JASA EXPRESS LETTERS

https://doi.org/10.1121/10.0001179

Perception of musical melody and rhythm as influenced by native language experience

Linjun Zhang,[1,a)] Songcheng Xie,[1] Yu Li,[2] Hua Shu,[3] and Yang Zhang[4]

[1]*Beijing Advanced Innovation Center for Language Resources and College of Chinese Studies, Beijing Language and Culture University, Beijing 100083, China*

[2]*Department of Applied Psychology, Beijing Normal University–Hong Kong Baptist University United International College (UIC), Zhuhai, Guangdong 519087, China*

[3]*National Key Laboratory of Cognitive Neuroscience and Learning, Beijing Normal University, Beijing 100875, China*

[4]*Department of Speech-Language-Hearing Sciences and Center for Neurobehavioral Development, University of Minnesota, Minneapolis, Minnesota 55455, USA*

zhanglinjunb@sina.com, absenteedec@163.com, charlesyuli@sina.com, shuhua@bnu.edu.cn, zhang470@umn.edu

Abstract: This study used the Musical Ear Test [Wallentin, Nielsen, Friis-Olivarius, Vuust, and Vuust (**2010**). Learn. Indiv. Diff. **20**, 188–196] to compare musical aptitude of native Japanese and Chinese speakers. Although the two groups had similar overall accuracy, they showed significant differences in subtest performance. Specifically, the Chinese speakers outperformed the Japanese counterparts on the melody subtest, but the reverse was observed on the rhythm subtest. Within-group comparisons revealed that Chinese speakers performed better on the melody subtest than the rhythm subtest, while Japanese speakers showed an opposite trend. These results indicate that native language pitch and durational patterns of the listener can have a profound effect on the perception of music melody and rhythm, respectively, reflecting language-to-music transfer of learning.

[Editor: Brenda L. Lonsbury-Martin] Pages: EL385–EL390

Received: 2 March 2020 **Accepted:** 10 April 2020 **Published Online:** 4 May 2020

1. Introduction

Speech and music are distinctive and universal elements of human cultures with both composed of complex sound sequences. Two main acoustic features shared by speech and music are structured pitch and durational patterns. Pitch patterning is fundamental to speech prosody and music melody. In speech, pitch plays language-universal roles in marking pragmatic meanings such as sentence modality, prosodic structure, and emotion (Ladd, 1996; Tong *et al.*, 2005), and furthermore, it is also used in tonal (e.g., Chinese) and pitch-accent (e.g., Japanese) languages to distinguish lexical meanings from otherwise identical strings of phonemes (Wang, 1973; Beckman and Pierrehumbert, 1986). In music, each note corresponds to a specific frequency value (e.g., A4 and A5 corresponds to 440 and 880 Hz, respectively) and successive notes form a melody. Durational patterning forms the basis of rhythmic organization of both speech and music. In speech, durational variability in adjacent vowel and consonant sequences are the primary acoustic correlates of rhythm, according to which languages are usually classified into three distinct categories, i.e., stress-timed (e.g., Dutch and English), syllable-timed (e.g., French and Chinese), and mora-timed (e.g., Japanese and Tamil) (Ramus *et al.*, 1999; Grabe and Low, 2002; Lee *et al.*, 2006; Lin and Wang, 2007). In music, contrastive patterns of successive long and short notes form the rhythmic structures regardless of absolute duration of each note.

Because of the similarities in acoustic features, prosodic arrangements of pitch and duration at the suprasegmental level can be regarded as "musical" aspects of the speech signal. Naturally, the relationship between language and music is of inherent interest to many scholars across a number of disciplines from linguistics to musicology with increasing evidence in support of some overlapping mechanisms for music and language processing (Palmer and Kelly, 1992; Wallin *et al.*, 2001; Patel and Daniele, 2003; Falk, 2004; Patel *et al.*, 2006). For example, musical experience contributes to the development of language and literacy skills in children and second language learners (Anvari *et al.*, 2002; Slevc and Miyake, 2006; Schön *et al.*, 2008), and tonal language experience increases the prevalence of absolute pitch among musicians (Deutsch *et al.*,

[a)]Author to whom correspondence should be addressed.

