



## OPEN Enduring musician advantage among former musicians in prosodic pitch perception

Xin Ru Toh<sup>1</sup>, Shen Hui Tan<sup>1</sup>, Galston Wong<sup>2</sup>, Fun Lau<sup>1</sup> & Francis C. K. Wong<sup>1</sup>

Musical training has been associated with various cognitive benefits, one of which is enhanced speech perception. However, most findings have been based on musicians taking part in ongoing music lessons and practice. This study thus sought to determine whether the musician advantage in pitch perception in the language domain extends to individuals who have ceased musical training and practice. To this end, adult active musicians ( $n = 22$ ), former musicians ( $n = 27$ ), and non-musicians ( $n = 47$ ) were presented with sentences spoken in a native language, English, and a foreign language, French. The final words of the sentences were either prosodically congruous (spoken at normal pitch height), weakly incongruous (pitch was increased by 25%), or strongly incongruous (pitch was increased by 110%). Results of the pitch discrimination task revealed that although active musicians outperformed former musicians, former musicians outperformed non-musicians in the weakly incongruous condition. The findings suggest that the musician advantage in pitch perception in speech is retained to some extent even after musical training and practice is discontinued.

Musical training has been associated with various cognitive enhancements<sup>1</sup>, making it an attractive enrichment and intervention activity. In the language domain, one notable finding is that musical training is linked to an advantage in speech perception. For instance, musicians are better than non-musicians at perceiving speech in noisy conditions even in older adulthood<sup>2,3</sup>. In particular, a vast amount of literature has documented positive music-to-language cross-domain transfer effects in pitch processing, which undergirds Patel's OPERA hypothesis<sup>4</sup>. The OPERA hypothesis describes how musical training benefits the neural encoding of speech when five requirements are satisfied: there is an *overlap* in the brain networks employed to process an acoustic feature common to music and speech; the processing of the shared acoustic feature occurs at higher *precision* in music than in speech; and the musical activities evoke strong positive *emotion*, have frequent *repetition*, and encompass focused *attention*. The OPERA hypothesis may account for the superior pitch processing abilities in speech seen in musically trained individuals, as pitch is a basic acoustic property found in both music and speech. While pitch differences are used to form melodies in music, they are used to convey contrastive meaning via lexical tones, stress, and intonation in speech. To augment the OPERA hypothesis, this study seeks to explicate whether the musician advantage persists in former musicians who have ceased musical training and practice.

The OPERA hypothesis is well supported by empirical studies comparing musicians and non-musicians. Studies have found that among individuals with no tone language experience, musicians outperform non-musicians in lexical tone perception<sup>5–15</sup>. In addition, musicians without tone language experience show enhanced brainstem and cortical encoding when listening to lexical tones<sup>14,16,17</sup>. Yet, it is uncertain whether the musician advantage in lexical tone perception also exists among tone language speakers. It was previously found that for English or French speakers, musicians outperform non-musicians in Cantonese tone discrimination, whereas for Cantonese speakers, musicians and non-musicians both show ceiling effects<sup>18</sup>. On the other hand, it was also found that Cantonese musicians outperform Cantonese non-musicians in the discrimination and identification of merging Cantonese tone pairs, especially the most difficult Tone 2/Tone 5 contrast<sup>19</sup>. More recently, Toh et al.<sup>20</sup> found that even among speakers of a tone language, those who have received musical training outperform non-musicians in non-native lexical tone perception. Apart from lexical tone perception, studies have found that among individuals with no tone language experience, musicians are better than non-musicians at perceiving stress, which is indicated by a combination of pitch, duration, and intensity variations<sup>21,22</sup>. However, it remains unclear whether the musician advantage in stress perception also applies to tone language speakers. Among English speakers, musicians outperform non-musicians in English stress perception, while Cantonese-English bilingual musicians

<sup>1</sup>Linguistics and Multilingual Studies, School of Humanities, Nanyang Technological University, Singapore, Singapore. <sup>2</sup>School of Brain and Behavioral Sciences, The University of Texas at Dallas, Dallas, TX, USA. ✉email: franciswong@ntu.edu.sg

and non-musicians perform equally well<sup>23</sup>. Tone language experience has been linked to enhanced pitch perception abilities in speech<sup>6,24–30</sup>. As such, one possible explanation for the conflicting findings is that the musician advantage for pitch perception in speech applies across speakers irrespective of language background, but the more subtle effect among tone language speakers is likely to be masked by ceiling-like performance in tasks that are not sufficiently sensitive.

Besides lexical tones and stress, another aspect in which there is mounting evidence for the musician advantage is prosody. A series of studies have consistently found that musicians outperform non-musicians in detecting pitch contour expectancy violations<sup>31–35</sup>. This research paradigm, first created by Schön et al.<sup>31</sup>, is designed by manipulating the fundamental frequency of either the final notes of musical phrases or final words of linguistic phrases. In particular, the weakly incongruous condition entails a small pitch change which is difficult to detect, and hence evaluates pitch perception in a more fine-grained manner. Through both behavioural and electrophysiological measures, they found that adult musicians detected these pitch variations better than non-musicians in not only music but also their native language, thereby lending support for a domain-general pitch processing mechanism. This finding was reinforced in follow-up cross-sectional and longitudinal studies<sup>32,33</sup>, in which they found similar group differences among 8-year-old children, despite the fact that the children musicians received a shorter duration of musical training than the adult musicians in the original study. Their finding was also expanded in follow-up studies introducing unfamiliar language contexts<sup>34,35</sup>, in which they found that participants across groups found it more difficult to detect pitch changes in a non-native language or pseudolanguage than in their native language. The researchers posited that understanding the semantic content and being familiar with intonational contours in sentences might help with anticipating and detecting pitch changes in one's native language. That said, the researchers found that musicians held an advantage over non-musicians in detecting prosodic pitch violations across native and non-native language contexts. Moreover, behavioural studies have found that musicians outperform non-musicians in matching spoken utterances to their intonation melodies<sup>36</sup> and identifying emotional prosody in speech<sup>37–39</sup>. Interestingly, similar results were seen in a longitudinal study with 6-year-old children, with those who were randomly assigned to receive 1 year of musical training in the form of keyboard or vocal lessons outperforming those who received no lessons when tested on the identification of emotional prosody in speech<sup>37</sup>. Collectively, these studies substantiate the notion that musical training facilitates speech perception at not only the segmental but also supra-segmental level.

Furthermore, neurological studies suggest that musical training is linked to structural and functional differences in the brain<sup>40–50</sup>. Notably, the effects of musical training on brain development seem to be causal in nature<sup>33,45–47</sup>. For instance, Hyde et al.<sup>46,47</sup> randomly assigned 6-year-old children without any behavioural or brain differences in pre-tests to receive either 15 months of musical training or no training. They found that only those who received musical training showed structural brain changes in motor and auditory areas which were correlated with behavioural improvements on melodic and rhythmic discrimination tests. These studies suggest that there may be musical training-induced brain plasticity effects that could potentially translate to long-lasting cognitive impacts. While the data on ageing and musicianship remains scant, there is emerging evidence that an age-related decline in auditory perception may be mitigated by musical training among lifelong musicians who maintain regular musical practice. Older and younger adult musicians outperform non-musicians in various auditory processing abilities, such as detecting speech-in-noise and mistuned harmonics, assessed using neurophysiological and behavioural measures<sup>51–53</sup>.

