INVESTIGATING THE ROLE OF MUSICAL EXPERIENCE IN LEXICAL TONE PERCEPTION: NON-MUSICIANS AND AMATEUR MUSICIANS' PERCEPTION OF MANDARIN TONE

Xiao Fu & Bronwen G. Evans

Department of Speech, Hearing and Phonetic Sciences, University College London, UK
xiao.fu.15@ucl.ac.uk, bronwen.evans@ucl.ac.uk

ABSTRACT

Previous studies have found that musicians typically discriminate Mandarin tones better than non-musicians. However, the relationship between musical experience and tone perception is unclear. In the current study, 39 monolingual native English speakers with no previous experience of tone languages and a range of musical backgrounds (non-/ amateur musicians) completed 6 tasks, including lexical tone identification, working memory, L1/L2 segmental perception and the Goldsmiths Musical Sophistication Index which measures musical ability/ experience. Path analysis showed that pitch interval discrimination (PID), but not musical ability or musical training, directly predicted tone identification. There was no relationship with working memory or L1/L2 segmental perception. Follow-up mixed effect models showed that Tone1 and Tone4 identification was associated with PID, musical ability and musical training, whereas Tone2 and Tone3 identification was only associated with musical training and PID respectively. Overall, musical training appears to be linked to PID, which in turn leads to better tone identification.

Keywords: Mandarin tone identification, amateur musicians, pitch interval discrimination

1. INTRODUCTION

It is well-established that for non-tonal language speakers, e.g., English native speakers, learning a tone language like Mandarin Chinese can be very difficult (e.g., [11, 12]). For example, Mandarin has 4 lexical tones, a high level (Tone1), rising (Tone2), dipping (Tone3), and falling (Tone4) tone. Although similar pitch patterns are used in English intonation to signal pragmatic meaning, learning to use pitch at the lexical level is particularly challenging for learners, e.g., [11].

Of course, pitch also plays an important role in music [23], and previous research has suggested that learners with musical training may be able to transfer their ability with pitch in music, in particular musical intervals [2], not only to their perception of pitch in a native language [28], but also to a second language (L2; [18]). Indeed, neuroimaging studies have shown that there are both structural [15] and functional [10, 17] brain differences between musicians and non-musicians. For example, musicians are better at tracking linguistic pitch changes than are non-musicians; musicians' brainstem responses show more faithful representation of the F0 contours and more robust neural phase-locking particularly for Tone3, the most complex tone contour [31, 23]. One possibility is that this lower level processing advantage, facilitated by musical training, enhances tone identification, in particular for Tone2 and Tone3, which are the most confusing pair among the four Mandarin tones [32].

Indeed, musical training has also been shown to predict the ability to perceive tone. For example, recent research [6] measured lexical tone discrimination in non-musicians and professional musicians with no previous experience with Mandarin, using Mandarin monosyllables. Musicians significantly outperformed non-musicians, indicating that musical expertise and melodic proficiency predict lexical tone discrimination. Likewise, musicians who have been playing for longer, are better at identifying lexical tone than those with fewer years' training [4].

However, how musical ability, musical training and pitch interval discrimination function together to benefit musicians in lexical tone identification remains unclear. Previous research has suggested that musical training does not just lead to better linguistic pitch processing [31], but that it also leads to wider changes in general speech processing [9, 21] and cognitive processing, such as attention and auditory working memory (WM; [8]). Given that these abilities are also closely linked to L2 segmental perception and learning more generally (e.g., [20]), it might be these changes rather than any advantage for pitch itself that means musicians are at an advantage when learning Mandarin tones.

