Phonological processing in Mandarin speakers with congenital amusia

Xiao Wang and Gang Peng[a)]
CUHK-PKU-UST Joint Research Centre for Language and Human Complexity, and Department of Linguistics and Modern Languages, The Chinese University of Hong Kong, Hong Kong SAR, China

(Received 6 January 2014; revised 14 July 2014; accepted 7 October 2014)

Although there is an emerging consensus that both musical and linguistic pitch processing can be problematic for individuals with a developmental disorder termed congenital amusia, the nature of such a pitch-processing deficit, especially that demonstrated in a speech setting, remains unclear. Therefore, this study tested the performance of native Mandarin speakers, both with and without amusia, on discrimination and imitation tasks for Cantonese level tones, aiming to shed light on this issue. Results suggest that the impact of the phonological deficit, coupled with that of the domain-general pitch deficit, could provide a more comprehensive interpretation of Mandarin amusics' speech impairment. Specifically, when there was a high demand for pitch sensitivity, as in fine-grained pitch discriminations, the operation of the pitch-processing deficit played the more predominant role in modulating amusics' speech performance. But when the demand was low, as in discriminating naturally produced Cantonese level tones, the impact of the phonological deficit was more pronounced compared to that of the pitch-processing deficit. However, despite their perceptual deficits, Mandarin amusics' imitation abilities were comparable to controls'. Such selective impairment in tonal perception suggests that the phonological deficit more severely implicates amusics' input pathways.
[http://dx.doi.org/10.1121/1.4900559]

PACS number(s): 43.75.Cd [MAH] Pages: 3360–3370

I. INTRODUCTION

Music is not simply a cultural product; like language, it is "a universal human trait that transcends time, place and culture" (Brown, 1991; Peretz, 2006). But due to a receptive musical disorder termed "congenital amusia" (amusia hereafter; Ayotte *et al.*, 2002; Jiang *et al.*, 2010), 3%–5% of the population cannot acquire basic musical skills (Kalmus and Fry, 1980; Nan *et al.*, 2010; Wong *et al.*, 2012), exhibiting lifelong difficulties with the acoustic and the structural features of music (Peretz, 2006). Accumulating evidence has shown that neither normal exposure to music nor the proficient use of lexical tones is sufficient to restore this congenital receptive disorder in music processing (e.g., Ayotte *et al.*, 2002; Nan *et al.*, 2010; Mignault-Goulet *et al.*, 2012; Wong *et al.*, 2012). Nevertheless, difficulties with music are not the only characteristics of individuals with amusia (amusics). Studies have shown that the perception and production of pitch changes in the linguistic domain may likewise be problematic for amusics, although to a less pronounced extent (e.g., Patel *et al.*, 2008; Liu *et al.*, 2010; Tillmann *et al.*, 2011a; Liu *et al.*, 2012; Thompson *et al.*, 2012; Liu *et al.*, 2013). However, the precise nature of amusia, or rather, of the deficits that underlie amusia, is not well understood.

While most previous studies suggest that a pitch-processing disorder could account for amusics' impaired speech perceiving and producing performance, deficient phonological processing reported by Jones *et al.* (2009a) provides another angle to interpret these speech-processing difficulties exhibited by amusics. But evidence supporting the latter case is scarce. In particular, it remains unclear whether a phonological deficit similarly affects amusics who speak tonal languages. To bridge the gap, this study administered Cantonese level tone discrimination and imitation tasks to Mandarin amusics and controls, with the aim of delineating the nature of the deficit(s) responsible for amusics' impaired pitch-processing abilities in the linguistic domain.

A. An alternative explanation for deficient speech processing

In light of the positive transfer effects between music and language, researchers have proposed that the music-processing deficit of amusia would affect speech perception (Patel *et al.*, 2008; Nguyen *et al.*, 2009; Tillmann *et al.*, 2011a; Tillmann *et al.*, 2011b; Liu *et al.*, 2012). Accordingly, a domain-general point of view has been used to explain amusics' pitch perception deficits in both music and language. Indeed, the abilities of amusics to perceive pitch and to control pitch production are poorer than their non-amusics counterparts, especially when the pitch variation is comparatively small, e.g., intonation contours (Patel *et al.*, 2008; Liu *et al.*, 2010; Hutchins and Peretz, 2012; Liu *et al.*, 2012), lexical tones (Nguyen *et al.*, 2009; Nan *et al.*, 2010; Tillmann *et al.*, 2011a; Liu *et al.*, 2012; Liu *et al.*, 2013), as well as the prosodic cues associated with specific emotional states (Thompson *et al.*, 2012).

[a)]Author to whom correspondence should be addressed. Also at Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, 1068 Xueyuan Boulevard, Shenzhen, 518055, China. Electronic mail: gpengjack@gmail.com

However, amusia is not synonymous with pitch difficulty, and amusics' performance does not always fit the concept of the domain-general pitch deficit. For example, by tapping into the phonological processing skills of English-speaking subjects, Jones *et al.* (2009a) found that those with amusia scored poorer than controls in various tasks such as syllable segmentation, word and non-word blending, elision, phoneme reversal, etc. Meanwhile, deficits in memory and attention have also been reported in the amusic population (Jones *et al.*, 2009b; Albouy *et al.*, 2013) beyond what could be predicted from a pitch-processing deficit (Ramus and Ahissar, 2012). It seems that, similar to dyslexia and other types of developmental disorders, amusia is likewise a syndromic disorder frequently accompanied by deficiencies of other kinds (Jones *et al.*, 2009a; Jones *et al.*, 2009b).

Of particular relevance to this study is the finding that amusics have deficient phonological awareness (Jones *et al.*, 2009a). Poor phonological awareness is one of the three major dimensions reliably identified in relation to the phonological deficit (e.g., Ramus and Szenkovits, 2008). The other two are impaired lexical access and poor verbal short-term memory (Wagner and Torgesen, 1987; Ramus and Szenkovits, 2008). Different theories have been proposed to specify the nature of the phonological deficit, two of which stand out as a result of their potentially strong explanatory power. One theory suggests that the phonological deficit may derive from degraded (i.e., of low resolution, or fuzzier) phonological representations (e.g., Snowling, 2000; Ramus and Szenkovits, 2008), while the other posits that compromised or inadequate access to relevant phonological representations is the more fundamental cause (Ramus and Szenkovits, 2008; Ramus and Ahissar, 2012). Despite differences in their details, both theories imply that the phonological deficit operates on existing phonological representations and processes (Ramus and Szenkovits, 2008). Thus, in tasks that frequently tap into one's phonological knowledge and representations (Ramus, 2001; Ramus and Szenkovits, 2008), such as native speech perception and production, the negative influence of the phonological deficit may explain amusics' impaired speech processing. In other words, amusic subjects may have greater difficulty decomposing the structure of the target stimuli or extracting sufficiently accurate pitch information from their phonological representations.

B. Relating Mandarin amusics to phonological deficit

The claim that a phonological deficit affects speech perception may be best evidenced by the observation that segmental content has a significant bearing on Mandarin amusics' lexical tone discrimination and identification performance. According to Nan *et al.* (2010), Mandarin speakers with amusia were impaired in detecting native tonal contrasts carried by different syllables, but unimpaired when similar tonal contrasts were carried by the same syllable within a trial. They also noticed that Mandarin tones in disyllabic words were recognized with significantly lower accuracy by amusics than tones in monosyllabic words.

