeners (proficient in English as an L2) in Chen et al. (2016) whose native language did not rely on pitch in contrasting lexical meaning, the Turkish advanced L2 learners in the current study did not obtain higher scores in MET ($M = 36.57$, $SD = 4.97$ for the Dutch participants, and $M = 35.40$, $SD = 4.21$ for the Turkish advanced L2 learners). L2 proficiency, on the other hand, facilitated musical pitch processing for both the Chinese and the Turkish speakers. Therefore, L2 proficiency played an additive effect on top of the native language advantage in musical melodic processing. Although the participants had limited musical training, it still showed significant effect, suggesting that the training effect was well captured by the MET. In particular, although the Turkish listeners had longer music training than the Chinese listeners, $F(1, 60) = 14.71$, $p < .001$, they still had lower MET scores than the Chinese listeners. Hence, native language seems to have a stronger impact than music experience on music processing, at least when such experience is limited.

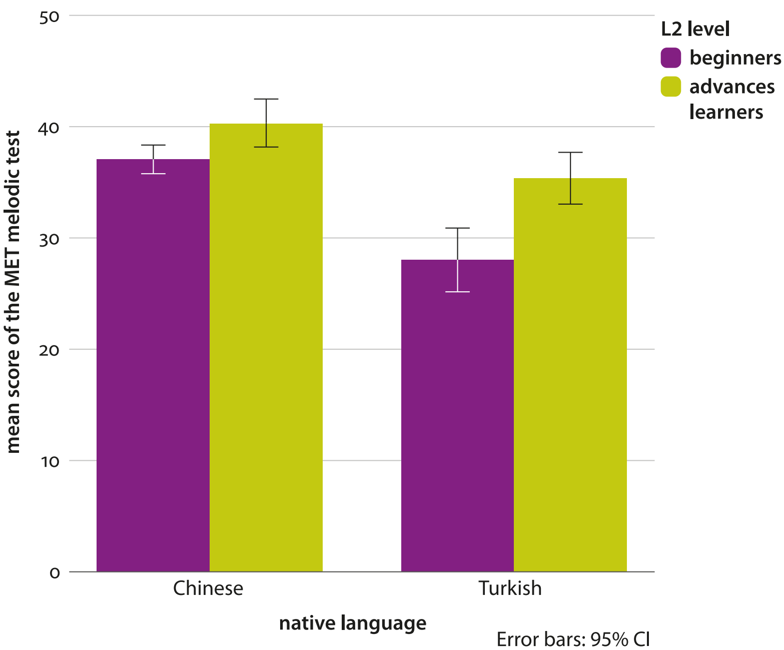

Figure 1. Mean score of the MET melodic aptitude test of the Chinese and Turkish beginners and advanced learners

The MANOVA conducted with the Mottier and BDS score showed a significant main effect of native language but not L2 experience. $F_{\text{mottier native language}}$ $F(1, 58) = 23.14$, $p < .001$, partial $\eta^2 = 0.29$, $F_{\text{mottier L2 experience}}$ $F(1, 58) = 1.70$, $p = .20$, $F_{\text{BDS native language}}$ $(1, 58) = 52.26$, $p < .001$, partial $\eta^2 = .47$, $F_{\text{BDS L2 experience}}$ $(1, 58) = 1.79$, $p = .19$. For both the Mottier and the BDS tasks, the Chinese participants outperformed the Turkish participants, regardless of L2 experience. Interaction between native language and L2 experience was only significant for BDS, $F(1,58) = 4.24$, $p = .044$, but not for Mottier, $F(1,58) = 1.29$, $p = .26$. As can be seen from Table 1., the Turkish advanced learners obtained higher scores than the Turkish beginners in the BDS tasks, while the Chinese beginners and Chinese advanced learners showed comparable performance. These findings were inconsistent with some previous studies (Bialystok, Craik, Green, & Gollan, 2009; Paap & Greenberg, 2013; Paap, Johnson, & Sawi, 2015); we did not find an L2 benefit in phonological or working memory tasks. Instead, speaking Chinese natively showed a facilitatory effect in such tasks. Nevertheless, neither the performance in phonological nor working memory tasks was significantly related to the accuracy of the MET melodic test. Meanwhile, although our participants only received limited musical training, the years of musical training were still significantly related to the MET melodic test performance. Such a finding was expected seeing that MET has been demonstrated to be sensitive at capturing the difference in musical expertise as well as the amount of practice (Wallentin et al., 2010).

