Neuropsychologia 48 (2010) 2630–2639

Contents lists available at ScienceDirect

Neuropsychologia

journal homepage: www.elsevier.com/locate/neuropsychologia

# Processing melodic contour and speech intonation in congenital amusics with Mandarin Chinese

Cunmei Jiang[a,c], Jeff P. Hamm[b,*], Vanessa K. Lim[b], Ian J. Kirk[b], Yufang Yang[a,**]

[a] *Institute of Psychology, Chinese Academy of Science, Beijing 100101, China*
[b] *Research Centre for Cognitive Neuroscience, University of Auckland, Auckland, New Zealand*
[c] *Institute of Music, Fujian Normal University, Fuzhou, China*

ARTICLE INFO

*Article history:*
Received 15 January 2010
Received in revised form 26 April 2010
Accepted 1 May 2010
Available online 13 May 2010

*Keywords:*
Congenital amusia
Mandarin Chinese
Melodic contour
Speech intonation
Discrimination
Identification
Tonal languages

ABSTRACT

Congenital amusia is a disorder in the perception and production of musical pitch. It has been suggested that early exposure to a tonal language may compensate for the pitch disorder (Peretz, 2008). If so, it is reasonable to expect that there would be different characterizations of pitch perception in music and speech in congenital amusics who speak a tonal language, such as Mandarin. In this study, a group of 11 adults with amusia whose first language was Mandarin were tested with melodic contour and speech intonation discrimination and identification tasks. The participants with amusia were impaired in discriminating and identifying melodic contour. These abnormalities were also detected in identifying both speech and non-linguistic analogue derived patterns for the Mandarin intonation tasks. In addition, there was an overall trend for the participants with amusia to show deficits with respect to controls in the intonation discrimination tasks for both speech and non-linguistic analogues. These findings suggest that the amusics' melodic pitch deficits may extend to the perception of speech, and could potentially result in some language deficits in those who speak a tonal language.



## 1. Introduction

Music and speech both feature structured melodic patterns (Patel, 2006). Pitch changes can be described in terms of the pattern of rises and falls in pitch and in the actual pitches at each point in time (Foxton, Brown, Chambers, & Griffiths, 2004). What is referred to as a melodic contour in music is called "intonation" in speech (Patel & Peretz, 1997). Despite the similarities between melodic contours and intonation contours, there is a paucity of research investigating the cognitive and neural resources shared between the two processes (Patel, 2006; Patel & Peretz, 1997).

Congenital amusia (hereafter amusia) is a disorder in the perception and production of pitch. Previous research has shown that people with amusia who speak non-tonal languages have difficulty processing different musical tones, melodic contours, and songs (Ayotte, Peretz, & Hyde, 2002; Foxton, Dean, Gee, Peretz, & Griffiths, 2004; Gosselin, Jolicoeur, & Peretz, 2009; Hyde & Peretz, 2004; Peretz et al., 2002; Tillmann, Schulze, & Foxton, 2009). However, whether or not the musical pitch deficits associated with amusia extend into speech processing is still a matter of debate. Some suggest that those with amusia do not have compromised discrimination and identification of western speech intonation (Ayotte et al., 2002; Peretz et al., 2002). It has been suggested that speech intonation processing may be spared because speech intonation contours usually involve relatively large pitch movements whilst melodies generally use smaller pitch intervals (Peretz, 2002). Such an explanation suggests the disorder experienced by amusic participants is not specific to the musical domain, but merely music-relevant (Peretz, 2002; Peretz & Hyde, 2003).

In contrast, it has been shown that musically tone-deaf individuals have difficulty discriminating intonation contours extracted from speech (Patel, Foxton, & Griffiths, 2005). In addition some amusics (30%) have been shown to have difficulty discriminating statements from questions based upon the final fall or rise in pitch (Patel, Wong, Foxton, Lochy, & Peretz, 2008). Individuals with amusia also perform worse than controls on tests of sensitivity to emotional prosody (Thompson, 2007). Based upon these findings, it appears that amusia may interfere with the processing of intonation contours, at least in speakers of non-tonal languages. Spared performance with verbal material, therefore, may be based upon inference from the semantic context rather than be due to the larger tonal intervals of the speech signals per se.

The previous studies concentrate on language processing in individuals with amusia who have a non-tonal language, English or

* Corresponding author at: Research Centre for Cognitive Neuroscience, Department of Psychology, 10 Symonds Street, University of Auckland, Auckland 1142, New Zealand. Tel.: +64 9 3737599 88519.
** Corresponding author at: Institute of Psychology, Chinese Academy of Sciences, Beijing, 100101, China. Tel: +86 10 64888629.
*E-mail addresses:* j.hamm@auckland.ac.nz (J.P. Hamm), yangyf@psych.ac.cn (Y. Yang).

0028-3932/$ – see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.neuropsychologia.2010.05.009

French, as their first language. However, the semantic comprehension in tonal languages, such as Mandarin and Cantonese, depends on recognition of lexical tone to a greater degree than the non-tonal languages. Thus, it is expected that (Western) individuals with amusia might have difficulties learning a tone language if their pitch perception is impaired for speech related stimuli (Peretz, 2008). To test this hypothesis, Nguyen, Tillmann, Gosselin, and Peretz (2009) employed Mandarin tones as speech stimuli. It was found that the amusic group as a whole performed significantly below the control group, although there was considerable overlap in performance between the groups.

In general, and with respect to the language backgrounds of amusic individuals, a cross-language study investigating speech suggests processing pitch in the auditory brainstem is sensitive to language experience (Krishnan, Xu, Gandour, & Cariani, 2005). More specifically, the neural substrates of pitch perception in the processing of lexical tones are shaped by long-term experience with a tonal language (Chandrasekaran, Krishnan, & Gandour, 2007a; Krishnan et al., 2005). This experience-dependent neural plasticity is not only specific to speech pitch contours, but also to non-speech analogues (Bent, Bradlow, & Wright, 2006; Chandrasekaran, Krishnan, & Gandour, 2007b; Chandrasekaran, Krishnan, & Gandour, 2009; Xu, Gandour, & Francis, 2006). Furthermore, genetic analyses have indicated a relationship between language and genes. Among a large database (983 alleles and 26 linguistic features in 49 populations), it has been found that linguistic tone was related to two growth-related genes (ASPM and microcephalin) at the population level (Dediu & Ladd, 2007). It suggests that the lower population frequency of the adaptive haplogroups of these two genes tends to occur in those who speak a tonal language even after accounting for geographical and historical factors. Although this relationship is correlational, it increases the plausibility of suggesting that some properties of language have cognitive, and ultimately genetic causes (Ladd, Dediu, & Kinsella, 2008).

Given that early exposure to a tonal language may compensate for the pitch disorder (Peretz, 2008), it seems reasonable to question whether there are any amusic individuals who speak a tonal language, and if so, whether or not there is a language related pitch deficit in these individuals. In the current study we investigated whether or not the deficits associated with amusia extend into the language domain for those who speak Mandarin as their first language. Although this does indicate there are amusics whose first language is a tonal language, we did not perform a formal investigation into the prevalence of amusia in first language Mandarin speakers.

In psychological models of pitch processing, contour processing is an initial step before actual pitch details are analyzed (Dowling, 1978; Foxton, Brown, et al., 2004; List, Justus, Robertson, & Bentin, 2007; Peretz, 1990). From this perspective, melodic contour and Mandarin intonation experiments were investigated in the current study. Both contour-violated and contour-preserved melodies were employed to assess the processing in pitch contour and pitch size intervals in the melodic contour task. All subjects were tested with discrimination and identification tasks of melodic contour and Mandarin intonation.