How can the above mixed results relate to the phonological deficit? One plausible explanation is that in the above case, the variation of phonetic segments generated a greater demand for phonological processing, as the subjects had to filter out irrelevant segmental variations in pitch discrimination and identification tasks (Nan *et al.*, 2010; Liu *et al.*, 2012). Therefore, affected by their phonological deficit, or more precisely, their impaired phonological awareness (Liu *et al.*, 2012), Mandarin amusics demonstrated significant decrements in performance compared to matched controls. On the other hand, the processing of pitch information may have interference from irrelevant variations of segments (e.g., Lee and Nusbaum, 1993; Tong *et al.*, 2008). Both explanations might provide a more plausible explanation that Mandarin amusics lack adequate phonological awareness, viz., the conscious access of and attention to the sound structures of their native language (Wagner and Torgesen, 1987; Ramus and Szenkovits, 2008).

Liu *et al.* (2012), on the other hand, controlled for irrelevant segmental variations in the lexical tone discrimination task. They also employed pitch excursions that were much smaller in size (1.5–4.7 semitones) than those used by Nan *et al.* (2010) (2.0–16.9 semitones). The speech-processing deficit of Mandarin amusics was reliably observed in their study, even when variation of segments was controlled. However, whether these findings are robust enough to imply that the nature of amusics' speech-processing deficit is pitch-based merits further investigation. In particular, these results may be equally interpretable in terms of the phonological deficit when it comes to amusics who speak a tonal language.

Moreover, the phonological system of tonal languages has an extra layer of complexity. Both segmental and tonal information are linguistically significant in tonal languages, such that both kinds of information are represented in tonal language speakers' mental lexicon (e.g., Gandour, 1983; Lee and Nusbaum, 1993; Ramus, 2001; Peng, 2006). It is thus a feasible hypothesis that degraded tonal representations or impaired access to these phonological representations is another fundamental cause for amusics' pitch-processing difficulties in the linguistic domain, in addition to the widely accepted pitch deficit.

Hence, for both the segmental and the tonal levels, the potential impact of the phonological deficit has called into question the explanatory power of the domain-general pitch deficit proposed in the literature. It has also presented a challenge for specifying the nature of the deficit underlying amusics' speech-processing impairment.

C. The relation between speech input and output

To specify the diagnostic patterns and ultimately the locus of the phonological deficit has long been a major challenge because of the hierarchical structures of the phonological representations and the varying forms of the phonological deficit (e.g., Wagner and Torgesen, 1987; Ramus and Szenkovits, 2008; Ramus and Ahissar, 2012). Questions about how phonological deficit operates in amusics—whether a deficit of this sort implicates both input and output pathways—can be addressed by studying the relation between the speech perception (input) and production (output) performance of affected individuals.

Like speech perception, speech production involves the retrieval of phonological representations (Baddeley *et al.*, 1984; Ramus and Szenkovits, 2008). It is often the case that individuals' perception and production abilities are well correlated (Loui *et al.*, 2008; Loui *et al.*, 2011; Hutchins and Peretz, 2012). But a mismatch can happen in individuals with amusia, where intact production performance can be demonstrated with impaired perception (e.g., Loui *et al.*, 2008; Nan *et al.*, 2010). Studies have frequently reported that amusics have the ability to produce what they cannot consciously perceive (Hutchins and Peretz, 2012), notably the falling/rising patterns of tonal intervals (Loui *et al.*, 2008), the intonational contours (Liu *et al.*, 2010; Hutchins and Peretz, 2012), and the F0 movement of lexical tones (Nan *et al.*, 2010). An issue that remains unclear is whether the ability to produce pitch height differences is likewise preserved in amusics.

In Liu *et al.* (2013), Mandarin amusics were tested on both speech and music imitation tasks. Using phonologically discrete Mandarin tones (high, low, and neutral tone), Liu *et al.* (2013) showed that the pitch matching abilities of Mandarin amusics were conspicuously compromised compared with controls'. At first glance, these results are seemingly irreconcilable with those obtained by Hutchins and Peretz (2012). According to Hutchins and Peretz (2012), French-speaking amusics had intact pitch height as well as pitch interval imitation abilities. But such conflicting findings may be explained by the fact that variations in the pitch height dimension have differential phonological status in Mandarin and French, so that different processing mechanisms might have been used in speech imitation. In other words, the pitch heights of the tonal stimuli were phonologically distinct and lexically highly significant in Liu *et al.* (2013); whereas in Hutchins and Peretz (2012), shifting the pitch height of the test syllable elicited no reliable changes in sentential meaning. Therefore, it could be that the pitch information of the target stimuli was processed phonologically by Mandarin speakers but phonetically by French subjects. Under the influence of the phonological deficit, it was thus in Mandarin amusics that impaired speech imitation performance was observed. But questions remain about whether this explanation holds and whether or not Mandarin amusics would demonstrate intact speech production abilities when changes in pitch height are not associated with differences in lexical meaning.

D. Research aims and questions

The purpose of the current study is twofold: (a) to provide insights into the nature of the pitch-processing difficulties that Mandarin amusics demonstrate in the linguistic domain: Whether they are primarily caused by a domain-general pitch deficit as suggested by previous literature, or whether they can be attributed to a phonological deficit; (b) to shed light on the mechanisms of the phonological deficit in Mandarin amusics: whether it operates by means of compromising the processing of both segmental and tonal representations, and whether it operates by impairing just the input pathways, or both the input and the output pathways. Specifically, the current study explores Mandarin amusics' speech-processing deficits from the following perspectives:

First, given that amusia is very likely a syndromic disorder (Jones *et al.*, 2009a; Jones *et al.*, 2009b), it is necessary to distinguish the consequences of the phonological deficit from those of the pitch-processing deficit. Therefore, in contrast to previous studies where Mandarin amusics were presented with only native speech materials (e.g., Nan *et al.*, 2010; Liu *et al.*, 2010; Liu *et al.*, 2012), language familiarity was a major variable being manipulated in the current study. The hypothesis was that in tasks tapping the phonological processing abilities of subjects, comparatively deficient speech discrimination and/or identification performance could arise from either the lack or the impairment of relevant phonological representations (Perrachione *et al.*, 2011; Kuhl, 2011). In other words, relative to controls, Mandarin amusics were expected to demonstrate impaired performance when processing speech stimuli that were native-like and thus perceptually highly familiar, but normal performance when attending to stimuli that were foreign and unfamiliar. However, if the speech-processing impairment demonstrated by amusics persists independently of speech familiarity (i.e., unconstrained by their phonological knowledge or the availability of linguistic representations), the likelihood is that the nature of the deficit is predominantly pitch-based.

Second, to explore the mechanisms of the phonological deficit in Mandarin speaking amusics, the familiarity of the speech materials used in this study was designed to differ at either the segmental or the tonal level (in fully crossed manner). To that end, Cantonese was used as the testing material as it has richer inventories of phonological segments and tones, vowels and level tones in particular, than Mandarin Chinese (e.g., Gandour, 1983; Lee *et al.*, 2002; Peng, 2006).

Third, to ascertain the influence of the phonological deficit that operates on the access to or the processing of relevant phonological representations, the way tonal information is encoded by Mandarin subjects during a speech discrimination task (phonetically or phonologically) needs to be determined first. Thus, in addition to the use of compressed tonal contrasts to reveal the speech-processing impairment of Mandarin amusics, the current study extended the design of Liu *et al.* (2012) by using more fine-grained tonal differences formed by the three Cantonese level tones and their shifted level tone counterparts. A perceptual assimilation effect may occur due to subjects' existing knowledge about the Mandarin tonal system (e.g., So and Best, 2010). Likely, Cantonese high-level tone and its shifted counterparts would be encoded phonologically, mapping almost perfectly onto the Mandarin high-level tone, whereas Cantonese low- and mid-level tones and their shifted counterparts would be encoded phonetically, leading to distinct discrimination performance across the three level tone categories.