With regard to the lexical tone discrimination of the Turkish listeners, the univariate ANOVA with the accuracy of lexical tone discrimination as a dependent variable and the L2 experience as independent variable with the years of musical training, the Mottier score, and BDS score as covariates showed a significant main effect of L2 experience, $F(1,24) = 7.14$, $p = .013$. Years of musical training were also significantly related to the lexical tone discrimination accuracy, $F(1,24) = 10.37$, $p = .004$. BDS and Mottier score, on the other hand, was not significantly related to the lexical tone discrimination accuracy, $F_{\text{Mottier}}\ (1,24) = 3.40$, $p = .078$, $F_{\text{BDS}}\ (1,24) = 1.92$, $p = .18$. Similar to what we found in the MET melodic test for the Turkish listeners, the advanced L2 learners outperformed the beginners. However, such enhancement did not seem to be explained by the improved working memory of the advanced learners.

In addition, although the musical training only lasted a very short time (hence, without much individual variation) among the Turkish listeners, it still related to the lexical tone discrimination significantly. As in Chen, Liu, & Kager (2016), we further examined whether the performance in the lexical tone discrimination and MET melodic test was correlated among the Turkish listeners. A significant Pearson's r was found between the accuracy of lexical tone discrimination and the score of MET melodic test, $r = .84$, $p < .001$. Figure 2 plots the lexical tone discrimination accuracy against the MET melodic test score for each individual Turkish participant. Taken together, musical training was positively related to lexical tone discrimination accuracy, and the performance in the lexical tone discrimination task significantly correlated with that of the MET melodic test. These findings indicate unified processing of the lexical tones and the musical melodies among the Turkish listeners. In other words, unlike Chinese listeners who exhibit a split between the discrimination of the lexical tones and that of the musical melodies, the Turkish listeners are more like non-tone language (Dutch) listeners in showing shared processing between musical melody and lexical tones.

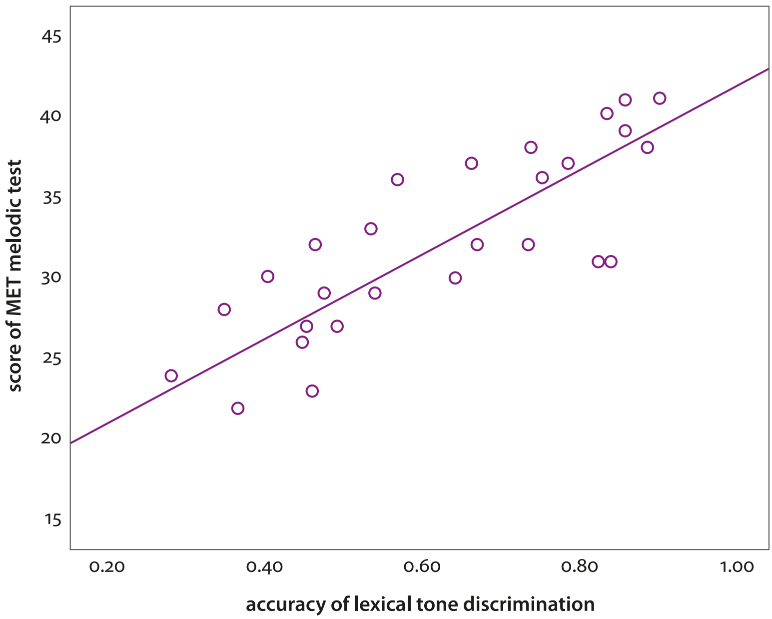

Figure 2. Correlation between the score of the MET melodic test and the accuracy of lexical tone discrimination of the Turkish listeners