## 2. General method

### 2.1. Participants

All the participants were undergraduates or postgraduate students, who were recruited by means of advertisements in the bulletin board system of universities in Beijing. For the experimental participants, they self-reported that they had difficulty in carrying a tune when singing, and in detecting the difference between pitches. The controls were those who self-reported that they could sing in tune. Ethical approval was attained from the Institute of Psychology, Chinese Academy of Sciences, and informed consents were obtained from all of the participants.

The musical abilities of all the participants were tested by the Montreal Battery of Evaluation of Amusia (MBEA; Peretz, Champod, & Hyde, 2003). Out of 33 individuals who reported difficulties with music, 14 individuals exhibited the deficits of music perception that resulted in scores below the MBEA cut-off score of 78%.[1] This cut-off corresponds to 2 standard deviations below the available norms, although it should be noted the norms are based upon Western speakers of a non-tonal language. Of those 14, 11 were willing to take part in the further melody and speech tasks (see Table 1 for details).

Moreover, a detailed questionnaire was conducted to gather further information about the participants. None had received extra curriculum music training. None reported a history of neurological, psychiatric diseases, or difficulty in speech perception. All the individuals classified with amusia reported that at least one other member in their family also sang out of tune. With the exception of one of the amusic participants, all were right-handed as assessed by the Edinburgh Handedness Inventory (Oldfield, 1971). In addition, all had normal hearing at mean hearing level of 20 dB HL or less (19.5 and 20 dB HL for the right and the left ear, respectively) as measured by pure tone audiometry at 125, 250, 500, 1000, and 2000 Hz.

### 2.2. Data analysis

In keeping with previous studies (Ayotte et al., 2002; Bonnel et al., 2003), a hit was defined as a different pair responded to as different, and a false alarm was defined as an identical pair responded to as different. A well-known bias-free measure of sensitivity from signal-detection theory is $d'$ or $d_a$ and higher values for $d_a$ indicate higher sensitivity and, hence, better discrimination (Green & Swets, 1966; Lim, Bradshaw, Nicholls, & Altenmuller, 2003). The $d_a$ measure was employed in all sets of data analysis in the current study.

The jackknife method avoids the pitfalls normally associated with pooling ratings that come from participants with different sensitivities and response biases (Irwin, Hautus, & Stillman, 1992). To reduce statistical bias, the jackknife method was employed during statistical testing in the current study. The following formula was used to measure the $i$th pseudovalues (Dorfman & Berbaum, 1986):

$$\hat{\theta}_{*_i} = \hat{t}\theta_{\text{all}} - (t-1)\hat{\theta}_{[i]}.$$

In the above $t$ is the number of observers, $i = 1, 2, 3, \ldots, t\theta_{[i]}$ denotes the estimate of the signal-detection parameter of interest derived from the solution to $\nabla \ln L[i] = 0$ (Dorfman & Berbaum, 1986). The pseudovalues can be analyzed as if they are observed data with standard statistical tests (Arvesen & Salsburg, 1975).

## 3. Experiment 1: melodic contour perception

### 3.1. Stimuli and procedure

There were 48 pairs of five pure tone sequences presented randomly for the melodic contour discrimination task. The notes in the sequences ranged from C4 (262 Hz) to E5 (659 Hz). Half the pairs were identical with the other half being different pairs. On a different pair, the third tone was modified. Out of 24 different pairs, half of the changed tone sequences preserved the contour with only an interval change, and the other half involved tone changes that violated not only the interval but also the contour. Each tone lasted 350 ms, and there was 100 ms between tones. The average interval sizes preceding and following the changed tones was 2.0 and 3.17 semitones for contour-violated melodies, and 1.92 and 4.58 semitones for contour-preserved melodies. Although the mean change was slightly larger in the contour-preserved melodies than in the contour-violated, the difference was not significant ($F_{(1,22)} = 2.63$, $p > 0.05$). Feedback was provided during six practice trials but not during the experimental trials.

In the melodic contour identification task, 48 sequences of five pure tones were included with the same pitch range as the melodic contour discrimination task. Similarly, each tone lasted 350 ms with 100 ms inter-tone interval. There were two contour models. Each

[1] Based upon MBEA norms available on the internet via the link: http://www.brams.umontreal.ca/plab/publications/article/57#downloads the norms for a similar aged group (193 participants between 18 and 30 years old) would be 75% (22.5/30) based upon the information downloaded on April 22, 2010, New Zealand time. Using this criterion does not change the findings.

**Table 1**
Participants' characteristics and individual scores from the MBEA.

| Participant | Age | Gender | Education | Scale | Contour | Interval | Rhythm | Meter | Memory | Average |
|---|---|---|---|---|---|---|---|---|---|---|
| A1 | 30 | M | 20 | 21 | 18 | 16 | 19 | 20 | 25 | 20 |
| A2 | 30 | M | 20 | 20 | **25** | 18 | 24 | 21 | 19 | 21 |
| A3 | 23 | M | 16 | 22 | 19 | 18 | 19 | 23 | 18 | 20 |
| A4 | 24 | M | 17 | 22 | 19 | 18 | 16 | 25 | 27 | 21 |
| A5 | 26 | F | 19 | 12 | 16 | 15 | 15 | 19 | 19 | 16 |
| A6 | 25 | M | 18 | 16 | 19 | 17 | 26 | 15 | 15 | 18 |
| A7 | 23 | F | 15 | **23** | **23** | 21 | 18 | 21 | 18 | 21 |
| A8 | 21 | F | 15 | 14 | 16 | 15 | 15 | 17 | 17 | 16 |
| A9 | 18 | F | 12 | 19 | **22** | 23 | 25 | 24 | 21 | 22 |
| A10 | 26 | M | 18 | 18 | 17 | 21 | 27 | 23 | 21 | 21 |
| A11 | 22 | F | 14 | 17 | 16 | 17 | 17 | 21 | 15 | 17 |
| Mean | 24 | | 17 | 19 | 19 | 18 | 20 | 21 | 20 | 19 |
| SD | 3.6 | | 2.6 | 3.5 | 3.0 | 2.6 | 4.5 | 3.0 | 3.8 | 2.3 |
| C1 | 23 | M | 17 | 30 | 29 | 28 | 29 | 28 | 30 | 29 |
| C2 | 23 | M | 15 | 26 | **23** | 25 | 27 | 23 | 28 | 25 |
| C3 | 24 | F | 18 | 28 | 27 | 26 | 27 | 23 | 30 | 27 |
| C4 | 25 | M | 17 | **23** | 27 | 27 | 26 | 25 | 29 | 26 |
| C5 | 22 | M | 16 | 28 | 26 | 27 | 27 | 27 | 28 | 27 |
| C6 | 25 | F | 18 | 27 | **25** | 26 | 24 | 26 | 28 | 26 |
| C7 | 22 | F | 16 | 27 | 27 | 28 | 26 | 29 | 29 | 28 |
| C8 | 23 | M | 15 | 26 | 27 | 24 | 25 | 25 | 22 | 25 |
| C9 | 22 | F | 16 | 28 | **24** | 26 | 27 | 26 | 29 | 27 |
| C10 | 18 | F | 14 | 27 | **21** | 25 | 27 | 24 | 25 | 25 |
| C11 | 24 | F | 18 | 25 | 28 | 29 | 25 | 25 | 28 | 27 |
| Mean | 23 | | 16 | 27 | 26 | 26 | 26 | 26 | 28 | 26 |
| SD | 1.9 | | 1.4 | 1.8 | 2.4 | 1.5 | 1.4 | 1.9 | 2.4 | 1.2 |

*Note*: A: amusic; C: control; F: female; M: male.

model was depicted using a visual figure in the response box (see Table 2). These two models differed in the complexity of contour but had the same mean pitch distance between the lowest and the highest tone.