Fourth, amusics' ability to detect and exploit regularities of speech sounds is also of interest to this study. The phonological level of representation also pertains to linguistic regularities (Ramus, 2001), and the remarkable ability to capitalize on the statistical patterns of the auditory signal is operative even in the infant stage (Kuhl, 2011). The working memory model also presumes that language learning operates by abstracting and registering the patterns of speech sounds which are more reliable and constant, rather than by storing every sound distinction perceivable to the human ears (Baddeley and Hitch, 1994; Baddeley *et al.*, 1998). Based on these views, the inability to grasp the linguistic regularities may prevent amusics from establishing sufficiently detailed representations of native speech sounds (Ramus, 2001), thereby accounting for their explicit deficits in phonological processing. Jiang *et al.* (2010) observed that the regular occurrence of reference notes exerted no reliable influence on amusics' discrimination performance in a pure tone setting, but it remains unclear whether the null effect of the reference tone also holds in a speech setting. To address this question, context was provided in this study in the form of the high-rising reference tone, from which a speakers' pitch range could be extracted to aid the estimation of level tone heights (Moore and Jongman, 1997; Zhang *et al.*, 2012). Mandarin amusics were expected to show poorer performance than controls in detecting and exploiting the regular occurrence of the reference tone.

Finally, the intactness of Mandarin amusics' tone height imitation abilities remains controversial. Therefore, subjects were also tested for their abilities to imitate Cantonese level tones. Taken together with their discrimination performance, a preliminary picture of how the phonological deficit operates in Mandarin amusics may be revealed in this study.

II. METHOD

A. Participants

Twelve native Mandarin speakers with amusia and eighteen normal controls were recruited for this study. All were born and raised in non-Cantonese-speaking regions of China and spoke Mandarin as their native language. Despite that they were working at the Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, or studying at the Chinese University of Hong Kong at the time of the experiment, these Mandarin subjects could not yet understand or converse in basic Cantonese. None of the subjects reported any histories of hearing disorders or cognitive injuries. To identify Mandarin amusics, the online Montreal Battery of Evaluation of Amusia (MBEA hereafter; Peretz *et al.*, 2008) was used in the screening stage, with the cutoff score set to 71% (Nan *et al.*, 2010). Subjects' musical training backgrounds were also controlled in accordance with previous amusia studies. A brief summary of the demographic characteristics of these subjects is given in Table I.

TABLE I. Demographic characteristics of Mandarin subjects. SD = standard deviation. N.S. = non-significant. *P* value was obtained from two-tailed independent *t*-tests.

	Amusic (N = 12)	SD	Control (N = 18)	SD	*t*-test
Male/Female	5/7	N/A	10/8	N/A	N/A
Mean age (range)	23.08 (18–26)	2.39	22.50 (18–27)	2.90	N.S.
Mean scores of MBEA	64.17%	4.63	88.17%	3.24	$P < 0.001$

B. Stimuli

To assess the tonal discrimination and production abilities of Mandarin amusics and controls, naturally produced speech samples of Cantonese were used in this study. All were simple (consonant−vowel combinations), isolated, and monosyllabic tonal syllables selected from the CU Corpora (Lee *et al.*, 2002). They were read and recorded by two native Cantonese speakers, one female and one male, in a natural and coherent way.

To examine the effects of familiarity on pitch discrimination, the three naturally produced Cantonese level tones that were acoustically either highly similar to or distinct from their Mandarin counterparts were selected as basic testing materials. Cantonese high-level tone (T1) was considered as familiar due to its acoustic similarity to Mandarin high-level tone, while Cantonese mid- (T3) and low-level tones (T6) were presumed to be unfamiliar for they have no direct phonological counterparts in Mandarin (e.g., Peng, 2006; So and Best, 2010). On average, T1 syllables were 3.7 semitones higher than T3 syllables whose mean pitch height was in turn 2.2 semitones above that of T6 syllables. Following a similar logic, Cantonese syllables /fu/, /ji/ were presumed to be familiar segmentally to Mandarin speakers, whereas syllables /si/ and /se/ were considered as unfamiliar (note that /si/ is different from /sɿ/). No significant difference was found between the heights of the level tones carried by familiar and unfamiliar syllables (one-tailed Mann-Whitney *U* non-parametric tests, all $p > 0.05$).

Moreover, shifted level tone counterparts whose F0 trajectories were either one-semitone higher or lower than each of the three Cantonese level tones were generated using Praat (Boersma and Weenink, 2011). Based on previous observations, the tonal contrasts used in this study (1–5.2 semitones) were overall larger than amusics' thresholds for a simple detection of pitch change (Tillmann *et al.*, 2011b; Liu *et al.*, 2012).

Additionally, in order to evaluate amusics' ability to extract linguistic regularities in speech, contextual cues were incorporated into the current experiment in the form of the high-rising reference tone. For example, a naturally produced high-rising tonal syllable (e.g., /fu2/) would be presented as the context immediately before the target level tone syllables (e.g., /ji1/) in the context-present condition. Whereas in the context-absent condition, only the target level tone syllables were presented. Altogether 72 target stimuli [3 tones $\times$ 3 conditions of F0 trajectories (unshifted, one semitone up, and one semitone down) $\times$ 4 syllables $\times$ 2 context conditions] were generated for the discrimination task. Afterward, stimuli used for production assessment were then generated using similar procedures as described above. Considering that differences in stimuli characteristics may be a confounding factor (Liu *et al.*, 2012), duration and intensity normalization were performed on all tonal syllables used in the experiment.

C. Procedures

The procedures of this study were approved by the Survey and Behavioral Research Ethics Committee of the Chinese University of Hong Kong. Written informed consent was provided by subjects before the experiment. Both level tone discrimination and level tone production tasks were conducted in a sound-attenuated room and controlled through DMDX (version 4.1.2.0, the Monash University and the University of Arizona). This program ensured that the stimuli were binaurally presented over the headphones, and that the presentations of the trials were randomized across subjects. Additional practice items with feedback on the screen were provided prior to individual tasks. Moreover, subjects were explicitly informed that a language other than Mandarin would be presented. And throughout the experiment, the male and female subjects listened respectively to speakers of the same gender, so that the F0 range of the tokens heard and produced by individual participants was more comparable.

The discrimination task followed the two-alternative forced-choice paradigm. Subjects were required to judge whether the two stimuli presented in a single trial had the same tone or not. In the current experiment, stimuli differing in context and familiarity conditions were presented in the same block and occurred with equal probability. It is worth noting that with conditions of segmental familiarity and context altering on a trial-to-trial basis, tonal contrasts served as the primary cue for discrimination within each trial. Subjects were instructed to respond by pressing buttons on the computer keyboard labeled with a tick (✓) which indicated same tone pairs and a cross (×) which indicated different tone pairs as quickly and accurately as possible. The inter-stimulus-interval was set to 300 ms, and the maximum response time to 3000 ms.

Consistent with previous amusia studies, two kinds of trials were presented in the discrimination task: same-token trials and different-token trials. In the same-token trials, level tone stimuli were paired with themselves (e.g., /si1/-/si1/), and subjects were instructed to make "yes" responses by pressing the "tick" button. In the different-token trials, stimuli carrying level tones that were naturally produced by Cantonese speakers were paired either with each other in counterbalanced order (e.g., /se1/-/se3/), or with their shifted counterparts to form fine-grained tonal distance. "No" response was defined as the correct answer for different-token trials.