5. General discussion

In the current study, we tested Chinese and Turkish advanced and beginning L2 learners of English on their discrimination of musical melodic phrases. We found that for the MET melodic test, the Chinese listeners outperformed the Turkish listeners, and the advanced L2 learners outperformed the beginners without interaction between the two factors. We further tested the Turkish listeners on their discrimination of the Chinese lexical tones, and again, the advanced L2 learners outperformed the beginning learners. In addition, a significant positive correlation between the score of the MET melodic test and the accuracy of lexical tone discrimination was observed for the Turkish listeners.

Mandarin Chinese is a typical tone language, and for each syllable, pitch alone (i.e., without co-occurring with other acoustical cues) is sufficient to distinguish lexical meaning. The Mandarin Chinese lexical tones are obligatory for all the syllables and are defined by multiple acoustical features, such as fundamental frequency (fo) level, fo contour, fo direction, fo turning point, etc. (Gandour, J. & Harshman, 1978; Moore & Jongman, 1997). In Turkish, by contrast, lexical accent was marked by presence or absence of pitch elevation (Levi, 2005). Therefore, for both Mandarin Chinese and Turkish, pitch is contrastive lexically, but Mandarin Chinese lexical tones make use of more phonetic features than Turkish, and pitch has a higher functional load in Chinese (Shaefer & Darcy, 2014). If the lexical contrastivity of pitch plays a determinant role in shaping musical processing, then the Turkish listeners should have performed comparably to the Chinese listeners in the MET melodic test. Yet, this is not what we found. Instead, the Turkish listeners had significantly lower scores than the Chinese listeners, and their performance was no better than the Dutch listeners in Chen, Liu, & Kager (2016), whose native language does not rely on pitch to contrast lexical meaning. These findings suggest different levels of pitch sensitivity as the result of different native tone languages. In addition, for the Turkish listeners, we found a significant positive correlation between the score of the MET melodic test and the accuracy of lexical tone discrimination, which is again similar to the Dutch rather than the Mandarin listeners (Chen, Liu, & Kager, 2016). Therefore, the music processing advantage does not seem to rely on *whether* pitch is used contrastively in one's native language, but *how* it is used. Compared to canonical tone languages such as Mandarin Chinese, the pitch accent in Turkish only distinguishes between presence and absence of elevated pitch. It is plausible that obligatory pitch for syllables as observed in canonical tone language is necessary for developing pitch processing expertise, which could transfer to music. Alternatively, phonetic realization of the lexical tones may play a role. Speakers of languages with a rich tonal inventory (e.g., both contour and level tones) may benefit more than speakers of those with phonetically simple tones (e.g., only level tones). Similarly, how pitch variation is categorized phonologically (e.g., many or few tone categories) in one's native language may also influence pitch processing in music. It would be interesting for future studies to investigate how different realization of pitch at the lexical level impact musical pitch processing.