In light of the above findings, musical training does appear to facilitate speech perception, providing empirical evidence for the OPERA hypothesis<sup>4</sup>. A critical question to consider is whether the OPERA hypothesis can be extended to former musicians. Studies on the effects of musical training typically characterise musicians as individuals with  $\geq 6$  years of musical training and ongoing instrumental practice for  $\geq 1$  h a week<sup>54</sup>. However, such professional musicians may not represent the general population in which many individuals who take up music lessons in childhood eventually do not commit to it<sup>55,56</sup>. Although there has been extensive research on professional active musicians, more research needs to be done with individuals who choose not to pursue musicianship professionally but nonetheless have had some musical experience. Of particular interest is whether cognitive benefits such as in speech perception persist even after musical training and practice is discontinued. Qualifying the extent of the influence of musical training among individuals who have undergone music attrition will serve to not only provide insight on the generalisability of the OPERA hypothesis, but also inform the effectiveness of musical training as a means of improving cognitive and linguistic abilities in the long-term, as well as protecting against age-related cognitive decline.

As noted by Costa-Giomi<sup>56</sup>, few studies to date have investigated whether cognitive advantages exist in the long term after musical training and practice is discontinued. Costa-Giomi and Ryan<sup>58</sup> (as cited in Costa-Giomi<sup>57</sup>) conducted a longitudinal study in which children in the experimental group received 3 years of piano lessons. Seven years after musical training was discontinued, the researchers found no differences in IQ or memory between the adults who had and had not received childhood musical training, suggesting that musical training does not result in permanent cognitive benefits. Nevertheless, the researchers postulated that the lack of long-lasting cognitive improvements may have been due to low attendance and time spent practising the musical instrument<sup>55,58</sup>. In contrast, two behavioural studies found improved performance in various cognitive tasks such as IQ<sup>59</sup> and executive functions<sup>60</sup> in adulthood even after musical training and practice had ceased, suggesting that musical training has long-term benefits and contributes to the establishment of a cognitive reserve. However, the measures used in these behavioural studies have focused on general cognitive abilities rather than speech perception abilities specifically.

In terms of auditory perception, two brain imaging studies have found that musical training in early childhood provide sustained enhanced neural processing of auditory stimuli in adulthood after musical training and practice had ceased. Skoe and Kraus<sup>61</sup> found that young adults who had received musical training in childhood showed more robust signal-to-noise ratio brainstem responses to pure tones, as compared to non-musicians.

White-Schwoch et al.<sup>62</sup> found that older adults with a greater number of years of musical training in childhood or young adulthood showed faster neural timing in response to consonant–vowel transitions in speech syllables presented in quiet and noise, compared to older adults with fewer number of years of musical training or no musical training at all. Although these two brain imaging studies suggest that music-related neuroplasticity is maintained even after music attrition, studies have yet to investigate if these neural traces translate to a clear behavioural advantage in acoustic processing of speech stimuli. This is a critical research gap that the present study aims to bridge.

In sum, although the literature has generally established that professional musicians have an advantage over non-musicians in pitch perception abilities in the language domain, it remains inconclusive whether this musician advantage would also be observed behaviourally among individuals who have ceased musical training and practice. That being the case, the overarching aim of our study is to add to the OPERA hypothesis and elucidate whether a potential music-to-language transfer effect exists among former musicians. To this end, our study compared active musicians, former musicians, and non-musicians in their ability to perform a well-replicated experimental task—detecting linguistic prosodic pitch violations.

Although there has been a burgeoning number of studies revolving around various types of pitch perception in speech, such as lexical tones and stress, we were theoretically motivated to focus on prosody for several reasons. Firstly, prosody is often described as “the music of speech”<sup>63</sup>, thereby making it an obvious candidate for the present study on music-to-language transfer. Patel himself has called attention to the fact that both melody in music and prosody in speech rely primarily on the same acoustic parameter of pitch contour, with the former necessitating more precise acoustic processing than the latter<sup>64</sup>. This overlap in neural resources has been demonstrated in the studies outlined above, in which musicians tend to surpass non-musicians in prosodic pitch perception. On top of that, Patel and other researchers have shown that individuals on the other end of the spectrum with a musical disorder known as amusia exhibit deficits in perceiving speech prosody<sup>64–67</sup>. Accordingly, speech prosody is of exceptional relevance to the OPERA hypothesis. Secondly, unlike lexical tones which are only of pertinence to tone languages, speech prosody is an important aspect of all languages, thereby making it a universal topic of interest with great practical significance. Broadly speaking, prosody signals speaker intention and meaning, imparting crucial information pertaining to syntax and pragmatics<sup>68</sup>. Research on first language acquisition has found that prosodic sensitivity is related to literacy skills<sup>69</sup>, reading comprehension<sup>70,71</sup>, and speech comprehension<sup>72–74</sup>. In a similar vein, research on second language acquisition in children and adults has found that prosodic sensitivity might facilitate the learning of word order and new vocabulary<sup>75,76</sup>, while exposure to prosodic features of the target language apparently improves second language proficiency and fluency<sup>77,78</sup>. The findings yielded from this study will therefore have important pedagogical implications for language and literacy skills as well as foreign language learning.

In order to study prosodic pitch perception, we chose to adopt the well-replicated prosodic pitch contour expectancy violation task, as it has consistently demonstrated the musician advantage in different age groups and languages with robust findings. Given that previous studies revealed a trend in which participants, regardless of musicianship, showed superior performance in detecting prosodic pitch violations in a familiar language relative to an unfamiliar language<sup>34,35</sup>, two different language contexts were implemented in the present study. We included a non-native language context in part to help circumvent a problem we anticipated; namely, that tasks using native language stimuli might not be adequately sensitive to tease apart group differences<sup>18,23</sup>, especially for tone language speakers. Furthermore, by introducing both native and non-native language contexts, we hoped to examine music-to-language transfer effects both with and without the top-down influence from other types of linguistic processing, allowing us to better assess the generalisability of the effects. Finally, the two language contexts mirror first language competence and second language learning respectively, shedding light on the practical application of the enduring music-to-language transfer effects in former musicians, if any.

## Method

**Participants.** Participants were recruited to take part in the study via an online screening questionnaire. They were between 19 and 42 years old ( $M = 23.04$ ,  $SD = 3.90$ ), with normal hearing based on an audiometric test (25 dB HL for octave frequencies from 500 to 4000 Hz). All of the participants were either native Singaporeans or had lived in Singapore for at least 10 years to ensure that they were familiar with the local accented variety of English. They had no formal exposure to the French language, the non-native speech stimuli used in this study.

A total of 127 individuals participated in this study. Data from 31 participants was excluded due to the following cases: (a) participants with self-reported exposure to French ( $n = 8$ ); (b) participants who had between 2 and 6 years of musical training experience ( $n = 23$ ).

The final dataset consisted of 96 participants. They were classified into three groups based on information obtained from a self-report questionnaire on their language and music background. In this study, active musicians consisted of those who had had at least 6 years of musical training and were still currently maintaining a consistent practice schedule of at least 3 h per week in the past 2 years ( $n = 22$ ). On the other hand, former musicians referred to those who similarly had at least 6 years of musical training but had stopped maintaining a regular practice schedule for at least 2 years ( $n = 27$ ). Finally, non-musicians referred to those who had had less than 2 years of musical training ( $n = 47$ ). Those with musical training predominantly had experience in string, wind, and vocal musical training. None of the participants were musicians by profession. Reflecting the diversity of multilingualism in the local population, the majority of the participants were proficient in English and Mandarin Chinese ( $n = 86$ ), while several were proficient in English and a second language other than Mandarin Chinese, specifically Malay ( $n = 3$ ), Tamil ( $n = 5$ ), Tagalog ( $n = 1$ ), and Burmese ( $n = 1$ ). The representation of non-Mandarin Chinese speakers was similar across groups,  $\chi^2(2) = 2.355$ ,  $p = 0.308$ .

To validate the grouping, participants' general musical abilities were assessed using the Musical Ear Test (MET)<sup>79</sup>. The MET consisted of two components: the melody subtest and the rhythm subtest. For each subtest, participants listened to 52 pairs of phrases, and had to judge whether the second phrase was the "same" or "different" compared to the first phrase. Half of the trials were "same" trials and the other half were "different" trials. The "different" trials involved a pitch violation in the melody subtest and a rhythmic change in the rhythm subtest. The MET stimuli were delivered via headphones, and participants gave their responses on an accompanying answer sheet. All participants completed the melody subtest followed by the rhythm subtest. Table 1 shows the final sample and descriptive information of each participant group.