As for the production task, subjects were instructed to imitate the speech sounds they heard as accurately and closely as possible, including the tone and the duration of the stimuli. Level tone stimuli used in this section were presented one at a time over the headphones, and subjects' vocal responses were recorded for further analyses. Questionnaires were distributed afterward to collect feedback. Subjects were instructed to listen to each of the recordings used in the experiment and to report, intuitively but voluntarily, if the stimuli reminded them of any words or characters in Mandarin Chinese.

D. Data extraction and statistical analyses

Analyses were separately carried out for trials with fine-grained and cross-category tonal contrasts. Following previous studies (Ayotte *et al.*, 2002; Peretz *et al.*, 2008; Nan *et al.*, 2010; Tillmann *et al.*, 2011a; Tillmann *et al.*, 2011b; Hutchins and Peretz, 2012), subjects' discrimination accuracy was measured by hits (number of correct responses for different-token trials/number of different-token trials) minus false alarms (FAs; number of incorrect responses for same-token trials / number of same-token trials).

For the production analyses, normalization was first carried out for all level tone imitations. Using a custom written Praat script (Zhang *et al.*, 2012), F0 information was extracted from 21 sampling points, positioned at equal distance along the F0 trajectories of the vowel segments. The first and last 10% of the sampling points were discarded to reduce F0 perturbation; so were the recorded imitations with stuttering and wrongly pronounced vowel segments (2%). Considering the intrinsic pitch of vowels, level tone intervals (T1–T3, T1–T6, and T3–T6) were further calculated between imitations sharing the same syllables. Two scales were used for statistical analyses: semitone and log z-score, to which raw F0 data (Hertz) were converted respectively for each individual subject. Semitone scale was mainly used to explore whether amusics had a greater tendency to compress interval sizes between two level tones (Liu *et al.*, 2013), while log z-score was used to examine the distinctness of level tone productions after controlling for individual variability in speaking F0 and cases where subjects systematically adjust the heights of the target stimuli to their own comfortable registers. Log z-score was also more informative in the sense that it reflected the relative distributions of the three level tones instead of the absolute distances between them. If Mandarin amusics were accurate in their tonal production of Cantonese, then the sizes and the distributions of the intervals thus obtained should approximately equal to those produced by control subjects, regardless of the methods used.

Both perception and production data were analyzed using the repeated-measures ANOVAs with Bonferroni's adjustments at the significance level of 0.05. Greenhouse-Geisser correction was applied to correct for violations of sphericity. In *post hoc* analyses, Welch's and Games-Howell's correction were adopted as appropriate. Given that the behavioral data of amusics and controls did not always follow normal distributions, Mann-Whitney *U* non-parametric tests (one-tailed unless otherwise specified) were used for confirmation of the statistical results. Where correlation analyses were appropriate, non-parametric Spearman's test (two-sided) was also applied to examine the relations between variables.

III. RESULTS

A. Post-experimental questionnaires

In the questionnaires that instructed subjects to report voluntarily if the stimuli reminded them of any characters in Mandarin Chinese, there was a clear-cut distinction between level tones carried by familiar and unfamiliar syllables. Namely, subjects generally found it easy to associate Chinese characters with stimuli consisted of familiar syllables (e.g., /pa1/ "八" "eight", or /fu1/ "夫", / "husband"), but hard to recollect familiar words for level tone stimuli carried by unfamiliar segments. But, different from expectations, neither amusics nor controls seemed to differentiate between level tones, and most of the level tone stimuli were identified as characters having a high-level tone, rather than the other three tonal categories (rising, dipping/low, and falling) in the Mandarin inventory.

B. Discrimination of fine-grained tonal contrasts

As described in the preceding text, the aim of using fine-grained tonal discriminations (i.e., one-semitone) is to examine the effects of familiarity at the tonal level. A four-way repeated-measures analysis of variance (ANOVA) with *Group* (amusic, control) as the between-subjects variable, while *Segmental familiarity* (familiar, unfamiliar), *Context* (present, absent), and *Tonal Category* (T1-centered, T3-centered, T6-centered) as the within-subjects variables found significant main effects of *Group* [$F(1, 28) = 5.52$, $p < 0.05$], and of *Context* [$F(1, 28) = 16.91$, $p < 0.001$]. Meanwhile, significant two-way interactions were found for *Segmental familiarity* by *Group* [$F(1, 28) = 5.51$, $p < 0.05$], and *Context* by *Group* [$F(1, 28) = 4.63$, $p < 0.05$]. No statistical significance was found for *Tonal Category* ($p = 0.14$); its interactions with other variables, such as *Group* ($p = 0.45$), also fell short of significance. The null effect of *Tonal Category* was confirmed by non-parametric tests performed on all tone pair conditions (all $p > 0.05$).

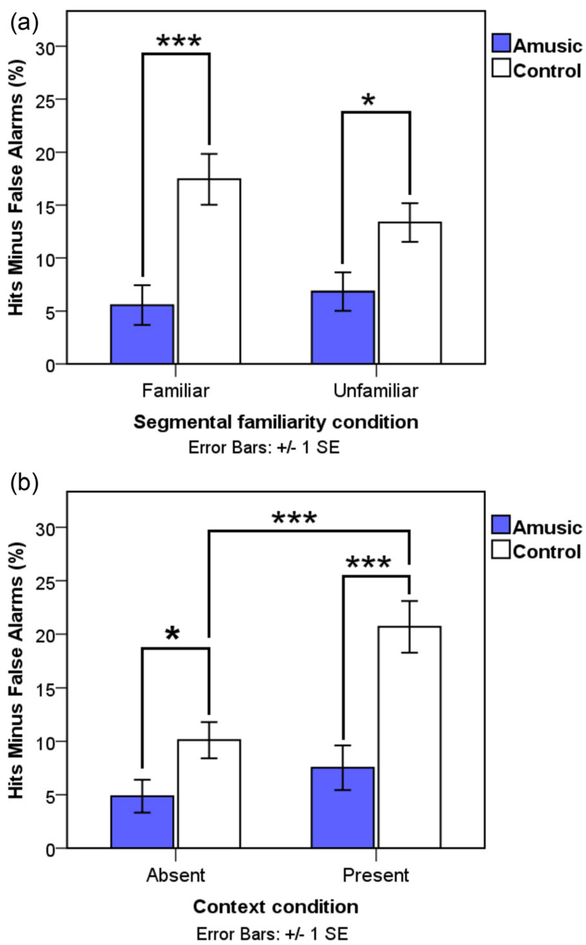

FIG. 1. (Color online) (a) Scores of hits minus false alarms obtained from fine-grained tonal discrimination trials. Results were obtained by averaging across conditions of context and tonal categories, and organized by segmental familiarity. ***$p < 0.001$, *$p < 0.05$. (b) Scores of hits minus false alarms obtained from fine-grained tonal discrimination trials. Results were obtained by averaging across conditions of segmental familiarity and tonal categories, and organized by context conditions. ***$p < 0.001$, *$p < 0.05$.