For contour model 1, half of the sequences started with an upward pitch movement with the other half starting with a downward pitch movement. For contour model 2, half of the sequences started with and ended with an upward pitch movement, whilst the other half started and ended with a downward pitch movement (see Table 2). Six sequences for practice trials were included and feedback was provided only during the practice trials.

In the discrimination tasks the participants were asked to respond by pressing one of six buttons. There were six response options: 1 = definitely same, 2 = probably same, 3 = possibly same, 4 = possibly different, 5 = probably different, and 6 = definitely different. As noted above, the musical contour identification task contained two different models. For identification of contour model 1 the keys were coded as follows, 1 = definitely “ ”, 2 = probably “ ”, 3 = possibly “ ”, 4 = possibly “ ”, 5 = probably “ ”, and 6 = “definitely “ ”. For identification of contour model 2, the keys were coded similarly, after substituting the appropriate contour symbols. In the melodic identification tasks, a hit was defined as “ ” and “ ” responded to as the “ ” and “ ”, and a false alarm was defined as “ ” and “ ” responded to as the“ ” and “ ” in the contour model 1 and 2 tasks, respectively.

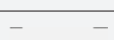
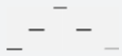
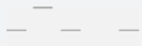
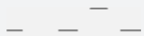
The stimuli were presented binaurally through Philips SHM1900 headphones in a sound-proofed room at an intensity level of 70 dB SPL-A. Previous research has suggested that contour-violated tasks are easier than contour-preserved tasks (Deruelle, Schon, Rondan, & Mancini, 2005; Peretz & Morais, 1987). Therefore, in this study, contour-violated tasks were presented first before the contour-preserved tasks.

### 3.2. Results

The ROCs of melodic contour discrimination and identification for the two groups were plotted in *z* coordinates (see Fig. 1). As can be seen, the ROCs for the amusic group are lower than those of the controls in both the melodic contour discrimination and identification tasks. The values of the sensitivity index $d_a$ are presented in Fig. 2. Moreover, the hits minus false alarms (%) for each group in the melodic tasks are plotted in Fig. 3 to ease comparison with previous studies (Ayotte et al., 2002; Gosselin et al., 2009; Peretz et al., 2002).

A mixed factor two-way ANOVA was performed on the $d_a$ scores, with group (amusics versus controls) as the between-subjects variable, and sequence condition (contour-violated and contour-preserved) was the within-subjects variable. There was a main effect of group ($F_{(1,20)} = 9.17$, $p < 0.05$), a main effect of sequence condition ($F_{(1,20)} = 28.77$, $p < 0.05$), with no significant interaction ($F_{(1,20)} = 0.77$, $p > 0.05$) during the contour discrimination task. Not surprisingly, an analysis of hits minus false alarm (%) also showed the same results with a significant main effect of group ($F_{(1,20)} = 11.70$, $p < 0.05$), a main effect of sequence condition ($F_{(1,20)} = 64.07$, $p < 0.05$), with no significant interaction ($F_{(1,20)} = 0.30$, $p > 0.05$). This indicated that the amusic group performed worse in both contour-violated and contour-preserved sequences relative to the controls. For both groups, contour-violated sequences yielded better discrimination sensitivity than contour-preserved sequences.

**Table 2**
Description of melodic contour identification task.

| Model | Stimulus | Corresponding figure |
|---|---|---|
| Contour Model 1 |  |  |
| |  |  |
| Contour Model 2 |  |  |
| |  |  |

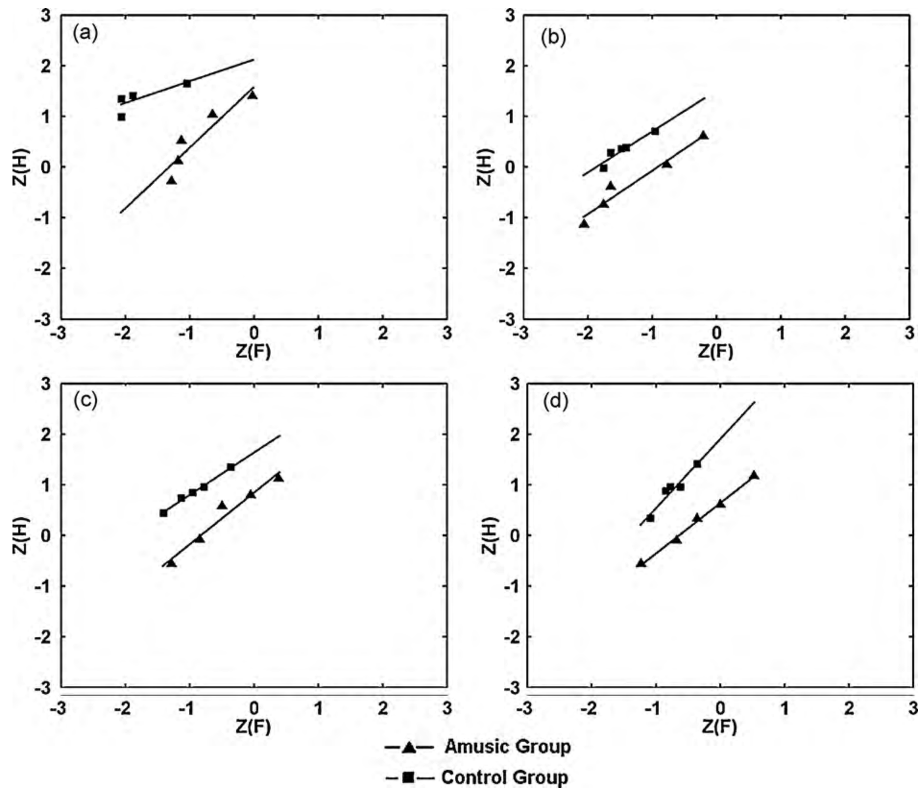



**Fig. 1.** ROC curves observed for each group in the melodic contour tasks: (a) contour-violated discrimination task; (b) contour-preserved discrimination task; (c) contour model 1 identification task; (d) contour model 2 identification task.

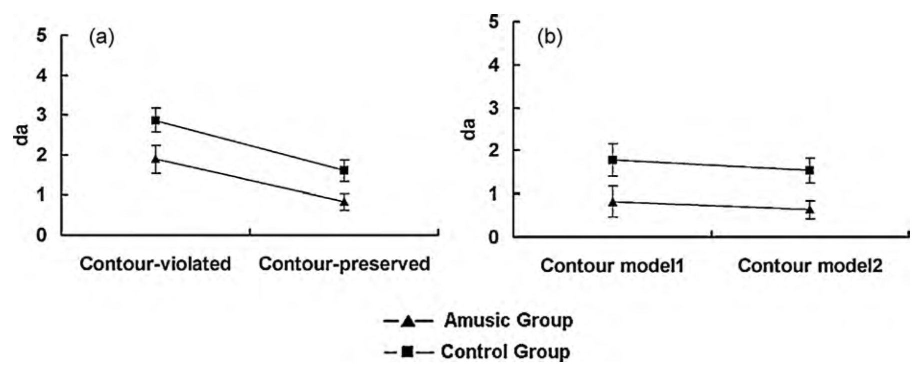



**Fig. 2.** Sensitivity index ($d_a$) for each group in the melodic contour tasks: (a) melodic contour discrimination task; (b) melodic contour identification task. Error bars indicate standard error.