In the current study, we measured musicians and non-musicians' performance on a range of tasks to further investigate how musical training is linked to tone language learning. Native English speakers with no previous experience with Mandarin, and from a range of musical backgrounds, completed a lexical tone identification task, a pitch interval discrimination (PID) task, verbal & melodic WM tasks, and L1 and L2 consonant categorization and discrimination tasks. Musical ability was measured using the Goldsmiths Musical Sophistication Index questionnaire (Gold-MSI; [22]), a psychometric tool for the measurement of musical attitudes, behaviours, and skills. This enabled us to explore the effects of musical experience in more depth than in previous studies, which have typically focused on the number of years of instrumental training. Based on previous studies, we predicted (1) that there would be differences between participants in terms of lexical tone identification, and (2) that these differences would at least in part, be related to musical experience, such that Tone1 and Tone4 would be relatively easy to identify for all participants, but that musicians might have an advantage for Tone2 and Tone3.

2. METHODS

2.1. Participants

Thirty-nine monolingual native English-speaking participants (age 18-35, M=24.51, SD=4.44) were tested. Participants had no history of speech, hearing or language impairment, and no experience with tonal languages. We aimed to recruit participants with a range of musical experience including formal lessons and informal practice on musical instruments and singing.

2.2. Stimuli and Materials

2.2.1. Reading Span

Verbal WM was measured using a reading span task [30]. Participants are asked to read aloud blocks of 2-6 sentences and at the end recall all the final words of the sentences within 7 secs. There are 100 sentences in total, and the measure is the number of words correctly recalled. Stimuli were presented using Psychopy2 [24] and the answers were recorded manually by the researcher.

2.2.2. Melodic WM Task

Melodic WM was measured using a melodic memory test [13]. There were 20 trials in total. On each trial, participants heard 3 versions of an unfamiliar melody, 3-16 notes in length, played in different keys but with only one containing an altered note (i.e., a different interval). Listeners' task was to spot which melody was the "odd one out". Participants gave their response by clicking the corresponding number (1, 2 or 3) on the computer, running RStudio [27]. The final score was calculated using Item Response Theory [5].

2.2.3. L1 Consonant Categorization

Participants completed a forced-choice identification task, in which they heard synthetic continua varying in 50 equal steps from pea-bee [14] and coat-goat [25]. A modified Levitt procedure was used to estimate the points on the continuum where the stimuli were labeled as one word of the pair [19]. The initial step size was 10 ms, reduced linearly to 4 ms over the first 3 reversals. The task ended after 7 reversals or a maximum of 40 trials. The measure used here was participants' average slope across both the /b-p/ and /k-g/ continua.

2.2.4. L2 Consonant Discrimination

To investigate participants' ability with unfamiliar L2 contrasts, participants completed a forced-choice 3-way oddity task in which they were tested in their ability to discriminate uvular, /q/, and palatal, /c/, plosives from their native velar plosive, /k/. These were selected because English listeners typically assimilate both uvular and palatal sounds to their native velar consonants. Participants heard plosives produced by 3 female phoneticians, in 3 different VCV contexts, /iCi/, /ɑCɑ/ or /uCu/. Participants heard all 6 combinations of each minimal pair presented 3 times, giving a total of 72 trials. Trials were presented in a randomized order using Praat [3]. The measure was proportion correct.

2.2.5. Tone Identification (ToneID) Task

Participants identified the 4 Mandarin tones embedded in 5 different monosyllabic Mandarin words, produced by a male and female Mandarin native speaker. They identified each word for each speaker only once, giving a total of 40 responses (10 per tone). As the participants had no experience with any tonal language, they were briefly introduced to the Mandarin tones and the corresponding tone marks used to give responses. The experiment ran in Praat [3]. The measure was proportion correct.

2.2.6. Pitch Interval Discrimination Task (PID)

The stimuli in this test were pure tones, 400ms in duration and ranging in frequency from 392Hz (G4) to 416Hz (G#4) in steps of 2Hz. This gave a total of 13 tones. Each tone was paired with each other in both a rising and falling combination, with an ISI of 100 ms, giving a total of 24 pairs. On each trial participants reported whether the 2nd tone was higher or lower than the 1st by clicking "higher" or "lower" on the computer screen. They completed 3 repetitions of each pair, presented in a random order. Performance was calculated with hitting rate minus false alarm rate [29].

2.3 Procedure

Participants completed the tasks in the order in which they are described above, followed by the Gold-MSI [22]. Testing took place in a sound attenuated booth and lasted 80 mins, including breaks.