Our finding is consistent with many previous studies in showing the tone language advantage in music processing (Bidelman, Gandour, & Krishnan, 2011; Bidelman et al., 2013; Deutsch, Henthorn, Marvin, & Xu, 2006; Pfordresher & Brown, 2009). The new finding of the current study is that, on top of the tone language benefit, there was an L2 benefit. Even for the Chinese listeners, and even if their L2, English, is a non-tone language, the advanced L2 learners outperformed the beginning learners in the MET melodic test. Therefore, with regard to the transfer effect from language to music, pitch expertise does not seem to be the only factor at play. The reverse transfer from music to language has been commonly found. Musicians outperform non-musicians in various speech processing tasks, whether the stimuli had a musical counterpart (e.g., pitch or rhythmic variation) or not (e.g., phonemes) (Delogu, Lampis, & Belardinelli, 2006; Marie et al., 2011; Milovanov, Huotilainen, Välimäki, Esquef, & Tervaniemi, 2008). One hypothesis about the music-to-language transfer effect involves shared acoustic features of the two domains, such as frequency, rhythm and spectral variation; expertise with these features in music may bring general auditory benefits and enhance speech processing (Kempe et al., 2015). Such a hypothesis may also hold true for the language-to-music transfer. Proficiency in an L2 requires representation and discrimination of tens of L2 phonemes that may or may not overlap with the native ones. Confronted with a more crowded acoustical space, advanced L2 learners may become more capable of capturing acoustic differences between speech sounds than monolingual or beginning learners. The need to discriminate between the L2 sounds can lead to structural and functional reorganization of the brain (Callan et al., 2003; Mechelli et al., 2004; Osterhout et al., 2008; Schlegel et al., 2012), and the "better equipped" brain may perform other auditory tasks, such as music, more easily. In addition, L2 prosody differs from the native one both at word and phrasal levels, and advanced L2 learners encounter more linguistic pitch variations than beginners. Such experience may facilitate pitch processing. In the current study, we found that L2 induced enhancement in musical melody processing does not necessarily originate from learning (more) *lexical tones* other than those present in one's L1. Instead, a non-tone L2 is sufficient to trigger such enhancement. Yet it should be acknowledged that although English is not a tone language, it exhibits rich pitch patterns (e.g., lexical stress and multiple nucleus tones, (Pierrehumbert, 1980)), and learning an L2 with plenty of pitch variation may be beneficial for music processing. It would be interesting for future study to test whether learning a tonal L2 would be more advantageous than learning a non-tonal L2 for pitch processing.

The fact that advanced L2 learners outperformed the beginners in the MET melodic test cannot be attributed to enhanced working memory of the former. First, L2 experience failed to show a significant effect for either the Mottier or the BDS scores, suggesting comparable phonological and working memory between the advanced and beginning L2 learners. Second, neither the Mottier score nor the BDS score was significantly related to the MET score. Although short term memory must be at play for the MET since listeners could only compare and discriminate between the two melodies if they were able to hold both in storage, we did not find evidence that the phonological and working memory tested with the current tasks were related to MET performance in a significant way. Similar to the findings of Bidelman et al. (2013), we also found that native Chinese listeners outperformed their Turkish counterparts in Mottier and BDS. The better performance of the Chinese listeners might be due to Chinese phonology. Syllable plays a highly salient role in Mandarin Chinese (in Cantonese as well). Syllables in Mandarin Chinese are phonetically simple: consonant clusters are not allowed and the coda can only be /n/ or /ŋ/ if present. In addition, individual syllables correspond to morphemes, which are the basic units for forming words and phrases, and orthographically, syllables corresponds to characters, which are the basic units for writing. Such clear and writing-consistent syllabification might work as a helpful rhythmic cue for the listeners to recall auditory strings. Alternatively, as some digits in Turkish are bisyllabic while all the digits are monosyllabic in Mandarin Chinese, the larger number of syllables in Turkish BDS task may have led to the worse performance of the Turkish listeners.

Another hypothesis for the L2 advantage in music processing might relate to general aptitude, which may be auditorily based. Those who are born with "good ears" to perceive and discriminate sounds may learn a second language more easily than those without, hence have a better chance to reach a high proficiency in their L2. Consequently, they might be more likely to participate in educational programs taught in their L2. Such high auditory acuity may also be reflected in musical processing. Aptitude can also involve attention, memory, ease with learning, openness to experience, and/or other cognitive abilities and personalities. People with such endowments may be more prone to engage in L2 learning and reach high proficiency, and their aptitude might exhibit in music processing as well, even without much musical training.

To conclude, pitch expertise that can be employed in music processing does not seem to be sufficiently triggered by learning natively where pitch is used constrastively at the lexical level. Tonal richness, instead, seems to be essential. Yet at the same time, experience with an L2 that does not use pitch contrastively improved music processing, suggesting a general auditory benefit as a result of L2 learning that is not restricted to speech. Different mechanisms may be at play with regard to how native language and L2 influences music processing. It would be interesting for future studies to compare tone and non-tone language listeners on their discrimination of tones of a third tone language, and investigate whether the pitch benefit can be observed in speech as well.