## Materials and procedures

The research procedures were approved by the Institutional Review Board at the Nanyang Technological University. All research methods were performed in accordance with the relevant guidelines and regulations. Written informed consent was obtained from all participants and/or their legal guardians before participation.

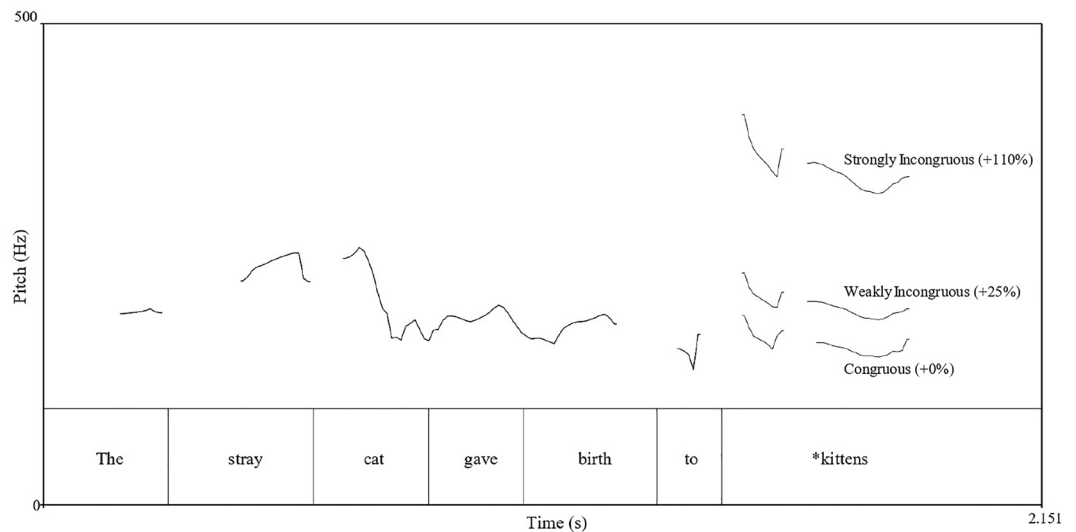
After providing their written informed consent, participants were seated comfortably in a soundproof booth, and undertook two experimental tasks. Firstly, the participants completed a two-choice speech pitch discrimination task. The English and French language blocks were counterbalanced across participants, with half of the participants presented with the English set first, and the other half with the French set first. Secondly, the participants completed a general musical abilities test, i.e., the MET. Short breaks were given between tasks to prevent fatigue. The total length of time for participation was approximately 1 h, and the participants were monetarily compensated for their time upon successful completion of the experiment.

Participants' linguistic perception abilities were assessed using a pitch violation discrimination task that has been well-replicated in the literature<sup>31–35</sup>. For the pitch discrimination task, 40 spoken declarative sentences in English and French respectively were recorded to form the experimental speech stimuli (see Supplementary Tables 1 and 2). The sentences were compiled and modified from a combination of sources, including the Harvard Sentences database<sup>80</sup> for the English stimuli and Smith's paper<sup>81</sup> for the French stimuli, with the final word in each sentence being disyllabic as in Marques et al.'s study<sup>34</sup>. Two female speakers, one native in Singapore English and the other in French, voiced the English and French sentences respectively at a normal speaking rate. The recorded sentences were then digitised (sampling at 44.1 kHz and 16 bit) using Audacity<sup>®</sup> Version 2.0.5.0<sup>82</sup>.

For each language, there were three different auditory conditions, and 40 sentences were presented in each auditory condition, thus leading to a total of 120 sentences. The final word of each sentence was either prosodically congruous, weakly incongruous, or strongly incongruous. In the prosodically incongruous conditions, the pitch (F0) of the final words was increased using Praat<sup>83</sup>, such that there was a local pitch manipulation on the final words (+25% in the weakly incongruous condition, +110% in the strongly incongruous condition) while maintaining the original natural global pitch contour (Fig. 1). The pitch increases used in the present study differ from those used in past studies (+35% in the weakly incongruous condition, +120% in the strongly incongruous condition)<sup>31–35</sup>. Preliminary pilot testing using conventional pitch increase values revealed a ceiling effect among our Singaporean participants, likely because enhanced pitch perception abilities in speech have been associated with bilingualism<sup>84</sup> and tone language experience<sup>6,24–30</sup>. As such, we reduced the pitch incongruity

	Participant group			Group differences on musical abilities (MET scores) as revealed by one-way ANOVA
	Active musicians (AM)	Former musicians (FM)	Non-musicians (NM)	
N	22	27	47	N/A
Age	21.45 (2.87)	23.07 (3.93)	23.77 (4.13)	N/A
No. of years of musical training	11.00 (3.80)	9.37 (3.12)	0.09 (0.28)	$F(2,93) = 202.564, p < 0.001$ AM > NM ( $p < 0.001$ ) AM = FM ( $p = 0.069$ ) FM > NM ( $p < 0.001$ )
No. of practice hours per week within the past two years	5.91 (2.51)	0.70 (0.77)	N/A	N/A
No. of years since musical training was discontinued	0.09 (0.29)	5.26 (3.44)	N/A	N/A
MET melody (%)	85.58 (7.43)	77.99 (7.92)	69.23 (7.38)	$F(2,93) = 37.299, p < 0.001$ AM > NM ( $p < 0.001$ ) AM > FM ( $p = 0.002$ ) FM > NM ( $p < 0.001$ )
MET rhythm (%)	78.76 (7.93)	73.29 (7.77)	66.98 (9.49)	$F(2,93) = 14.579, p < 0.001$ AM > NM ( $p < 0.001$ ) AM = FM ( $p = 0.093$ ) FM > NM ( $p = 0.010$ )
Language background	20 English-Mandarin bilingual 1 English-Malay bilingual 1 English-Tagalog bilingual	26 English-Mandarin bilingual 1 English-Tamil bilingual	40 English-Mandarin bilingual 2 English-Malay bilingual 4 English-Tamil bilingual 1 English-Burmese bilingual	

**Table 1.** Participant group demographics. Mean values and standard deviations (in parentheses) are given for age, musical background, and musical abilities. For the group differences on musical abilities, one-way ANOVA and pairwise comparisons are Bonferroni corrected.



**Figure 1.** Fundamental frequency (F0 in Hz) for a sample sentence in the three prosodic conditions.

in order to increase the difficulty of the task, and preliminary pilot testing using our modified pitch increase values obtained pitch discrimination accuracy rates across the experimental conditions which were similar to those found by Marques et al.<sup>34</sup>.

Participants listened to the speech stimuli via headphones. They were briefed that they would be listening to either English or French sentences, and that comprehension of the sentences was not required. In each trial, participants were asked to judge whether the final word of each sentence sounded normal (congruous condition) or strange (weakly incongruous or strongly incongruous conditions). Responses were recorded via a keyboard press, “N” or “S” respectively. Participants were asked to provide a response within 3 s. The practice phase consisted of 6 trials, with feedback provided at the end of each trial to indicate if the participants had answered correctly. The experimental phase consisted of 120 trials, broken up into four blocks of 30 sentences each. Sentence blocks were counterbalanced across participants; half of the participants in each group heard blocks one and two first, while the other half heard blocks three and four first. Sentences from each experimental condition occurred equally frequently within each block and in pseudorandom order. Up to three consecutive “strange” trials were allowed within each block, while pitch-manipulated variants of the same sentence were not allowed to occur within the same block.