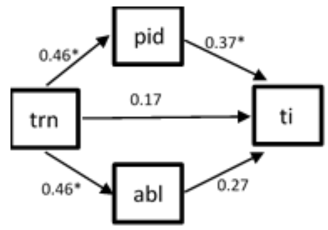
3. RESULTS

Initial independent t-tests compared the performance of participants with more musical training (N=19) and those with less musical training (N=20) grouped according to their Gold-MSI musical training score. Those with a score higher than the median (Median=19) were assigned to the more musical training group and those with a score equal to or lower than the median, to the less musical training group. The results showed that there were differences in performance between the two groups in terms of ToneID ($t_{(37)}$=-2.61, p<0.05), PID ($t_{(28.07)}$=-3.14, p<0.05), and musical ability ($t_{(30.66)}$=-2.96, p<0.05). Participants with more musical training outperformed those with less musical training in ToneID and PID, and as expected, scored higher in terms of musical ability as measured using the Gold-MSI [22]. There were no significant differences in any other tasks.

To investigate if performance on these tasks predicted ToneID performance, path analysis was used. Preliminary Pearson correlations between ToneID and all other tasks, i.e., predictor variables, showed that ToneID was significantly correlated with all measures (R=0.37 to 0.51, p<0.05) except for L1 categorization and melodic memory. Among these 5 predictors, PID, musical training and musical ability were forwarded to the path analysis due to their high correlation coefficients (R=0.51, 0.47 and 0.43 respectively, p<0.05).

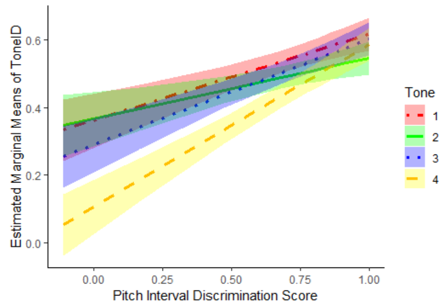
The path model (Fig. 1) was fitted using lavaan [26], an R package for structural equation modelling. The model showed a good fit; $\chi2_{(1)}$=0.07, p=0.93, CFI = 1.00, RMSEA =0.00, SRMR =0.004. However, only PID had a significant direct influence on ToneID (β=0.37, p=0.01). Neither training (β=0.17, p=0.28) nor ability (β=0.27, p=0.056) had a significant direct influence on ToneID, although training appeared to significantly influence musical ability (β=0.46, p<0.01) and had an indirect influence on ToneID, with PID as a mediating variable (β=0.46, p<0.01). Follow-up mixed effect models were built in R [27], using the lmer function of the lme4 package [7] to investigate how identification of the different Mandarin tone types interacted with PID (Model1), musical ability (Model2) and musical training (Model3) to affect ToneID. In each model, tone type together with either PID, ability or training respectively were coded as fixed effects, with ToneID as the dependent variable and participant as a random effect.

Figure 1: The path model used toneID (ti), PID (pid) and musical ability (abl) as endogenous variables, and musical training (trn) as exogenous variable.



A type II ANOVA was applied to each model. Model1 indicated that there were significant main effects of PID ($F_{(37)}$=13.07, p<0.01) and tone type ($F_{(1515)}$=4.41, p<0.01), and a significant interaction of PID and tone type ($F_{(1515)}$=2.62, p<0.05). Similarly, there were significant main effects of training ($F_{(37)}$=10.20, p<0.01) and tone type ($F_{(1515)}$=6.60, p<0.01), and a significant interaction of training and tone type ($F_{(1515)}$=4.49, p<0.05) in Model3. However, in Model2, the interaction between ability and tone type was not significant ($F_{(1515)}$=2.41, p=0.065), although the main effect of the two fixed effects was significant (ability: $F_{(37)}$=8.44, p<0.01; tone type: $F_{(1515)}$=2.67, p <0.05).

Figure 2: Graph to show the relationship between the estimated marginal means of tone identification score and pitch interval discrimination. The ribbon represents for the SE from the model output.