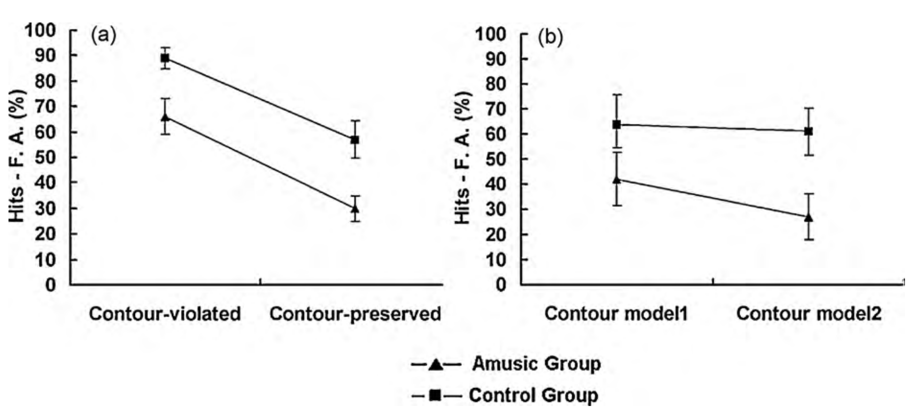



**Fig. 3.** Hits minus false alarm (%) for each group in the melodic tasks: (a) melodic contour discrimination task; (b) melodic contour identification task. Error bars indicate standard error.

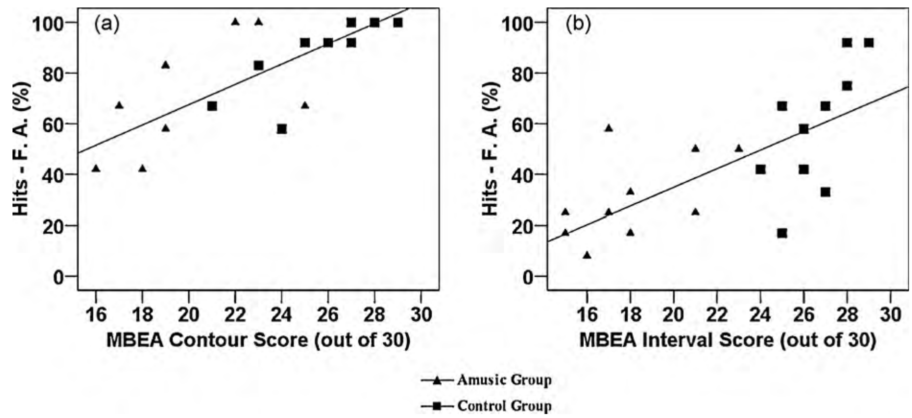



**Fig. 4.** Individual scores on melodic tasks as a function of scores on the MBEA (a) contour and (b) interval test (amusics = triangles; controls = squares; note some participants have equal scores and so the number of visible symbols is less than the number of participants).

To assess whether or not performance on the melodic contour discrimination tasks was related to the score obtained from the contour and interval subtests of the MBEA, correlation analyses between individual hits – false alarms (%) scores and the subtest score (maximum score of 30 as the catch trial is not scored) were performed. Performances on contour-violated sequences were correlated with the scores obtained in the contour subtest on the MBEA, when computed across both groups ($r_{(20)} = 0.79$, $p < 0.05$, see Fig. 4a) and was also significant for the amusic group ($r_{(9)} = 0.69$, $p < 0.05$). Furthermore, the performance on contour-preserved sequences was correlated with the MBEA interval subtest ($r_{(20)} = 0.71$, $p < 0.05$, see Fig. 4b) as computed over both groups. Although this failed to reach significance for the amusic group alone ($r_{(9)} = 0.52$, $p < 0.1$) under a two-tailed test, it would be significant as a one-tailed test as the relationship is in the expected direction.

For the identification task, a two-way ANOVA taking group as the between-subjects factors and contour model (model 1 versus model 2) as the within-subjects factors revealed significant main effect of the group ($F_{(1,20)} = 6.52$, $p < 0.05$). There was no significant main effect of the contour model ($F_{(1,20)} = 0.96$, $p > 0.05$) nor was there a group by contour model interaction ($F_{(1,20)} = 0.03$, $p > 0.05$). These results indicated that the amusic group had difficulty in identifying both model 1 and model 2 relative to the controls. Furthermore, analysis of percentage correct also showed a significant main effect of group ($F_{(1,20)} = 4.96$, $p < 0.05$). There was no significant main effect of contour model, nor was there a group by contour model interaction ($F_{(1,20)} = 1.89$, $p > 0.05$, $F_{(1,20)} = 0.72$, $p > 0.05$, respectively). These findings also suggested that the amusic group showed lower scores in identifying both model 1 and model 2 relative to the controls (see Fig. 5).

## 4. Experiment 2: Mandarin intonation perception

### 4.1. Stimuli and procedure

As noted above, in tonal languages the semantic comprehension depends on the recognition of lexical tone to a great degree. However, the boundary of the tone plays a critical role for distinguishing between a question and a statement in Mandarin Chinese (Lin, 2004, 2006). In this light, the speech materials in the current study contained two syllables, which were verb–object construction with the subject of the sentence omitted. All speech materials were spoken by a male native speaker of Mandarin Chinese and were selected from familiar everyday expressions. For example,

bei$^4$ke$^4$./? (prepare lessons./?)
kan$^4$shu$^1$./? (read book./?)
shua$^1$ya$^2$./? (brush teeth./?)

There were 64 pairs of two-syllable verb–object construction in the Mandarin intonation discrimination task. Half of the pairs were identical (both statements or both questions), whilst the other half contained different pairs (statement/question or question/statement). Each member of a pair lasted 850 ms and the interval between two members was 1500 ms. In comparison with those in (Patel et al., 2008), the speech materials in the current study were spoken with slower rates. The stimuli were presented randomly. All pairs employed the same verb–object words, with only the fundamental frequency (F0) curve of the second syllable differing on different trials. For example, bei$^4$ke$^4$ was used in the intonation discrimination task. The mean F0 of bei$^4$ke$^4$ was 228.6 and 185.8 Hz for each syllable, respectively in the statement, and 228.6 and 267.5 Hz in the question. Although the mean F0s are slightly different between the two tasks (see Table 3), no significant differences were observed ($t_{(31)} = -1.324$, $-0.598$, and $-0.617$ for the first syllable, final syllable in statement, and final syllable in question, respectively, all $p > 0.05$).

The speech materials were edited using a cross-splicing technique to ensure that the first syllables were acoustically identical and the timings of the second syllables were roughly equal, to ensure the discrimination was based upon the F0 curve of the second syllable. The differences between the acoustic waveform amplitudes for the second syllable were within ±5 dB in each pair. There were four different Mandarin tones employed in the second syllables.