2006). Theoretical accounts based on a plethora of behavioral and neural data have been proposed to explain why musical training with increased sensitivity to pitch and rhythm information is beneficial for many linguistic as well as socio-cognitive skills (e.g., Koelsch, 2011; Patel, 2011).

At the perceptual level, the language-to-music transfer effects have been much less investigated but nonetheless documented in some empirical studies involving tonal versus non-tonal language comparisons. For example, Giuliano *et al.* (2011) demonstrated superiority of Mandarin speakers in detecting small pitch changes and interval distances over the control group, which was marked by an earlier event-related potential response to the detection of pitch changes. Wong *et al.* (2012) examined the perception of music melody and rhythm by native Cantonese speakers compared with Canadian French and English speakers by the use of Montreal Battery of Evaluation of Amusia (MBEA) and found that Cantonese speakers had better melody perception ability than the other two groups. The superiority of tonal language speakers in music melody perception is confirmed by Bradley (2016) as well as our recent study (Zhang *et al.*, 2019), which, respectively, used the Musical Ear Test (MET) (Wallentin *et al.*, 2010) to examine musical aptitude of native Yoruba and Chinese speakers. Perception of music rhythm was also examined in Wong *et al.* (2012) and our previous study (Zhang *et al.*, 2019) with no significant difference found between speakers of languages with distinct rhythmic properties.

Thus far, the language-to-music transfer effects seem confined to pitch perception with reported advantages for tonal language users [but see Peretz *et al.* (2011) for a disadvantage in tonal-language speakers for discriminating falling pitches in tone sequences as compared with non-tonal language speakers]. Intriguingly, studies on musicians with different language backgrounds indicated that speech rhythm and pitch patterns of a composer's native language influence the rhythmic and melodic properties of their music, respectively (Patel and Daniele, 2003; Patel *et al.*, 2006). As previous studies on pitch and rhythm perception have primarily compared tonal versus intonational and stressed-timed versus syllable-timed languages, there is a need for testing languages with pitch-accent and mora-timed linguistic prosodic properties in the comparative approach to gain insights about how native language experience may alter musical melody and rhythm perception.

In the current study, we recruited native Chinese and Japanese speakers and employed MET to compare their melody and rhythm discrimination abilities. Although both Chinese and Japanese make use of pitch patterns to signify word meanings, there are only high versus low pitch patterns in Japanese in contrast to Chinese in which pitch direction and slope play important roles in lexical tone perception (Wang, 1973; Massaro *et al.*, 1985; Lee *et al.*, 2006). With regards to speech rhythm, the two languages belong to distinct categories: Chinese is syllable-timed and Japanese is mora-timed (Ramus *et al.*, 1999; Grabe and Low, 2002; Lin and Wang, 2007). We aimed to determine whether the two selected groups of participants who spoke languages with distinct pitch and durational prosodic properties would show significant differences in musical melody and rhythm perception. In particular, we were interested in examining possible between-group differences in each subtest as well as within-group differences between the subtest scores, which could provide corroborative evidence for the language-specific influences on music perception.

2. Methods

2.1 Participants

Thirty-one native Chinese (13 females; age range: 19–21 years) and 28 native Japanese (12 females; age range: 18–20 years) participants with no formal musical training took part in the present study. The native Chinese speakers were undergraduate students from Beijing Language and Culture University. The native Japanese speakers were freshmen who just enrolled in the same university for a 4-year bachelor's degree program and did not have any Chinese learning experience at the time of the perceptual tests for the study. Both native Chinese and Japanese speakers were late English learners (onset age >8 years) with self-reported intermediate proficiency level, and none had lived in an English-speaking country for over 2 weeks. They all passed a hearing screening at 25 dB hearing level for octave frequencies between 125 and 4000 Hz. No significant difference was found between the two groups in the hearing thresholds at any frequency. None of the participants had a history of neurological, psychiatric, or neuropsychological disorders. Written informed consent was obtained from each participant. This study was approved by the Institutional Review Board of Beijing Language and Culture University.