A follow-up pairwise t-test was conducted for each model, using tone type and either PID, musical ability or musical training as the independent variable, and ToneID score for different tone types as the dependent variable (Figs 2-4). As displayed in Fig. 2, participants with higher PID scores performed significantly better at identifying Tone1 ($t_{(99.47)}$=2.34, p<0.05), Tone3 ($t_{(99.47)}$=2.87, p<0.05) and Tone4 ($t_{(99.47)}$=4.41, p<0.01) than those with lower PID scores. Participants with better musical ability also performed better with Tone1 ($t_{(92.55)}$=3.13, p<0.01) and Tone4 ($t_{(92.55)}$=3.27, p<0.01), but the ability did not significantly affect identification of Tone2 or Tone3 (Fig. 3). The more musical training participants had, the better they were at identifying Tone1 ($t_{(94.88)}$=2.38, p<0.05), Tone2 ($t_{(94.88)}$=2.03, p<0.05) and Tone4 ($t_{(94.88)}$=4.61, p<0.05). However, more musical training did not lead to better identification of Tone3 (Fig. 4).

Figure 3: Graph to show the relationship between the estimated marginal means of tone identification score and musical ability score. The ribbon represents for the SE from the model output.

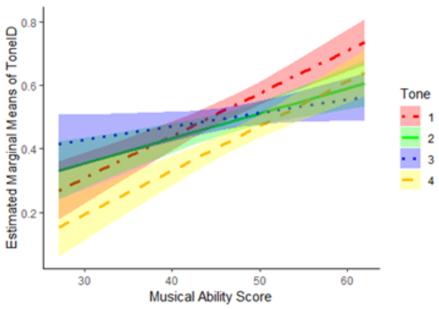
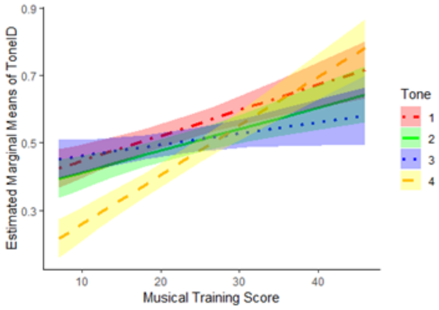


Figure 4: Graph to show the relationship between the estimated marginal means of tone identification score and musical training score. The ribbon represents for the SE from the model output.



4. DISCUSSION

To investigate the relationship between music and lexical tone perception, we measured participants' ability in a variety of tasks, including tests of WM, L1 & L2 segmental speech perception, lexical tone and pitch processing. Musical experience was measured using the Gold-MSI [22], a self-report questionnaire which provides a nuanced measure of engagement with music, separating out measures of musical ability from musical training (i.e., experience of learning an instrument or studying singing).

Consistent with previous research [e.g., 1], musical training affected participants' identification of lexical tone. In addition, there was also a significant influence of musical training in terms of PID, musical ability and training, as expected. However, the path model indicated that pitch interval discrimination (i.e., performance on the PID task), rather than either musical training or ability, directly predicted tone identification. That said, the relationship is complex; musical training has a strong and significant direct influence on pitch interval discrimination. It is thus not the case that just having high pitch interval discrimination necessarily leads to better tone identification. Indeed, inspection of the data revealed that some non-musicians performed similarly to musicians on the PID task, but despite this, did not perform as well as musicians with an equivalent PID score on the ToneID task. Similarly, although more musical training leads to better musical ability, as measured in the Gold-MSI [22], musical ability alone does not lead to better tone identification.