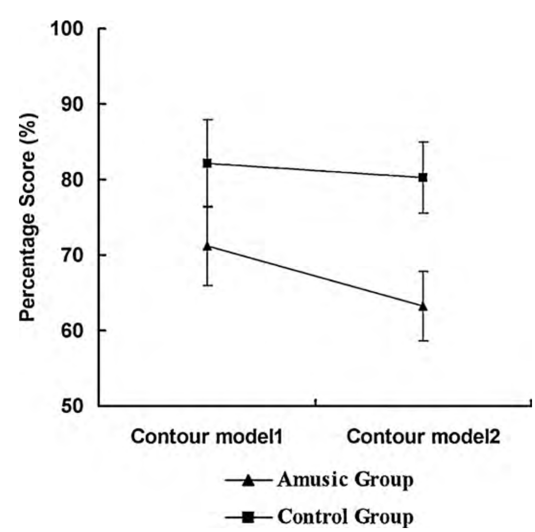



**Fig. 5.** Percentage correct for each group in the melodic identification task. Error bars indicate standard error.

**Table 3**
The mean pitch values for each trial used in the statement and question discrimination and identification tasks.

| | Discrimination stimuli | | Identification stimuli | |
|---|---|---|---|---|
| | Statement | Question | Statement | Question |
| Mean FO of first syllable | 90(5.3) | 90(5.3) | 91.7(6.0) | 91.7(6.0) |
| Mean FO of final syllable | 86.9(6.0) | 93.9(3.3) | 88.2(3.9) | 94.4(3.0) |
| Size of final pitch glide | −0.9(7.4) | 8.4(12.6) | −2.8(9.9) | 9.0(14.0) |
| Rate of final pitch glide (st/s) | −4.1(36.0) | 31.5(59.6) | −12.0(52.6) | 38.9(70.7) |

**Table 4**
The mean size and rate of the final pitch glide for each tone used in the statement and question identification task.

| | Size of final pitch glide | | Rate of final pitch glide (st/s) | |
|---|---|---|---|---|
| | Statement | Question | Statement | Question |
| Tone 1 | −1.5 (0.6) | 4.1 (1.9) | −8.9 (3.2) | 26.8 (14.6) |
| Tone 2 | 11.7 (1.5) | 20.6 (2.5) | 69.2 (6.9) | 82.1 (8.6) |
| Tone 3 | −12.5 (5.8) | 22.3 (1.7) | −44.9 (16.5) | 113.4 (11.4) |
| Tone 4 | −9.0 (2.7) | −10.9 (3.9) | −63.4 (10.7) | −66.1 (22.6) |

Non-linguistic analogues were created by extracting the F0 from the speech stimuli using Praat software (Boersma, 2001). In so doing the acoustic parameters were preserved whilst no linguistic information remained. We individually generated the non-linguistic version of all three syllables, the first syllable and the second syllables for the question and the statement from each pair. Some spoken files were manually edited to remove artifacts prior to their conversion to the non-linguistic analogue. These were sampled at 44,100 Hz. To get roughly equal intensity, the non-linguistic analogues were adjusted to have the same average intensity as the original speech sound. During each section of the experiment, the participants were asked to adjust the loudness of headphones to a comfortable level. The discrete F0 pattern for each syllable in each pair was then combined together and edited to ensure the same duration for each member of a pair and for the interval between two members. There was an equal number of speech and non-linguistic analogue stimuli used in the discrimination task.

The Mandarin intonation identification task was constructed by creating 32 pairs of two syllables of verb–object construction. Unlike during the discrimination task only one member the pair was presented on a given trial thus forming 64 experimental trials, each lasting 850 ms. Similarly, by using cross-splicing, the first syllables were acoustically identical and the timings of the final syllables were edited to be roughly equal. There was also the same number of speech and non-linguistic analogue stimuli used in the identification task.

The average pitch values for each speech trial used in the statement and question of both discrimination and identification are presented in Table 3. Generally speaking, an upward pitch glide corresponds to a question whilst a downward pitch glide corresponds to a statement. However, this correspondence is incomplete due to Mandarin lexical tones. The pitch glides of all lexical tones go up for questions except for Tone 4, which goes down. Similarly, the pitch glides of all lexical tones go down for a statement except for Tone 2, where it goes up (see Table 4, identification speech stimuli).[2] These pitch values for the four tones suggest that each lexical tone still remains in its original form even when it is embedded in a question and a statement for Mandarin tones (Lin, 2004, 2006). As shown in Table 4, although the average magnitude and rate of the final glide were slightly larger in the identification than in the discrimination tasks, the difference was not significant ($F_{(1,62)} = 0.54$, $p > 0.05$ for statement, and $F_{(1,62)} = 0.17$, $p > 0.05$ for question). Furthermore, in comparison with those of Patel et al. (2008), the speech materials in the current study were uttered with smaller final glide sizes and rates.

In the intonation discrimination tasks, the participants were required to respond by pressing one of six buttons. There were the same options on the response box as in the melodic contour discrimination tasks. However, in the tasks of intonation identification, the six options were coded as follows, 1 = definitely statement, 2 = probably statement, 3 = possibly statement, 4 = possibly question, 5 = probably question, and 6 = definitely question. For the analyses, a statement response assigned to the statement was defined as a hit and as a false alarm when assigned to a question in the intonation identification task. The stimuli were presented binaurally through Philips SHM1900 headphones in a sound-proofed room as in Experiment 1. The participants could adjust the volume to suit their listening preference.

Previous research has suggested that amusics perform better in speech tasks than in non-speech tasks (Ayotte et al., 2002; Patel et al., 2005). Therefore, in this study, the speech stimuli were presented before the non-linguistic analogue.

### 4.2. Results

The ROCs for each group in the Mandarin intonation tasks are presented in Fig. 6. Generally speaking, it shows that the ROCs for the amusic group are lower relative to the controls in all four tasks.

For the intonation discrimination task the sensitivity indices ($d_a$) were computed. These can be seen in Fig. 7a. A mixed factor two-way ANOVA taking stimuli type (speech versus non-linguistic analogue) as the within-subjects factor and group as the between-subjects factor showed that the main effect of group reached a marginally significant level ($F_{(1,20)} = 3.12$, $p = 0.09$, which would be significant under a one-way directional test since it is in the expected direction). There was neither a significant main effect of stimuli type nor was there a group by stimuli type interaction ($F_{(1,20)} = 0.39$, $p > 0.05$, $F_{(1,20)} = 1.92$, $p > 0.05$, respectively) for the intonation discrimination task. This indicated that there was a trend for the amusic participants to score worse than the control participants in the intonation discrimination tasks for both speech and non-linguistic analogues.

The $d_a$ values were computed for the intonation identification task and they can be seen in Fig. 7b. A mixed factor two-way ANOVA was conducted and showed a significant main effect of group ($F_{(1,20)} = 7.03$, $p < 0.05$) and a main effect of stimuli type ($F_{(1,20)} = 128.14$, $p < 0.001$). The group by stimuli type interaction was not significant ($F_{(1,20)} = 0.09$, $p > 0.05$). The results indicated that the amusic group performed worse in the intonation identification tasks for both speech and non-linguistic analogues relative to the controls, and both groups scored lower for non-linguistic analogues than for speech stimuli.

[2] To examine whether or not Tone 2 and Tone 4 were more difficult to discriminate than the other tones, a one-way ANOVA of hits minus false alarm (%) was conducted in the discrimination tasks, and showed no significant difference, either when computed across both groups, or computed only for amusic group (all $p > 0.05$). For the identification tasks, percentage correct values were submitted to a one-way ANOVA and again no significant differences were observed with respect to the tones, either when computed across both groups, or computed only for amusic group (all $p > 0.05$).