Funding

The study was partly supported by National Office for Philosophy and Social Sciences of China under project 17CYY041 granted to the first author.

Acknowledgements

We thank Drikus Roor for his help on data collection, and we thank Piet van Tuijl and Likan Zhan for their help on statistical analysis. We also thank the reviewers for their valuable comments on the study.

References

Bauer, R. S., & Benedict, P. K. (1997). *Modern Cantonese phonology*. Mouton de Gruyter. https://doi.org/10.1515/9783110823707

Bhatara, A., Yeung, H. H., & Nazzi, T. (2015). Foreign language learning in french speakers is associated with rhythm perception, but not with melody perception. *Journal of Experimental Psychology: Human Perception and Performance*, *41*(2), 277–282.

Bialystok, E., Craik, F. I. M., Green, D. W., & Gollan, T. H. (2009). Bilingual minds. *Psychological Science in the Public Interest*, *10*(3), 89–29. https://doi.org/10.1177/1529100610387084

Bidelman, G., Gandour, J., & Krishnan, A. (2011). Cross-domain effects of music and language experience on the representation of pitch in the human auditory brainstem. *Journal of Cognitive Neuroscience*, *23*, 425–434. https://doi.org/10.1162/jocn.2009.21362

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. https://doi.org/10.1371/journal.pone.0060676

Bradley, E. D. (2016). Phonetic dimensions of tone language effects on musical melody perception. *Psychomusicology: Music, Mind, and Brain*, *26*(4), 337. https://doi.org/10.1037/pmu0000162

Bradlow, A. R., Akahane-Yamada, R., Pisoni, D. B., & Tohkura, Y. (1999). Training Japanese listeners to identify English /r/and /l/: Long-term retention of learning in perception and production. *Perception & Psychophysics*, *61*(5), 977–985. https://doi.org/10.3758/BF03206911

Callan, A. M., Callan, D. E., Tajima, K., Kubo, R., Masaki, S., & Akahane-Yamada, R. (2003). Learning-induced neural plasticity associated with improved identification performance after training of a difficult second-language phonetic contrast. *Neuroimage*, *19*(1), 113–124. https://doi.org/10.1016/S1053-8119(03)00020-X

Chen, A., Liu, L., & Kager, R. (2016). Cross-domain correlation in pitch perception, the influence of native language. *Language, Cognition and Neuroscience*, *31*(6), 751–760. https://doi.org/10.1080/23273798.2016.1156715

Chen, A., Peter, V., Wijnen, F., Schnack, H., & Burnham, D. (2018). Are lexical tones musical? native language's influence on neural response to pitch in different domains. *Brain and Language*, *180–182*, 31–41. https://doi.org/10.1016/j.bandl.2018.04.006

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. https://doi.org/10.1007/s10339-006-0146-7

Deutsch, D., Henthorn, T., Marvin, E., & Xu, H. (2006). Absolute pitch among American and Chinese conservatory students: Prevalence differences, and evidence for a speech-related critical period. *Journal of Acoustic Society of America*, *119*(2), 719–722. https://doi.org/10.1121/1.2151799

Flege, J. E., Bohn, O.-S., & Jang, S. (1997). Effects of experience on non-native speakers' production and perception of English vowels. *Journal of Phonetics*, *25*(4), 437–470. https://doi.org/10.1006/jpho.1997.0052

Gandour, J., & Harshman, R. (1978). Cross-language difference in tone perception: A multidimensional scaling investigation. *Language and Speech*, *21*, 1–33. https://doi.org/10.1177/002383097802100101

Gandour, J. (1983). Tone perception in far eastern languages. *Journal of Phonetics*, *11*, 149–176. https://doi.org/10.1016/S0095-4470(19)30813-7