Among the 4 Mandarin tones, Tone1 and Tone4 are commonly considered to be easier to identify for listeners with no prior experience with a tonal language [16], whilst Tone3 is considered the most confusable [31]. This was reflected in our results, and again, there were also links with PID and musical training. Specifically, the mixed effect models showed that Tone1 and Tone4 were significantly easier for those who had higher PID and musical ability scores regardless of whether or not they had completed any musical training. However, musical ability alone was not sufficient for successful identification of Tone2 and Tone3. Although identification of Tone2 appeared to rely primarily on musical training, successful identification of Tone3 was related to pitch interval discrimination. Thus, even with musical training, learners who performed poorly on the PID struggled to identify this tone reliably. Given the finding that musical training also leads to improvements in PID and musical ability, this further suggests that musical training likely directly or indirectly leads not only to better identification of Tone2 and Tone3, but can also improve identification of Tone1 and Tone4. This may explain why amateur musicians performed better than non-musicians on the ToneID task overall, even though PID was the only direct predictor.

In sum, the current study contributes to our understanding of how listeners with no previous experience of tone languages process lexical tone, and how this is affected by musical experience. Our results extend the findings of previous studies showing differences in tone identification between musicians and non-musicians, demonstrating that both formal and informal musical training and practice, influences lexical tone identification, although the relationship is not direct. Moreover, pitch interval discrimination, musical ability, and musical training appear to affect identification of different tones differently. The results further underline the complex nature of the relationship between music and speech. In particular, although a certain amount of musical training does indeed appear to benefit lexical tone perception, musical training alone does not appear to lead directly to better tone identification. Rather, musical training appears to enable learners to transfer their ability with pitch to processing lexical tone.

5. REFERENCES

[1] Alexander, J. A, Wong, P. C. M., Bradlow, A. R., Il, U. S. 2005. Lexical Tone Perception in Musicians and Non-musicians. *Proc Interspeech,* Lisbon, 397–400.

[2] Besson, M., Schön, D., Moreno, S., Santos, A., Magne, C. L. 2007. Influence of musical expertise and musical training on pitch processing in music and language. *Restorative Neurology and Neuroscience,* 25(3–4), 399–410.

[3] Boersma, Paul. 2001. Praat, a system for doing phonetics by computer. *Glot International* 5:9/10, 341-345.

[4] Chua, A., Brunt. 2014. Effect of musical experience on tonal language perception. *The Journal of the Acoustical Society of America,* 135(4), 2356.

[5] De Ayala. R.J. 2009. The Theory and Practice of Item Response Theory . *Psychometrika*, 75(4), 778-779.

[6] Delogu, F., Lampis, G., Olivetti Belardinelli, M. 2006. Music-to-language transfer effect: May melodic ability improve learning of tonal languages by native nontonal speakers? *Cog Proc*, 7(3), 203–207.

[7] Bates, D., Mächler, M., Bolker, B., Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *J. Statistical Software,* 67(1), 1-48.

[8] Ferreri, L., Aucouturier, J.-J., Muthalib, M., Bigand, E., Bugaiska, A. 2013. Music improves verbal memory encoding while decreasing prefrontal cortex activity: an fNIRS study. *Frontiers in Microbiology*, 7, 1–9.

[9] Flaugnacco, E., Lopez, L., Terribili, C., Montico, M., Zoia, S., Schön, D. 2015. Music training increases phonological awareness and reading skills in developmental dyslexia: A randomized control trial. *PLoS ONE,* 10(9), 1–17.

[10] Gaser, C., Schlaug, G. 2003. Brain structures differ between musicians and non-musicians. *J Neurosci*, 23(27), 9240-9245.

[11] Guion, S. G., Pederson, E. 2007. *Investigating the role of attention in phonetic learning,* 57–77. In O.-S. Bohn & M. Munro (Eds.) Language Experience in Second Language Speech Learning: In honor of James Emil Flege. Amsterdam: John Benjamins.

[12] Hallé, P., Best, Chang. 2002. Perception of Mandarin Chinese tones by Mandarin versus French listeners. *The J Acoust Soc Am*, 112(5), 2357.

[13] Harrison, P. M., Müllensiefen, D., Collins, T. 2017. Applying modern psychometric techniques to melodic discrimination testing: Item response theory, computerised adaptive testing, and automatic item generation. *Scientific Reports*, 7(1).