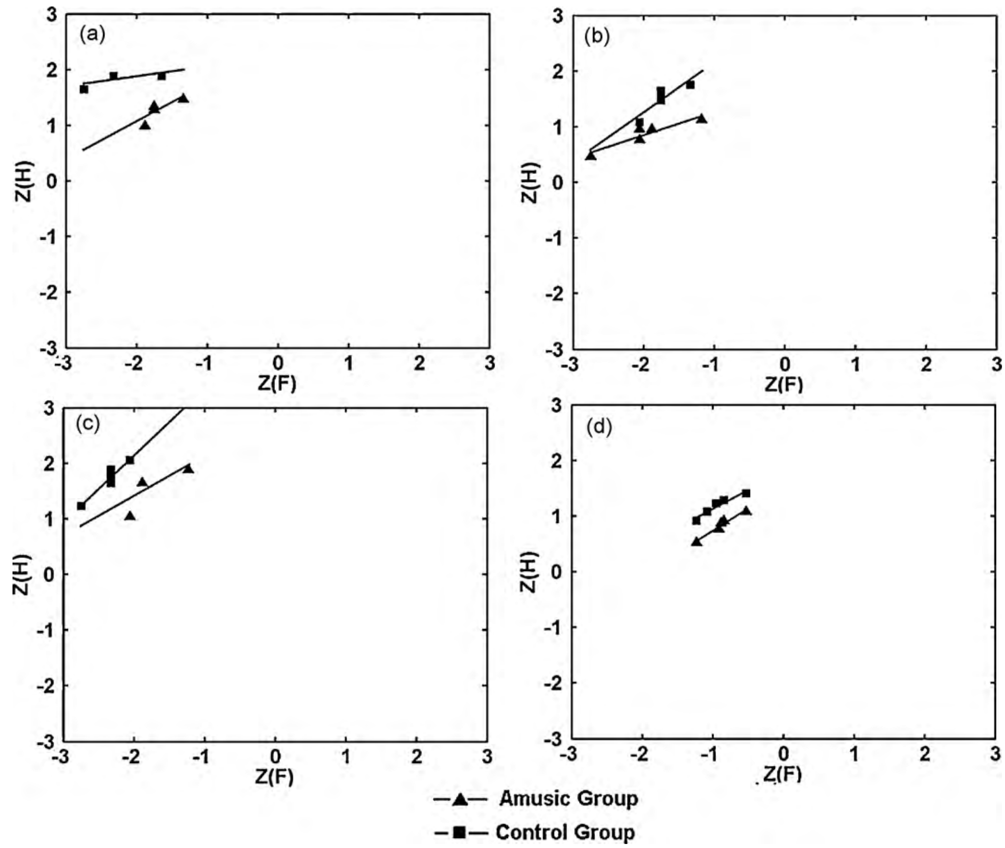



**Fig. 6.** ROC curves observed for each group in the Mandarin intonation tasks: (a) intonation discrimination for speech pattern; (b) intonation discrimination for non-linguistic analogue; (c) intonation identification for speech pattern; (d) intonation identification for non-linguistic analogue.

A previous study has shown about 30% of those with amusia have substantially lower performance on speech versus non-linguistic analogue discriminations (Patel et al., 2008). In the current study, there were also three participants with amusia (~27%) who obtained 11% lower performance on the speech stimuli than on the non-linguistic analogues (data in bold in Table 5). However, individual analysis did not show this to be significantly different for any of the participants, and when the responses of these three participants were pooled, no significant difference was found between the proportions of correct and incorrect responses between tasks ($\chi^2 = 2.312$, $n = 384$, $p > 0.05$). Interestingly, there was the same number of control participants who showed the same pattern of a higher or equal score for the non-linguistic analogues compared to the speech stimuli. The mean difference for the controls was 4% in terms of percent hits minus percent false alarms (data in bold in Table 5). In addition, only one amusic participant demonstrated better performance in the identification task with non-linguistic analogues compared to speech stimuli.

In order to investigate whether or not performances on Experiment 1 and Experiment 2 were related, correlations were calculated between individual hits − false alarm (%) scores. There were significant correlations between performances on the two experiments under the following conditions when the two groups were combined (for the discrimination tasks, $r_{(20)} = 0.60$, $p < 0.05$, $r_{(20)} = 0.44$, $p < 0.05$, and $r_{(20)} = 0.58$, $p < 0.05$ for speech versus non-linguistic analogues, melody versus speech pattern, and melody versus non-linguistic analogues, respectively; for the identification tasks, $r_{(20)} = 0.66$, $p < 0.05$, $r_{(20)} = 0.43$, $p < 0.05$, and $r_{(20)} = 0.48$, $p < 0.05$ for speech versus non-linguistic analogues, melody versus speech pattern, and melody versus non-linguistic analogues, respectively). When computed within the amusic group alone, there were marginally significant or significant correlations between speech

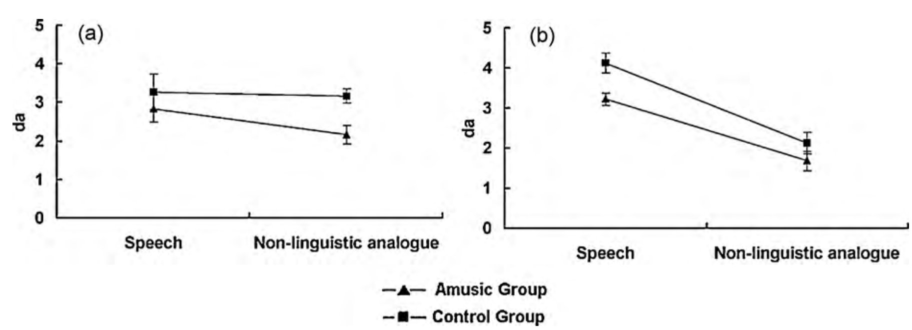



**Fig. 7.** Sensitivity index ($d_a$) for each group in the Mandarin intonation tasks: (a) intonation discrimination task; (b) intonation identification task. Error bars indicate standard error.

**Table 5**
Hits minus false alarm (%) for all participants in the Mandarin intonation tasks.

| Participant | Discrimination (speech) | Discrimination (non-linguistic) | Identification (speech) | Identification (non-linguistic) |
|---|---|---|---|---|
| A1 | 75 | 59.4 | 93.8 | 59.4 |
| A2 | 100 | 90.6 | **87.5** | **93.8** |
| A3 | 75 | **78.1** | 87.5 | 34.4 |
| A4 | 100 | 90.6 | 96.9 | 75 |
| A5 | 90.6 | 68.8 | 90.6 | 53.1 |
| A6 | 84.4 | 78.1 | 93.8 | 78.1 |
| A7 | **84.4** | **93.8** | 96.9 | 34.3 |
| AS | **65.4** | **81.3** | 93.8 | 71.9 |
| A9 | 100 | 90.6 | 93.8 | 78.1 |
| A10 | 81.3 | 59.4 | 81.3 | 34.3 |
| A11 | 100 | 93.8 | 93.8 | 71.9 |
| Mean | 86.9 | 80.4 | 91.8 | 62.2 |
| SD | 12.2 | 13.0 | 4.7 | 20.7 |
| C1 | 100 | 90.6 | 96.9 | 87.5 |
| C2 | **81.3** | **87.5** | 93.8 | 84.4 |
| C3 | 93.8 | 81.3 | 87.5 | 21.9 |
| C4 | 96.9 | 90.6 | 100 | 59.4 |
| C5 | 96.9 | 81.3 | 96.9 | 65.6 |
| C6 | 100 | 90.6 | 96.9 | 84.3 |
| C7 | **93.8** | **100** | 100 | 81.3 |
| C8 | **96.9** | **96.9** | 96.9 | 87.5 |
| C9 | 100 | 90.6 | 100 | 71.9 |
| C10 | 94.8 | 87.5 | 90.6 | 59.4 |
| C11 | 100 | 93.8 | 93.8 | 81.3 |
| Mean | 95.9 | 90.0 | 95.8 | 71.3 |
| SD | 5.4 | 5.7 | 4.0 | 19.6 |

*Note*: A: amusic; C: control; F: female; M: male.

and non-linguistic analogues in both the discrimination and identification tasks ($r_{(9)} = 0.56$, $p < 0.08$ for discrimination tasks, which would be significant under a one-tailed test, and $r_{(9)} = 0.67$, $p < 0.05$ for identification task). However, none of relationships under the other two conditions (melody versus speech pattern, and melody versus non-linguistic analogues) were significant for the amusic group in either the discrimination or identification tasks ($r_{(9)} = 0.19$, $r_{(9)} = 0.43$ for melody versus speech pattern, and melody versus non-linguistic analogues, respectively in the discrimination task, and $r_{(9)} = 0.29$, $r_{(9)} = -0.02$ for melody versus speech pattern, and melody versus non-linguistic analogue, respectively in the identification task, all $p > 0.05$).