2.2 Musical aptitude measurement

Musical aptitude was measured by administering MET, a standardized test of music perception developed by Wallentin *et al.* (2010). MET focuses on the two fundamental aspects of music,

https://doi.org/10.1121/10.0001179

i.e., melody and rhythm with either subtest consisting of 52 trials (pairs of melodic/rhythmic phrases) (Fig. 1). Each melodic phrase is played with sampled piano sounds containing three to eight notes, and each rhythmic phrase is played with wood block sounds containing 4 to 11 beats. All phrases have a duration of one measure and are presented at a rate of 100 beats per minute (refer to the supplementary multimedia files for example sounds[1]).

Participants were tested individually in a quiet room with ambient noise level below 45 dB(A). The stimuli were presented via headphones at 70 dB sound pressure level with equal phase and intensity at both ears. Participants judged and reported on the answer sheet whether the two phrases in a trial were identical. Instructions were provided in participants' native languages. Phrases within a trial were separated by three beats (or about 1.8 s), and participants had the same interval to respond. The presentation order of the rhythm and melody subtests was counterbalanced across participants. Before each subtest participants were given two example trials with feedback. However, no feedback was given during the test itself.

3. Results

Percent-correct values were obtained based on the raw scores (number of correct responses) and a 2×2 repeated measures analysis of variance (ANOVA) was then carried out with participant group (Chinese versus Japanese) as the between-subject factor and musical aspect (melody versus rhythm) as the within-subject factor (see Fig. 2). Neither main effect was significant [participant group: $F(1,\ 57) = 0.072$, $p = 0.789$, partial $\eta^2 = 0.001$; musical aspect: $F(1,\ 57) = 0.266$, $p = 0.608$, partial $\eta^2 = 0.005$], indicating that the two groups performed equally well on music perception as a whole and that rhythm and melody were discriminated with similar accuracies.

However, ANOVA tests revealed a significant interaction effect between participant group and musical aspect [$F(1,\ 57) = 14.818$, $p < 0.001$, partial $\eta^2 = 0.206$], indicating that rhythm and melody were discriminated disproportionately by the two groups. Further, simple effect analyses revealed that native Chinese participants outperformed Japanese counterparts on melody discrimination [$F(1,\ 57) = 5.583$, $p = 0.022$, partial $\eta^2 = 0.089$], but a reverse pattern was observed for rhythm discrimination [$F(1,\ 57) = 5.288$, $p = 0.025$, partial $\eta^2 = 0.085$]. Moreover, within-group comparisons showed that melody was better discriminated than rhythm in native Chinese speakers [$F(1,\ 57) = 10.039$, $p = 0.002$, partial $\eta^2 = 0.150$], but a reverse pattern was observed in native Japanese speakers with rhythm discriminated better than melody [$F(1,\ 57) = 5.287$, $p = 0.025$, partial $\eta^2 = 0.085$].

4. Discussion

In the current study, we investigated the effects of native language experience with distinct prosodic properties on the perception of musical melody and rhythm in native Chinese and Japanese speakers. The lack of significant main effects of participant group and musical aspect demonstrated that the two groups did not differ in overall musical aptitude with similar composite scores for rhythm and melody discrimination. Most importantly, the significant interaction between participant group and musical aspect revealed that prosodic properties of native language had substantial differential effects on music melody and rhythm perception.

As one of the primary acoustic features commonly used in both speech and music, pitch has been the focus of previous studies on the relationship between the two domains. A growing

Fig. 1. Example trials from the MET. A pair of musical phrases with same or different melodic/rhythmic patterns are presented in (A) and (B), respectively.