[14] Hazan, V., Messaoud-Galusi, S., Rosen, S., Nouwens, S., Shakespeare, B. 2009. Speech perception abilities of adults with dyslexia: Is there any evidence for a true deficit? *J Speech, Lang, & Hearing Res*, 52, 1510–1529. doi:10.1044/1092-4388(2009/08-0220).

[15] Hyde, K. L., Lerch, J., Norton, A., Forgeard, M., Winner, E., Evans, A. C., Schlaug, G. 2009. Musical Training Shapes Structural Brain Development. *J Neuroscience*, 29(10), 3019–3025.

[16] Kirkham, J., Lu, S., Wayland, R., Kaan, E. 2011. Comparison of Vocalists and Instrumentalists on Lexical Tone Perception and Production Tasks. *Proc. 17th ICPhS, Hong Kong*, 1098-1101.

[17] Lappe, C., Herholz, S., Trainor, L., Pantev, C. 2008. Cortical plasticity induced by short-term unimodal & multimodal musical training. *J Neurosci* 28, 9632-9.

[18] Marques, C., Moreno, S., Castro, S. L., Besson, M. 2007. Musicians detect pitch violation in a foreign language better than non-musicians: behavioural and electrophysiological evidence. *J. Cogn. Neurosci*. 19, 1453–1463.

[19] McCarthy. K. M., Mahon. M., Rosen. S., Evans. B. G. 2014. Speech perception and production by sequential bilingual children: A longitudinal study of voice onset time acquisition. *Child Dev*. (85), pp. 1965–1980.

[20] Meador, D., Flege, J. E., MacKay, I. R. A. 2001. The identification of English consonants by native speakers of Italian. *Phonetica*, 58(1–2), 103–125.

[21] Moreno, Sylvain, Marques, Carlos, Santos, Andreia, Santos, Manuela, Castro, São Luís, Besson, Mireille. 2009. Musical Training Influences Linguistic Abilities in 8-Year-Old Children: More Evidence for Brain Plasticity. *Cerebral Cortex*, 19(3), 712-723.

[22] Müllensiefen, D., Gingras, B., Musil, J. Stewart, L. 2014. The musicality of non-musicians: An index for assessing musical sophistication in the general population. *PLoS One* 9, 1–23.

[23] Plack CJ, Oxenham AJ, Fay R, Popper AN, eds. 2005. *Pitch: neural coding and perception*. New York: Springer.

[24] Peirce, JW. 2007. PsychoPy - Psychophysics software in Python (v. 1. 84.1). *J Neurosci Methods, 162:8-13*

[25] Ramus, F., Rosen, S., Dakin, S., Day, B. L., Castellote, J. M., White, S., Frith, U. 2003. Theories of developmental dyslexia: Insights from a multiple case study of dyslexic adults. *Brain*, 126, 841–865.

[26] Rosseel. Y. 2012. lavaan: An R Package for Structural Equation Modeling. *Journal of Statistical Software,* 48(2), 1-36.

[27] RStudio Team. 2016. RStudio: Integrated Development for R (version 1.1.453). RStudio, Inc., Boston, MA.

[28] Schön, D., Magne, C., Besson, M. 2004. The music of speech: Music training facilitates pitch processing in both music and language. *Psychophysiology* 41: 341–349.

[29] Stevens, C. J., Keller, P. E., & Tyler, M. D. (2013). Tonal language background and detecting pitch contour in spoken and musical items. *Psychology of Music*, *41*(1), 59–74. doi.org/10.1177/0305735611415749

[30] Van Den Noort, M., Bosch, P., Haverkort, M., Hugdahl, K. 2008. A standard computerized version of the reading span test in different languages. European *Journal of Psychological Assessment*, 24(1), 35–42.

[31] Wong, P. C. M., Skoe, E., Russo, N. M., Dees, T., Kraus, N. 2007. Musical experience shapes human brainstem encoding of linguistic pitch patterns. *Nature Neuroscience*, 10(4), 420–422.

[32] Wang, Y., Spence, Jongman, & Sereno. (1999). Training American listeners to perceive Mandarin tones: Transfer to production. The Journal of the Acoustical Society of America*, 105*(2), 1095.