Figure 1 displays the mean scores of hits-FAs as a function of segmental familiarity [Fig. 1(a)] and context conditions [Fig. 1(b)]. As is shown in Fig. 1(a), amusics' tonal discriminations were slightly worse in the familiar condition than in the unfamiliar condition, and the reverse seemed to be true for controls. But statistically, neither of these differences reached significance (amusics, $p = 0.63$; controls, $p = 0.18$). Nonetheless, *post hoc* analyses revealed that group differences persisted across familiarity conditions—controls were significantly more accurate than amusics in fine-grained tonal discriminations, although more so in the familiar condition [$F(1, 177.57) = 15.22$, $p < 0.001$] compared to the unfamiliar condition [$F(1, 171.45) = 6.43$, $p < 0.05$].

To disentangle the *Context* by *Group* interaction, *post hoc* analyses were likewise carried out. *Group* effect remained statistically significant across context conditions, but more so in the context-present condition [$F(1, 177.41) = 16.99$, $p < 0.001$] than in the context-absent condition [$F(1, 175.95) = 5.28$, $p < 0.05$]. Analyses also showed that while the presence of the context led to a significant increment (10.57%) in controls' accuracy relative to the no context condition [$F(1, 191.59) = 12.85$, $p < 0.001$], such an effect fell short of significance in the amusic group ($p = 0.31$). Thus, the main effect of *Context* observed above was driven mostly by control subjects, not by amusics.

To sum up, controls were better overall than amusics in discriminating fine-grained tonal contrasts. The reference tone better facilitated performance in controls than in amusics. But contrary to the predictions, *Tonal category* showed no significant effect. This result was more consistent with the feedback collected from the post-experimental questionnaires, suggesting that the influence of Mandarin tonal system, namely familiarity at the tonal level, played no significant role during fine-grained tonal discriminations.

C. Discrimination of cross-category tonal contrasts

Since the differences between the three level tones were much larger than the discrimination thresholds of amusics and controls, almost all subjects obtained above-chance-level hit rates. The FAs were very low for amusics as they were for controls ($<3\%$), and no significant group difference was observed in this respect (all $p > 0.5$ according to non-parametric tests). Considering that in the above analyses of fine-grained pitch differences, subjects' performance was not significantly biased by their existing knowledge about the Mandarin tonal system, the factor of *Tonal Category* was disregarded in the following analyses.

A three-way repeated-measures ANOVA was first employed to examine the effect of *Group*, *Segmental familiarity*, and *Context* on subjects' cross-category discrimination performance. The main effect of *Group* was significant [$F(1, 28) = 6.65$, $p < 0.05$], as was the interaction of *Group* by *Segmental familiarity* [$F(1, 28) = 4.75$, $p < 0.05$]. Figure 2 displays the discrimination performance of amusics and controls as a function of segmental familiarity. There was a

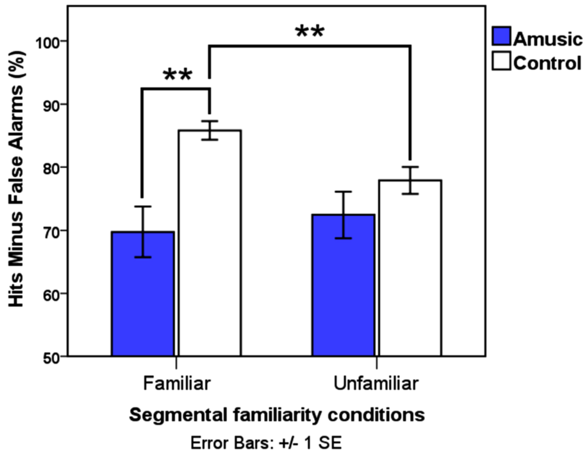

FIG. 2. (Color online) Hits minus false alarms in cross-category tonal discrimination trials. Results were obtained by averaging across context conditions. $^{**}p < 0.01$.

significant familiarity advantage in the control group [$F(1, 62.11) = 9.37$, $p < 0.01$]—the performance controls obtained in the familiar condition was much superior to that obtained in the unfamiliar condition, whereas amusics exhibited approximately equal performance across familiarity conditions ($p = 0.62$).

More important (see Fig. 2), *post hoc* analyses revealed that whereas amusics' performance was evidently poorer than controls' in the familiar condition [$F(1, 29.30) = 14.25$, $p < 0.01$], their performance was no more impaired than controls' in the unfamiliar condition ($p = 0.21$). Similar results were yielded by non-parametric tests performed on all measures. It was also in the familiar condition that a significant positive correlation was found between subjects' musical abilities (represented by MBEA scores) and their discrimination accuracy [$r(30) = 0.59$, $p < 0.01$]; in the unfamiliar condition, no reliable correlation was observed between the two variables ($p = 0.36$).

Table II displays in more detail the performance of individual amusic subjects. Consistent with previous observations (e.g., Jones *et al.*, 2009b; Nan *et al.*, 2010; Liu *et al.*, 2012), not all amusics were impaired to a significant degree in tonal discrimination tests and their performance showed a large degree of variability. But as many as 50% of amusics performed below the normal range in the familiar condition—they scored two standard deviations (2SDs) below the mean scores of matched controls (controls' mean in the familiar condition = 85.83%, SD = 7.15; controls' mean in the unfamiliar condition 77.89%, SD = 11.96). In contrast, only one amusic subject showed similar signs of impairment in the unfamiliar condition. Therefore, the main effect of *Group* observed in the three-way analysis was mainly attributed to the familiar condition, and was driven mostly by the five amusic subjects who showed selective impairment in the familiar condition as opposed to the unfamiliar condition.

Thus, in cross-category tonal discriminations, amusics exhibited selective impairment in discriminating tonal contrasts carried by familiar segments. This finding is inconsistent with the hypothesis of a domain-general pitch deficit, but is more compatible with the idea of a phonological deficit which operates on language-specific phonological representations.

D. Level tone productions

Production analysis was first performed using a four-way repeated-measures ANOVA, with *Group* (amusic, control) defined as the between-subjects variable, while *Segmental familiarity* (familiar, unfamiliar), *Context* (present, absent), and *Interval category* (T1–T3, T1–T6, T3–T6) as the within-subjects variables. Significant main effects of *Interval category* [$F(1.50, 42.04) = 97.25$, $p < 0.001$] and *Segmental familiarity* were found [$F(1, 28) = 10.86$, $p < 0.01$], as was the interaction between them [$F(1.35, 37.83) = 58.04$, $p < 0.001$]. These were expected results as the mean heights of the level tone stimuli used for production assessment differed between the two familiarity conditions. Non-parametric test showed that the mean size of the

TABLE II. Hits-FAs obtained by amusic subjects in cross-category speech discrimination trials. Results were calculated by averaging across context conditions. Scores 2SDs below controls' means are marked out by ↓.