## 5. General discussion

The current study assessed pitch sensitivity of 11 adults with amusia whose first language is Mandarin in the melodic contour and Mandarin intonation tasks. To the best of our knowledge this is the first study to examine amusic individuals whose first language is a tonal language. Moreover, the relative ease with which these participants were located suggests that tonal languages may not provide much, if any, compensation for amusia. However, this suggestion must be viewed as tentative as the current sample size is insufficient to draw firm conclusions on this point and a larger, more formal investigation into this question is required.

To assess music perception abilities with respect to tone and language based stimuli the rating method of signal detection was employed in the current study. This method avoids response bias and other problems associated with a two alternative forced choice method (Irwin et al., 1992). To the best of our knowledge, this is the first study to employ this method to investigate pitch performance of amusic individuals. The current study showed the amusics with Mandarin as their first language was impaired in discriminating and identifying melodic contours. Moreover, these abnormalities were also detected in identifying both speech and non-linguistic analogues for the Mandarin intonation tasks, with a similar trend found in discrimination tasks.

For the melodic contour tasks, the amusics with Mandarin as their first language had difficulty discriminating both contour-violated and contour-preserved sequences. This finding supports the previous studies of amusics with non-tonal language who have deficits of musical contour discrimination (Ayotte et al., 2002; Foxton, Dean, et al., 2004; Peretz et al., 2002). Infants aged 8–11 months can discriminate between non-violated and violated melody contours (Trehub, Bull, & Thorpe, 1984). From this perspective, it might be argued that the neural substrates involved in pitch processing may be compromised in amusics from birth. In the melodic contour identification task, the deficits were also detected in identifying melodic contour even though the corresponding visual figures were provided. This may be due to the abnormality of processing high or low pitch in amusical individuals (Janata, 2007) and of determining pitch direction (Foxton, Dean, et al., 2004). Furthermore, the amusics' discrimination performance on contour-violated and contour-preserved sequences was correlated with their score on the contour and interval subtests of the MBEA, respectively (Fig. 4a and b). This result not only supports the validity of the MBEA, but also implies that the criterion of MBEA from the Canadian sample appears to be applicable to a Chinese population.

As noted earlier, there are contradictory results regarding whether or not pitch deficits in individuals with amusia are music-specific. The current findings are in line with those of Ayotte et al. (2002) who demonstrated that amusics with non-tonal languages have deficits in discriminating non-speech intonation patterns. However, the current findings extend this to suggest that amusics have abnormal discrimination of actual speech intonations. As far as speech stimuli are concerned, speech sounds may contain extra cues aiding perception, such as intensity and rhythm (Griffiths, 2008), compared to non-speech sounds. However, it has been found that about 30% of amusics scored worse on speech than on non-linguistic analogues (Patel et al., 2008), which suggests that pitch direction perception deficits in amusia can extend to speech. The current study employed speech stimuli which had smaller final glide sizes and slower spoken rates than in Patel et al. (2008) and provides further evidence of a trend for the amusic participants to score lower relative to controls during discrimination of Mandarin intonation for both non-linguistic analogues and speech stimuli. These abnormalities may be attributed to the deficits in melodic contour perception of the amusic participants. In the intonation identification tasks, these deficits were detected for both speech and non-linguistic analogues, which differs from the data of Peretz et al. (2002) and Ayotte et al. (2002), who found that most amusics could judge a statement and a question based on the final pitch change.

Overall, although the amusic participants do not report any speech impairments, our findings suggest that the melodic pitch impairments of amusic participants may extend to subtle speech perception deficits as well. As noted earlier, the correlations between the performances on speech and non-linguistic analogues for the amusic group imply that there may be a common pitch mechanism underlying performance in speech and non-speech intonation processing. However, further neuropsychological studies on pitch processing in music and language for amusics are needed to clarify this issue.

It seems that tonal language experience provides little compensation for the melodic deficits of amusics whose first language is a tonal language. However, whether or not language experience influences the incidence of amusia requires a formal large scale cross-language study. Furthermore, based on the finding that all the amusic participants in the current study reported at least one other member in their family also sang out of tune, further genetic research may be worth pursuing as well.

As a final note, to examine whether or not amusia exists among speakers whose first language is Mandarin, the MBEA was employed and the cut-off score of 2 SD below the norms of Western participants (Peretz et al., 2003) was used in the current study. Although this criterion is based on the performance of a large pool of Canadian participants and the test material is based upon Western musical conventions, it appears a similar cut-off may be utilized in a Chinese population. As can be seen from Table 1, among the three key subtests including scale, contour, and interval, use of this cut-off criterion for the MBEA produces either no overlap between the amusic and control participants (interval) or there is little overlap (scale and contour; data in bold in Table 1). For the amusic participants the means of the three subtests in the current study were below the cut-off scores defined by 2 SD from the mean of a similar cohort (19 versus 21, 19 versus 21, and 18 versus 20, for scale, contour, and interval, respectively). For the controls, the means and standard deviations of the three subtests in the current study are similar to or within those of the MBEA norms (Peretz et al., 2003). On the other hand, to test if the overlap data influenced the melodic and intonation results in our study, we dropped all the data of the overlap participants, and found that the interpretation of the data remains the same, suggesting that the small number of participants who overlap on a few of the subscales, did not appear to have a great influence in the overall results of our study. In the long term, however, although the criterion from the Canadian sample appears to provide good separation between the groups in our study, a larger scale study will be required to determine appropriate norms for Chinese population due to cultural differences in music exposure and styles.

## 6. Conclusion

The current behavioral studies examine the processing of melodic contour and speech intonation in amusics who speak Mandarin as their first language. The findings suggest that the participants with amusia have difficulty processing melodic contour, and this deficit extends to speech processing. This could potentially result in other experience-related impairments in language in those who speak a tonal language. Future neuropsychological and genetic studies on amusia may shed light on the nature of the deficit underlying this disorder and provide a new way to explore the processing of music and language.

## Acknowledgments

We thank Dr. Michael Hautus (University of Auckland) for the computer program to calculate the sensitivity index and for his valuable comments on the data analysis. We also thank the editor, Prof. Sabine Kastner, and Prof. Isabelle Peretz, and one anonymous reviewer for their valuable comments and suggestions on a previous version of the manuscript. For correspondence concerning this article please contact either Dr. Yufang Yang (yangyf@psych.ac.cn) or Dr. Jeff P. Hamm (j.hamm@auckland.ac.nz).

## References

Arvesen, J. N., & Salsburg, D. S. (1975). Approximate tests and confidence intervals using the jackknife. In R. M. Elashoff (Ed.), *Perspectives in biometrics*. New York: Academic Press.