## Results

A  $2 \times 3 \times 3$  mixed ANOVA was conducted with pitch discrimination accuracy as the dependent variable, language (native vs. non-native) and prosodic congruity (congruous vs. weakly incongruous vs. strongly incongruous) as the within-subject factors, and music group (active musicians vs. former musicians vs. non-musicians) as the between-subject factor. As Mauchly’s Test indicated that the assumption of sphericity had been violated for the prosodic congruity effect,  $\chi^2(2) = 242.497$ ,  $p < 0.001$ , and the language by prosodic congruity effect,  $\chi^2(2) = 153.767$ ,  $p < 0.001$ , Greenhouse–Geisser correction was applied,  $\epsilon = 0.519$  and  $\epsilon = 0.552$  respectively. As Box’s M Test indicated that the assumption of equality of covariance had been violated, Box’s  $M = 249.251$ ,  $F = 5.307$ ,  $p < 0.001$ , Pillai’s Trace was used. For all pairwise comparisons, Bonferroni correction was applied.

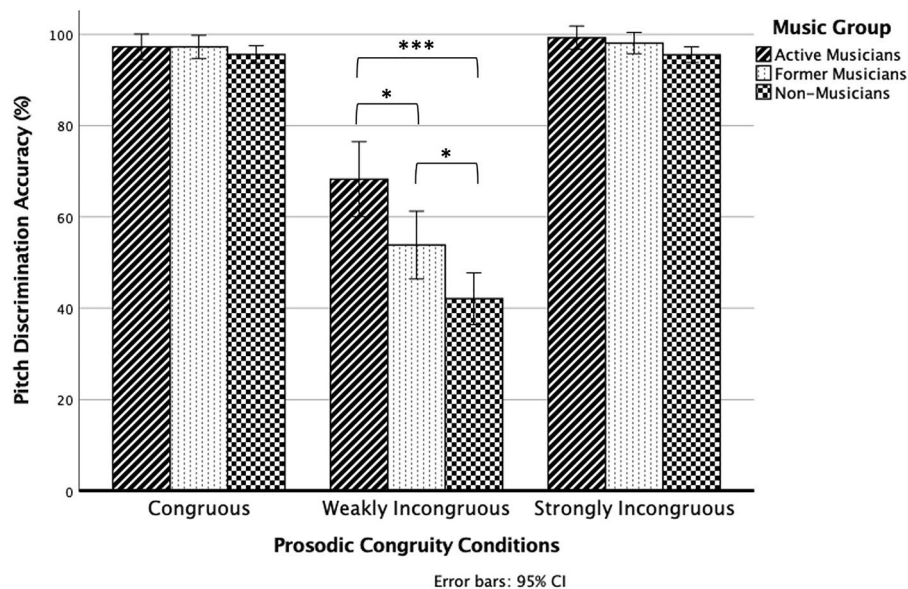
The interaction effect between prosodic congruity and music group was statistically significant,  $F(2.074, 96.456) = 10.124$ ,  $p < 0.001$  (Fig. 2). The main source of the interaction effect as revealed by simple effect analyses was from the weakly incongruous condition,  $F(2, 93) = 13.877$ ,  $p < 0.001$ ; the mean pitch discrimination accuracy was significantly different between active musicians (68%), former musicians (54%), and non-musicians (42%). Pairwise comparisons revealed that active musicians outperformed former musicians ( $p = 0.034$ ) and non-musicians ( $p < 0.001$ ), while former musicians also outperformed non-musicians ( $p = 0.042$ ).

The effect of music group was significant for the strongly incongruous condition,  $F(2, 93) = 3.293$ ,  $p = 0.042$ ; the mean pitch discrimination accuracy was significantly different between active musicians (99%), former musicians (98%), and non-musicians (96%). However, pairwise comparisons revealed no significant differences between groups after Bonferroni correction. Active musicians did not differ from non-musicians ( $p = 0.060$ ), and former musicians differed from neither active musicians ( $p = 1.000$ ) nor non-musicians ( $p = 0.257$ ). Meanwhile, the effect of music group was not significant for the congruous condition,  $F(2, 93) = 0.767$ ,  $p = 0.467$ .

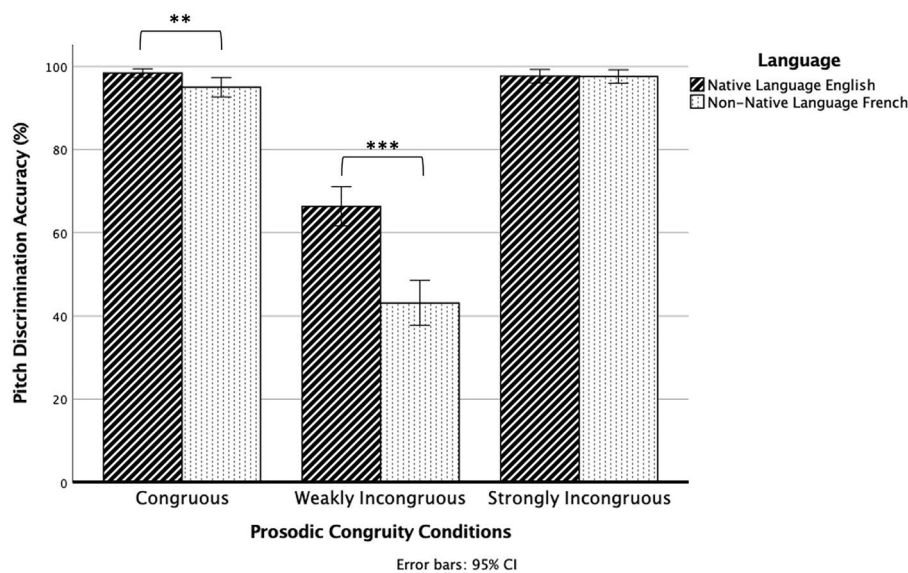
The three-way interaction between language and prosodic congruity and music group was not significant,  $F(2.207, 102.648) = 0.335$ ,  $p = 0.737$ ; neither was the interaction between language and music group,  $F(2.000, 93.000) = 0.615$ ,  $p = 0.543$ .

There was also a significant interaction effect between language and prosodic congruity,  $F(1.104, 102.648) = 39.450$ ,  $p < 0.001$  (Fig. 3). The effect of language was significant for the weakly incongruous condition,  $F(1.000, 93.000) = 63.833$ ,  $p < 0.001$ , where participants showed higher pitch discrimination accuracy





**Figure 2.** Pitch discrimination accuracy of active musicians, former musicians, and non-musicians in the three prosodic conditions. Error bars denote standard error. \* $p < 0.05$ , \*\*\* $p < 0.001$ .



**Figure 3.** Pitch discrimination accuracy for the native language English and non-native language French in the three prosodic conditions. Error bars denote standard error. \*\* $p < 0.01$ , \*\*\* $p < 0.001$ .

in their native language English (66%) than in their non-native language French (43%). The effect of language was also significant for the congruous condition,  $F(1,000,93.000) = 8.816$ ,  $p = 0.004$ , where participants showed higher pitch discrimination accuracy in their native language English (98%) than in their non-native language French (95%). However, the effect of language was not significant for the strongly incongruous condition,  $F(1,000,93.000) = 0.008$ ,  $p = 0.931$ .

Significant main effects were found for language,  $F(1,000,93.000) = 85.129$ ,  $p < 0.001$ , prosodic congruity,  $F(1,037,96.456) = 376.582$ ,  $p < 0.001$ , and music group,  $F(2,93) = 14.275$ ,  $p < 0.001$ . Participants showed higher pitch discrimination accuracy in their native language English (87%) than in their non-native language French (79%). The weakly incongruous condition (52%) was the most difficult to detect compared to the congruous condition (96%) and strongly incongruous condition (97%). Active musicians (88%) and former musicians (83%) showed higher pitch discrimination accuracy compared to non-musicians (78%).