			Familiar segment condition		Unfamiliar segment condition	
Subject	Age	MBEA score (%)	Hits-FAs (%)	2SDs ↓	Hits-FAs (%)	2SDs ↓
1	25	62	73.48		63.33	
2	24	65	43.33	↓	77.78	
3	18	65	78.41		79.17	
4	26	69	62.50	↓	62.50	
5	24	63	66.67	↓	66.67	
6	22	71	91.67		62.50	
7	24	62	82.95		83.33	
8	24	58	37.50	↓	45.83	↓
9	25	64	58.33	↓	62.50	
10	23	67	63.33	↓	69.70	
11	19	55	87.50		95.83	
12	23	69	91.29		100.00	
Mean	23.08	64.17	69.75		72.43	
SD	2.39	4.63	17.82		15.45	

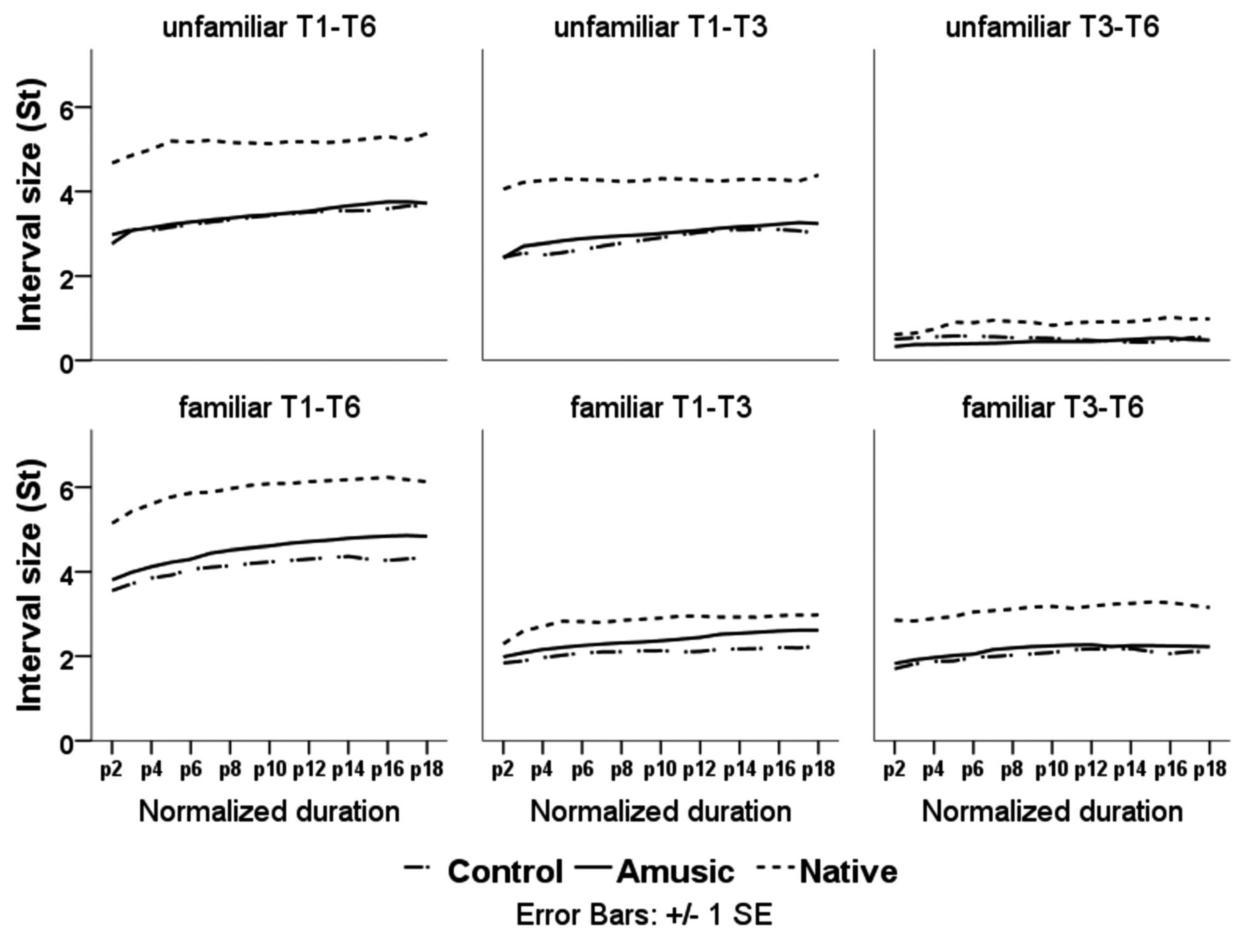

FIG. 3. Level tone intervals (in semitones) produced by Mandarin amusics and controls. Cantonese speakers' F0 trajectories are provided as a reference. Familiar = familiar segmental condition, unfamiliar = unfamiliar segmental condition, T1–T6 = intervals obtained between T1 and T6 imitations, T1–T3 = intervals obtained between T1 and T3 imitations, T3–T6 = intervals obtained between T3 and T6 imitations, p2, p4, p6… = the second, fourth, sixth…sampling points.

T3–T6 intervals was significantly larger in the familiar condition than in the unfamiliar condition, but not for other level tone intervals. No other results, including the four-way interaction between these variables, approached significance.

Subjects' level tone productions are demonstrated in Fig. 3 divided by segmental familiarity and interval categories (with Cantonese speakers' tonal trajectories being included as a reference). Both amusics and controls seemed to compress interval sizes during production. But, there was neither a main effect of *Group* ($p = 0.63$) nor significant interactions between *Group* and other variables. Subsequent analyses using the log z-score method replicated these results, implying that the three level tones produced by Mandarin amusic and control subjects were distributed in similar manners and were approximately equally distinct from one another. The null effect of *Group* was confirmed by non-parametric tests performed on all measures. It seemed that the ability of Mandarin amusics to imitate the heights of the level tones was somehow preserved.

IV. DISCUSSION

Current results suggest that both a phonological deficit and a pitch-processing deficit can explain why Mandarin amusics show impaired speech processing. However, which of them plays the more predominant role in the task at issue depends further on the demand for pitch sensitivity. When the demand was relatively low (as in the case of cross-category discriminations), the impact of the phonological deficit turned out to be more consequential than that of the pitch-processing deficit. But when faced with fine-grained trials, it was the domain-general deficit in pitch-processing that was mainly responsible for amusics' impaired speech discrimination performance. Sections IV A–IV C below provide further discussions about these points. Another major finding of this study is that Mandarin amusics were not impaired compared to controls at producing level tone stimuli, with respect to both imitation accuracy and the distinctness of level tone productions. Questions about how this relates to their perception performance and the phonological deficit are discussed in greater detail in Sec. IV D.

A. The phonological deficit—evidence from cross-category discriminations

Similar to Liu *et al.* (2012), this study demonstrates that when attending to cross-category tonal differences that are above their pitch thresholds by a moderate amount, Mandarin amusics can have manifest difficulties in pitch discrimination compared with controls. But by further manipulating the familiarity of the speech materials, this study shows that amusics may have additional deficits in phonological processing. Moreover, just as what Jones *et al.* (2009a, 2009b) observed in English-speaking amusics, current observations suggest that the phonological deficit operates in Mandarin amusics by means of impacting the access to or the resolution of segmental level representations.

It was hypothesized at the beginning that with other conditions being equal, amusics' compromised speech performance would be constrained by language familiarity if they were influenced by a deficit that is phonological in nature. As shown in Fig. 2, this was indeed the case: The task requirements were the same across the two familiarity conditions, and there was no significant difference between the sizes of the tonal contrasts being presented, and yet the effect of *Group* only surfaced in the familiar condition. These results are clearly at variance with the predictions of the domain-general pitch deficit. It appears that Mandarin amusics, compared to non-amusic controls, lack sufficient phonological resources to efficiently decompose the structures of speech sounds. Consequently, they are unable accurately to extract and compute the pitch differences between native-like speech stimuli.

Evidence bearing on the phonological deficit also comes from amusics' lacking of familiarity advantage (see Fig. 2) and the high prevalence rate of amusics' speech-processing impairment in the familiar condition in comparison to the unfamiliar condition (see Table II). It seems that similar to the pitch-processing deficit and other widespread auditory and attentional deficits already reported in amusics (e.g., Jones *et al.*, 2009a; Jones *et al.*, 2009b; Nan *et al.*, 2010; Liu *et al.*, 2012), the phonological deficit observed in this study and by Jones *et al.* (2009a) also represents a continuum, ranging from severe to more subtle and ambiguous conditions.