https://doi.org/10.1121/10.0001179

JASA EXPRESS LETTERS

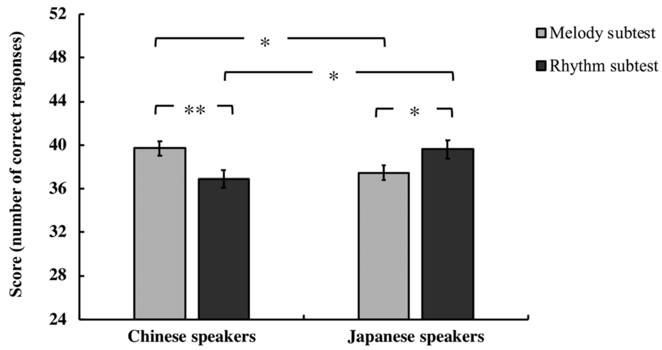

Fig. 2. Scores of the melody and rhythm subtests for native Chinese and Japanese speakers. **significant at $p < 0.01$, *significant at $p < 0.05$. Error bars represent standard deviation across subjects.

body of research has suggested bidirectional transfer effects at the perceptual level (Patel, 2011; Asaridou and McQueen, 2013). On the one hand, musical training improves the perception of linguistic pitch. For example, Marques *et al.* (2007) found that musicians detected pitch violation in a foreign language better than non-musicians and Wu *et al.* (2015) revealed better performance in musicians than non-musicians for discriminating within-category stimuli pairs of a lexical tone continuum. On the other hand, tonal language experience enhances pitch perception in music. For example, Wong *et al.* (2012), Bradley (2016), Chen *et al.* (2016), and our recent study (Zhang *et al.*, 2019) all found that tonal language speakers outperformed non-tonal counterparts in discriminating music melody. However, these studies uniformly compared tonal and non-tonal language speakers, and pitch-accent languages were not included in the comparisons. Our current study bridged this gap in the literature, and the results showed that native Chinese speakers did better in discriminating melody than rhythm and, more importantly, they outperformed native Japanese counterparts on melody discrimination. These results are consistent with previous findings (Wong *et al.*, 2012; Bradley, 2016; Chen *et al.*, 2016; Zhang *et al.*, 2019). The data further revealed that the language-to-music transfer effect in pitch perception did not simply reflect a tonal versus intonational language distinction, but reflected the specific tonal features of particular languages. Specifically, a tonal language like Chinese in which both pitch direction and slope play important roles in tone perception enhances sensitivity to music melody more than a pitch-accent language like Japanese in which pitch direction is the most important phonetic cue to accent perception. It needs to be pointed out that the pitch perception superiority of tonal language speakers over non-tonal language speakers and Chinese speakers over Japanese speakers is not from different pitch perception thresholds because the pitch violations (at least one semitone) between melodic phrases are above the perceptual thresholds of all the listeners irrespective of native languages. Rather, it is from higher-level pitch perception abilities, e.g., sensitivity to pitch patterns and working memory for pitch (Bidelman *et al.*, 2013).

Compared with the number of studies on pitch perception, studies on rhythm perception across language and music are sparse. To the best of our knowledge, Wong *et al.* (2012) is the only previous study that directly compared the effects of linguistic rhythmic properties on musical rhythm perception, but no significant difference was found between native Cantonese and English speakers. In this regard, the current study contributed novel findings by demonstrating that native Japanese speakers outperformed Chinese counterparts in musical rhythm perception, indicating that language-to-music transfer effects also apply to rhythm perception. Japanese mora-timed rhythm is primarily realized by contrastive vowel length, which results in greater durational variability in successive vocalic intervals. By contrast, Chinese syllable-timed rhythm is characteristic of low duration variability (Ramus *et al.*, 1999; Grabe and Low, 2002; Lee *et al.*, 2006). Therefore, it is very likely that the linguistic rhythmic properties enhance sensitivity of native Japanese speakers to music rhythm more than Chinese speakers. It is worth noting that although stressed-timed rhythms in English and Dutch are also characteristic of high vocalic interval variability, better proficiency with English and Dutch did not lead to enhanced musical rhythm perception in comparison with tonal language speakers of Cantonese and Mandarin (Wong *et al.*, 2012; Zhang *et al.*, 2019). Several reasons might underlie the discrepancies, including different musical aptitude tests used in these studies and various phonological factors responsible for high vocalic interval variability in different languages. First, MBEA used in the study of Wong *et al.* (2012) was designed to screen individuals with amusia, while MET used in Zhang *et al.* (2019) and the current study was designed to measure musical aptitude of typical nonmusicians and musicians. Some English and Chinese speakers tended to score near the ceiling in