## Discussion

The present study is one of the first to ascertain whether individuals who have discontinued musical training and practice retain a behavioural advantage over non-musicians in pitch perception abilities in speech. Our key finding is that there was a significant interaction effect between prosodic congruity and music group. In the weakly incongruous condition where pitch deviations were small and difficult to detect, our results showed a stepwise progression in pitch discrimination accuracy, with active musicians having better performance than former musicians, who in turn had better performance than non-musicians.

Our finding of an advantage by musicians over non-musicians in pitch discrimination echoes past findings that musical training facilitates pitch perception in the language domain<sup>5–23,31–39</sup>, thereby pointing towards a common domain-general pitch processing mechanism in music and speech perception as described in the OPERA hypothesis<sup>4</sup>. Our finding also coheres with findings of long-lasting neural changes from past musical training in young adults<sup>61</sup> and older adults<sup>62</sup>, as well as findings of improved cognitive performance in adulthood even after ceasing musical training<sup>59,60</sup>. Taken together, it appears that former musicians share similar neural enhancement as active musicians, and that the sharpened neural processing translates to perceptual benefits behaviourally. One explanation is that musical training requires individuals to attend to subtle sound contrasts, such as in pitch and duration. Consequently, musicians become more sensitive to such subtle acoustic cues, which has a positive spillover effect when discriminating similar contrasts in speech. Musical training contributes to the establishment of cognitive enhancement, such that there are some enduring cross-domain transfer benefits of musical training on the discrimination of subtle speech contrasts even after musical training and practice is discontinued.

More importantly, our finding that former musicians differed from active musicians qualifies the extent of the positive music-to-language transfer effects. Drawing on previous studies showing a clear behavioural advantage held by adult musicians over non-musicians in a pitch contour violation task similar to that used in the present study<sup>31,34</sup>, it appears that former musicians retain some musician advantage, but that such advantage may fade over time after musical training and practice is discontinued. Parallel results were seen in the data from the musical abilities tests. There were significant group differences in MET melody scores (Table 1), where active musicians had better musical abilities than former musicians, and former musicians in turn had better musical abilities than non-musicians. The difference in pitch discrimination and musical abilities between active musicians and former musicians cannot be attributed to the length of musical training, which was similar for the two groups as seen in the pairwise comparison (Table 1). Therefore, our data suggests that music attrition in pitch perception manifests in both the music and language domains among former musicians. One possible explanation is that subtle acoustic cues may no longer be behaviourally relevant in former musicians' everyday auditory environment, such that positive music-to-language transfer benefits may diminish over time. This interpretation is corroborated by exploratory Pearson correlations conducted to assess the relationship between the number of years since discontinuing musical practice and other factors. As reported in the Supplementary Results, in the sample of musicians both active and former, the number of years since discontinuing musical practice was significantly negatively correlated with pitch discrimination accuracy for the weakly incongruous prosodic condition in the native language English and non-native language French (see Supplementary Figs. 1 and 2), as well as with MET melody and rhythm subtest performance. The longer the period since discontinuing musical training and practice, the poorer one is at discriminating subtle violations in speech and music, indicating that there might be a gradual attenuation in finer acoustic discrimination abilities among former musicians. Future research can be conducted with older adults as participants, including former musicians who have discontinued musical training and practice for a longer period, to examine the issue further.

We also found a significant main effect for language and a significant interaction effect between language and prosodic congruity. Apart from a native language context with multiple possible sources of information that might assist in pitch processing, we implemented a non-native language context without additional information for participants to rely on. The interaction effect revealed that in both the congruous and weakly incongruous conditions, participants were more accurate in prosodic pitch discrimination in their native language than in a non-native language, consistent with previous findings by Marques et al.<sup>34</sup> and Deguchi et al.<sup>35</sup>. There are two possible explanations for this native language advantage, which Deguchi et al.<sup>35</sup> investigated by introducing jabberwocky sentences that preserved the intonational contours of the native language but consisted of meaningless legal pseudowords. They found that participants were better at detecting pitch changes in their native language than in jabberwocky, but were also better at detecting pitch changes in jabberwocky than in the non-native language. This suggests that participants were familiar with typical intonational contours in their native language, and were consequently better able to detect pitch changes in the native language and jabberwocky speech stimuli but not in the non-native language speech stimuli. At the same time, participants could understand the meaning of their native language, and for that reason were better able to anticipate when the final word carrying the pitch variation would occur in the native language speech stimuli but not in the jabberwocky and non-native language speech stimuli. The native language advantage observed in our study can thus be explained by the fact that participants were making use of both prosodic and semantic information to complete the pitch incongruity detection task. Nonetheless, regardless of the language used for the speech stimuli, group differences were seen in the weakly incongruous condition. Active and former musicians were more accurate than non-musicians in detecting prosodic pitch violation no matter whether they had prior knowledge of the language tested. In the non-native language context which better isolated the prosodic pitch dimension without influence from other types of linguistic processing, participants would not have been able to exploit additional top-down processing frameworks and would have been relying solely on bottom-up pitch perceptual sensitivity to acoustic cues. In the native language context, group differences were also seen despite the fact that all participants were able to draw on additional linguistic resources. The present study hence extends the OPERA hypothesis<sup>4</sup> by underscoring that—presumably due to their prior musical training with melodic pitch patterns—former musicians retain

enhanced underlying pitch processing abilities, which generalise to the perception of prosodic pitch contours in speech for both native and non-native languages.

A caveat to keep in mind, though, is that although there is a large body of evidence in which musicians—be it active and former—outperform non-musicians in various pitch perception tasks, it may not be straightforward to conclude that the so-called musician advantage is a result of musical training. It is plausible that the results reported in this present study may be driven by a third, unexplored factor, such as general intelligence, education background, or socioeconomic status. On top of that, as with most of the previously reported literature, this present study adopted a cross-sectional design comparing different population groups at a specific point in time. In recent years, some researchers have propounded the idea that inherent musical abilities, rather than musical training, might be linked to enhanced speech perception<sup>30,85,86</sup>. Ergo, the music-to-language transfer effects that we speak of may be a consequence of pre-existing differences and self-selection, as opposed to a consequence of musical training per se. In other words, individuals pre-disposed with superior auditory or pitch processing abilities to begin with may be more inclined to pick up and continue musical training, such that the differences observed between active, former, and non-musicians later in life may not be a direct outcome of musical training in and of itself.

However, as highlighted in the introduction, there is some compelling evidence in the existing literature that musical training has a causal influence on brain development and pitch perception. Participants initially matched in musical aptitude, general intelligence, and socioeconomic status have been shown to demonstrate group differences in neurological and behavioural post-tests related to pitch perception depending on the training they are randomly assigned to<sup>33,45–47</sup>. Of particular relevance to our study, Moreno et al.<sup>33</sup> conducted a longitudinal experimental study with 8-year-old children without any prior musical training. Pre-tests confirmed that the children were initially matched in pitch perception performance, general cognitive abilities, as well as socioeconomic status. These children were then randomly assigned to receive 6 months of either musical training or painting training. The researchers recorded both electrophysiological and behavioural measures for a pitch violation discrimination task similar to that used in this present study. They found that children who received musical training, but not those who received painting training, showed improved prosodic pitch discrimination abilities in speech. Along the same lines, Nan et al.<sup>45</sup> randomly assigned 4- to 5-year-old children with tone language experience to receive 6 months of piano training, reading training, or no training. Although the children were initially matched in general cognitive abilities and socioeconomic status, and although all groups showed improvements in general cognitive abilities, only those who received piano training showed enhanced cortical responses to pitch changes in music and speech which were correlated with behavioural performance. These findings suggest that musical training can indeed cause experience-dependent transfer effects that cannot be attributed to external factors or pre-existing differences, while our study further suggests that some transfer effects may be retained even after musical training and practice is discontinued. Having said that, future research can strengthen our finding by performing an intervention study with longitudinal randomised controlled trials to track and compare the effects of long-term, short-term, and no musical training among individuals who are otherwise matched on other variables.