Furthermore, results of this study have demonstrated that the processing of segmental and suprasegmental information are intertwined or integral during a speech perception task (Lee and Nusbaum, 1993; Tong *et al.*, 2008; Perrachione *et al.*, 2011), even though segmental information was irrelevant to the task at issue and occurred merely as between-trial variability as opposed to within-trial conflict in Nan *et al.* (2010).

B. Core deficit in fine-grained pitch discrimination

The situation between Mandarin amusics and controls, however, is slightly different when they were faced with more fine-grained pitch differences as opposed to cross-category tonal contrasts. It turned out that under such circumstances, controls could reliably outperform amusics by a significant margin, regardless of whether or not the segments are familiar. Therefore, it seems that when there is a higher demand for pitch sensitivity, amusics' pitch-processing deficiency is the predominant factor that modulates their performance. As such, results of this study support that a domain-general pitch-processing deficit is characteristic of individuals with amusia (e.g., Jiang *et al.*, 2010; Tillmann *et al.*, 2011b), but it likely operates on more fine-grained pitch differences than suggested by previous studies on Mandarin amusics (Nan *et al.*, 2010; Liu *et al.*, 2012).

C. The effect of the context on pitch processing

Context effect surfaced in the current study during fine-grained tonal discriminations, differing from what Jiang *et al.* (2010) observed in pure tone settings. Specifically, control subjects performed much better in the context-present condition compared to the context-absent condition. For control subjects, resorting to available cues could be a way of optimizing their discrimination performance. However, similar to Jiang *et al.* (2010), the current study found no indications that contextual cues significantly benefited subjects with amusia. Given that the Cantonese high-rising tone has a large pitch excursion size and bears a strong resemblance to the corresponding category in Mandarin (Peng, 2006; So and Best, 2010), the null effect of the context on Mandarin amusics' speech discriminations is unlikely a result of amusics' inability to perceive pitch cues per se. Rather, as hypothesized, this may be a reflection that amusics are less able to grasp regularities in sound patterns, in which case their representations of the speech sounds may be less detailed than controls'. But due to the null effect of *Tonal category* during fine-grained pitch discriminations, the current study provided no evidence for the phonological encoding of pitch information among Mandarin subjects, thus failing to provide more concrete evidence that Mandarin amusics may have degraded tonal representations that could account for their difficulties with linguistic pitch processing. Future studies may investigate this possibility by manipulating the familiarity of the contour information, which is the more reliable dimension by which Mandarin tones are identified (e.g., Gandour, 1983; Lee *et al.*, 2002).

Alternatively though, amusics' inability to use reference cues may be explained by their lack of access to relevant tonal representations—they may be less efficient in retrieving and analyzing additional tonal information within the limited time frame. But whether unable or less efficient in accessing and processing contextual cues, amusics' behavioral manifestations observed in this study fit the predictions of the phonological deficit (e.g., Ramus, 2001; Ramus and Szenkovits, 2008; Ramus and Ahissar, 2012; Perrachione *et al.*, 2011).

D. Preserved speech imitation abilities

Whereas the tendency of Mandarin subjects to compress pitch interval sizes during speech imitation is consistent with the findings of Liu *et al.* (2013), the current study observed no significant difference in performance that could distinguish amusics from controls. This discrepancy can be interpreted as resulting from the phonetic encoding of pitch information elicited in this study as opposed to the phonological encoding strategy elicited in Liu *et al.* (2013). Hence, in line with Hutchins and Peretz (2012), this study shows that subjects with amusia not necessarily have impaired pitch imitation abilities, especially if the pitch differences at issue have no phonological status in their native language system. When taken together with existing findings about amusics' impaired discrimination abilities, it appears to be the case that the input pathways of Mandarin amusics are more severely and frequently implicated in the phonological deficit than their output pathways.

Although the current observations cannot yet reveal the specific mechanisms by which the phonological deficit operates in Mandarin amusics, they nevertheless showed that selective impairment to speech input does not necessarily entail difficulties in speech output. These findings corroborate the proposals made by Loui *et al.* (2008) and Hutchins and Peretz (2012). They both proposed that the neural pathways involved in perceptual and vocal performance are to some extent distinct, such that the pitch information of the auditory stimuli is encoded into different representations: vocal-motor codes and abstract perceptual codes. Given the abundant evidence that auditorily presented items are temporarily stored in the form of articulatory codes (e.g., Baddeley *et al.*, 1984; Baddeley and Hitch, 1994; Baddeley *et al.*, 1998), the above suggestion that distinct representations are used to support speech imitation and perception tasks is highly plausible (Hutchins and Peretz, 2012).

It is worth mentioning that the hypothesized nature of the phonological deficit, namely impaired access to or poorer resolution of the phonological representations (e.g., Snowling, 2000; Ramus and Szenkovits, 2008; Ramus and Ahissar, 2012), does coincide with the neurological findings reported in amusia studies. In particular, they are compatible with the reduced intrinsic connectivity in the auditory cortices, compromised structural integrity of the arcuate fasciculus, and other anomalies in the fronto-temporal pathways that have been found to be featured in the amusic brain (e.g., Albouy *et al.*, 2013; Loui *et al.*, 2009). Considering that developmental disorders with similar behavioral manifestations, such as amusia and dyslexia, may have a shared, although not identical neural basis (Loui *et al.*, 2011), future studies may explore the exact causal role played by the phonological deficit by examining the commonalities of these developmental disorders at the behavioral level, and by comparing their structural and functional similarities at the neurological level.

V. CONCLUSION

By manipulating speech familiarity at the segmental and at the suprasegmental levels in fully crossed manner, this study has demonstrated that language familiarity can modulate amusics' speech discrimination performance, thereby providing more concrete evidence for the presence of the phonological deficit in amusics who speak a tonal language. Current observations also suggest that the constraint of language familiarity on amusics' speech-processing impairment relates further to the specific task design: whether pitch-processing abilities are tapped into or whether the phonological awareness abilities are involved. However, the current study is only a preliminary investigation. Our findings cannot unambiguously lead to the conclusion that a phonological deficit contributes to the perceptual impairment of Mandarin amusics. Future experiments with more rigorous screening procedures and larger sample sizes are warranted to explore the causal role played by the phonological deficit: What constrains it, how it interacts with amusics' pitch-processing deficit, and what its neurological correlates are in the amusic brain.

ACKNOWLEDGMENTS

This study is supported in part by grants from the National Basic Research Program of the Ministry of Science and Technology of China (973 Grant: 2012CB720700), and Key Program of National Natural Science Foundation of China (NSFC: 61135003). We thank Dr. Caicai Zhang, Dr. James Minett, Dr. Tom Schoenemann, and Professor William S-Y. Wang for their valuable comments on this study, and thank Fei Chen for his tremendous help in data collection.

Albouy, P., Mattout, J., Bouet, R., Maby, E., Sanchez, G., Aguera, P. E., Daligault, S., Delpuech, C., Bertrand, O., Caclin, A., and Tillmann, B. (**2013**). "Impaired pitch perception and memory in congenital amusia: The deficit starts in the auditory cortex," Brain **136**, 1639–1661.

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

Baddeley, A., Lewis, V., and Vallar, G. (**1984**). "Exploring the articulatory loop," Q. J. Exp. Psychol. **36**, 233–252.

Baddeley, A. D., Gathercole, S. E., and Papagno, C. (**1998**). "The phonological loop as a language learning device," Psychol. Rev. **105**, 158–173.