Ayotte, J., Peretz, I., & Hyde, K. (2002). Congenital amusia: A group study of adults afflicted with a music-specific disorder. *Brain*, *125*, 238–251.

Bent, T., Bradlow, A. R., & Wright, B. A. (2006). The influence of linguistic experience on the cognitive processing of pitch in speech and nonspeech sounds. *Journal of Experimental Psychology: Human Perception and Performance*, *32*, 97–103.

Boersma, P. (2001). Praat, a system for doing phonetics by computer. *Glot International*, *5*(9/10), 341–345.

Bonnel, A., Mottron, L., Peretz, I., Trudel, M., Gallun, E., & Bonnel, A. M. (2003). Enhanced pitch sensitivity in individuals with autism: A signal detection analysis. *Journal of Cognitive Neuroscience*, *15*, 226–235.

Chandrasekaran, B., Krishnan, A., & Gandour, J. T. (2007a). Mismatch negativity to pitch contours is influenced by language experience. *Brain Research*, *1128*, 148–156.

Chandrasekaran, B., Krishnan, A., & Gandour, J. T. (2007b). Experience-dependent neural plasticity is sensitive to shape of pitch contours. *Neuroreport*, *18*, 1963–1967.

Chandrasekaran, B., Krishnan, A., & Gandour, J. T. (2009). Relative influence of musical and linguistic experience on early cortical processing of pitch contours. *Brain and Language*, *108*, 1–9.

Dediu, D., & Ladd, D. R. (2007). Linguistic tone is related to the population frequency of the adaptive haplogroups of two brain size genes, ASPM and Microcephalin. *Proceedings of the National Academy of Sciences USA*, *104*, 10944–10949.

Deruelle, C., Schon, D., Rondan, C., & Mancini, J. (2005). Global and local music perception in children with Williams syndrome. *Neuroreport*, *16*, 631–634.

Dorfman, D. D., & Berbaum, K. S. (1986). RSCORE-J: Pooled rating-method data: A computer program for analyzing pooled ROC curves. *Behavior Research Methods, Instruments, & Computers*, *18*, 452–462.

Dowling, W. J. (1978). Scale and contour: Two components of a theory of memory for melodies. *Psychological Review*, *85*, 341–354.

Foxton, J. M., Brown, A. C., Chambers, S., & Griffiths, T. D. (2004). Training improves acoustic pattern perception. *Current Biology*, *14*, 322–325.

Foxton, J. M., Dean, J. L., Gee, R., Peretz, I., & Griffiths, T. D. (2004). Characterization of deficits in pitch perception underlying 'tone deafness'. *Brain*, *127*, 801–810.

Gosselin, N., Jolicoeur, P., & Peretz, I. (2009). Impaired memory for pitch in congenital amusia. *Annals of the New York Academy of Science*, *1169*, 270–272.

Green, D. M., & Swets, J. A. (1966). *Signal Detection Theory and Psychophysics*. New York: Wiley.

Griffiths, T. D. (2008). Tone deafness: A model complex cortical phenotype. *Clinical Medicine*, *8*, 592–595.

Hyde, K. L., & Peretz, I. (2004). Brains that are out of tune but in time. *Psychological Science*, *15*, 356–360.

Irwin, R. J., Hautus, M. J., & Stillman, J. A. (1992). Use of the receiver operating characteristic in the study of taste perception. *Journal of Sensory Studies*, *7*, 291–314.

Janata, P. (2007). The highs and lows of being tone deaf. *Nature Neuroscience*, *10*, 810–812.

Krishnan, A., Xu, Y., Gandour, J., & Cariani, P. (2005). Encoding of pitch in the human brainstem is sensitive to language experience. *Brain Research Cognitive Brain Research*, *25*, 161–168.

Ladd, D. R., Dediu, D., & Kinsella, A. R. (2008). Languages and genes: Reflections on biolinguistics and the nature-nurture question. *Biolinguistics*, *2*, 114–126.

Lim, V. K., Bradshaw, J. L., Nicholls, M. E., & Altenmuller, E. (2003). Perceptual differences in sequential stimuli across patients with musician's and writer's cramp. *Movement Disorders*, *18*, 1286–1293.

Lin, M. C. (2004). Chinese intonation and tone. *Applied Linguistics*, *3*, 57–67.

Lin, M. C. (2006). Interrogative mood and boundary tone in Chinese. *Chinese Language*, *4*, 364–379.

List, A., Justus, T., Robertson, L. C., & Bentin, S. (2007). A mismatch negativity study of local-global auditory processing. *Brain Research*, *1153*, 122–133.

Nguyen, S., Tillmann, B., Gosselin, N., & Peretz, I. (2009). Tonal language processing in congenital amusia. *Annals of the New York Academy of Science*, *1169*, 490–493.

Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*, *9*, 97–113.

Patel, A. D. (2006). An empirical method for comparing pitch patterns in spoken and musical melodies: A comment on J.G.S. Pearl's "Eavesdropping with a master: Leos Janáček and the music of speech". *Empirical Musicology Review*, *1*, 166–169.

Patel, A. D., Foxton, J. M., & Griffiths, T. D. (2005). Musically tone-deaf individuals have difficulty discriminating intonation contours extracted from speech. *Brain and Cognition*, *59*, 310–313.

Patel, A. D., & Peretz, I. (1997). Is music autonomous from language? A neuropsychological appraisal. In I. Deliege, & J. A. Sloboda (Eds.), *Perception and cognition of music*. Philadelphia: Psychology Press.

Patel, A. D., Wong, M., Foxton, J. M., Lochy, A., & Peretz, I. (2008). Speech intonation perception deficits in musical tone deafness. *Music Perception*, *25*, 357–368.

Peretz, I. (1990). Processing of local and global musical information by unilateral brain-damaged patients. *Brain*, *113*, 1185–1205.

Peretz, I. (2002). Brain specialization for music. *Neuroscientist*, *8*, 372–380.

Peretz, I. (2008). Musical disorders: From behavior to genes. *Current Directions in Psychological Science*, *17*, 329–333.

Peretz, I., Ayotte, J., Zatorre, R. J., Mehler, J., Ahad, P., Penhune, V. B., et al. (2002). Congenital amusia: A disorder of fine-grained pitch discrimination. *Neuron*, *33*, 185–191.

Peretz, I., Champod, A. S., & Hyde, K. (2003). Varieties of musical disorders. The Montreal Battery of Evaluation of Amusia. *Annals of the New York Academy of Science*, *999*, 58–75.

Peretz, I., & Hyde, K. L. (2003). What is specific to music processing? Insights from congenital amusia. *Trends in Cognitive Science*, *7*, 362–367.

Peretz, I., & Morais, J. (1987). Analytic processing in the classification of melodies as same or different. *Neuropsychologia*, *25*, 645–652.

Thompson, W. F. (2007). *Exploring variants of amusia: Tone deafness, rhythm impairment, and intonation insensitivity*. Paper presented at the International Conference on Music Communication Science.

Tillmann, B., Schulze, K., & Foxton, J. M. (2009). Congenital amusia: A short-term memory deficit for non-verbal, but not verbal sounds. *Brain and Cognition*, *71*, 259–264.

Trehub, S. E., Bull, D., & Thorpe, L. A. (1984). Infants' perception of melodies: The role of melodic contour. *Child Development*, *55*, 821–830.

Xu, Y., Gandour, J. T., & Francis, A. L. (2006). Effects of language experience and stimulus complexity on the categorical perception of pitch direction. *Journal of the Acoustical Society of America*, *120*, 1063–1074.