In conclusion, our study shows that musical training confers positive cross-domain benefits in speech perception, adding to the body of literature on music-to-language transfer and suggesting that there is a common pitch processing mechanism underlying pitch perception in the two domains. More importantly, our results further show that these benefits may be retained to some extent over time, such that former musicians show some behavioural advantage over non-musicians even after the discontinuation of musical training and practice. Situated within the OPERA hypothesis<sup>4</sup>, it appears that musical training alters the shared neural networks for music and speech in a long-lasting manner, such that the musician advantage applies not only to active musicians but to former musicians as well. Moreover, this advantage in prosodic pitch perception is seen with both native and non-native languages. Possible future directions for research include using neurological and behavioural measures to compare active musicians, former musicians, and non-musicians' pitch perception abilities in the language domain in other areas such as the perception of lexical tones, stress, and emotional prosody. Our findings have real-life implications for boosting first language acquisition and foreign language learning, as well as protecting against age-related cognitive and auditory decline in the ageing population. It appears that musical training and practice can serve as an effective enrichment activity and intervention method to improve speech perception, and that individuals can reap some long-lasting cognitive benefits throughout their lifespan even after musical training and practice is discontinued.

### Data availability

The dataset generated during and/or analysed during the current study is included in the Supplementary Information file.

Received: 19 September 2022; Accepted: 9 February 2023

Published online: 14 February 2023

### References

- Schellenberg, E. G. & Weiss, W. M. Music and cognitive abilities. In *The Psychology of Music* (ed. Deutsch, D.) 499–550 (Academic Press, 2013).
- Parbery-Clark, A., Anderson, S., Hittner, E. & Kraus, N. Musical experience offsets age-related delays in neural timing. *Neurobiol. Aging* **33**(1483), e1–4. <https://doi.org/10.1016/j.neurobiolaging.2011.12.015> (2012).
- Parbery-Clark, A., Skoe, E., Lam, C. & Kraus, N. Musician enhancement for speech-in-noise. *Ear. Hear.* **30**, 653–661. <https://doi.org/10.1097/AUD.0b013e3181b412e9> (2009).



4. Patel, A. D. Why would musical training benefit the neural encoding of speech? The OPERA hypothesis. *Front. Psychol.* <https://doi.org/10.3389/fpsyg.2011.00142> (2011).
5. Alexander, J. A., Wong, P. C. M. & Bradlow, A. R. Lexical tone perception in musicians and non-musicians. In *Proceedings of Proc. Annual Conference of the International Speech Communication Association Interspeech* (2005).
6. Burnham, D., Brooker, R. & Reid, A. The effects of absolute pitch ability and musical training on lexical tone perception. *Psychol. Music.* **43**, 881–897. <https://doi.org/10.1177/0305735614546359> (2015).
7. Choi, W. The selectivity of musical advantage: Musicians exhibit perceptual advantage for some but not all Cantonese tones. *Music Percept.* **37**, 423–434. <https://doi.org/10.1525/MP.2020.37.5.423> (2020).
8. Delogu, F., Lampis, G. & Belardinelli, M. O. From melody to lexical tone: Musical ability enhances specific aspects of foreign language perception. *Eur. J. Cogn. Psychol.* **22**, 46–61. <https://doi.org/10.1080/09541440802708136> (2010).
9. Gottfried, T. L. & Riester, D. Relation of pitch glide perception and Mandarin tone identification. *J. Acoust. Soc. Am.* **108**, 2604. <https://doi.org/10.1121/1.4743698> (2000).
10. Gottfried, T. L., Staby, A. M. & Ziemer, C. J. Musical experience and Mandarin tone discrimination and imitation. *J. Acoust. Soc. Am.* **115**, 2545. <https://doi.org/10.1121/1.4783674> (2001).
11. Han, Y., Goudbeek, M., Mos, M. & Swerts, M. Mandarin tone identification by tone-naïve musicians and non-musicians in auditory-visual and auditory-only conditions. *Front. Commun.* **4**, 1–14. <https://doi.org/10.3389/fcomm.2019.00070> (2019).
12. Hung, T.-H. & Lee, C.-Y. Processing linguistic and musical pitch by English-speaking musicians and non-musicians. In *20th North American Conference on Chinese Linguistics* (2008).
13. Lee, C.-Y. & Hung, T.-H. Identification of Mandarin tones by English-speaking musicians and nonmusicians. *J. Acoust. Soc. Am.* **124**, 3235–3248. <https://doi.org/10.1121/1.2990713> (2008).
14. Marie, C. L., Delogu, F., Lampis, G., Belardinelli, M. O. & Besson, M. Influence of musical expertise on segmental and tonal processing in Mandarin Chinese. *J. Cogn. Neurosci.* **23**, 2701–2715. <https://doi.org/10.1162/jocn.2010.21585> (2011).
15. Wayland, R. P., Herrera, E. & Kaan, E. Effects of musical experience and training on pitch contour perception. *J. Phon.* **38**, 654–662. <https://doi.org/10.1016/j.wocn.2010.10.001> (2010).
16. Bidelman, G. M., Gandour, J. T. & Krishnan, A. Cross-domain effects of music and language experience on the representation of pitch in the human auditory brainstem. *J. Cogn. Neurosci.* **23**, 425–434. <https://doi.org/10.1162/jocn.2009.21362> (2011).
17. Wong, P. C. M., Skoe, E., Russo, N. M., Dees, T. & Kraus, N. Musical experience shapes human brainstem encoding of linguistic pitch patterns. *Nat. Neurosci.* **10**, 420–422. <https://doi.org/10.1038/nn1872> (2007).
18. Mok, P. P. K. & Zuo, D. The separation between music and speech: Evidence from the perception of Cantonese tones. *J. Acoust. Soc. Am.* **132**, 2711–2720. <https://doi.org/10.1121/1.4747010> (2012).
19. Ong, J. H., Wong, P. C. M. & Liu, F. Musicians show enhanced perception, but not production, of native lexical tones. *J. Acoust. Soc. Am.* **148**, 3443. <https://doi.org/10.1121/10.0002776> (2020).
20. Toh, X. R., Lau, F. & Wong, F. C. K. Individual differences in nonnative lexical tone perception: Effects of tone language repertoire and musical experience. *Front. Psychol.* **13**, 940363. <https://doi.org/10.3389/fpsyg.2022.940363> (2022).
21. Kolinsky, R., Cuvelier, H., Goetry, V., Peretz, I. & Morais, J. Music training facilitates lexical stress processing. *Music Percept.* **26**, 235–246. <https://doi.org/10.1525/mp.2009.26.3.235> (2009).
22. Choi, W. Towards a native OPERA hypothesis: Musicianship and English stress perception. *Lang. Speech* **65**, 697–712. <https://doi.org/10.1177/00238309211049458> (2022).
23. Choi, W. What is “music” in music-to-language transfer? Musical ability but not musicianship supports Cantonese listeners’ English stress perception. *J. Speech Lang. Hear. Res.* **65**, 4047–4059. [https://doi.org/10.1044/2022\\_JSLHR-22-00175](https://doi.org/10.1044/2022_JSLHR-22-00175) (2022).
24. Lee, Y.-S., Vakoč, D. A. & Lee, H. W. Tone perception in Cantonese and Mandarin: A cross-linguistic comparison. *J. Psycholinguist. Res.* **25**, 527–542. <https://doi.org/10.1007/BF01758181> (1996).
25. Morett, L. M. The influence of tonal and atonal bilingualism on children’s lexical and non-lexical tone perception. *Lang. Speech* **63**, 221–241. <https://doi.org/10.1177/0023830919834679> (2020).
26. Qin, Z. & Mok, P. P. K. Discrimination of Cantonese tones by speakers of tone and non-tone languages. *Kans. Work. Pap. Linguist.* **34**, 26–42. <https://doi.org/10.17161/KWPL.1808.12864> (2013).
27. Schaefer, V. & Darcy, I. Lexical function of pitch in the first language shapes cross-linguistic perception of Thai tones. *Lab. Phonol.* **5**, 489–522. <https://doi.org/10.1515/lp-2014-0016> (2014).
28. Schaefer, V. & Darcy, I. Applying a newly learned second language dimension to the unknown: The influence of second language Mandarin tones on the naive perception of Thai tones. *Psychol. Lang. Commun.* **24**, 90–123. <https://doi.org/10.2478/plc-2020-0007> (2020).
29. Wayland, R. P. & Guion, S. G. Training English and Chinese listeners to perceive Thai tones: A preliminary report. *Lang. Learn.* **54**, 681–712. <https://doi.org/10.1111/j.1467-9922.2004.00283.x> (2004).
30. Wayland, R. P. & Li, B. Effects of two training procedures in cross-language perception of tones. *J. Phon.* **36**, 250–267. <https://doi.org/10.1016/j.wocn.2007.06.004> (2008).
31. Schön, D., Magne, C. & Besson, M. The music of speech: Music training facilitates pitch processing in both music and language. *Psychophysiology* **41**, 341–349. <https://doi.org/10.1111/1469-8986.00172.x> (2004).
32. Magne, C., Schön, D. & Besson, M. Musician children detect pitch violations in both music and language better than nonmusician children: Behavioral and electrophysiological approaches. *J. Cogn. Neurosci.* **18**, 199–211. <https://doi.org/10.1162/089892906775783660> (2006).
33. Moreno, S. *et al.* Musical training influences linguistic abilities in 8-year-old children: More evidence for brain plasticity. *Cereb. Cortex* **19**, 712–723. <https://doi.org/10.1093/cercor/bhn120> (2009).
34. Marques, C., Moreno, S., Castro, S. L. & Besson, M. Musicians detect pitch violation in a foreign language better than nonmusicians: Behavioral and electrophysiological evidence. *J. Cogn. Neurosci.* **19**, 1453–1463. <https://doi.org/10.1162/jocn.2007.19.9.1453> (2007).
35. Deguchi, C. *et al.* Sentence pitch change detection in the native and unfamiliar language in musicians and non-musicians: Behavioral, electrophysiological and psychoacoustic study. *Brain Res.* **1455**, 75–89. <https://doi.org/10.1016/j.brainres.2012.03.034> (2012).
36. Thompson, W. F., Schellenberg, E. G. & Husain, G. Perceiving prosody in speech: Effects of music lessons. *Ann. N. Y. Acad. Sci.* **999**, 530–532. <https://doi.org/10.1196/annals.1284.067> (2003).
37. Thompson, W. F., Schellenberg, E. G. & Husain, G. Decoding speech prosody: Do music lessons help?. *Emotion* **4**, 46–64. <https://doi.org/10.1037/1528-3542.4.1.46> (2004).
38. Farmer, E., Jicol, C. & Petrini, K. Musicianship enhances perception but not feeling of emotion from others’ social interaction through speech prosody. *Music Percept.* **37**, 323–338. <https://doi.org/10.1525/mp.2020.37.4.323> (2020).
39. Lima, C. F. & Castro, S. L. Speaking to the trained ear: Musical expertise enhances the recognition of emotions in speech prosody. *Emotion* **11**, 1021–1031. <https://doi.org/10.1037/a0024521> (2011).
40. Pantev, C. *et al.* Increased auditory cortical representation in musicians. *Nature* **392**, 811–814. <https://doi.org/10.1038/33918> (1998).
41. Pantev, C., Engelien, A., Candia, V. & Elbert, T. Representational cortex in musicians: Plastic alterations in response to musical practice. *Ann. N. Y. Acad. Sci.* **930**, 300–314. <https://doi.org/10.1111/j.1749-6632.2001.tb05740.x> (2001).
42. Schlaug, G. The brain of musicians: A model for functional and structural adaptation. *Ann. N. Y. Acad. Sci.* **930**, 281–299. <https://doi.org/10.1111/j.1749-6632.2001.tb05739.x> (2001).