Wong *et al.* (2012), but no participants approached ceiling in the MET test. It is possible that the ceiling effects failed to distinguish the two groups in Wong *et al.* (2012), but the between-group difference was quantified successfully by MET. Second, phonological factors that contribute to high variability of vocalic intervals vary across the three languages. Contrastive vowel length plays a primary role in Japanese, but tense vowels and vowel reduction play roles in English and Dutch. Duration is the primary acoustic cue for the perception of long versus short vowel in Japanese, but is the secondary acoustic cue for the perception of tense versus lax and full versus reduced vowels. It is possible that better sensitivity to phonetic duration variability in Japanese listeners relative to listeners of stressed-timed languages improves musical rhythm perception. Considering the limited number of comparative studies on rhythm perception in language and music, future research needs to include more language samples to verify whether the language-to-music transfer effects apply to musical rhythm perception, especially how phonological factors underlying linguistic rhythm perception influence sensitivity to music rhythm.

In conclusion, differences between native Japanese and Chinese speakers in musical melody and rhythm perception confirm that linguistic prosodic characteristics of native languages affect music perception. Follow-up studies on the Japanese speakers are under way to examine how learning a foreign language with distinct tonal and rhythmic properties may affect the perception of music melody and rhythm.

Acknowledgments

This research was supported by grants from the Social Science Fund of Beijing (Grant No. 17YYA004), the Discipline Team Support Program (Grant No. JC201901), and the Science Foundation of Beijing Language and Culture University (Fundamental Research Funds for the Central Universities) (Grant No. 18PT09) to L.J.Z.

References and links

[1]See supplementary material at https://doi.org/10.1121/10.0001179 multimedia files for example sounds.

Anvari, S. H., Trainor, L. J., Woodside, J., and Levy, B. A. (**2002**). "Relations among musical skills, phonological processing, and early reading ability in preschool children," J. Exp. Child Psychol. **83**(2), 111–130.

Asaridou, S. S., and McQueen, J. M. (**2013**). "Speech and music shape the listening brain: Evidence for shared domain-general mechanisms," Front. Psychol. **4**, 321.

Beckman, M. E., and Pierrehumbert, J. B. (**1986**). "Intonational structure in Japanese and English," Phonol. Yearbook **3**, 255–309.

Bidelman, G. M., Hutka, S., and Moreno, S. (**2013**). "Tone language speakers and musicians share enhanced perceptual and cognitive abilities for musical pitch: Evidence for bidirectionality between the domains of language and music," PLoS One **8**(4), e60676.

Bradley, E. D. (**2016**). "Phonetic dimensions of tone language effects on musical melody perception," Psychomusicol. Music Mind Brain **26**(4), 337–345.

Chen, A., Liu, L., and Kager, R. (**2016**). "Cross-domain correlation in pitch perception, the influence of native language," Lang. Cogn. Neurosci. **31**(6), 751–760.

Deutsch, D., Henthorn, T., Marvin, E., and Xu, H. (**2006**). "Absolute pitch among American and Chinese conservatory students: Prevalence differences, and evidence for a speech-related critical period," J. Acoust. Soc. Am. **119**(2), 719–722.

Falk, D. (**2004**). "Prelinguistic evolution in early hominins: Whence motherese," Behav. Brain Sci. **27**, 491–503.

Giuliano, R. J., Pfordresher, P. Q., Stanley, E. M., Narayana, S., and Wicha, N. Y. (**2011**). "Native experience with a tone language enhances pitch discrimination and the timing of neural responses to pitch change," Front. Psychol. **2**, 146.

Grabe, E., and Low, E. L. (**2002**). "Durational variability in speech and the rhythm class hypothesis," in *Laboratory Phonology 7*, edited by C. Gussenhoven and N. Warner Mouton de Gruyter (Mouton de Gruyter, Berlin, Germany), pp. 515–546.