Garbin, G., Costa, A., Sanjuan, A., Forn, C., Rodriguez-Pujadas, A., Ventura, N., … Ávila, C. (2011). Neural bases of language switching in high and early proficient bilinguals. *Brain and Language*, *119*(3), 129–135. https://doi.org/10.1016/j.bandl.2011.03.011

Kempe, V., Bublitz, D., & Brooks, P. J. (2015). Musical ability and non-native speech-sound processing are linked through sensitivity to pitch and spectral information. *British Journal of Psychology*, *106*(2), 349–366. https://doi.org/10.1111/bjop.12092

Kraus, N., & Chandrasekaran, B. (2010). Music training for the development of auditory skills. *Nature Reviews Neuroscience*, *11*(8), 599–605. https://doi.org/10.1038/nrn2882

Krizman, J., Marian, V., Shook, A., Skoe, E., & Kraus, N. (2012). Subcortical encoding of sound is enhanced in bilinguals and relates to executive function advantages. *Proceedings of National Academy of Science USA*, *109*(20), 7877. Retrieved from http://www.pnas.org/content/109/20/7877.abstract. https://doi.org/10.1073/pnas.1201575109

Lee, C. Y., & Hung, T. H. (2008). Identification of mandarin tones by English-speaking musicians and nonmusicians. *Journal of the Acoustical Society of America*, *124*, 3235–3248. https://doi.org/10.1121/1.2990713

Levi, S. V. (2005). Acoustic correlates of lexical accent in Turkish. *Journal of the International Phonetic Association*, *35*(1), 73–97. https://doi.org/10.1017/S0025100305001921

Li, C., & Thompson, S. A. (1989). *Mandarin Chinese: A functional reference grammar*. University of California Press.

Lively, S. E., Pisoni, D. B., Yamada, R. A., Tohkura, Y., & Yamada, T. (1994). Training Japanese listeners to identify English /r/ and /l/. III. long-term retention of new phonetic categories. *The Journal of the Acoustical Society of America*, *96*(4), 2076–2087. https://doi.org/10.1121/1.410149

Marian, V., Blumenfeld, H. K., & Margarita, K. (2007). The language experience and proficiency questionnaire (LEAP-Q): Assessing language profiles in bilinguals and multilinguals. *Journal of Speech, Language, and Hearing Research*, *50*(4), 940–967. https://doi.org/10.1044/1092-4388(2007/067)

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. https://doi.org/10.1162/jocn.2010.21585

Mechelli, A., Crinion, J. T., Noppeney, U., O'Doherty, J., Ashburner, J., Frackowiak, R. S., & Price, C. J. (2004). Structural plasticity in the bilingual brain. *Nature*, *431*(7010), 757. https://doi.org/10.1038/431757a

Milovanov, R., Huotilainen, M., Välimäki, V., Esquef, P.A.A., & Tervaniemi, M. (2008). Musical aptitude and second language pronunciation skills in school-aged children: Neural and behavioral evidence. *Brain Research*, *1194*, 81–89. https://doi.org/10.1016/j.brainres.2007.11.042

Mok, P.K.P., & Zuo, D. (2012). The separation between music and speech: Evidence from the perception of cantonese tones. *The Journal of the Acoustical Society of America*, *132*(4), 2711–2720. https://doi.org/10.1121/1.4747010

Moore, C.B., & Jongman, A. (1997). Speaker normalization in the perception of mandarin chinese tones. *The Journal of the Acoustical Society of America*, *102*(3), 1864–1877. Retrieved from http://link.aip.org/link/?JAS/102/1864/1. https://doi.org/10.1121/1.420092

Osterhout, L., Poliakov, A., Inoue, K., McLaughlin, J., Valentine, G., Pitkanen, I., … Hirschensohn, J. (2008). Second-language learning and changes in the brain. *Journal of Neurolinguistics*, *21*(6), 509–521. https://doi.org/10.1016/j.jneuroling.2008.01.001

Paap, K.R., & Greenberg, Z.I. (2013). There is no coherent evidence for a bilingual advantage in executive processing. *Cognitive Psychology*, *66*(2), 232–258. https://doi.org/10.1016/j.cogpsych.2012.12.002

Paap, K.R., Johnson, H.A., & Sawi, O. (2015). Bilingual advantages in executive functioning either do not exist or are restricted to very specific and undetermined circumstances. *Cortex*, *69*, 265–278. https://doi.org/10.1016/j.cortex.2015.04.014

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

Pierrehumbert, J.B. (1980). The phonology and phonetics of English intonation. [Doctoral dissertation]. Massachusetts Institute of Technology.