43. Bermudez, P. & Zatorre, R. J. Differences in gray matter between musicians and nonmusicians. *Ann. N. Y. Acad. Sci.* **1060**, 395–399. <https://doi.org/10.1196/annals.1360.057> (2005).
44. Gaser, C. & Schlaug, G. Brain structures differ between musicians and non-musicians. *J. Neurosci.* **23**, 9240–9245. <https://doi.org/10.1523/JNEUROSCI.23-27-09240.2003> (2003).
45. Nan, Y. *et al.* Piano training enhances the neural processing of pitch and improves speech perception in Mandarin-speaking children. *Proc. Natl. Acad. Sci. USA* **115**, 6630–6639. <https://doi.org/10.1073/pnas.1808412115> (2018).
46. Hyde, K. L. *et al.* The effects of musical training on structural brain development. *Ann. N. Y. Acad. Sci.* **1169**, 182–186. <https://doi.org/10.1111/j.1749-6632.2009.04852.x> (2009).
47. Hyde, K. L. *et al.* Musical training shapes structural brain development. *J. Neurosci.* **29**, 3019–3025. <https://doi.org/10.1523/JNEUROSCI.5118-08.2009> (2009).
48. Kraus, N., Skoe, E., Parbery-Clark, A. & Ashley, R. Experience-induced malleability in neural encoding of pitch, timbre, and timing. *Ann. N. Y. Acad. Sci.* **11691**, 543–557. <https://doi.org/10.1111/j.1749-6632.2009.04549.x> (2009).
49. Wan, C. Y. & Schlaug, G. Music making as a tool for promoting brain plasticity across the lifespan. *Neuroscientist* **16**, 566–577. <https://doi.org/10.1177/1073858410377805> (2010).
50. Neves, L., Correia, A. I., Castro, S. L., Martins, D. & Lima, C. F. Does music training enhance auditory and linguistic processing? A systematic review and meta-analysis of behavioral and brain evidence. *Neurosci. Biobehav. Rev.* **140**, 104777. <https://doi.org/10.1016/j.neubiorev.2022.104777> (2022).
51. Zengel, B. R. & Alain, C. Musicians experience less age-related decline in central auditory processing. *Psychol. Aging* **27**, 410–417. <https://doi.org/10.1037/a0024816> (2012).
52. Zengel, B. R. & Alain, C. The influence of lifelong musicianship on neurophysiological measures of concurrent sound segregation. *J. Cogn. Neurosci.* **25**, 503–516. [https://doi.org/10.1162/jocn\\_a\\_00329](https://doi.org/10.1162/jocn_a_00329) (2013).
53. Alain, C., Zengel, B. R., Hutka, S. & Bidelman, G. M. Turning down the noise: The benefit of musical training on the aging auditory brain. *Hear. Res.* **308**, 162–173. <https://doi.org/10.1016/j.heares.2013.06.008> (2014).
54. Zhang, J. D., Susino, M., McPherson, G. E. & Schubert, E. The definition of a musician in music psychology: A literature review and the six-year rule. *Psychol. Music* **48**, 389–409. <https://doi.org/10.1177/0305735618804038> (2020).
55. Costa-Giomi, E. Music instruction and children's intellectual development: The educational context of music participation. In *Music, Health, and Wellbeing* (eds MacDonald, R. *et al.*) 339–355 (Oxford University Press, 2012).
56. Costa-Giomi, E. The long-term effects of childhood music instruction on intelligence and general cognitive abilities. *Update Appl. Res. Music Educ.* **33**, 20–26. <https://doi.org/10.1177/8755123314540661> (2015).
57. Costa-Giomi, E. & Ryan, C. The benefits of music instruction: What remains years later. *Symp. Res. Music Behav.* **20**, 25 (2007).
58. Costa-Giomi, E. The effects of three years of piano instruction on children's cognitive development. *J. Res. Music Educ.* **47**, 198–212. <https://doi.org/10.2307/3345779> (1999).
59. Schellenberg, E. G. Long-term positive associations between music lessons and IQ. *J. Educ. Psychol.* **98**, 457–468. <https://doi.org/10.1037/0022-0663.98.2.457> (2006).
60. Strong, J. V. & Midden, A. Cognitive differences between older adult instrumental musicians: Benefits of continuing to play. *Psychol. Music* **48**, 67–83. <https://doi.org/10.1177/0305735618785020> (2020).
61. Skoe, E. & Kraus, N. A little goes a long way: How the adult brain is shaped by musical training in childhood. *J. Neurosci.* **32**, 11507–11510. <https://doi.org/10.1523/JNEUROSCI.1949-12.2012> (2012).
62. White-Schwoch, T., Carr, K. W., Anderson, S., Strait, D. L. & Kraus, N. Older adults benefit from music training early in life: Biological evidence for long-term training-driven plasticity. *J. Neurosci.* **33**, 17667–17674. <https://doi.org/10.1523/JNEUROSCI.2560-13.2013> (2013).
63. Wennerstrom, A. *The Music of Everyday Speech: Prosody and Discourse Analysis* (Oxford University Press, 2001).
64. Patel, A. D., Wong, M., Foxton, J., Lochy, A. & Peretz, I. Speech intonation perception deficits in musical tone deafness (congenital amusia). *Music Percept.* **25**, 357–368. <https://doi.org/10.1525/mp.2008.25.4.357> (2008).
65. Hutchins, S., Gosselin, N. & Peretz, I. Identification of changes along a continuum of speech intonation is impaired in congenital amusia. *Front. Psychol.* **1**, 236. <https://doi.org/10.3389/fpsyg.2010.00236> (2010).
66. Jiang, C., Hamm, J. P., Lim, V. K., Kirk, I. J. & Yang, Y. Processing melodic contour and speech intonation in congenital amusia with Mandarin Chinese. *Neuropsychologia* **48**, 2630–2639. <https://doi.org/10.1016/j.neuropsychologia.2010.05.009> (2010).
67. Liu, F., Patel, A. D., Fourcin, A. & Stewart, L. Intonation processing in congenital amusia: Discrimination, identification and imitation. *Brain* **133**, 1682–1693. <https://doi.org/10.1093/brain/awq089> (2010).
68. Monrad-Krohn, G. H. The prosodic quality of speech and its disorders. *Acta Psychiatr. Scand.* **22**, 255–269. <https://doi.org/10.1111/j.1600-0447.1947.tb08246.x> (1947).
69. Zhang, J. & McBride-Chang, C. Auditory sensitivity, speech perception, and reading development and impairment. *Educ. Psychol. Rev.* **22**, 323–338. <https://doi.org/10.1007/s10648-010-9137-4> (2010).
70. Holliman, A. J. *et al.* Beginning to disentangle the prosody-literacy relationship: A multi-component measure of prosodic sensitivity. *Read. Writ.* **27**, 255–266. <https://doi.org/10.1007/s11145-013-9443-6> (2014).
71. Groen, M. A., Veenendaal, N. J. & Verhoeven, L. The role of prosody in reading comprehension: Evidence from poor comprehenders. *J. Res. Read.* **42**, 37–57. <https://doi.org/10.1111/1467-9817.12133> (2019).
72. Cutler, A., Dahan, D. & van Donselaar, W. Prosody in the comprehension of spoken language: A literature review. *Lang. Speech* **40**, 141–201. <https://doi.org/10.1177/002383099704000203> (1997).
73. Hellbernd, N. & Sammler, D. Prosody conveys speaker's intentions: Acoustic cues for speech act perception. *J. Mem. Lang.* **88**, 70–86. <https://doi.org/10.1016/j.jml.2016.01.001> (2016).
74. Hupp, J. M., Jungers, M. K., Hinerman, C. M. & Porter, B. L. Cup! Cup? Cup: Comprehension of intentional prosody in adults and children. *Cogn. Dev.* **57**, 100971. <https://doi.org/10.1016/j.cogdev.2020.100971> (2021).
75. Campfield, D. E. & Murphy, V. A. The influence of prosodic input in the second language classroom: Does it stimulate child acquisition of word order and function words? *Lang. Learn. J.* **45**, 81–99. <https://doi.org/10.1080/09571736.2013.807864> (2017).
76. Saksida, A., Fló, A., Guedes, B., Nespór, M. & Peña Garay, M. Prosody facilitates learning the word order in a new language. *Cognition* **213**, 104686. <https://doi.org/10.1016/j.cognition.2021.104686> (2021).
77. Saito, Y. & Saito, K. Differential effects of instruction on the development of second language comprehensibility, word stress, rhythm, and intonation: The case of inexperienced Japanese EFL learners. *Lang. Teach. Res.* **21**, 589–608. <https://doi.org/10.1177/1362168816643111> (2017).
78. Yenkimaleki, M. Prosody training benefits in perception vs production skills in simultaneous interpreting: An experimental study. *Dutch J. Appl. Linguist.* <https://doi.org/10.51751/dujal9888> (2021).
79. Wallentin, M., Nielsen, A. H., Friis-Olivarius, M., Vuust, C. & Vuust, P. The Musical Ear Test, a new reliable test for measuring musical competence. *Learn. Individ. Differ.* **20**, 188–196. <https://doi.org/10.1016/j.lindif.2010.02.004> (2010).
80. Rothauer, E. H. *et al.* IEEE recommended practice for speech quality measures. *IEEE Trans. Audio Electroacoust.* **17**, 225–246. <https://doi.org/10.1109/TAU.1969.1162058> (1969).
81. Smith, C. L. Prosodic finality and sentence type in French. *Lang. Speech* **45**, 141–178. <https://doi.org/10.1177/00238309020450020301> (2002).
82. Audacity: Free Audio Editor and Recorder v. 2.3.2 (2018).
83. Boersma, P. Praat, a system for doing phonetics by computer. *Glott Int.* **5**, 341–345 (2001).