Koelsch, S. (**2011**). "Toward a neural basis of music perception—a review and updated model," Front. Psychol. **2**, 110.

Ladd, D. R. (**1996**). *Intonational Phonology* (Cambridge University Press, Cambridge).

Lee, B., Guion, S. G., and Harada, T. (**2006**). "Acoustic analysis of the production of unstressed English vowels by early and late Korean and Japanese bilinguals," Stud. Second Lang. Acquis. **28**, 487–513.

Lin, H., and Wang, Q. (**2007**). "Mandarin rhythm: An acoustic study," J. Chin. Lang. Comput. **17**(3), 127–140.

Marques, C., Moreno, S., Castro, S. L., and Besson, M. (**2007**). "Musicians detect pitch violation in a foreign language better than nonmusicians: Behavioral and electrophysiological evidence," J. Cogn. Neurosci. **19**(9), 1453–1463.

Massaro, D. W., Cohen, M. M., and Tseng, C. (**1985**). "The evaluation and integration of pitch height and pitch contour in lexical tone perception in Mandarin Chinese," J. Chin. Ling. **13**, 267–290.

Palmer, C., and Kelly, M. H. (**1992**). "Linguistic prosody and musical meter in song," J. Mem. Lang. **31**(4), 525–542.

Patel, A. D. (**2011**). "Why would musical training benefit the neural encoding of speech? The OPERA hypothesis," Front. Psychol. **2**, 142.

Patel, A. D., and Daniele, J. R. (**2003**). "An empirical comparison of rhythm in language and music," Cognition **87**(1), B35–B45.

Patel, A. D., Iversen, J. R., and Rosenberg, J. C. (**2006**). "Comparing the rhythm and melody of speech and music: The case of British English and French," J. Acoust. Soc. Am. **119**(5 Pt 1), 3034–3047.
Peretz, I., Nguyen, S., and Cummings, S. (**2011**). "Tone language fluency impairs pitch discrimination," Front. Psychol. **2**, 145.
Ramus, F., Nespor, M., and Mehler, J. (**1999**). "Correlates of linguistic rhythm in the speech signal," Cognition **73**(3), 265–292.
Schön, D., Boyer, M., Moreno, S., Besson, M., Peretz, I., and Kolinsky, R. (**2008**). "Songs as an aid for language acquisition," Cognition **106**(2), 975–983.
Slevc, L. R., and Miyake, A. (**2006**). "Individual differences in second-language proficiency: Does musical ability matter," Psychol. Sci. **17**(18), 675–681.
Tong, Y., Gandour, J., Talavage, T., Wong, D., Dzemidzic, M., Xu, Y., Li, X., and Lowe, M. (**2005**). "Neural circuitry underlying sentence-level linguistic prosody," Neuroimage **28**(2), 417–428.
Wallentin, M., Nielsen, A. H., Friis-Olivarius, M., Vuust, C., and Vuust, P. (**2010**). "The Musical Ear Test, a new reliable test for measuring musical competence," Learn. Individ. Differ. **20**(3), 188–196.
Wallin, N. L., Merker, B., and Brown, S. (**2001**). *The Origins of Music* (MIT Press, Cambridge, MA).
Wang, W. S. Y. (**1973**). "The Chinese language," Sci. Am. **228**, 50–60.
Wong, P. C. M., Ciocca, V., Chan, A. H. D., Ha, L. Y. Y., Tan, L.-H., and Peretz, I. (**2012**). "Effects of culture on musical pitch perception," PLoS One **7**(4), e33424.
Wu, H., Ma, X., Zhang, L., Liu, Y., Zhang, Y., and Shu, H. (**2015**). "Musical experience modulates categorical perception of lexical tones in native Chinese speakers," Front. Psychol. **6**, 436.
Zhang, L., Jiang, W., Shu, H., and Zhang, Y. (**2019**). "Congenital blindness enhances perception of musical rhythm more than melody in Mandarin speakers," J. Acoust. Soc. Am. **145**(5), EL354–EL359.