Baddeley, A. D., and Hitch, G. J. (**1994**). "Development in the concept of working memory," Neuropsychology **8**, 485–493.

Boersma, P., and Weenink, D. (**2011**). "Praat: Doing phonetics by computer (Version 5.3.34) [Computer program]," http://www.praat.org (Last viewed May 13, 2014).

Brown, D. E. (**1991**). *Human Universals* (Temple University Press, Philadelphia, PA).

Gandour, J. T. (**1983**). "Tone perception in Far Eastern languages," J. Phon. **11**, 149–175.

Hutchins, S., and Peretz, I. (**2012**). "Amusics can imitate what they cannot discriminate," Brain Lang **123**, 234–239.

Jiang, C., Hamm, J. P., Lim, V. K., Kirk, I. J., and Yang, Y. (**2010**). "Fine-grained pitch discrimination in congenital amusics with Mandarin Chinese," Music Percept. **28**, 519–526.

Jones, J., Lucker, J., Zalewski, C., Brewer, C., and Drayna, D. (**2009a**). "Phonological processing in adults with deficits in musical pitch recognition," J. Commun. Disord. **42**, 226–234.

Jones, J., Zalewski, C., Brewer, C., Lucker, J., and Drayna, D. (**2009b**). "Widespread auditory deficits in tune deafness," Ear. Hear. **30**, 63–72.

Kalmus, H., and Fry, D. B. (**1980**). "On tune deafness (dysmelodia): frequency, development, genetics and musical background," Ann. Hum. Genet. **43**, 369–382.

Kuhl, P. K. (**2011**). "Who's talking," Science **333**, 529–530.

Lee, L., and Nusbaum, H. C. (**1993**). "Processing interactions between segmental and suprasegmental information in native speakers of English and Mandarin Chinese," Percept. Psychophys. **53**, 157–165.

Lee, T., Lo, W. K., Ching, P. C., and Meng, H. (**2002**). "Spoken language resources for Cantonese speech processing," Speech Commun. **36**, 327–342.

Liu, F., Jiang, C., Pfordresher, P. Q., Mantell, J. T., Xu, Y., Yang, Y., and Stewart, L. (**2013**). "Individuals with congenital amusia imitate pitches more accurately in singing than in speaking: Implications for music and language processing," Atten. Percept. Psychophys. **75**, 1783–1798.

Liu, F., Jiang, C., Thompson, W. F., Xu, Y., Yang, Y., and Stewart, L. (**2012**). "The mechanism of speech processing in congenital amusia: Evidence from Mandarin speakers," PLoS One **7**, e30374.

Liu, F., Patel, A. D., Fourcin, A., and Stewart, L. (**2010**). "Intonation processing in congenital amusia: Discrimination, identification, and imitation," Brain **133**, 1682–1693.

Loui, P., Alsop, D., and Schlaug, G. (**2009**). "Tone deafness: A new disconnection syndrome?," J. Neurosci. **29**, 10215–10220.

Loui, P., Guenther, F. H., Mathys, C., and Schlaug, G. (**2008**). "Action–perception mismatch in tone-deafness," Curr. Biol. **18**, R331–R332.
Loui, P., Kroog, K., Zuk, J., Winner, E., and Schlaug, G. (**2011**). "Relating pitch awareness to phonemic awareness in children: Implications for tone-deafness and dyslexia," Front. Psychol. **2**, 111.
Mignault-Goulet, G., Moreau, P., Robitaille, N., and Peretz, I. (**2012**). "Congenital amusia persists in the developing brain after daily music listening," PLoS ONE **7**, e36860.
Moore, C. B., and Jongman, A. (**1997**). "Speaker normalization in the perception of Mandarin Chinese tones," J. Acoust. Soc. Am. **102**, 1864–1877.
Nan, Y., Sun, Y., and Peretz, I. (**2010**). "Congenital amusia in speakers of a tone language: Association with lexical tone agnosia," Brain **133**, 2635–2642.
Nguyen, S., Tillmann, B., Gosselin, N., and Peretz, I. (**2009**). "Tonal language processing in congenital amusia," Ann. N.Y. Acad. Sci. **1169**, 490–493.
Patel, A. D., Wong, M., Foxton, J., Lochy, A., and Peretz, I. (**2008**). "Speech intonation perception deficits in musical tone deafness (congenital amusia)," Music Percept. **25**, 357–368.
Peng, G. (**2006**). "Temporal and tonal aspects of Chinese syllables: A corpus-based comparative study of Mandarin and Cantonese," J. Chin. Linguist. **34**, 134–154.
Peretz, I. (**2006**). "The nature of music from a biological perspective," Cognition **100**, 1–32.
Peretz, I., Gosselin, N., Tillmann, B., Cuddy, L. L., Gagnon, B., Trimmer, C. G., Paquette, S., and Bouchard, B. (**2008**). "On-line identification of congenital amusia," Music Percept. **25**, 331–343.
Perrachione, T. K., Del Tufo, S. N., and Gabrieli, J. D. E. (**2011**). "Human voice recognition depends on language ability," Science **333**, 595.
Ramus. F. (**2001**). "Outstanding questions about phonological processing in dyslexia," Dyslexia **7**, 197–216.
Ramus, F., and Ahissar, M. (**2012**). "Developmental dyslexia: The difficulties of interpreting poor performance, and the importance of normal performance," Cogn. Neuropsychol. **29**, 104–122.
Ramus, F., and Szenkovits, G. (**2008**). "What phonological deficit?," Q. J. Exp. Psychol. **61**, 129–141.
Snowling, M. J. (**2000**). *Dyslexia* (Blackwell, Oxford, UK), pp. 1–101.
So, C. K., and Best, C. T. (**2010**). "Cross-language perception of non-native tonal contrasts: Effects of native phonological and phonetic influences," Lang. Speech **53**, 273–291.
Thompson, W. F., Marin, M. M., and Stewart, L. (**2012**). "Reduced sensitivity to emotional prosody in congenital amusia rekindles the musical protolanguage hypothesis," Proc. Natl. Acad. Sci. U.S.A. **109**, 19027–19032.
Tillmann, B., Burnham, D., Nguyen, S., Grimault, N., Gosselin, N., and Peretz, I. (**2011a**). "Congenital amusia (or tone-deafness) interferes with pitch processing in tone languages," Front. Psychol. **2**, 120.
Tillmann, B., Rusconi, E., Traube, C., Butterworth, B., Umiltà, C., and Peretz, I. (**2011b**). "Fine-grained pitch processing of music and speech in congenital amusia," J. Acoust. Soc. Am. **130**, 4089–4096.
Tong, Y., Francis, A. L., and Gandour, J. T. (**2008**). "Processing dependencies between segmental and suprasegmental features in Mandarin Chinese," Lang. Cognit. Process **23**, 689–708.
Wagner, R. K., and Torgesen, J. K. (**1987**). "The nature of phonological processing and its causal role in the acquisition of reading skills," Psychol. Bull. **101**, 192–212.
Wong, P. C. M., Ciocca, V., Chan, A. H. D., Ha, L. Y. Y., Tan, L. H., and Peretz, I. (**2012**). "Effects of culture on musical pitch perception," PLoS ONE **7**, e33424.
Zhang, C-C., Peng, G., and Wang, W. S-Y. (**2012**). "Unequal effects of speech and nonspeech contexts on the perceptual normalization of Cantonese level tones," J. Acoust. Soc. Am. **132**, 1088–1099.