84. Krizman, J., Marian, V., Shook, A., Skoe, E. & Kraus, N. Subcortical encoding of sound is enhanced in bilinguals and relates to executive function advantages. *Proc. Natl. Acad. Sci. USA* **109**, 7877–7881. <https://doi.org/10.1073/pnas.1201575109> (2012).
85. Mankel, K. & Bidelman, G. M. Inherent auditory skills rather than formal music training shape the neural encoding of speech. *Proc. Natl. Acad. Sci. USA* **115**, 13129–13134. <https://doi.org/10.1073/pnas.1811793115> (2018).
86. Swaminathan, S. & Schellenberg, E. G. Musical ability, music training, and language ability in childhood. *J. Exp. Psychol. Learn. Mem. Cogn.* **46**, 2340–2348. <https://doi.org/10.1037/xlm0000798> (2020).

### Acknowledgements

This study was supported by research grants from the Ministry of Education (MOE), Singapore (MOE2019-SSRTG-016, MOE-T2EP402A20-0003). We thank Dr. Alice H. D. Chan for her insight on the design of the study. We are also grateful to all participants for their contribution to the study.

### Author contributions

S.T. and G.W. contributed to the conception, design, and implementation of the study. X.T., S.T., G.W., F.L., and F.W. performed the statistical analysis and interpreted the data. X.T., S.T., and G.W. wrote drafts of the manuscript. X.T., F.L., and F.W. revised and finalised the manuscript. All authors reviewed and approved the submitted manuscript.

### Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-023-29733-3>.

**Correspondence** and requests for materials should be addressed to F.C.K.W.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023