Roncaglia-Denissen, M., Bouwer, F.L., & Honing, H. (2018). Decision making strategy and the simultaneous processing of syntactic dependencies in language and music. *Frontiers in Psychology*, *9*, 38. https://doi.org/10.3389/fpsyg.2018.00038

Roncaglia-Denissen, M., Roor, D., Chen, A., & Sadakata, M. (2016). The enhanced musical rhythmic perception in second language learners. *Frontiers in Human Neuroscience*, *10*, 288. https://doi.org/10.3389/fnhum.2016.00288

Roncaglia-Denissen, M., Schmidt-Kassow, M., Heine, A., Vuust, P., & Kotz, S.A. (2013). Enhanced musical rhythmic perception in turkish early and late learners of german. *Frontiers in Psychology*, *4*, 645. https://doi.org/10.3389/fpsyg.2013.00645

Roncaglia-Denissen, M., Schmidt-Kassow, M., & Kotz, S.A. (2013). Speech rhythm facilitates syntactic ambiguity resolution: ERP evidence. *Plos One*, *8*(2), e56000. https://doi.org/10.1371/journal.pone.0056000

Salthouse, T.A. (2009). When does age-related cognitive decline begin? *Neurobiology of Aging*, *30*(4), 507–514. https://doi.org/10.1016/j.neurobiolaging.2008.09.023

Schellenberg, E.G. (2015). Music training and speech perception: A gene-environment interaction. *Annals of the New York Academy of Sciences*, *1337*, 170–177. https://doi.org/10.1111/nyas.12627

Schlegel, A.A., Rudelson, J.J., & Tse, P.U. (2012). White matter structure changes as adults learn a second language. *Journal of Cognitive Neuroscience*, *24*(8), 1664–1670. https://doi.org/10.1162/jocn_a_00240

Shaefer, V., & Darcy, I. (2014). Lexical function of pitch in the first language shapes cross-linguistic perception of Thai tones. *Laboratory Phonology*, *5*, 489–522. https://doi.org/10.1515/lp-2014-0016

Slevc, L. R., & Miyake, A. (2006). Individual differences in second-language proficiency: Does musical ability matter? *Psychological Science*, *17*(8), 675–681. https://doi.org/10.1111/j.1467-9280.2006.01765.x

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. https://doi.org/10.1016/j.lindif.2010.02.004

Wong, C. M. P., Skoe, E., Russo, N. M., Dees, T., & Kraus, N. (2007). Musical experience shapes human brainstem encoding of linguistic pitch patterns. *Nature Neuroscience*, *10*, 420–422. https://doi.org/10.1038/nn1872

Address for correspondence

Ao Chen
School of Communication Science
Beijing Language and Culture University
Utrecht Institute of Linguistics
Utrecht University
room 1235, Xueyuan Road 15
100083 Beijing
China
irischen71@hotmail.com

Co-author information

Melis Çetinçelik
Institute for Logic, Language and Computation
University of Amsterdam
Max Planck Institute for Psycholinguistics
meliscetincelik@gmail.com

M. Paula Roncaglia-Denissen
Department of Cognitive Science & Artificial Intelligence, Tilburg School of Humanities and Digital Sciences
Tilburg University
M.P.Roncaglia@tilburguniversity.edu

Makiko Sadakata
Institute for Logic, Language and Computation
Musicology department
University of Amsterdam
m.sadakata@uva.nl

Publication history

Date received: 27 April 2020
Date accepted: 7 June 2021
Published online: 16 July 2021